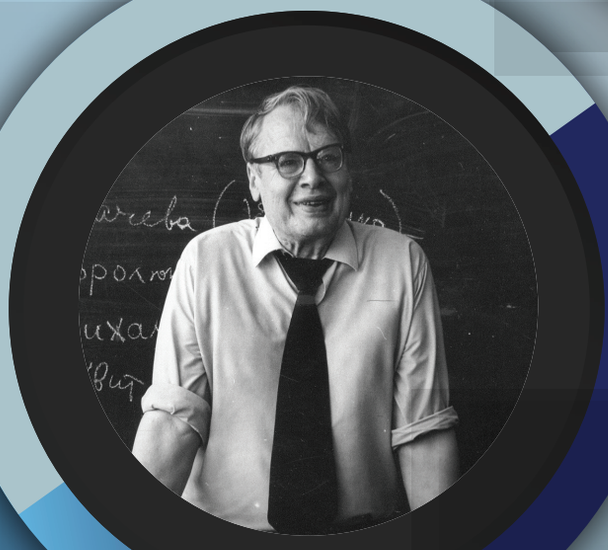
RTA

ISSN 1932-2321

JOURNAL IS REGISTERED IN THE LIBRARY OF THE U.S. CONGRESS

RELIABILITY: THEORY & APPLICATIONS

INTERNATIONAL GROUP ON RELIABILITY



RELIABILITY
RISK ANALYSIS
MAINTENANCE
SAFETY

GNEDENKO FORUM PUBLICATIONS

#2

(84) VOL.20

JUNE

2025

SAN DIEGO



ISSN 1932-2321

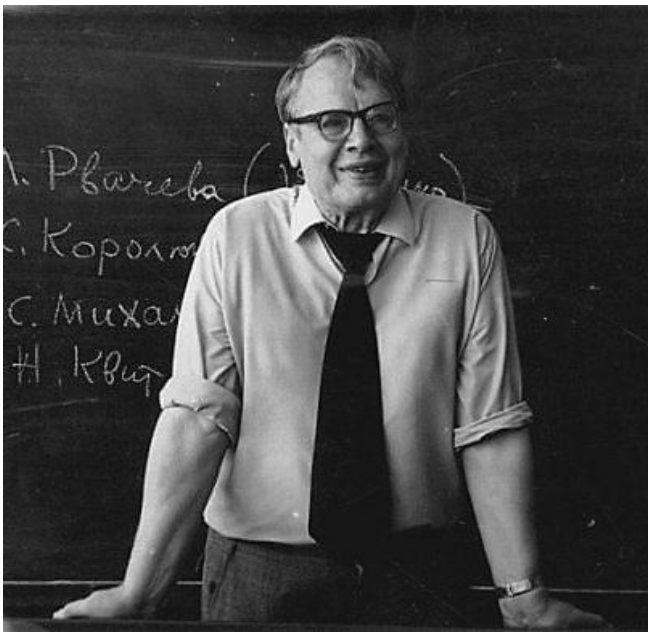
© "Reliability: Theory & Applications", 2006, 2007, 2009-2025
© " Reliability & Risk Analysis: Theory & Applications", 2008

© I.A. Ushakov
© A.V. Bochkov, 2006-2025
© Kristina Ushakov, Cover Design, 2024
http://www.gnedenko.net/Journal/index.htm

All rights are reserved

The reference to the magazine "Reliability: Theory & Applications"
at partial use of materials is obligatory.





RELIABILITY:

THEORY & APPLICATIONS

Vol.20 No.2 (84),
June 2025

San Diego
2025

Editorial Board

Editor-in-Chief

Rykov, Vladimir (Russia)
Doctor of Sci, Professor, Department of Applied Mathematics & Computer Modeling, Gubkin Russian State Oil & Gas University, Leninsky Prospect, 65, 119991 Moscow, Russia.
e-mail: vladimir_rykov@mail.ru

Managing Editors

Bochkov, Alexander (Russia)
Doctor of Technical Sciences, Scientific Secretary JSC NIIAS, Scientific-Research and Design Institute Informatization, Automation and Communication in Railway Transport, Moscow, Russia, 107078, Orlikov pereulok, 5, building 1
e-mail: a.bochkov@gmail.com

Gnedenko, Ekaterina (USA)
PhD, Lecturer Department of Economics Boston University, Boston 02215, USA
e-mail: gnedenko@bu.edu

Bushinskaya, Anna (Russia)
Candidate of Tech. Sci. Leading Research Fellow of the Sci & Engng Center of the Russian Academy of Sciences, Ekaterinburg
e-mail: bushinskaya@gmail.com

Sazonov, Aleksey (Russia)
Leading Specialist of the Standardization Department, JSC NIIAS (Joint Stock Company "Design &Research Institute for Information "Technology, Signaling and Telecommunications on Railway Transport), Bild 1, 5 Orlikov Pereulok, Moscow, Russia, 107078
e-mail: sazonoff2007@gmail.com

Deputy Editors

Dimitrov, Boyan (USA)
Ph.D., Dr. of Math. Sci., Professor of Probability and Statistics, Associate Professor of Mathematics (Probability and Statistics), GMI Engineering and Management Inst. (now Kettering)
e-mail: bdimitro@kettering.edu

Gnedenko, Dmitry (Russia)
Doctor of Sci., Assos. Professor, Department of Probability, Faculty of Mechanics and Mathematics, Moscow State University, Moscow, 119899, Russia
e-mail: dmitry@gnedenko.com

Kashtanov, Victor A. (Russia)
PhD, M. Sc (Physics and Mathematics), Professor of Moscow Institute of Applied Mathematics, National Research University "Higher School of Economics" (Moscow, Russia)
e-mail: VAKashtan@yandex.ru

Krishnamoorthy, Achyutha (India)
M.Sc. (Mathematics), PhD (Probability, Stochastic Processes & Operations Research), Professor Emeritus, Department of Mathematics, Cochin University of Science & Technology, Kochi-682022, INDIA.
e-mail: achyuthacusat@gmail.com

Recchia, Charles H. (USA)
PhD, Senior Member IEEE Chair, Boston IEEE Reliability Chapter A Joint Chapter with New Hampshire and Providence, Advisory Committee, IEEE Reliability Society
e-mail: charles.recchia@macom.com

Shybinsky Igor (Russia)
Doctor of Sci., Professor, Division manager, VNIIAS (Russian Scientific and Research Institute of Informatics, Automatics and Communications), expert of the Scientific Council under Security Council of the Russia
e-mail: igor-shubinsky@yandex.ru

Yastrebenetsky, Mikhail (Ukraine)
Doctor of Sci., Professor. State Scientific and Technical Center for Nuclear and Radiation Safety
e-mail: ma.yastreb2013@gmail.com

Associate Editors

Aliyev, Vugar (Azerbaijan)
Doctor of Sci., Professor, Chief Researcher of the Institute of Physics of the National Academy of Sciences of Azerbaijan, Director of the AMIR Technical Services Company
e-mail: prof.vugar.aliyev@gmail.com

Balakrishnan, Narayanaswamy (Canada)
Professor of Statistics, Department of Mathematics and Statistics, McMaster University
e-mail: bala@mcmaster.ca

Carrión García, Andrés (Spain)
Professor Titular de Universidad, Director of the Center for Quality and Change Management, Universidad Politécnica de Valencia, Spain
e-mail: acarrion@eio.upv.es

Chakravarthy, Srinivas (USA)
Ph.D., Professor of Industrial Engineering & Statistics, Departments of Industrial and Manufacturing Engineering & Mathematics, Kettering University (formerly GMI-EMI) 1700, University Avenue, Flint, MI48504
e-mail: schakrav@kettering.edu

Cui, Lirong (China)
PhD, Professor, School of Management & Economics, Beijing Institute of Technology, Beijing, P. R. China (Zip:100081)
e-mail: lirongcui@bit.edu.cn

Finkelstein, Maxim (SAR)
Doctor of Sci., Distinguished Professor in Statistics/Mathematical Statistics at the UFS. Visiting researcher at Max Planck Institute for Demographic Research, Rostock, Germany and Visiting research professor (from 2014) at the ITMO University, St Petersburg, Russia
e-mail: FinkelM@ufs.ac.za

Kaminsky, Mark (USA)
PhD, principal reliability engineer at the NASA Goddard Space Flight Center
e-mail: mkaminskiy@hotmail.com

Krivtsov, Vasiliy (USA)
PhD. Director of Reliability Analytics at the Ford Motor Company. Associate Professor of Reliability Engineering at the University of Maryland (USA)
e-mail: VKrivtso@Ford.com, krivtsov@umd.edu

Lemeshko Boris (Russia)
Doctor of Sci., Professor, Novosibirsk State Technical University, Professor of Theoretical and Applied Informatics Department
e-mail: Lemeshko@ami.nstu.ru

Lesnykh, Valery (Russia)
Professor, Doctor of Sci., Adviser to Director General, LLC Gazprom gaznadzor, Novocheryo-mushkinskaya Street, 65, Moscow, 117418, Russia
e-mail: vvlesnykh@gmail.com

Levitin, Gregory (Israel)
PhD, The Israel Electric Corporation Ltd. Planning, Development & Technology Division. Reliability & Equipment Department, Engineer-Expert; OR and Artificial Intelligence applications in Power Engineering, Reliability.
e-mail: levitin@iec.co.il

Limnios, Nikolaos (France)
Professor, Université de Technologie de Compiègne, Laboratoire de Mathématiques, Appliquées Centre de Recherches de Royallieu, BP 20529, 60205 COMPIEGNE CEDEX, France
e-mail: Nikolaos.Limnios@utc.fr

Papic, Ljubisha (Serbia)
PhD, Professor, Head of the Department of Industrial and Systems Engineering Faculty of Technical Sciences Cacak, University of Kragujevac, Director and Founder the Research Center of Dependability and Quality Management (DQM Research Center), Prijevor, Serbia
e-mail: dqmcenter@mts.rs

Ram, Mangey (India)
Professor, Department of Mathematics, Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun, India. Visiting Professor, Institute of Advanced Manufacturing Technologies, Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russia.
e-mail: mangeyram@gmail.comq

Timashev, Sviatoslav (Russia)
Doctor of Sci., Professor, Director and principal scientist the Sci & Engng Center of the Russian Academy of Sciences, Ekaterinburg
e-mail: timashevs@cox.net

Zio, Enrico (Italy)
PhD, Full Professor, Direttore della Scuola di Dottorato del Politecnico di Milano, Italy.
e-mail: Enrico.Zio@polimi.it

e-Journal ***Reliability: Theory & Applications*** publishes papers, reviews, memoirs, and bibliographical materials on Reliability, Quality Control, Safety, Survivability and Maintenance.

Theoretical papers must contain new problems, finger practical applications and should not be overloaded with clumsy formal solutions.

Priority is given to descriptions of case studies.
General requirements for presented papers.

1. Papers must be presented in English in MS Word or LaTeX format.
2. The total volume of the paper (with illustrations) can be up to 15 pages.
3. A presented paper must be spell-checked.
4. For those whose language is not English, we kindly recommend using professional linguistic proofs before sending a paper to the journal.

The manuscripts complying with the scope of journal and accepted by the Editor are registered and sent for external review. The reviewed articles are emailed back to the authors for revision and improvement.

The decision to accept or reject a manuscript is made by the Editor considering the referees' opinion and considering scientific importance and novelty of the presented materials. Manuscripts are published in the author's edition. The Editorial Board are not responsible for possible typos in the original text. The Editor has the right to change the paper title and make editorial corrections.

The authors keep all rights and after the publication can use their materials (re-publish it or present at conferences).

Publication in this e-Journal is equal to publication in other International scientific journals.

Papers directed by Members of the Editorial Boards are accepted without referring.
The Editor has the right to change the paper title and make editorial corrections.

The authors keep all rights and after the publication can use their materials (re-publish it or present at conferences).

Send your papers to Alexander Bochkov, e-mail: a.bochkov@gmail.com

Table of Contents

STRUCTURAL STRENGTH AND STABILITY ASSESSMENT, RISKS AND POTENTIAL HAZARDS IN THE ARCTIC ENVIRONMENT 31

Alena Rotaru

The paper deals with the assessment of the strength and stability of structures in Arctic soil conditions. The main attention is paid to the unique properties of such soils, which affect the reliability of structures. These include extremely low temperatures, permafrost, high humidity and the presence of permafrost. It is noted that the existing methods for assessing strength and stability require further development and adaptation. The specifics of Arctic soils should be taken into account when choosing research methods and approaches. The operation of structures in the Arctic is associated with risks of deformation and destruction due to loads associated with soil features. Problems may also occur due to changes in temperature and humidity. To reduce risks and improve the reliability of structures, it is proposed to develop and implement comprehensive security measures. They should take into account the specifics of Arctic soils at all stages-from design to operation of infrastructure facilities. This may include the use of special materials, the adaptation of design solutions, and the development of new methods for monitoring the state of structure. The results of the study can be useful in the design and operation of facilities in the Arctic, as well as in the development of regulatory documents and standards for construction in difficult ground conditions. This will ensure the safety and efficiency of structures in extreme climatic conditions.

ASSESSING TRANSFORMATION RATIO OF VOLTAGE TRANSFORMER UNDER NON-SINUSOIDAL SUPPLY VOLTAGE CONDITIONS 41

Nurali Yusifbayli, Huseyngulu Guliyev, Naib Hajiyev and Nijat Huseynov

Voltage transformers are rarely subjected to periodic inspection and testing while in operation, which is the reason for the difference between actual transformation ratios and their rated values. The frequency characteristics of voltage transformers are checked neither. Therefore, there can be large errors in the case of non-sinusoidal supply voltage. Existing studies on the study of voltage transformer faults reflect the results of the study of the equivalent circuit based on traditional lumped-parameter circuits. However, it should be noted that for high-frequency circuits and at the same time from the point of view of taking into account the structure of the windings, increasing the accuracy of the research results is an urgent issue. The case of non-sinusoidal supply voltage is of great scientific interest in this regard. Therefore, considering the voltage transformer as a distributed-parameter circuit, its fault under non-sinusoidal supply voltage was investigated and studied. This work explores errors of voltage transformers in the presence of non-sinusoidal voltage. To this end, a mathematical model of the transformation ratio is built considering the equivalent circuit of transformer as a distributed-parameter circuit.

ENHANCING SENSITIVITY OF MTTF AND NETWORK RELIABILITY STABILITY OF K-OUT-OF-N ACYCLIC TRANSMISSION NETWORK 52

Hemalatha.G, Vijayalakshmi.G

Networks are used in many areas of engineering, including computers and communication, transportation and electric transmission. Component importance measures are applicable to improve the system design. The reliability characteristics of k-out-of-n acyclic transmission network have been evaluated using universal generating function (UGF). This paper explores and analyses Mean Time To Failure (MTTF), Birnbaum importance measurement, critical component importance, component risk growth factor, average risk growth factor and MTTF sensitivity for acyclic transmission network. Finally, The stability illustrates the impact of component failure on the reliability of the network system. The network reliability stability for acyclic transmission network is computed. In this work, we have evaluated reliability characteristics and of the acyclic transmission network using UGF method. The network reliability stability for k-out-of-n acyclic transmission network is computed. Furthermore, numerical illustrations are provided.

TWO PARAMETRIC GENERALIZED FUZZY ENTROPY MEASURE WITH SOURCE CODING 67

Tabasum Fatima, M. A. K. Baig

In this manuscript, we introduce a new generalized fuzzy information measure for a fuzzy set. We establish its validity as a fuzzy entropy measure. Also, we define a new generalized fuzzy average code-word length for a fuzzy set, and explore its relationship with the fuzzy information measure. We prove coding theorems for discrete noiseless channels. The measures outlined in this communication are not only novel, but also encompass certain well-established measures from the existing literature on fuzzy information theory. We also show the proposed measure through a dataset and and observe its monotonicity through tabular and graphical forms.

SHREEKANT DISTRIBUTION WITH PROPERTIES AND APPLICATIONS TO MODEL STRESS AND STRENGTH DATA FROM ENGINEERING 78

Rama Shanker

Modeling and analyzing real lifetime data that are stochastic in nature with the existing lifetime distributions is really a challenging task for researchers. During the last decade several one parameter lifetime distributions have been proposed in statistics but almost each of them has been found having some problems regarding their goodness of fit on certain datasets. This might have been due to their distributional properties or due to stochastic nature of datasets. In this paper, a new lifetime distribution named Shreekant distribution has been proposed which has been found to give improved fits to lifetime datasets than those by other existing ones such as exponential, Lindley, Shanker, Akash, Sujatha and Uma distributions. This has been shown by presenting its fittings to two datasets relating to stress and strength data from engineering. Its various structural and other properties and estimation of parameters by different methods of estimation have also been studied.

ON THE FAILURE RATE FUNCTION OF A SYSTEM UNDER THE δ-SHOCK MODEL 96

Reza Farhadian, Habib Jafari, Hamed Lorvand

An interesting type of shock model in reliability theory is the delta-shock model, which is defined based on the length of time between successive shocks. Whenever this length is less than or equal to a prefixed threshold called delta, the system will fail. In this paper, the failure rate function of a system subjected to random shocks under the delta-shock model is discussed. To be more precise, a simple approximate formula for the failure rate function is established. Numerical studies are performed under the assumption that the arrival of shocks follows a Poisson process. Based on the numerical results, the accuracy of the approximation is evaluated, and it is also shown that reducing the mean of inter-arrival times between successive shocks leads to an increase in the failure rate of the system. Conclusions are presented at the end.

COMPARATIVE ANALYSIS OF BAYESIAN ESTIMATION TECHNIQUES FOR GAMMA DISTRIBUTIONS 101

Vijay Kumar Lingutla, Naganani Nadiminti, R. K. Davala

This paper explores the estimation of the common scale parameter for two Gamma populations using three Bayesian methods: Approximate Bayesian Computation (ABC), Hamiltonian Monte Carlo (HMC), and Metropolis-Hastings (MH). Unlike traditional methods like Maximum Likelihood Estimation (MLE), which can be sensitive to model misspecification, these Bayesian methods are more flexible in handling parameter uncertainty and complex data. ABC bypasses explicit likelihood calculations, HMC improves sampling efficiency through gradient information, and MH offers simplicity and adaptability. A simulation study compares these methods based on bias, mean squared error (MSE), and computational efficiency. Additionally, we apply these methods to real glucose level data from male and female individuals with diabetes, highlighting their strengths and limitations. The results offer practical recommendations for real-world applications involving Gamma distributions, contributing to the advancement of Bayesian estimation techniques in complex parameter estimation problems.

STOCHASTIC ANALYSIS THROUGH MATLAB OF TWO-UNIT PARALLEL SYSTEM 112

Sonia, Shakuntla Singla

This paper presents a reliability modelling of a two parallel unit of manufacturing Industry. Original conservation figures of the manufacturing industry have been applied for this determination. Four kinds of failure were observed: developmental, electronic, motorized and infrastructural failures. Both units of manufacturing plant work in parallel and independently. Various reliability parameters of the plant such as availability, busy period for repair, expected number of repairs and profitability for each type of failure are examined. Graphical Technique of Regenerative Point, process of Markov and Genetic Algorithm are applied for analysis. Profit analysis for the plant is examined with respect to returns of the plant and repair rate along with a graphical representation.

PERFORMANCE ANALYSIS OF SOFTWARE SYSTEM SUBJECT TO PREVENTIVE MAINTENANCE AND SOFTWARE UP-GRADATION FACILITY 120

Kavita Thukral, Shiv Kumar Sharma

Every industrialist as well as scientist try to satisfy the customers requirement by providing reliable software systems and a software system is reliable and profitable if it is easily manageable. This study analyzes the criteria under which software performs correctly with software upgrade and preventive maintenance facility. It is observed that researcher threw light on negligible, major, and load faults cause software to fail, after which load recovery, major, and negligible upgrades are carried out, but there are no discussion on preventive maintenance concept. Sometimes, to enhance the software reliability and performance, preventive maintenance is utilized and it is better to upgrade the software before completely failure. The webull distributions are followed by all of these failures and upgrade rates. The malfunctioning program can always be fixed by a skilled technician. The software performance is analyzed by applying the regenerative point technique and the semi markov process. Software reliability metrics including mean time to software system failure, software availability, and software profit values are calculated using tables.

PERFORMANCE MEASURES AND PARAMETRIC ESTIMATION OF A PARALLEL-SERIES SYSTEM MODEL WITH WEIBULL FAILURE AND REPPAIR TIME DISTRIBUTIONS 128

Shallu Sharma, Pawan Kumar, Poonam Sharma

The present paper aims at the study of three non-identical unit parallel-series system model with Weibull failure and repair time distributions. The main focus is on the analysis of reliability characteristics and also estimation of parameters in Classical and Bayesian paradigms. The units in the system are named as A, B and C. In order that system works successfully, unit A and anyone of the units B or C should work. Upon failure a unit will immediately be taken up to repair facility for necessary remedial actions. A single repair facility is available to deal with any kind of failure/fault detected with any unit. First come first served service discipline is followed. The failure and repair time distributions for all the units are taken Weibull with varying parameters. Regenerative point technique is used to study various measure of effectiveness. MLE and Bayes estimates of various failure and repair parameters involved in the study have been obtained. A simulation study has also been undertaken to exhibit the behaviour of obtained characteristics in Classical and Bayesian setup and a comparison is made thereupon. Various conclusions have been drawn from the tables and graphs plotted for numerous performance measures for varying values of repair and failure parameters. Having obtained all the reliability characteristics, the profit incurred by the system has been obtained and studied graphically too.

EXPONENTIAL TYPE ESTIMATORS OF POPULATION MEAN-A REVIEW 141

Sajad Hussain, Vilayat Ali Bhat

Survey sampling is a crucial statistical approach for making inferences about population characteristics based on the analysis of a representative sample. This method offers a practical solution for saving time and resources while ensuring reliable outcomes. Auxiliary information plays a pivotal role in enhancing the accuracy of estimators used in finite population sampling. Traditional estimators, such as ratio, product, and regression types, have proven effective in improving the precision of population parameter estimation. The introduction of exponential-type transcendental functions marked a significant advancement in this field, further increasing the efficiency of ratio and product estimators. This study provides an in-depth review of the literature on exponential-type ratio and product estimators, emphasizing their relevance in survey sampling. It examines theoretical progress and practical implementations, showcasing their potential to reduce bias and improve estimation accuracy by leveraging complex relationships between study variables and auxiliary information. The insights from this review contribute to advancing research in finite population sampling and offer valuable directions for future studies in the field.

A COMPARATIVE ANALYSIS OF GAMMA AND WEIBULL DISTRIBUTIONS IN TAMIL CINEMA DATA 156

A. Vanathu Suresh, R Subramani, S Vijayan

This paper aims to systematically investigate the utility of the Gamma and Weibull distributions, focusing on their application to lifetime datasets and clarifying their mathematical and statistical properties. By analyzing lifetime data across various disciplines, the research emphasizes the effectiveness and flexibility of these distributions in capturing the complexities inherent in such data. It underscores the importance of parameters such as standard error, log-likelihood, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan–Quinn Information Criterion (HQIC) in value estimation. The findings suggest that both distributions provide valuable insights into the underlying data, with practical implications for reliability engineering and failure analysis. Moreover, the study demonstrates that the Weibull distribution offers a better fit to the given data than the Gamma distribution due to its adaptability, which yields superior results. A key contribution of this study is the proposal of a model based on estimating the Conditional Weibull distribution for feature parameters, which accurately predicts a finite mixture of two-parameter Weibull distributions initially verified on datasets.

LEAST SQUARES, MAXIMUM LIKELIHOOD, AND ANDERSON-DARLING TYPES ESTIMATORS FOR ENTROPY TRANSFORMED EXPONENTIAL DISTRIBUTION 163

Ikwuoche John DAVID, Stephen MATHEW, Isah Charles SAIDU

Aim: This research aim is to investigate the performance stability and determine the best estimator among Least Squares (LS), Maximum Likelihood (ML), and Anderson Darling (AD) type (right tail AD, left tail AD, second order right tail AD, and second order left tail AD), estimators for estimating the parameter of the Entropy Transformed Exponential distribution. Methods: A Monte Carlo simulation was conducted to evaluate the finite sample performance stability and efficiency at sample sizes of 20, 100, 250, 700, 1000, 1500, 2500, and 3000. The evaluation measures used for assessment are the Absolute Bias (AB) and Root Mean Square Error (RMSE) as performance metrics. Real life data analysis was performed using a field intensity data measured in microvolts recorded at 5-second intervals. The data is a positive right skewed data and highly peaked. Results: The simulation results demonstrated that the LS method consistently outperformed the ML and AD type estimators. This is based on the lowest values of AB and RMSE obtained across all the eight sample sizes. Additionally, the findings were validated through real world data, where the LS estimator exhibited the lowest standard error and the narrowest confidence interval length. Based on the results, the LS is ranked 1st, right tail AD ranked 2nd, second order right tail AD ranked 3rd, MLE ranked 4th, left tail AD ranked 5th, and second order left tail AD ranked 6th. However, results from t-test in determining the significance of the parameter estimate from each estimator, shows that MLE and second order left tail AD estimates are not significant at 5%. The graphical representation of the field intensity data by each estimators fit clearly depict the LS estimator best describe the data distribution. Conclusion: Based on these findings, the study concludes that the LS estimator is the most effective method for estimating the parameter of the Entropy Transformed Exponential distribution followed by the right tail AD, second order right tail AD, MLE, left tail AD, and second order left tail AD being the least effective estimator.

SOME ADDITIONAL PROPERTIES OF KUMARASWAMY LOG-LOGISTIC DISTRIBUTION BASED ON ORDER RANDOM VARIATES 175

Bavita Singh, Nazia Wahid, Mahfooz Alam, Aafaq A. Rather

The Exponentialed Kumaraswamy-Dagum distribution represents a significant advancement in the family of probability distributions. This family encompasses several well-established sub-models, including a variant of the Kumaraswamy-Fisk or Kumaraswamy log-logistic distribution, referred to as the Exponentialed Kumaraswamy-Dagum family distribution. In this paper, we characterize the Kumaraswamy log-logistic (KSLL) distribution by employing the moment features of generalized order statistics. Explicit expressions and recurrence relations are derived for both the single and product moments of generalized order statistics associated with the KSLL distribution. These formulations specifically address record values and moments of order statistics. Additionally, the distribution is further analyzed through conditional moments, and a recurrence relation for single moments of generalized order statistics is established. This rigorous investigation seeks to deepen the understanding of the KSLL distribution and its statistical properties as obtained by generalized order statistics.

COMPARATIVE ANALYSIS OF THE RELIABILITY OF VACUUM AND OIL CIRCUIT BREAKERS 188

Ilham RAHIMLI, Aliashraf BAKHTIYAROV

In modern power systems, switching equipment plays a crucial role in ensuring the safe and reliable connection and disconnection of electrical circuits. A failure of a circuit breaker can lead to serious consequences, including prolonged power outages and damage to expensive equipment. This paper presents a comparative analysis of the reliability of vacuum and oil circuit breakers, considering operational, technical, and economic factors. It examines the main causes of failures, methods of reliability assessment, and provides a comprehensive evaluation of the effectiveness of each type of breaker under various conditions. The novelty of the proposed solution lies in the in-depth analysis of operational characteristics and a multifaceted evaluation of the impact of external factors on equipment reliability. Modern diagnostic and monitoring methods, including thermography, ultrasonic diagnostics, and the integration of IoT systems for real-time monitoring, are used. These approaches improve the processes of equipment selection and operation, ultimately contributing to the overall reliability of power systems. The paper aims to support informed decision-making in equipment selection to enhance the overall reliability of power systems.

RELIABILITY ANALYSIS OF THE SHAFT SUBJECTED TO FLUCTUATING LOADS UNDER TORSION AND BENDING MOMENTS WHEN SHEAR STRESS FOLLOWS EXPONENTIAL DISTRIBUTION 194

Md. Yakoob Pasha, M. Tirumala Devi, T. Sumathi Uma Maheswari

The reliability of shafts subjected to fluctuating loads is essential for ensuring the safe and efficient operation of rotating machinery in various engineering applications, including automotive, aerospace, and industrial systems. Accurate reliability assessments are crucial to preventing failures and extending the lifespan of mechanical components. This study examines the reliability analysis of shafts subjected to combined torsional and bending moments, which create complex stress states that significantly affect shaft performance and failure behaviour. A statistical model is considered since the shear stress generated is random in nature. Analytical expressions for reliability are derived by changing the parameters twisting moment, bending moment, and shaft diameter.

ANALYSIS OF TWO VACATION POLICIES UNDER RETRIAL ATTEMPTS, MARKOVIAN ENCOURAGED ARRIVAL QUEUING MODEL 201

E. IsmailKhan, V. Narmadha, N. Yathavan

This paper presents a comprehensive analysis of a single-server Markovian queueing system incorporating two distinct vacation policies under a framework of encouraged arrivals and retrial attempts. The novelty of the model lies in its simultaneous treatment of several realistic operational dynamics: server vacations with acceleratory growth, customer impatience, server breakdown and repair, and encouraged arrivals, where potential customers are motivated to join the system depending on its state. To reflect real-world retrial scenarios, customers who do not receive immediate service reattempt after some time, modeled using a retrial queue mechanism. Two key vacation policies are compared: one governed by a first-come, first-served (FCFS) discipline and the other utilizing bulk service. The paper derives steady-state probabilities and evaluates system performance measures, such as mean queue length, server utilization, and the expected number of retrials.

CONSTRUCTING A SEPARATE RESERVE IN A CHAIN OF LOW-RELIABILITY ELEMENTS 210

Gurami Tsitsiashvili, Yury Kharchenko

In this paper, we construct upper and lower estimates of the minimum separate reserve of low-reliability chain elements of length m, at which the probability of operability of the entire chain becomes close to unity. The appeal to a separate reserve is caused by the properties of this reserve, established by Barlow and Proschan. It is shown that the upper and lower estimates of the minimum separate reserve are quite close. If the probability of operability of an individual element of the chain is 1/ln m, then the minimum reserve volume is of the order ln2 m.

EVALUATING THE IMPACT OF PARAMETER SENSITIVITY IN META-HEURISTICS ON RELIABILITY PERFORMANCE 214

Shakuntla Singla, Manisha Rani, Shilpa Rani, Umar Muhammad Modibbo

The optimizing reliability of a system is a crucial target for any industry to have good values and outcomes with the balanced cost of maintenance. The assessment of costs and reliability in intricate systems, such as Distributed Redundancy Architecture (DRA), is essential for optimal system design. Achieving this involves addressing various constraints to enhance the overall system's reliability. Researchers have explored system dependability and cost optimization challenges, emphasizing the evolution of metaheuristic approaches. Enhancing system resilience and optimizing costs are central objectives of this study, aiming to improve the reliability and cost-efficiency of DRA systems. The study's framework builds upon existing metaheuristic methods, including Moth Flame Optimization (MFO), Whale Optimization Algorithm (WOA), Dragonfly Algorithm (DA), Gazelle Optimization Algorithm (GOA), and Coati Optimization Algorithm (COA). The findings reveal that COA demonstrates superior performance compared to other methods. By employing the Coati Optimization Algorithm, this study effectively addresses cost and reliability-optimization challenges, providing highly efficient solutions. Key metrics such as reliability and cost were compared across the five metaheuristic approaches, showcasing COA's advantages. Notably, the Coati Optimization Algorithm delivers faster resolutions and outperforms its counterparts in reducing costs while enhancing system reliability. The results underscore COA's ability to optimize intricate system parameters more effectively than methods like MFO, WOA, GOA, and DA. This research highlights the value of using COA for improving reliability and minimizing expenses in complex systems. It offers a novel approach to addressing these challenges, establishing COA as a powerful tool for optimizing distributed system architectures. By analyzing comparative solutions, the study found that COA consistently achieved better outcomes, affirming its success in reliability and cost optimization. In conclusion, this research emphasizes the importance of advanced metaheuristic techniques, particularly the Coati Optimization Algorithm, in enhancing the performance of Distributed Redundancy Architectures. By presenting a numerical example, the comparison has been made and the outcomes are shown graphically and in the tabular form to have a clear understanding. The findings contribute significantly to the field by providing a robust method for achieving cost-efficient and reliable system designs.

ESTIMATING THE FRECHET DISTRIBUTION'S SCALE PARAMETER USING BAYESIAN METHODS FOR SYMMETRIC AND ASYMMETRIC LOSS FUNCTIONS 227

S. C. Premila, S. Jayabharathi, D. Kanagajothi, D. Pachiyappan

Comparing several estimators to estimate the scale parameter of the Frechet distribution based on complete data when the shape parameter is known is the focus of this research. Using the power function distribution as a prior distribution and the maximum likelihood estimator (MLE), we apply the Bayes estimators under the squared error loss function (SELF), linear exponential loss function (LLF), asymmetric precautionary loss function (APLF), and composite LINEX loss function (CLLF). A Monte Carlo simulation study served as the basis for the comparison. The performance of these estimators in relation to the mean square error (MSE) was compared through the simulation study.

MODELING AND ANALYSIS OF THE WEIGHTED POWER CHRIS JERRY DISTRIBUTION FOR RELIABILITY ENGINEERING 234

R. Sembiyan, Iyappan. M, D. Suresh, Kalaivanan. S, R. Saranraj

This paper introduces a new class of probability distribution, termed the Weighted Power Chris Jerry (WPCJ) distribution. This distribution is developed by applying a weighting technique to the baseline Power Chris Jerry distribution, resulting in a more flexible and robust statistical model. The mathematical and statistical properties of the WPCJ distribution, including its moments, reliability measures, and order statistics, are derived and analyzed in detail. The parameters of the proposed distribution are estimated using the Maximum Likelihood Estimation (MLE) method, ensuring effective application in real-world scenarios. To demonstrate the practical utility of the WPCJ distribution, it is applied to a real-life dataset. This application highlights the distribution's ability to model lifetime data effectively, showcasing its superiority over existing models. The weighted distribution approach offers significant advancements in distribution theory, providing enhanced flexibility for modeling complex data structures across various domains, including reliability engineering, biomedicine, and ecological studies. The introduction of weighted distributions is grounded in the foundational work of Fisher and Rao, which demonstrated their importance in situations where observations are influenced by ascertainment methods or weight functions. By extending these principles, the WPCJ distribution offers a novel tool for addressing conceptual challenges in data representation and model development. This study further contributes to the literature by illustrating the adaptability and effectiveness of weighted distributions in tackling diverse analytical problems.

BAYES ESTIMATES FOR THE PARAMETERS OF POISSON TYPE ONE PARAMETER RAYLEIGH CLASS SRGM USING GAMMA AND INVERTED GAMMA PRIORS 242

Rajesh Singh, Kailash R. Kale and Pritee Singh

Research on the reliability of both software and hardware commonly utilizes the Rayleigh distribution as a model. During software testing, the failure occurrence behavior can be used to evaluate the software's reliability. In this paper, the Poisson pattern of failure occurrence is assumed to study the parameters of the Rayleigh Class software reliability growth model. It has been studied how the Rayleigh Class Model's scale parameter and the quantity of software's intrinsic failures behave. Additionally, using the Gamma and Inverted Gamma Priors, respectively, Bayes estimates for the number of intrinsic failures and the scale parameter are suggested. The comparative analysis of the proposed Bayes estimators and accompanying MLEs is done using the relative efficiency that are acquired via the use of Monte Carlo Simulation.

COX'S REGRESSION MODEL FOR RELIABILITY ANALYSIS AND FAILURE TIME PREDICTION IN ENGINEERING SYSTEMS 250

Iyappan. M, Balaji. M, G. Sathya Priyanka, Jesmon Raj. N, P B Deepa

Accelerated Failure Time Models (AFTMs) are utilized to examine survival patterns, particularly in the context of diseases such as diabetes, where traditional models like Cox proportional hazards may not apply. In this study, we apply AFTMs to identify significant prognostic factors affecting survival in male diabetic patients, using parametric distributions such as Exponential, Weibull, Log-normal, and Log-logistic. The primary objective of this study is to gain

deeper insights into how key covariates influence survival times and enhance risk stratification, ultimately leading to better treatment outcomes. We evaluated the fit of several AFTMs by comparing Log-Likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) values. Among the models tested, the Weibull AFTM emerged as the best fit, demonstrating superior Log-Likelihood and lower AIC and BIC values compared to other distributions. The study identified several significant predictors of survival in male diabetic patients, including age, family history of diabetes, duration of diabetes, early use of insulin (within one year of diagnosis), retinopathy, use of diuretics, proteinuria, hypertension, and fasting plasma glucose levels. All of these covariates were found to significantly influence survival outcomes (P<0.05), with older age and longer diabetes duration correlating with poorer survival rates. These findings underscore the importance of using AFTMs to model survival patterns in diabetic patients and the potential for targeted interventions based on identified risk factors. While the small sample size limits the generalizability of the results, this study provides valuable insights into the survival characteristics of male diabetic patients and highlights the need for further research with larger datasets. Future studies should consider a wider range of covariates and explore the application of AFTMs in broader diabetic populations to refine survival prediction models and optimize patient care. Additionally, integrating machine learning techniques with AFTMs could enhance predictive accuracy and provide a more comprehensive understanding of survival trends in diabetic patients.

BIPOLAR NEUTROSOPHIC MINIMUM SPANNING TREE NETWORK USING UPPER TRIANGULAR WEIGHT/COST ALGORITHM 259

Shayathri Linganathan, Purusotham Singamsetty

A minimum spanning tree (MST) is a spanning tree in which the total weights of all of its edges are the lowest of all conceivable spanning trees for the network. MST is crucial in discipline of the operation research. Similarly, for the neutrosophic minimum spanning tree problem (NMST) the weights are unpredictable and inconsistent. The next form of NMST is considered the bipolar neutrosophic minimum spanning tree problem. Each edge of the standard NMST is assigned a bipolar weight based on the bipolar neutrosophic number (consisting of positive and negative membership degrees). This paper introduces the upper triangular weight algorithm for a given network and the optimal results are executed in MATLAB.

DESIGNING AN ECONOMIC REPETITIVE GROUP SAMPLING PLAN FOR LIFE TESTS BASED ON PERCENTILES OF EXPONENTIATED RAYLEIGH DISTRIBUTION 265

P. Umanaheswari, S. Suganya, K. Pradeepa Veerakumari

In this study, a unique Repetitive Group Sampling Plan (RGSP) for life tests is presented, emphasizing the use of percentiles obtained from the Exponentialed Rayleigh Distribution. To enhance the accuracy, effectiveness, and general reliability of life testing techniques, the suggested sample plan provides a strong foundation for assessing the longevity and reliability of goods in a variety of industries. By offering a strong framework for evaluating the robustness and lifespan of products in diverse industries, the suggested sample plan seeks to improve the efficiency and reliability of life testing methods. Through the incorporation of statistical techniques grounded in the Exponentiated Rayleigh Distribution, this research advances the fields of reliability assessment and quality control in industrial processes. Table values are calculated by using Python programming. The implementation of these statistical techniques not only enhances the accuracy of life testing but also enables businesses to make more informed decisions regarding product development and quality assurance. As industries increasingly rely on data-driven approaches, this research provides a valuable resource for optimizing performance and minimizing costs.

COMMON FIXED POINT THEOREMS IN COMPLEX VALUED B-METRIC SPACES AND USING WEAK COMPATIBLE MAPPINGS 271

Shivani Chourasiya, Harshit Khare, Kavita Shrivastava

Complex-valued metric spaces, where the distance between points is represented as a complex number, serve as a generalization of traditional metric spaces and present a variety of intriguing applications, particularly in advanced mathematics and theoretical physics, such as in quantum mechanics, quantum computing, functional analysis, operator theory, complex dynamics, fractal geometry, complex networks, and signal processing, as well as theoretical physics, relativity, geometry, topology, mathematical finance, and economics. Despite the increasing interest in

complex-valued metric spaces, the current literature on fixed point theorems within these contexts is constrained in several ways. The available literature has notable shortcomings; consequently, this article investigates the results and implications of contractive mappings in complex-valued b-metric spaces by utilizing weak compatible mappings, and it also employs control functions while solving a system of Urysohn integral equations based on our primary result. These findings contribute to this area of study.

ESTIMATION AND APPLICATION OF A NEW GENERALIZATION OF EXPONENTIAL DISTRIBUTION 282

D. Kumar, P. K. Chaurasia, P. Kumar, A. Sahoo

In statistical literature, several lifetime distributions exist for real phenomena. And one of the methods to find new lifetime distribution by existing baseline distribution such a method is known as the transformation method. In this article, we proposed a generalization of the existing transformation by introducing the additional shape parameter. Here, we considered baseline distributions as an exponential distribution. Various statistical properties of the new lifetime distribution, such as survival function, hazard rate function, cumulative hazard rate function, moments, quantile function, and order statistics, have been discussed. Demonstrate the applicability and suitability of the proposed distribution. Here, we focus only on the estimation of the parameters likes MLE, LSE and also to check long-run behaviour of the estimators.

PARABOLIC DENSE CLOUDY FUZZY MODEL WITH DEMAND DEPENDENT PRODUCTION RATE IN AN IMPRECISE PRODUCTION PROCESS 295

Upsana Rana, Tanuj Kumar

In this paper we develop a classical economic production lot size model of a manufactured item in a inadequate production process in which the ordering cost is fixed and shortage is not allowed. In this model the valuation of an item is considered as a parabolic cloudy fuzzy number and production rate is depends on demand. In the previous studies, fuzziness of parameters of the model is presented in the form of fuzzy numbers using the membership function only. Parabolic dense cloudy intuitionistic fuzzy numbers are found to be more generalized than fuzzy numbers and have the capability to express the fuzziness of parameters of the model in a more generalized way using both membership and non-membership functions. The presented model is solved in crisp, fuzzy and Intuitionistic fuzzy setup using proposed generalize and defuzzification method with minimum loss of data. The present model is solved by DBPSO algorithm to determine optimal values. The validity of the proposed model and parameter sensitivity analysis is presented to justify the idea by the numerical analysis.

RELIABILITY AVAILABILITY MAINTAINABILITY DEPENDABILITY (RAMD) OPTIMIZATION: A CASE STUDY OF MANUFACTURING PLANT 308

Deepika Garg, Dipesh Popli, Pradeep Kamboj, Nakul Vashishth

In the context of modern Industry practices, optimizing the performance of complex systems necessitates a thorough understanding of Reliability, Availability, Maintainability, and Dependability (RAMD). This paper dives into the application of two powerful optimization techniques, Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), in RAMD optimization. GA, a simulation of natural evolution, employs selection, crossover, and mutation operations to search for optimal solutions iteratively. In contrast, PSO replicates the collective behavior of a swarm, iteratively updating particle positions based on their best position and the swarm's best position. The research specifically focuses on RAMD optimization in terms of maintenance schedules, resource allocation, and system configurations. The results from both techniques are compared, with PSO demonstrating superior performance. The paper concludes by underscoring the need for further research and practical implementation of PSO and GA to fully exploit their potential in optimizing RAMD parameters across diverse domains.

ON THE PROTECTION AGAINST NOISE FOR MEASUREMENT-BASED QUANTUM COMPUTATION 324

Valentin Vankov Iliev

In this paper we establish conditions for some pairs of quantum logic gates which operate on one qubit to be protected against crosstalk. We represent these logic gates by self-adjoint operators acting on an unitary plane, consider the resembling bipartite quantum systems, and extend the operators naturally to a commuting pair. Thus the corresponding two measurements can be considered as simultaneous and we can define them as events in a classical probability space. In particular, we are able to study their dependence, the information noise thus produced, and to find necessary and sufficient conditions for being informationally independent events.

HIDDEN MARKOV MODELS FOR ESTIMATING IMPORT DYNAMICS WITH PROBABILITY DISTRIBUTION ANALYSIS 332

Vyshnavi. M, Muthukumar. M

This article examines the use of Hidden Markov Models and probability distributions to study agricultural import dynamics, with a focus on revealing hidden patterns in the data. The idea behind the study is to enhance our understanding of how transitions between different agricultural import states occur and to explore the most suitable probability distributions for modeling these hidden states. The purpose of the study is to identify the optimal distributional fit for hidden states by evaluating the Transition Probability Matrix, Emission Probability Matrix, and Initial Probability Vector (π) for each state. The research implements the Akaike Information Criterion and the Bayesian Information Criterion to select the best-fitting distribution for each scenario. Furthermore, the paper focuses on the practical implications of these discoveries, such as determining the most likely state sequence using the Viterbi path, which can help influence decision-making and forecasting. The analysis is carried out using R software, which provides information about the associations between probability distributions, stationarity tests, and the role of model selection criteria such as AIC and BIC. Graphical representations of AIC and BIC values over several probability distributions, and additionally a correlation matrix between selected distributions, help to highlight the findings. Overall, the paper enhances our understanding of probability distributions through HMM frameworks for agricultural import dynamics, providing recommendations for optimal model selection in various applications.

OPTIMIZED SCHEDULING FOR THREE-STAGE FLOW SHOP WITH PARALLEL MACHINES: APPLICATION IN WASTE MANAGEMENT AND RECYCLING FACILITIES 347

Sonia, Pooja kaushik, Deepak Gupta, Sonia Goel

Scheduling is one of the most important aspects of manufacturing sectors. Various factors in terms of performance evaluation, make business more competitive. Scheduling is the process of allocating scarce resources to tasks throughout time in order to accomplish one or more optimization goals. In fact, the methodical process of organizing, managing and maximizing work while ensuring the greatest possible use of both time and resources is the main aim of scheduling. This paper indicates the desirable and necessary steps to discover the optimum solution by efficient scheduling in waste recycling facilities which is essential for optimizing resources utilization and reducing operational delay. In this study, we have taken three stage flow shop scheduling with multiple machines at first two stages and single machine at last stage. This type of model has a number of applications not only in several industries but also in saving natural resources like water, electricity and reducing waste, idle time and unnecessary energy consumption. As, this research specially focused on the scenario of waste management and recycling facility where incoming waste undergoes three primary stages sorting, cleaning, and recycling. Here the scheduling evaluation is done by comparing the Branch and Bound method with Palmer's and Dannenbring method to determine an optimal job sequence that minimize the total processing time. The processing time is represented as a trapezoidal fuzzy number making an impact. The goal of this activity is to determine the optimal job plan in order to reduce the total amount of time that has processed and allocate the lead in the best feasible way to reduce the machine idle time, minimizes the water usage and energy consumption. This comparison demonstrated how well our suggested approach handled this challenging scheduling issue the findings contribute to the efficient management of recycling facilities, ensuring improved throughput. Because better scheduling strategies must be implemented in order to the growing demand for waste processing and recycling. So, the proposed

approach aids in streamlining waste management processes, ultimately supporting environmental conservation efforts. Here, Branch and Bound is the best approach as compared to Palmer and Dannenbring approach. Future work can be extending this methodology to the three-phase flow shop scheduling with parallel machines at all the stages.

RELIABILITY MEASURES AND CLASSICAL AND BAYESIAN PARAMETRIZATION OF TWO NON-IDENTICAL UNITS SYSTEM MODEL WITH ON-LINE/OFF-LINE REPAIRS OF REPAIR MACHINE 358

Poonam Sharma, Pawan Kumar, Shallu Sharma

The present paper deals with the reliability analysis of a system model which consists of two non identical units and a repair machine. One of the units initially, is in operative mode and other is kept in standby mode. Repair machine is initially in good condition, it may fails while repairing the failed units. In case of failure of repair machine, firstly it undergoes on-line repair and then to off-line repair if not repaired online. The failure time distributions of both the units and repair machine are assumed to be exponential while the repair time distributions are taken as general. By using regenerative point techniques, various measures of system effectiveness such as transition probabilities, MTSE, reliability, availability, busy period etc. have been obtained. Also, some of them have been studied graphically. A simulation study is also carried out to analyse the considered system model both in Classical and Bayesian setups. From the graphs and tables, various important conclusions are drawn.

DEVELOPMENT OF AN ATTRIBUTE CONTROL CHART BASED ON THE INVERSE KUMARASWAMY DISTRIBUTION 372

Nagaraju, R., Palanivel, M., Sriramachandran, G. V.

This article presents the development of an attribute control chart designed for products with lifespans following the inverse Kumaraswamy distribution, using the distribution's mean as the quality metric. The study assumes one distribution parameter is known while varying the other to design the control chart. The chart's performance is evaluated using Average Run Length (ARL), with adjustments to parameter values determining ARL in out-of-control conditions. The proposed control chart is assessed through simulations. A comparative analysis with other similar charts is conducted based on ARL values, and the findings are discussed in detail.

ASSESSMENT OF POSSIBLE SEISMIC HAZARD AND RISK USING THE EXAMPLE OF ALMATY 382

Gennadii Nigmetov, Andrey Savinov, Temir Nigmetov, Syrym Gabbasov

In January 2024, residents of Almaty felt the seismic impact of an earthquake that occurred on the border of China and Kyrgyzstan. The city was not prepared for such an event. How could the city residents prepare for this devastating earthquake? In order to properly prepare the population for a destructive earthquake, it is necessary to have a system for assessing the individual seismic risk of the city, which uses data for risk assessment: on predicted possible earthquake sources (PES) dangerous for the city for the next 10 years, on the seismic resistance of buildings and the city's population. Based on these data, the possible individual risk for the city population is assessed. In order to assess the individual risk as accurately as possible, in addition to data on the PES and seismic resistance of buildings, data on the macroseismic field from the PES forecast are needed. At present, there are all the necessary scientific and technical capabilities for creating monitoring systems for seismic protection of cities.

AN APPLICATIONS-BASED STOCHASTIC MODELING OF MIXTURE DISTRIBUTION FOR CANCER DATA 394

Pooyitha R and Pandiyan P

This study proposed a new stochastic model by the mixture of Gamma and Zeghdoudi distribution (MGZ Distribution). The Reliability Analysis and statistical features such as moments, Moment generating function, order statistics, stochastic ordering, entropy, and the Maximum Likelihood Estimation of the model parameters were derived.

Two real-time datasets of cancer patients' survival times were used to demonstrate the model's goodness-of-fit using the AIC, BIC, and AICC model selection methods. The new MGZ model has been compared with the other classical models like Exponential, Lindley and Shanker. The result shows that the proposed model is more flexible than the other models with the lowest values of AIC, BIC and AICC.

**HIGH-TEMPERATURE HEAT-INSULATING MATERIALS:
A COMPROMISE BETWEEN THERMAL CONDUCTIVITY AND RELIABILITY** 408

N.M. Piriyeva, R.K. Karimova, G.K. Abdullayeva

This article examines heat transfer mechanisms in high-temperature thermal insulation materials, addressing key operational challenges such as thermal degradation, mechanical stability, and chemical resistance. The novelty of the study lies in an integrated approach that includes optimizing the micro- and nanostructure to minimize radiative and conductive heat transfer, developing advanced multilayer and composite systems, applying predictive computational modeling for property assessment, and exploring self-healing coatings and intelligent materials. These innovations ensure reduced thermal conductivity without compromising mechanical strength, enhance resistance to thermal cycling and aggressive chemical exposure, and significantly extend the service life of insulation systems, making them highly promising for aerospace, energy, metallurgy, and other industries requiring reliable thermal protection. Additionally, the proposed solutions contribute to increased energy efficiency, reduced material consumption, and improved environmental sustainability, paving the way for next-generation thermal insulation technologies with broader industrial applicability.

**EVALUATING MULTI SERVER QUEUING SYSTEM EFFICIENCY:
A COMPARATIVE STUDY BETWEEN FUZZY QUEUING MODEL
AND INTUITIONISTIC FUZZY QUEUING MODEL
WITH INFINITE CAPACITY CONSTRAINTS** 414

S. Aarthi

In this study, we suggest an innovative analytical approach employing triangular fuzzy and triangular intuitionistic fuzzy numbers to ascertain the membership functions governing significant service-execution proportions within multiserver queuing models. Departing from conventional methodologies, both the inter-entry and completion rates are characterized by fuzzy natures, introducing a novel dimension to the analysis. Through numerical validation, our model demonstrates viability in the multiserver queuing context, bolstering its credibility. Furthermore, we conduct a contextual investigation, juxtaposing individual fuzzy metrics, thereby revealing the superior categorization potential of intuitionistic fuzzy queuing models over their fuzzy counterparts. Expanding the horizons of traditional fuzzy queuing theory, our integration of intuitionistic fuzzy environments promises enhanced queuing model implementation efficacy. Our primary objective lies in evaluating the productivity of a multi-server queuing model with limitless capacity, leveraging both fuzzy queuing and intuitionistic fuzzy queuing theories. Embracing a framework where arrival and service rates are characterized by triangular and intuitionistic triangular fuzzy numbers, we undertake a thorough assessment to establish evaluation criteria, adhering to a design protocol that preserves fuzzy values without converting them into crisp values. Additionally, we address two statistical challenges to ascertain the method's validity.

**FORECASTING VOLATILITY IN INDIAN NATIONAL STOCK EXCHANGE USING A
MARKOV SWITCHING GARCH MODEL** 429

D. Pachiyappan, S. Jayabharathi, Y. Geeyha Evanjalin, S. Kavitha

This paper analyzes various GARCH models in terms of their effectiveness in predicting volatility in the national Stock Exchange (NSE). This paper examines GARCH models that feature both Gaussian and fat-tailed residual conditional distributions, and evaluates their ability to characterize and predict volatility over time horizons ranging from 1 to 22 days. The AR(2)-MR-SGARCH-GED model demonstrates superior performance compared to other models at a one-day horizon. The AR(2)-MR-SGARCH-GED and AR(2)-MRSGARCH+ models surpass other models at the 5-day horizon. Within a ten-day timeframe, three AR(2)-MR-SGARCH models surpass the performance of alternative

models. Regarding the 22-day forecast horizon, the results demonstrate no distinctions between MR-SGARCH models and standard GARCH models. When looking at risk management outside of samples (95% VaR), a number of models seem to give accurate and reasonable VaR estimates for a one-day period, with a coverage rate close to the nominal level. According to the risk management loss functions, no model is universally the most accurate. This variability suggests that model performance is context-dependent, influenced by the specific characteristics of the data and the underlying assumptions of each approach. Therefore, practitioners should consider employing a combination of models to enhance robustness in their risk assessments and decision-making processes.

MULTI REFERENCE SKIP-LOT SAMPLING PLAN: A NOVEL APPROACH FOR QUALITY CONTROL 447

Moby Joseph Manoj, Azarudheen S

Skip-lot sampling plans have become significant in modern quality control due to rising production volumes and the demand for cost-effective inspection methods that will yield high-quality outputs. When inspecting a submitted lot, a skip-lot plan is economically favourable and guarantees high quality. Thus, this approach benefits both producers and consumers. The skip-lot sampling plan generally utilizes the same sampling plan as the reference plans for both skipping and normal inspection. However, using the same plan in both phase favours either the producer or the consumer in the most essential situations. This article introduces a novel approach, the Multi Reference Skip-Lot Sampling Plan with the provision of having two different reference plans in the normal and skipping phases of the skip-lot plan. The paper explores the efficacy of this approach by deriving performance measures using a power series approach. To evaluate the proposed plan, a comparison is made with existing skip-lot sampling plans that use single sampling plans or double sampling plans as reference plans. This comparison is based on operational characteristics and average sample number values, accompanied by graphical representations. The comparative analysis demonstrates that the new plan effectively balances the satisfaction of both producers and consumers. Additionally, the study offers a strategy for selecting the plan parameters using the unity value approach, supported by a table providing unity values.

MARSHALL - OLKIN TWO PARAMETER SUJATHA DISTRIBUTION AND ITS APPLICATIONS 456

V. Sreejith, Ajitha Sasi, B. Vineshkumar

Sujatha distribution, often discussed in the context of reliability engineering and survival analysis, is a probability distribution that is useful for modeling lifetimes of objects and systems. In this paper, we introduce a three parameter distribution called the Marshall - Olkin two parameter Sujatha distribution. Marshall - Olkin generalization is a novel method of enlarging a known family of distributions. It leads to a rich class of distributions that can capture different shapes and behaviours and thus allows for better fitting to empirical data, thereby establishing its importance in reliability engineering. We derive many useful statistical properties such as the hazard rate function, reverse hazard rate function, order statistics, moments, measure of skewness and measure of kurtosis of the Marshall - Olkin two parameter Sujatha distribution. The stress - strength reliability of this distribution has also been derived. The application of this new distribution is established using real life data sets. We also compare the performance of this distribution with that of some existing distributions.

A NEW GENERALIZATION OF PARETO DISTRIBUTION: PROPERTIES AND APPLICATIONS 466

Tabasum Ahad, S.P. Ahmad

This manuscript introduces a new extension of the Pareto distribution using the SMP transformation technique, known as SMP Pareto distribution (SMPP). The SMP technique is named after (Shamshad, Murtiza, Parvaiz) who pioneered this approach to enhance the flexibility and applicability of statistical models. This new distribution provides a better fit for data than many existing models. The shapes of the density function exhibit great flexibility. It can support different hazard shapes, such as increasing, decreasing and constant shapes. Various statistical properties of the proposed distribution such as moments, quantile function, entropy and order statistics were presented. The maximum likelihood technique is used to estimate the model parameters. An extensive simulation study is carried out to illustrate

the behavior of MLEs. In addition, two real-world datasets are analyzed to showcase the applicability of the proposed approach. The results indicate that the SMP Pareto distribution (SMPP) is more flexible and offers a superior fit for describing data compared to several existing forms of the Pareto distribution.

FUZZY DATA-DRIVEN ESTIMATION FOR DUAL WEIBULL POPULATIONS: EM AND BAYESIAN METHODS WITH GIBBS SAMPLING 479

Abbarapu Ashok and Nadiminti Nagamani

In statistical analysis, accurately estimating parameters in the presence of uncertain and imprecise data is critical, particularly when dealing with complex, dual population models. Fuzzy data, which effectively represents real-world ambiguity, provides a framework for handling such uncertainties. While parameter estimation for single-population models with fuzzy data has been explored extensively, extending these methods to dual populations remains challenging. This study addresses this gap by developing estimation techniques for two Weibull populations that share a common shape parameter but differ in scale parameters, under fuzzy data conditions. We employ the Expectation-Maximization (EM) algorithm for Maximum Likelihood estimation and a Bayesian framework with TK approximation for parameter estimation. To refine our Bayesian estimates, we use Gibbs sampling to compute posterior densities. Through Monte Carlo simulations, for generated data as well as for a real dataset, we evaluate the accuracy and robustness of the proposed estimators, demonstrating their practical utility in applications where data imprecision is a significant factor. This research highlights the importance of robust methodologies for dual-population Weibull models, contributing to enhanced reliability in statistical analysis across diverse fields.

ACCELERATED FAILURE TIME MODELING FOR PROGNOSTIC FACTORS IN MALE DIABETES PATIENTS: A COMPARATIVE ANALYSIS OF SURVIVAL MODELS 493

Jeslin J, Radhika A, Haripriya M

Accelerated Failure Time Models (AFTMs) are utilized to examine survival patterns, particularly in the context of diseases such as diabetes, where traditional models like Cox proportional hazards may not apply. In this study, we apply AFTMs to identify significant prognostic factors affecting survival in male diabetic patients, using parametric distributions such as Exponential, Weibull, Log-normal, and Log-logistic. The primary objective of this study is to gain deeper insights into how key covariates influence survival times and enhance risk stratification, ultimately leading to better treatment outcomes. We evaluated the fit of several AFTMs by comparing Log-Likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) values. Among the models tested, the Weibull AFTM emerged as the best fit, demonstrating superior Log-Likelihood and lower AIC and BIC values compared to other distributions. The study identified several significant predictors of survival in male diabetic patients, including age, family history of diabetes, duration of diabetes, early use of insulin (within one year of diagnosis), retinopathy, use of diuretics, proteinuria, hypertension, and fasting plasma glucose levels. All of these covariates were found to significantly influence survival outcomes (\(P<0.05\)), with older age and longer diabetes duration correlating with poorer survival rates. These findings underscore the importance of using AFTMs to model survival patterns in diabetic patients and the potential for targeted interventions based on identified risk factors. While the small sample size limits the generalizability of the results, this study provides valuable insights into the survival characteristics of male diabetic patients and highlights the need for further research with larger datasets. Future studies should consider a wider range of covariates and explore the application of AFTMs in broader diabetic populations to refine survival prediction models and optimize patient care. Additionally, integrating machine learning techniques with AFTMs could enhance predictive accuracy and provide a more comprehensive understanding of survival trends in diabetic patients.

PERFORMANCE EVALUATION AND OPTIMIZATION OF CRYSTALLIZATION AND SUGAR HANDLING SUBSYSTEMS IN SUGAR MILL PLANT USING PSO 503

Sandeep Kumar, Vikas Modgil

An appropriate maintenance strategy is necessary for sugar mill plants and particularly focuses on the core processing system. This research creates an availability-based simulation model for a sugar mill with 25,500 tons sugarcane crushing capacity. A Markov performance evaluation approach is used, and differential equations generated from a transition diagram covering fully operational, partially operational, and failure modes are included. Among the major processing units, the Evaporator is identified as a core processing component, being directly involved in sugar concentration.

In contrast, the Rotary Dryer and the Centrifugal Machine (Batch or Continuous Centrifuge) are classified as auxiliary units. These support the production process by reducing moisture content and separating sugar from molasses, respectively, without altering the sugar chemically. The Crystallizer (e.g., A-Pan Crystallizer) is also a core processing component, directly responsible for crystal formation. PSO is used to improve performance by optimizing the failure and repair rates and enhancing the performance metrics. Experimental results showed the system achieved 88.84% availability. These results allow maintenance strategies that focus on the most critical subsystems to be processing developed. A case for focused maintenance is made based on the evidence provided for more effective scheduling, ultimately improving system performance.

ASSESSING THE PREDICTIVE RESILIENCE OF COMP
EXPERIMENTAL DESIGNS UNDER MISSING OB
THE INFLUENCE OF A ALPHA VALUE VA ... 518

A.R. Gokul and M. Pachamuth

The field of robust m *research. Regarding*
Response Surf *experimental design, missing*
observ *challenging. In Response Surface*
arily used. These are the Central Composite
ution V. This study employs a range of alpha values
apabilities in the context of missing data. These alpha values
en *novel value termed by practical efficiency alpha, intended to rival the*
pra *ency is gauged through the minimax loss criterion, while the predictive*
capa *on of Design Space plot. The research findings suggest that are performs effectively*
across *es, even without specific observations.*

Withdrawal

FIXED POINT THEOREMS FOR SEQUENCE OF MAPPINGS IN TRICOMPLEX VALUED METRIC SPACE 535

Shivani Chourasiya, Kavita Shrivastava

Fixed point theory plays a crucial role in mathematical analysis, with profound applications in differential equations, optimization, and dynamic systems. Over the years, researchers have extended classical fixed point results to more complex structures, such as tricomplex valued metric spaces, to address problems in multidimensional and hypercomplex settings. Despite the growing interest in tricomplex variable spaces, existing literature on fixed point theorems within these frameworks remains limited by various constraints. The existing literature exhibits several limitations; therefore, in this article, we utilize the concepts of continuity and weak compatibility to establish fixed point results for a novel class of contraction mappings by employing a sequence of mappings within a tricomplex valued metric space. If two mappings commute at their point of coincidence, they are said to be weakly compatible. These results not only deepen the comprehension of fixed point theory but also open up new possibilities for its application in more intricate and varied mathematical contexts. As a result, our research progresses in the field, providing a solid foundation for future investigations and potential uses across multiple scientific and engineering domains.

DISCOUNT PRICING POLICY WITH STOCHASTIC DEMAND AND PRESERVATION TECHNOLOGY FOR DETERIORATING ITEMS 547

Ramkrishna Tiwari, Ram Kumar Tiwari, Lalji Kumar, Pooran Lal Prajapati, Utram Kumar Khedlekar

Managing inventory for deteriorating items presents significant challenges, especially in environments where demand is stochastic and price-sensitive. Retailers face the dual pressure of minimizing losses from product deterioration while ensuring sufficient stock availability to meet customer demand. Additionally, backlogged shortages, if not managed strategically, can erode customer satisfaction and profitability. These complexities are further compounded by the need

to strike a balance between promotional efforts, preservation investments, and pricing strategies to remain competitive in dynamic markets. To address these challenges, this study develops a comprehensive mathematical model aimed at optimizing replenishment strategies for deteriorating items. The model integrates preservation technologies and strategic price discounts to enhance product demand while accounting for the intricate relationship between market potential, stock levels, promotional efforts, and selling prices. The objective is to determine optimal pricing policies, order quantities, promotional expenditures, preservation costs, and schedules for shortages and replenishment, ultimately maximizing total profit per unit time. The findings reveal that total profit per unit time is concave concerning price, shortage duration, and inventory holding period. Numerical analyses demonstrate the efficacy of preservation technologies in mitigating deterioration rates, enabling retailers to adopt competitive pricing and stimulate sales. Additionally, the demand pattern index is shown to significantly influence inventory policies, while promotional campaigns play a critical role in augmenting product sales. Quantity discounts, tied to demand and order volumes, are identified as effective tools for reducing procurement costs and enhancing retailer profitability. These insights highlight the strategic importance of investing in preservation technologies, adopting dynamic pricing strategies, and implementing targeted promotional campaigns. Together, these approaches provide a robust framework for optimizing inventory management, increasing sales, and maximizing profitability in highly competitive markets.

CAUSES OF ELECTROMECHANICAL RESONANCE IN WIND TURBINES AND METHODS FOR ITS ELIMINATION 562

N.S. Mammadov

Electromechanical resonance in wind turbines is a complex and multifaceted phenomenon that significantly affects the performance, service life, and reliability of wind turbines. This problem is especially relevant for modern high-power wind turbines operating under variable loads and high air turbulence. The article provides a detailed analysis of the mechanisms of electromechanical resonance, including the influence of design parameters, aerodynamic loads, electromagnetic interactions, and operating conditions. The physical foundations of resonance phenomena and their impact on the dynamic behavior of wind turbines are considered, which allows for a deeper understanding of the mechanisms of their occurrence and the development of methods for their prevention. Particular attention is paid to modern methods of diagnostics and prediction of resonance effects. The existing technologies for suppressing resonance oscillations are analyzed, such as the use of damping elements, active vibration dampers, adaptive control systems, and digital twins, which allow for real-time prediction and correction of generator behavior. In addition, the article proposes innovative engineering solutions aimed at increasing the stability of wind turbines to resonance phenomena. These solutions include optimizing design parameters, selecting new high-strength and damping materials, and improving automated control systems. Research into this problem and developing effective measures to eliminate it are key aspects of further development of wind energy, allowing for an increase in the service life of equipment, its reliability, and a reduction in operating costs.

MINIMIZATION OF ACTIVE POWER LOSSES IN ALTERNATING CURRENT MAGNETIC SYSTEMS 570

G.V. Mamedova

In magnetic systems of electromagnetic alternating current devices, active power losses occur in steel and copper, reducing which is an urgent task. In order to minimize the loss of active power, a method based on the geometric optimization of the magnetic system has been developed. The optimization criteria are the coefficients of the proportionality principle. The last requirement is necessary for the proper coordination of the designed device with other mechanisms and for improving the technical and economic performance of the devices. One of the proportionality principles for electromagnetic devices is the construction of designs with a given coefficient $n_0 = h\omega_{co}$, where $h\omega_{co}$ and are the height and thickness of the winding, respectively, as the overall dimensions of the device are mainly determined by these parameters. As a rule, the value of the coefficient n_0 can be ensured after the preliminary calculation of the magnetic circuit. On the other hand, there are complex interrelations between the coefficient n_0 and the temperature rise of the winding. Excessive increase or decrease of the coefficient n_0 disrupts this principle, reduces the current density in the winding, increases the cross-section of the material and the winding capacity, and fails to achieve the specified temperature rise of the winding, etc. In this case, the main task of optimizing the winding dimensions is to ensure the maximum current density in the winding while maintaining the given coefficient n_0, temperature rise, and the ampere-turn values obtained from the preliminary calculation.

In this study, for the most common electromagnetic design, considering the given ratios of winding height to its thickness, temperature rise, and ampere-turns, the optimal parameters and dimensions of the magnetic system have been determined. Loss minimization has been carried out for electromagnetic devices with voltage winding and current winding.

STEADY STATE ANALYSIS OF BULK ARRIVAL, FIXED BATCH SERVICE QUEUE WITH INSPECTION, REWORK AND MULTIPLE VACATIONS 577

S. Karpagam, R. Lokesh

A non-Markovian bulk queueing system receives bulk arrivals through a single server that handles fixed-size batch services. The server conducts quality checks after each service to identify defective items, which are then sent to rework for quality control purposes. The server takes random-length vacations until new arrivals reach the threshold "K" before returning to service. Through the supplementary variable technique, the key performance metrics are computed together with steady-state governing equations. Using graphical representations and numerical tables, the model demonstration shows essential information about system performance through calculated data.

ANALYSIS OF THE PERFORMANCE ON SINGLE SERVER BATCH SERVICE MULTIPLE WORKING VACATION QUEUING MODEL WITH COMPULSORY AND EXTENDED REPAIR 590

Lidiya P, K Julia Rose Mary

In this paper, the concept of single-server multiple-working vacations queuing model with compulsory and extended repair is analyzed. In this model, customers arrive at a service facility and form a queue to be served by a single server. The arrival follows the Poisson distribution, and the service follows the exponential distribution. Batches of customers are served under the General Bulk Service Rule. In GBSR, rather than the individual customer arriving in a queue one by one, the customers arrive in groups or batches. Thus, each batch of services contains a minimum of 'a' units and a maximum of 'b' units of customers. In this study, two types of repairs, that is, compulsory and extended repair, are considered. The steady-state equation, the performance of measures for the system, and particular cases of described model are derived. Finally, in the form of tables and graphs, numerical results have been analyzed.

A NEW EXTENSION OF TWO PARAMETER ARADHANA DISTRIBUTION WITH COMPREHENSIVE STATISTICAL PROPERTIES AND APPLICATIONS 605

Rashid A. Ganaie, Manzoor A. Khanday, Prem Lata Gautam, R. Shenbagaraja, T. Vivekanandan

In this study, we introduce a modified extension of standard two parameter Aradhana distribution called as the length biased two parameter Aradhana distribution. The length-biased distributions are a subclass of weighted distributions that adjust the probabilities based on lengths or sizes of observed phenomena, making them suitable for modelling certain practical scenarios. The length biased version is identified as a specific instance of a broader category of weighted distributions, which are widely used in statistical modelling when data collection mechanisms introduce some inherent bias. The distribution is thoroughly examined and analyzed with new structural properties and parameter estimation is performed using a robust and widely used technique of maximum likelihood estimation. The distribution's superiority and effectiveness are evaluated through a comparative analysis involving two real-world datasets. These datasets are used to highlight the practical applicability of the new distribution, demonstrate its flexibility and ability to fit real-world data better than existing models. This comprehensive approach highlights both theoretical advancements and real-world relevance, emphasizing the distribution's versatility and utility.

LAKIBUL - G FAMILY OF DISTRIBUTIONS: SPECIAL MODELS, PROPERTIES AND APPLICATIONS 617

Iozhar A. Lakibul

In this study, a new generated family of distributions is introduced, called the Lakhul - G Family of Distributions. The T-X family of distributions is used in the derivation of the proposed class. Three special models of the proposed family are derived, such as the Lakhul - Uniform, Lakhul-Kumaraswamy, and Lakhul-Rayleigh distributions. Some properties of the proposed special models such as moments, mean, variance, moment generating function, survival, and hazard functions are studied. The Maximum likelihood approach is used to estimate the parameters of proposed models. Results show the following: (i) for Lakhul-Uniform distribution, it is found that this proposed probability distribution produces better estimate for electronic dataset as compared with the Cubic Transmuted Uniform distribution; (ii) for Lakhul - Kumaraswamy distribution is found to give a good fit for the COVID-19 dataset compared with the Marshall-Olkin reduce Kies and the Transmuted Kumaraswamy distributions; and (iii) for the Lakhul - Rayleigh distribution, it is observed that this distribution provides a better fit for precipitation dataset as compared with the Exponentialed Transform Inverse Rayleigh and the Exponentialed Inverse Rayleigh distributions.

MULTIVARIATE IMPUTATION BY MAHALANOBIS DISTANCE OPTIMIZATION (MIMDO) 632

Geoverr John D. Labita, Altea S. Labita, Bernadette F. Tubo

This paper introduces a new method for missing data imputation based on an optimization approach and is now available as an R package called "mimdo". This method deals with imputing the missing values by computing the values that minimize the Mahalanobis distance between an observation and the overall mean. The effectivity of mimdo was demonstrated in both classification and regression tasks using popular benchmark datasets. From all experiments, it was found out that using mimdo for imputing the missing values in the dataset, on the average, the classification rate is more than 80% and an R-squared of more than 50%. Furthermore, the consistency of the results were validated through simulation studies.

ENHANCING PREDICTIVE ACCURACY WITH ROBUST LIU AND RANDOM FOREST REGRESSION 642

Muthukrishnan.R and Karthika Ramakrishnan

Robust and Random Forest regression procedures are powerful machine learning algorithms that enhance predictive accuracy and handle complex datasets with diverse characteristics. The Robust Liu regression algorithm is designed to address the limitations of conventional regression methods. The robust Liu regression method is considered more reliable in the presence of outliers and multicollinearity. On the other hand, Random Forest is an ensemble learning method that constructs multiple decision trees to improve classification and regression tasks by aggregating predictions from individual trees, thereby mitigating overfitting and increasing robustness to noise. Its ability to capture nonlinear relationships and handle high-dimensional data makes it suitable for many real-world applications. This study compares the performance of Least Squares, Liu, Robust Liu, and Random Forest regression methods using prediction error measures in real and simulated environments. This synergy offers new possibilities for researchers to improve prediction accuracy in the presence of heterogeneous non-normal, multicollinear data and outliers.

ELECTROMECHANICAL DEVICES WITH LEVITATION ELEMENTS FOR CONTROL OF NON-ELECTRICAL PARAMETERS 648

G.S. Kerimzade

The paper presents a mathematical model of electromechanical devices with levitation elements for automatic control of non-electrical parameters, as well as calculation stages for any devices. The mathematical model is based on the equations of levitation, currents, magnetomotive forces, overheating temperature of the windings, and magnetic induction. Numerical values of dimensionless quantities are determined, which can be used as reference data in the

development of electromechanical devices with levitation elements. The functional dependences of electromagnetic parameters on the control voltage and stroke of the levitation element are established, and expressions for calculating the force of gravity of the levitation element, lifting force, magnetomotive force and winding overheating temperature are obtained.

GEOMETRIC HAZARD-BASED PREVENTIVE MAINTENANCE AND REPLACEMENT (GHBPMAR) MODEL FOR AGING MECHANICALLY REPAIRABLE MACHINES 655

Nse Udoh, Iniobong Uko, Fredrick Ohaegbunem

This work aims to develop a geometric hazard-based preventive maintenance model to improve the sustainability of aging mechanically repairable machines particularly in developing countries. The goal is to optimize maintenance and replacement schedules to meet user needs effectively. The model utilizes a geometric process as a hazard rate adjustment factor to account for increasing deterioration rates of machines due to aging and prolong usage and applies same model to a cassava grinding machine across three phases of lifespan: average lifespan, beyond average lifespan, and beyond initial replacement age. The Weibull failure distribution is used to characterize the machine's increasing failure rate. The result reveals that the machine's operational time decreases following preventive maintenance, and maintenance costs increase as the machine progresses through many phases of its lifespan. This highlights the model's effectiveness in managing the cost and timing of maintenance activities. The model's implementation can significantly impact communities in developing countries by improving the reliability and lifespan of essential machinery like cassava grinding machines and other industrial machines which can lead to stable local economies. This approach helps in managing costs and improving the sustainability of machines, especially in resource-constrained environments over extended lifespan.

MOMENTS OF GENERALIZED ORDER STATISTICS FROM DOUBLY TRUNCATED MODIFIED MAKEHAM DISTRIBUTION AND ITS CHARACTERIZATION 663

Abu Bakar, Zuber Akhter and Mohd Amir

The Makeham distribution has been extensively used in modeling actuarial data and helps as a fundamental life distribution. It describes a mortality failure law in which the hazard rate is composed of two components: a non-aging failure distribution and an aging failure distribution, the latter characterized by an exponentially increasing failure rate. The concept of generalized order statistics (gos), introduced by Kamps in 1995, provides a unified model for combining several models of ordered random variables, including order statistics, record values, progressive type II right-censored order statistics and more. The development of recurrence relations for the moments of gos is an important field of research due to the model's accessibility and flexibility. These recurrence relations simplify the computation of higher-order moments and establish the associations between moments of different orders. In this article, we have derived the recurrence relations of gos for single and product moments from the doubly truncated modified Makeham distribution and have also reduced these relations for specific ordered random schemes, such as order statistics, k-th record values and sequential order statistics for selected different values of parameters. In the final section, we have presented a characterization theorem based on the recurrence relations of the moments. This theorem has identified the doubly truncated modified Makeham distribution through its moment structure within the gos framework. This shows how important these relationships are for setting this distribution apart from other probability models.

A COMPREHENSIVE STUDY OF LIFETIME DISTRIBUTION MODELS FOR AIR CONDITIONING SYSTEMS USING FAILURE TIME DATA 677

S. Mythily, S.C. Premila, P. Manigandan, K. Geetha, D. Pachiyappan and A. Radhika

The purpose of this paper was to fit the model for the air conditioning system with failure time for the normal, gamma, and generalized gamma distributions and to identify the optimal distribution model for estimating the air conditioning system failure time. The performance of the normal, gamma, and generalized gamma distributions on air conditioning system failure data was investigated. To determine the optimal model for calculating the air conditioning system failure time, the secondary data was fitted to the normal, gamma, and generalized gamma distribution models. To evaluate the models performance, the Kolmogorov-Smirnov, and Akaike Information Criterion, were used. The outcomes of the research indicated that the gamma model had the lowest AIC and KS values, measuring 359.108 and 0.47624; the

generalized gamma model came in second with AIC and KS values are 359.6058 and 0.32988; and the normal model had the highest AIC and KS values, measuring 3203.729 and 0.4298, respectively. Results show that the gamma model has better flexibility in handling real data over the conventional normal and generalized gamma distribution.

RELIABILITY ANALYSIS OF AAC BLOCK PLANT USING BOOLEAN FUNCTION AND PATH TRACING METHOD 685

Bahram, S.C. Malik, A.D. Yadav and S. Malik

This study offers a comprehensive reliability analysis of Autoclaved Aerated Concrete (AAC) block plants, employing Boolean function technique and path tracing method. The analysis addresses a non-repairable system comprising five interconnected subsystems arranged in series. Utilizing these mathematical approaches, we derive expressions for reliability and mean time to system failure (MTSF). The study meticulously evaluates system reliability, with a focus on the Weibull distribution, especially the exponential distribution case.

RELIABILITY- AVAILABILITY - MAINTAINABILITY OF A REPAIRABLE SYSTEM USING MARKOV APPROACH 694

A. D. Yadav, S.C. Malik, S. Malik and N. Nandal

This study presents a reliability-availability-maintainability (RAM) analysis of a repairable system consisting of two non-identical units, where one unit operates as the main unit while another unit is in cold standby mode. A Markov approach is applied to develop the system model and transition probability equations are derived using the Chapman-Kolmogorov technique. The system is deemed to have failed only when both units are inoperative.

GARCH-MIDAS ANALYSIS OF INFLATIONS IMPACT ON ASIAN STOCK RETURN VOLATILITY 705

S. C. Premila, R. Jeena, D. Pachiyappan, B. Ramesh Kumar and S. Kavitha

This study examines the impact of inflation on stock return volatility in five Asian countries using a GARCH-MIDAS approach with mixed-frequency data (daily stock returns and monthly inflation). Results reveal that higher inflation significantly increases volatility and depresses returns, underscoring inflation's critical role in shaping market dynamics.

MATHEMATIC SIMLUATION OF SR. CONF. SYSTEM UNDER P. AND C. MAINTENANCE USING PSO 715

Shakuntla Singla, Diksha Mangla, Shilpa Rani

This paper addressed the performance of a well-maintained two-unit system by calculating reliability metrics like the average duration of system failure and the available time to perform it by counting the effect of the collapse rate and its separation facilities. The goal of this paper is to maximize the overall profit using a nature inspired optimization tool, particle swarm optimization by counting the reliability metrics for the performance of a series framework subjected to the influence of routine upkeep and changing of components.

ATTRIBUTE CONTROL CHART APPROACH FOR LIFE TESTING USING EXPONENTIAL-POISSON DISTRIBUTION 727

Gokila B and Sheik Abdullah A

This article introduces a novel attribute np control chart for monitoring the median lifetime of products under a hybrid censoring plan, assuming an Exponential-Poisson lifetime distribution. Optimal control chart parameters are derived to minimize deviations from a target Average Run Length (ARL) for an in-control process.

ESTIMATE MISSING VALUE OF RANDOMIZED BLOCK DESIGN USING VARIABLE CONTROL CHARTS THROUGH RESPONSE SURFACE METHODOLOGY 736

Kokkilambigai S and Pachamuthu M

In this paper, we are introducing the new methods to estimate the missing value using variable control charts such as 3-Sigma and 6-Sigma control charts through Response Surface Methodology (RSM) with the numerical example. RSM can be used to find factor systems that produce a desired maximum, minimum, or optimal response.

RELIABILITY AND PROFIT ANALYSIS OF REPAIRABLE EDIBLE OIL PLANT 750

Ashish Kumar, and Geetanjali Sharma

This paper described the reliability, availability and profit values of the edible oil plant by using regenerative point graphical technique. Generally, it contains four units such that bleacher (B), deodorizer (D), thermo fluid boiler (C) and pressured filter (A). The whole refinery plant work properly when all units work properly.

TOUCH AND STEP VOLTAGE HAZARD ANALYSIS 758

E.S. Safiyev, S.Y. Shikhaliyeva

This article analyzes the risks associated with electric shock, specifically focusing on the effects of touch and step voltage, which can lead to severe injury or death. The study highlights key factors influencing shock severity, including current magnitude, duration, body resistance, and current path.

RELIABILITY ESTIMATION AND PARAMETERS FOR THE LIFESPAN DISTRIBUTION BASED ON FAILURE TIME DATA FOR ELECTRICAL COMPONENTS . 769

K. Geetha, S.N. Manoharan, A. Radhika, S. Mythily, and P. Manigandan

The objective of the present study is to analyse the failure times of electrical components in order to identify the optimal distribution model for their longevity. Fifteen components failure data were analysed using the Weibull, Lognormal, and Exponential lifespan distribution algorithms.

COMPARATIVE STUDY OF REPAIRABLE JUICE PLANTS USING RPGT 776

Mohit Yadav, Vandana Swami, Naveen Kumar, Puneet Garg

In this paper, comparative study of repairable juice plants is analyzed by using regenerative point graphical technique. Generally, juice contains calcium, vitamin, iron, etc. to give the refresh tests. There are multiple steps to store the juice at large levels such as storing, grinding, pasteurization, etc.

CLUSTERING ELECTRIC POWER SYSTEMS INTO ADEQUACY ASSESSMENT AREAS . 784

Dmitry Krupenev, Denis Boyarkin, Nikolay Belyaev

The article discusses the issue of clustering electric power systems (EPS) into adequacy assessment areas to form equivalent models (EM) of power systems, with such models designed to calculate adequacy metrics. The level of detail of equivalent models is backed by the consideration of two properties of power systems being reliability and economic feasibility.

ON CHARACTERIZATION OF 2KTH ORDER WEIGHTED MAXWELL BOLTZMANN DISTRIBUTION 792

Nuzhat Ahad, S.P. Ahmad, J.A. Reshi

This paper presents the characterization results of the 2Kth order weighted Maxwell Boltzmann distribution (KWMBD). The characterization results are based on the coefficient of variance, a straightforward relationship between two truncated moments, the conditional expectation of a specific function of the random variable, the hazard function, and the reverse hazard function.

CERTAIN STATIONARY POINT OUTCOMES FOR ASYMPTOTICALLY REGULAR MAPS IN INTERVAL- VALUED N-FUZZY METRIC SPACE 803

Heera Ahirwar, Kavita Shrivastava, Harshit Khare

In this study, we introduce interval valued N-fuzzy metric space (IVNFM) and we verify certain popular fixed point results on the structure of interval valued N-fuzzy metric space through asymptotically regular mappings.

OPTIMIZATION OF PERISHABLE INVENTORY IN TWO-WAREHOUSE SYSTEM USING GENETIC ALGORITHMS WITH LOG-GAMMA DECAY AND NONLINEAR DEMAND .. 813

Garima Sethi, Ajay Singh Yadav, Chaman Singh

This study presents a detailed inventory framework for managing perishable products distributed across two storage facilities: a capacity-constrained, company-owned primary warehouse and an auxiliary rented warehouse. The model accounts for inventory shortages with partial backlogging, where both the demand rate and backlogging rate are modeled as generalized exponentially decreasing functions of time and selling price.

DESIGNING AN ATTRIBUTE CONTROL CHART BASED ON EXPONENTIAL-RAYLEIGH DISTRIBUTION UNDER HYBRID CENSORING 824

T. Kavitha & M. Gunasekaran

In this paper, we present a new attribute up control chart aimed at monitoring the median lifetime of products within a hybrid censoring scheme, based on the premise that the lifetimes of these products follow an Exponential-Rayleigh distribution.

A NEW LENGTH BIASED EXPONENTIAL- EXPONENTIAL DISTRIBUTION AND ITS APPLICATIONS TO CANCER DATA 834

Latha J, M Vijayakumar, Shibu D S

In this paper we propose a new length biased form of the Exponential-Exponential distribution called Length Biased Exponential-Exponential distribution (LBEED) and derive its statistical properties. The statistical properties including the moments, moment generating function, characteristic function, reliability functions and entropy measures are discussed.

A NEW GENERALIZATION OF AREA-BIASED XGAMMA DISTRIBUTION WITH PROPERTIES AND ITS APPLICATION TO CANCER DATA 844

Athira D V, P. Pandiyan, and D.S. Shibu

In this paper, we proposed a new model, "Area Biased Xgamma Distribution", a generalization of the Xgamma distribution. The various distributional properties survival function, hazard function, mean residual life, statistical

properties, moments, moment generating function and characteristic function, Bonferroni and Lorenz curve, entropy of the new model were studied.

APPLICATION OF TYPE-2 FUZZY TOPSIS METHOD FOR ESTIMATING RENEWABLE ENERGY SOURCES 857

Kamala Aliyeva

This paper offers the integration of spectrum selection optimization algorithms based on the evaluation of the characteristics of various renewable energy options using the TOPSIS fuzzy method.

A STATISTICAL ANALYSIS OF VARIANCE FOR ONE- WAY CLASSIFICATION USING TRAPEZOIDAL FUZZY TECHNIQUES AND RANKING METHODS 868

Vanitha. R, Pachamuthu. M, Kirthik Vaïramariappan. A, Kokkilambigai. S and Sona. S

This paper presents a statistical analysis of variance for one-way classification using trapezoidal fuzzy techniques and the ranking method, demonstrated through a numerical example.

ON A CLASS OF INDEFINITE KENMOTSU MANIFOLDS ADMITTING QUARTER-SYMMETRIC METRIC CONNECTION 882

K. L. Sai Prasad, S. Sunitha Devi

In this present paper, a class of almost contact metric manifolds equipped with an indefinite metric, termed indefinite Kenmotsu manifolds (known as, ϵ-Kenmotsu), is considered that accepts a connection of quarter-symmetric.

PROPERTIES AND APPLICATION OF ALPHA LOGARITHM TRANSFORMED SINE WEIBULL DISTRIBUTION 890

A.A. Osi, O.O. Ishaq, B.A. Abdulsalam, Usman Abubakar

In this research, we introduced a new submodel of the alpha logarithm transform sine-X family of distributions, called the alpha logarithm transform sine-Weibull. We derived its statistical properties, including survival function, hazard function, moments, moment-generating function, and order statistics.

A BATHTUB ADDITIVE FAILURE RATE MODEL FOR ACTUARIAL RELIABILITY 903

M. A. Alhassan, A. Yahaya, O. O. Ishaq, B. Abba, Y. Hussaini, A. Bello

This paper introduces a lifetime model named Hybrid Weibull-Exponential Power model, based on the later phenomenon vis-a-vis the former approach. The failure rate (FR) of the proposed model displayed suitable compound trends such as increasing-decreasing and various complex forms of bathtub-curve, with and without the lingering useful life period.

RESEARCH ON SIGNATURE RECOGNITION METHOD BASED ON DEEP LEARNING TECHNIQUE IN PYTHON LANGUAGE 918

Phuc Hau Nguyen

This paper presents a comprehensive study on the construction of a signature recognition system using deep learning with Python. We focus on the preprocessing of signature images, the construction of a CNN model architecture combined with RNN to exploit image features and stroke sequences, and the performance comparison between models.

BI STAGE FSSP WITH TRIANGULAR INTUITIONISTIC FUZZY NUMBER AND COMPARISON BETWEEN HEURISTICS 929

Pooja Kaushik, Sonia, Deepak Gupta, Sonia Goel

This study explores a two-stage FSSP, where parallel machines execute jobs in the both stages. The processing times in the flow shop scheduling are represented by fuzzy numbers in this work. This paper gives comparative study to address the FSSP under an uncertain environment based on triangular intuitionistic fuzzy numbers (T+IFNs).

A NEW DISCRETE ENTROPIC MODEL AND ITS APPLICATION TO CONTINGENCY TABLE INFERENCE 944

Poonam Kumari, Nishi Kumari

This paper introduces a novel discrete information entropic model and explores its applications in statistical analysis. We have demonstrated the efficacy of this new model by applying it to the inference of contingency tables, a cornerstone of statistical analysis.

RT&A, No 2 (84)
Volume 20, June 2025

STRUCTURAL STRENGTH AND STABILITY ASSESSMENT, RISKS AND POTENTIAL HAZARDS IN THE ARCTIC ENVIRONMENT

Alena Rotaru

●

All-Russian Scientific Research Institute for Civil Defence
and Emergencies of the EMERCOM of Russia
alenarotaru@mail.ru

Abstract

The paper deals with the assessment of the strength and stability of structures in Artic soil conditions. The main attention is paid to the unique properties of such soils, which affect the reliability of structures. These include extremely low temperatures, permafrost, high humidity and the presence of permafrost. It is noted that the existing methods for assessing strength and stability require further development and adaptation. The specifics of Artic soils should be taken into account when choosing research methods and approaches. The operation of structures in the Artic is associated with risks of deformation and destruction due to loads associated with soil features. Problems may also occur due to changes in temperature and humidity. To reduce risks and improve the reliability of structures, it is proposed to develop and implement comprehensive security measures. They should take into account the specifics of Arctic soils at all stages-from design to operation of infrastructure facilities. This may include the use of special materials, the adaptation of design solutions, and the development of new methods for monitoring the state of structure. The results of the study can be useful in the design and operation of facilities in the Arctic, as well as in the development of regulatory documents and standards for construction in difficult ground conditions. This will ensure the safety and efficiency of structures in extreme climatic conditions.

Keywords: strength, calculations, construction, artic soils, stability, safety.

I. Introduction

In the conditions of the Far North, the construction and operation of engineering structures face a number of unique challenges related to the peculiarities of Arctic soils. One of the key aspects determining the success of projects in this region is the calculation of the strength and stability of structures, taking into account the specifics of interaction with soils. The purpose of this study is to carry out calculations aimed at assessing the strength and stability of engineering structures in Arctic soil conditions. To achieve this goal, calculation methods have been developed that take into account the specific characteristics of the region. As part of the study, calculations were carried out aimed at determining the main parameters affecting the strength and stability of structures when interacting with Arctic soils. Aspects such as the mechanical properties of soils, their deformation characteristics, as well as the influence of climatic factors on soil behavior were studied. The results obtained make it possible to more accurately assess the behavior of engineering structures in the Far North and develop measures to ensure their strength and stability. Further research in this area can contribute to more efficient use of the Arctic territories and reduce the risks associated with the operation of engineering structures in this region.

II. Methods

Methods for calculating the strength and stability of structures in Arctic soils [6; 7; 19]:

1) Preparatory stage:
- Study of technical documentation and data on soil characteristics in the construction area.
- Collecting information about previous research and construction experience in similar conditions.
- Determination of the type and dimensions of the structure, as well as its operating conditions.

2) Strength calculation:
- Determination of the effective loads on the structure [1; 2; 3].
- Calculation of stresses in the structural material [4; 5; 8].
- Comparison of the obtained values with the permissible strength values [4; 16].
- If necessary, adjust the size or material of the structure.

3) Stability calculation:
- Determination of soil parameters (adhesion, angle of internal friction, density, etc.).
- Calculation of soil stability under the influence of structural loads [6].
- Comparison of the obtained values with the stability criteria.
- If necessary, additional soil studies [6; 7].

4) Analysis of the results:
- Evaluation of the obtained values of strength and stability.
- Comparison with the results of previous research and construction experience.
- Identification of possible risks and problems.

5) Recommendations:
- Development of measures to ensure the strength and stability of the structure.
- Making changes to the project or construction technology.
- Monitoring compliance with recommendations at all stages of construction.

6) Further research:
- Conducting additional studies of the soil and structures in similar conditions.
- Collecting data on the behavior of structures during operation.
- Analysis of the results obtained and development of recommendations for future projects.

This method makes it possible to more accurately assess the behavior of engineering structures in the Far North and develop measures to ensure their strength and stability. Further research in this area can contribute to more efficient use of the Arctic territories and reduce the risks associated with the operation of engineering structures in this region. With the help of non-destructive testing, you can find out the strength of reinforced concrete structures of a building and for this purpose an ultrasonic method was used which measured the speed of ultrasound vibrations with “UK-1401” tester (see Figure 1, 2) [8; 10; 12; 14].



Figure 1: *Building in the Arctic zone and in which the measurements of the structures were carried out*



Figure 2: *Ultrasonic tester UK-1401 [9]*

After ultrasonic measurements, the measurement points of reinforced concrete structures are shown in the diagrams (see figure 3-4).

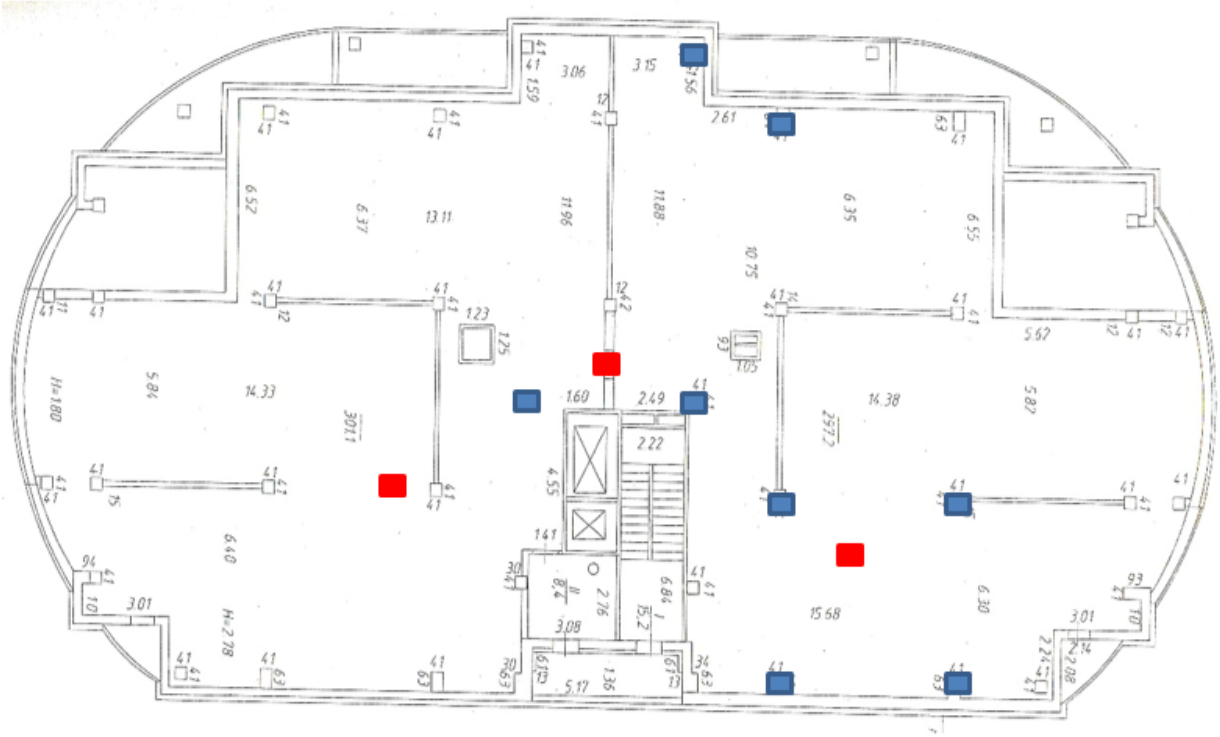


Figure 3: *The scheme of ultrasound examinations of the technical floor of the building (blue-reinforced concrete columns; red-reinforced concrete slabs)*

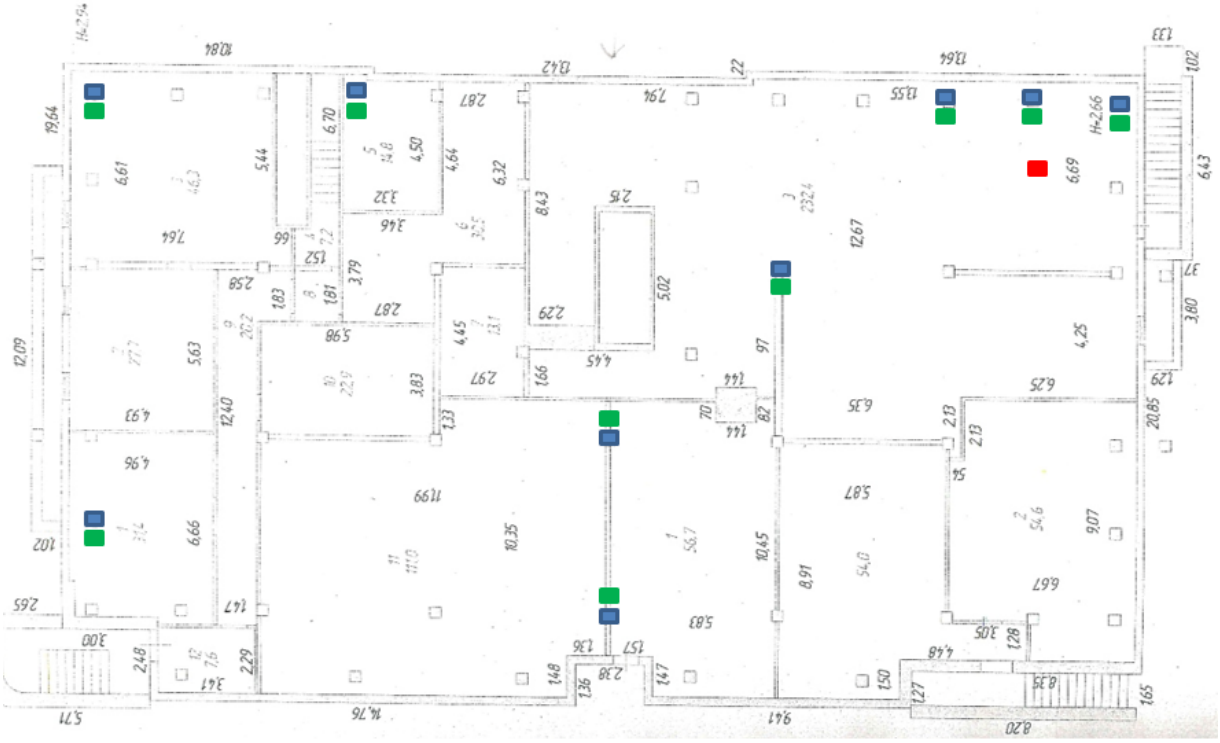


Figure 4: *The scheme of ultrasound examinations of the basement of the building (blue-reinforced concrete columns; green-reinforced concrete grilling's; red-reinforced concrete slabs)*

III. Results

To more accurately determine the strength of reinforced concrete structures, the obtained speed values are set in the formula

$$R=0,016xV-27,3 \quad (1)$$

According to the obtained data on the speed of ultrasound transmission, the formula can be used to estimate the strength of materials of structural elements of a building [1; 2]. The obtained strength values are converted to the concrete grade or class. The results of ultrasound examinations are presented in Table 1, 2.

Table 1: *The results of the ultrasound examination of the structural elements of the technical floor of the building (see mfigure 3)*

Segment	*Value*			*Arithmetic mean.*	*Sigma*	*0,95*	*Average, strength, MPa*
	1	*2*	*3*				
reinforced concrete column, point No 1							
1	2120	2130	2150	2133,33	15,28	2108	14,60
2	2090	2080	2090	2086,67	5,77	2077	14,39
3	2110	2120	2130	2120,00	10,00	2103	14,57
Mean value							**14,52**
reinforced concrete column, point No 2							
1	2560	2580	2590	2576,67	15,28	2551	17,70
2	2530	2510	2510	2516,67	11,55	2497	17,33
3	2050	2090	2080	2073,33	20,82	2038	14,12
Mean value							**16,38**
reinforced concrete column, point No 3							
1	2290	2230	2300	2273,33	37,86	2210	15,31
2	2300	2310	2330	2313,33	15,28	2288	15,86
3	2090	2080	2070	2080,00	10,00	2063	14,29
Mean value							**15,16**
reinforced concrete column, point No 4							
1	2110	2130	2150	2130,00	20,00	2096	14,52
2	2160	2170	2180	2170,00	10,00	2153	14,92
3	2230	2230	2210	2223,33	11,55	2204	15,28
Mean value							**14,91**
reinforced concrete column, point No 5							
1	2670	2630	2650	2650,00	20,00	2616	***18,16***
2	2610	2600	2610	2606,67	5,77	2597	18,03
3	2160	2180	2170	2170,00	10,00	2153	14,92
4	2130	2130	2140	2133,33	5,77	2124	14,71
Mean value							**16,46**
reinforced concrete column, point No 6							
1	2350	2370	2380	2366,67	15,28	2341	16,23
2	2300	2340	2330	2323,33	20,82	2288	15,87
3	2180	2180	2170	2176,67	5,77	2167	15,02
Mean value							**15,71**
reinforced concrete column, point No 7							
1	2180	2190	2200	2190,00	10,00	2173	15,06
2	2130	2140	2150	2140,00	10,00	2123	14,71
3	2130	2130	2170	2143,33	23,09	2104	14,58
Mean value							**14,78**
reinforced concrete column, point No 8							

Segment	*Value*			*Arithmetic mean.*	*Sigma*	*0,95*	*Average, strength, MPa*
	1	*2*	*3*				
1	1930	1970	1990	1963,33	30,55	1912	***13,23***
2	2020	2050	2050	2040,00	17,32	2011	13,92
3	2130	2130	2140	2133,33	5,77	2124	14,71
Mean value							**13,96**
reinforced concrete slabs, point No 1							
1	2170	2180	2170	2173,33	5,77	2164	14,99
2	2270	2230	2210	2236,67	30,55	2185	15,14
3	2500	2530	2580	2536,67	40,41	2469	17,13
4	2510	2510	2520	2513,33	5,77	2504	17,37
Mean value							**16,16**
reinforced concrete slabs, point No 2							
1	2310	2340	2380	2343,33	35,12	2284	15,84
2	2310	2290	2280	2293,33	15,28	2268	15,72
3	2130	2130	2140	2133,33	5,77	2124	14,71
4	2520	2510	2530	2520,00	10,00	2503	***17,37***
5	2410	2480	2430	2440,00	36,06	2379	16,50
6	2390	2370	2380	2380,00	10,00	2363	16,39
Mean value							**16,09**
reinforced concrete slabs, point No 3							
1	2180	2160	2170	2170,00	10,00	2153	14,92
2	2130	2150	2160	2146,67	15,28	2121	***14,69***
3	2230	2180	2170	2193,33	32,15	2139	14,82
Mean value							**14,81**

Conclusion:

- the strength of reinforced concrete columns has a range of values from 13,96 – 16,46 MPa (up to 15% of the maximum), the average values range from 13,23 – 18,16 MPa, which corresponds to concrete classes B10 – B12,5;

- the strength of reinforced concrete floor slabs has a range of values from 14.69 – 17.37 MPa (up to 15% of the maximum), the average values range from 14.81 – 16.16 MPa, which corresponds to concrete classes B10 – B12.5.

Table 2: *The results of the ultrasound examination of the structural elements of the basement of the building (see figure 4)*

Segment	*Value*			*Arithmetic mean.*	*Sigma*	*0,95*	*Average, strength, MPa*
	1	*2*	*3*				
reinforced concrete column, point No 1							
1	2530	2570	2540	2546,67	20,82	2512	17,43
2	2880	2890	2880	2883,33	5,77	2874	18,68
3	2700	2770	2730	2733,33	35,12	2674	18,57
Mean value							**18,22**
reinforced concrete column, point No 2							
1	2440	2360	2300	2366,67	70,24	2248	15,59
2	3020	3080	3030	3043,33	32,15	2989	20,53
3	2960	2940	2940	2946,67	11,55	2927	19,54
Mean value							**18,55**
reinforced concrete column, point No 3							
1	3030	3010	3050	3030,00	20,00	2996	***20,64***
2	2560	2570	2570	2566,67	5,77	2557	17,75
3	2540	2530	2410	2493,33	72,34	2371	16,45

Segment	*Value*			*Arithmetic mean.*	*Sigma*	*0,95*	*Average, strength, MPa*
	1	*2*	*3*				
Mean value							**18,28**
reinforced concrete column, point No 4							
1	2940	2950	2930	2940,00	10,00	2923	19,47
2	2880	2870	2830	2860,00	26,46	2815	17,75
3	2660	2640	2630	2643,33	15,28	2618	18,17
Mean value							**18,46**
reinforced concrete column, point No 5							
1	2200	2230	2240	2223,33	20,82	2188	15,17
2	2300	2310	2320	2310,00	10,00	2293	15,90
3	2340	2300	2310	2316,67	20,82	2282	15,82
Mean value							**15,63**
reinforced concrete column, point No 6							
1	2330	2340	2300	2323,33	20,82	2288	15,87
2	2350	2300	2220	2290,00	65,57	2179	15,10
3	2190	2180	2170	2180,00	10,00	2163	***14,99***
Mean value							**15,32**
reinforced concrete columns, point No 7							
1	2300	2310	2340	2316,67	20,82	2282	15,82
2	2320	2310	2310	2313,33	5,77	2304	15,97
3	2200	2180	2180	2186,67	11,55	2167	15,02
Mean value							**15,60**
reinforced concrete column, point No 8							
1	2850	2830	2840	2840,00	10,00	2823	17,87
2	2750	2740	2730	2740,00	10,00	2723	18,91
3	2700	2750	2720	2723,33	25,17	2681	18,61
Mean value							**18,46**
reinforced concrete column, point No 9							
1	2840	2830	2850	2840,00	10,00	2823	17,87
2	2400	2450	2470	2440,00	36,06	2379	16,50
3	2700	2750	2730	2726,67	25,17	2684	18,64
Mean value							**17,67**
reinforced concrete slabs, point No 1							
1	2300	2290	2300	2296,67	5,77	2287	15,86
2	2970	2950	2940	2953,33	15,28	2928	***19,54***
3	2280	2270	2270	2273,33	5,77	2264	***15,69***
Mean value							**17,03**
reinforced concrete grillwork, point No 1							
1	2370	2380	2330	2360,00	26,46	2315	16,06
2	2900	2970	2980	2950,00	43,59	2877	***18,72***
3	2350	2340	2330	2340,00	10,00	2323	16,11
Mean value							**16,96**
reinforced concrete grillwork, point No 2							
1	1950	1970	1970	1963,33	11,55	1944	13,46
2	2040	2040	2050	2043,33	5,77	2034	14,08
3	2240	2270	2240	2250,00	17,32	2221	15,39
Mean value							**14,31**
reinforced concrete grillwork, point No 3							
1	2110	2120	2120	2116,67	5,77	2107	14,60
2	2160	2130	2150	2146,67	15,28	2121	14,69
3	2020	2080	2060	2053,33	30,55	2002	13,86
Mean value							**14,38**

Segment	*Value*			*Arithmetic mean.*	*Sigma*	*0,95*	*Average, strength, MPa*
	1	*2*	*3*				
reinforced concrete grillwork, point No 4							
1	2140	2150	2170	2153,33	15,28	2128	14,74
2	2220	2300	2310	2276,67	49,33	2194	15,20
3	2350	2340	2330	2340,00	10,00	2323	16,11
Mean value							**15,35**
reinforced concrete grillwork, point No 5							
1	2460	2470	2480	2470,00	10,00	2453	17,02
2	2430	2430	2400	2420,00	17,32	2391	16,58
3	2110	2120	2150	2126,67	20,82	2092	14,49
Mean value							**16,03**
reinforced concrete grillwork, point No 6							
1	2140	2150	2170	2153,33	15,28	2128	14,74
2	2180	2150	2200	2176,67	25,17	2134	14,79
3	2220	2230	2220	2223,33	5,77	2214	15,34
Mean value							**14,96**
reinforced concrete grillwork, point No 7							
1	2090	2070	2080	2080,00	10,00	2063	14,29
2	2190	2170	2180	2180,00	10,00	2163	14,99
3	2200	2210	2220	2210,00	10,00	2193	15,20
Mean value							**14,83**
reinforced concrete grillwork, point No 8							
1	1940	1970	1950	1953,33	15,28	1928	***13,34***
2	2010	2020	2030	2020,00	10,00	2003	13,87
3	2050	2070	2030	2050,00	20,00	2016	13,96
Mean value							**13,72**
reinforced concrete grillwork, point No 9							
1	2140	2150	2170	2153,33	15,28	2128	14,74
2	2110	2120	2130	2120,00	10,00	2103	14,57
3	2170	2180	2190	2180,00	10,00	2163	14,99
Mean value							**14,77**

Conclusion:

- the strength of reinforced concrete columns ranges from 14.99 – 20.64 MPa (up to 27% of the maximum), with average values ranging from 15.32 – 18.55 MPa, which corresponds to concrete grades B12.5 – B15;

- the strength of reinforced concrete slabs has a range of values from 15.69 – 19.54 MPa (up to 20% of the maximum), the average value is 17.03 MPa, which corresponds to the concrete class B12.5;

- the strength of reinforced concrete grillings has a range of values from 13.34 – 18.72 MPa (up to 29% of the maximum), the average values range from 13.72 – 16.96 MPa, which corresponds to concrete classes B10 – B12.5.

IV. Conclusions

The following results were obtained during ultrasound method of examination (see Table 3) [13; 17; 18].

Table 3: *General results of ultrasound examination of building structural elements*

Segment	Structural element	Concrete class (taking into account the influence of the plaster layer)	Range of values in % of the maximum
Technical floor	reinforced concrete columns	B10-B12,5	To 15%
	reinforced concrete floor slabs	B10-B12,5	To 15%
Basement	reinforced concrete columns	B12,5-B15	To 27%
	reinforced concrete floor slabs	B12,5	To 20%
	reinforced concrete grillings	B10-B12,5	To 29%

According to the requirements of the norms p. 6.1.6 "For reinforced concrete structures, the concrete compressive strength class should be used not lower than B15" SP 63.13330.2018 "Concrete and reinforced concrete structures", the results of the survey show that the strength of structural elements exceeds the standard values [2; 11; 15].

As a result of calculations of calculations of structural strength and stability in Arctic soil conditions, the following conclusions can be drawn:

- Arctic soils have unique properties that significantly affect the strength and stability of structures. These properties must be taken into account when designing and building in the Arctic.
- Existing methods for assessing strength and stability in Arctic soil conditions have the following advantages: it has its own limitations and needs to be improved based on the specifics of the region.
- The operation of structures in the Artic is associated with a number of risks and potential hazards, including extreme climatic conditions, permafrost, cryogenic processes.
- To reduce risks and improve the reliability of structures, it is necessary to develop and implement comprehensive approaches that take into account the specifics of Artic soils. This may include the use of special materials, the adaptation of design solutions, and the development of new methods for monitoring and controlling the condition of structures.
- The results of the study can be used in the design and operation of facilities in the Arctic, as well as in the development of regulatory documents and standards for construction in difficult ground conditions. This will ensure the safety and efficiency of structures in the region.

References

[1] Maklakov, A.S. and Rotaru, A.N., The application of the ultrasonic method to assess the strength of structures made of different materials, in the collection: Construction and Architecture: theory and practice of seismic safety. Collection of scientific articles based on the results of the international scientific and practical conference dedicated to the memory of Doctor of Technical Sciences, Professor A.D. Abakarov, Makhachkala, 2023, 142-146.

[2] Rotaru, A.N., Measurements of bridge structures using non-destructive testing methods and their stability in wind gusts, Reliability: Theory & Applications, T. 18. No 4 (76), 2023, 1056-1066.

[3] Rotaru, A.N., Measurement and determination of strength of load-bearing structures materials by shear test method, Reliability: Theory & Applications, T. 19 No 4 (80), 2024, 919-923.

[4] Rotaru, A.N., Application of non-destructive testing methods and evaluation of condition of reinforced concrete framing, T. 19 No 2 (78), 2024, 400-407.

[5] Rotaru, A.N., Measuring and determining the strength characteristic of materials of load-bearing structures using the method of separation with chipping. Building materials, equipment, technologies of the XXI century, 5 (286), 2024, 31-33.

[6] Rotaru, A.N., and Maklakov, A.S. Monitoring of engineering and geological surveys in landslide areas to reduce risls to the public, Building materials, equipment, technologies of the XXI century, 3 (284), 2024, 59-65.

[7] Rotaru, A.N., Maklakov, A.S., Maklakova J. V.,Application of geophysical methods for operational assessment of ground conditions of residential buildings in permafrost conditions, Foundations, 4 (10), 2022, 25-27.

[8] Maierhofer, Ch., Reinhardt, H.-W., Dobmann, G., Non-destructive evaluation of reinforced concrete structures, Woodhead Publishing Limited, 2010.

[9] Kouche, El., Hassanein, H.S., Ultrasonic non-destructive testing using wireless sensor networks. Procedia Computer Science, 10, 136-143.

[10] Sanayei, M., Phelps, J.E., Sipple, J.D., Bell, E.S., Brenner, B.R., Instrumentation, nondestructive testing, and finite-element model updating for bridge evaluation using strain measurements, Journal of bridge engineering, 17(1), 2011, 130-138.

[11] Teplý, B., Keršner, Z., Rovnaník, P. and Chromá, M., Durability vs. Reliability of RC structures. Durability of Building Materials and Components, 2005, 17-20.

[12] Balaji Rao, K., Anoop, M.B., Lakshmanan, N., Gopalakrishnan, S. and Appa Rao, T.V.S.R., Risk-based remaining life assessment of corrosion affected reinforced concrete structural members, 2004.

[13] Mehta, P.K. and Gerwick, B.C., Concrete in service of modern world. Dundee, Scotland, UK: Univ. of Dundee, 1996, 1-28.

[14] Lang, Ch., Willmes, M., Non-destructive testing of reinforced concrete structures, Germany, 2018.

[15] Kairu, Wilson Macharia, Non-destructive of concrete structures using Schmidt Hammer and profometer 5+ , University of Nairobi, 2016, 32-41.

[16] Al-Mishhadanu S.A., Joni H.H., Radhi M.S., Effect of age on Non-destructive tests results for existing concrete, The Iraqi journal for mechanical and material engineering, vol. 12 (4), 2012.

[17] Akash J., Ankit K., Adarsh K., Yogesh V. and Krishna M., Combined Use of Non-Destructive Tests for Assessment of Strength of Concrete in Structure, Procedia Engineering 54, 2013, 241–251.

[18] ACI Committee 228, Report on Non-destructive Test Methods for Evaluation of Concrete in Structures, ACI 228.2R. American Concrete Institute, Farmington Hills, MI, 2013.

[19] Ngugi H.N., Effects of Sand Quality on Compressive Strength of Concrete: A Case of Nairobi County and Its Environs, Kenya Open Journal of Civil Engineering, 4, 2014, 255-273.

ASSESSING TRANSFORMATION RATIO OF VOLTAGE TRANSFORMER UNDER NON-SINUSOIDAL SUPPLY VOLTAGE CONDITIONS

Yusifbayli Nurali Adil[1], Guliyev Huseyngulu Bayram[2*],
Hajiyev Naib Ismihan[3], Huseynov Nijat Ramiz[4]

•

[1, 2*,3]Azerbaijan Technical University, Baku, Azerbaijan AZ1073, H. Javid avenue 25
[4]Sumgait State University, Sumgait, Azerbaijan AZ5008, Baku Street 1
[1]Emails: nyusifbeyli@area.gov.az, Orcid: 0000-0001-7948-4682
[2*]Emails: huseyngulu@mail.ru, Orcid: 0009-0006-7362-0619
[3]Emails: naib.haciyev.sdu@mail.ru
[4]Emails: nijat.huseyn.98@gmail.com

Abstract

Voltage transformers are rarely subjected to periodic inspection and testing while in operation, which is the reason for the difference between actual transformation ratios and their rated values. The frequency characteristics of voltage transformers are checked neither. Therefore, there can be large errors in the case of non-sinusoidal supply voltage. Existing studies on the study of voltage transformer faults reflect the results of the study of the equivalent circuit based on traditional lumped-parameter circuits. However, it should be noted that for high-frequency circuits and at the same time from the point of view of taking into account the structure of the windings, increasing the accuracy of the research results is an urgent issue. The case of non-sinusoidal supply voltage is of great scientific interest in this regard. Therefore, considering the voltage transformer as a distributed-parameter circuit, its fault under non-sinusoidal supply voltage was investigated and studied. This work explores errors of voltage transformers in the presence of non-sinusoidal voltage. To this end, a mathematical model of the transformation ratio is built considering the equivalent circuit of transformer as a distributed-parameter circuit.

Keywords: power quality, voltage transformer, lumped-parameter circuit, distributed- parameter circuit, non-sinusoidal voltage, transformation ratio

I. Introduction

One of the main requirements, which are imposed in power supply systems of industrial enterprises on results of experimental studies on metering electricity and monitoring its quality is the reliability of the results, i.e. correspondence of calculated and measured parameters to their real values. The reliability of calculation and measurement results depends both on the error of calculation methods (methodological error) and the error of information and measuring tools (measurement error). These issues are crucial when metering electricity to assess the distortion of the current and voltage curves [1-5]. GOST (State Standard) 3341-2013 establishes permissible error

standards for measurements of power quality indicators. This standard is ±10% for the total harmonic distortion and ±5% - for the harmonic factor of the n-th harmonic components [4–7].

The reliability of calculations of non-sinusoidality indicators based on instantaneous voltage (current) values and measurement results can be objectively assessed by comparing these indicators with their real values. In actuality, however, there is unknown information about amplitudes and phase shifts of individual harmonic voltage (current) components. At the same time, all instruments and computer systems for measuring power quality indicators are separated from the high-voltage network through voltage and current measuring transformers. The main difficulty encountered is that the voltage and current measuring transformers operating in substations are sources of measurement error themselves [8-12]. These errors are mainly due to the following factors:

- A transformer accuracy class;
- A mismatch between transformation ratio and its nominal value;
- A lack of information about transformation ratio for high harmonics.

The first two above-listed factors can be eliminated by initial testing and, if necessary, by replacing instrument transformers, whereas the third factor requires additional special research. High harmonic voltage (current) distorts the transformation ratios of measuring transformers, which also increases measurement errors. It must be borne in mind that inductive voltage transformers may experience resonance phenomena at frequencies above 1 kHz (harmonic 20 and higher), which is also an additional source of error [13-16].

Voltage transformers (VT) are rarely inspected and tested during operation, which leads to the discrepancy between real transformation ratios and their nominal values [17], [18-20]. In addition, it is also important to know the frequency characteristics of VT to assess the impact of the harmonic factors of the n-th harmonic components and the total harmonic distortion on VT transformation ratio. These VT characteristics have not been fully investigated. Therefore, the VT errors may exceed their permissible values not only for high but also for the fundamental harmonic voltage (current). As a result, the readings of measuring and metering instruments can be significantly distorted.

In the context of the above assumptions, the study on the VT frequency characteristics with a focus on the distortion of the supply voltage curve is a pressing issue.

II. Mathematical model of the transformation ratio for a VT lumped-parameter circuit

Some works address the assessment of the errors of instrument transformers in the electricity metering system [13], [16]. The error of the VT transformation ratio under a non-sinusoidal voltage is examined in [3], [4]. The VT equivalent circuit is considered as a lumped-parameter circuit. The following formula is obtained for the transformation ratio of the n-th harmonic (Figure 1):

$$K'_{t,n} = 1 + \frac{Z_{1n}}{Z_{mn}} + \frac{Z_{1n}Z'_{2n}}{Z_{mn}Z_{mn}} + \frac{Z_{1n} + Z'_{2n}}{Z'_{Hn}}, \quad (1)$$

where, $Z_{1n}; Z_{mn}$ are the complex impedance of the primary winding and magnetization for the n-th harmonic, respectively; Z_{2n} is the complex impedance of the secondary winding connected to the primary winding; Z'_{Hn} is the complex load impedance for the n-th harmonic ($n = \overline{2,40}$).

The study of VT transformation ratio considering voltage harmonic components based on the

VT lumped-parameter circuit indicates that in this case, large VT errors arise and it is impossible to use formulas (1) above the 13th harmonic. As a result, the information delivered when measuring distortion indicators and metering electricity in high-voltage circuits is incorrect.

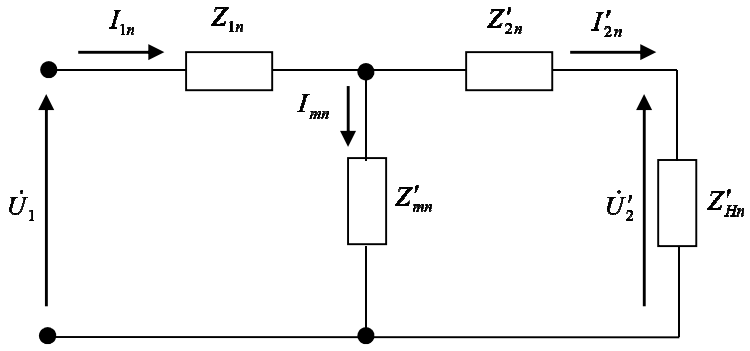


Figure 1: *Equivalent circuit of VT with lumped parameters*

III. Mathematical model of the transformation ratio for a VT distributed-parameter circuit.

To accurately determine the VT transformation ratio for high harmonics when dealing with a non-sinusoidal supply voltage, it is crucial to consider high voltage and high frequency circuits as distributed-parameter circuits. This means modeling VT under high frequencies and voltages as a transmission line with distributed parameters. This aspect of voltage transformers was not studied or was studied insufficiently. Therefore, comprehensive studies of VTs are required for high harmonic voltage with VT modeled as an element with distributed parameters.

The VT distributed-parameter circuit is shown in Figure 2.

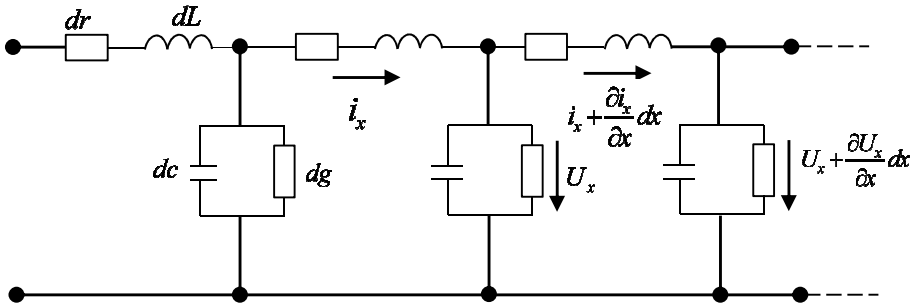


Figure 2: *The VT distributed-parameter circuit.*

As is known from the theory of electrical engineering, high-voltage and high-frequency circuits must factor in the wave nature of power transmission. In this case, to analyze the power transmission process, a line with length l is represented as a distributed-parameter circuit, where each element with an infinitesimal length dx has active resistance r_0dx, inductive reactance x_0dx, conductance g_0dx, and capacitance c_0dx, i.e., all parameters in the diagram in Fig. 1 are reduced to sections of infinitesimal length dx. Therefore, we can write the following equalities:

$$\left.\begin{aligned} dr &= r_0 dx \\ dL &= L_0 dx \\ dc &= c_0 dx \\ dg &= g_0 dx \end{aligned}\right\} \tag{2}$$

where r_0 is the linear active resistance of the line; L_0 is circuit inductance; c_0 capacitance between turns of windings; g_0 is conductance between the turns.

We assume that the VT parameters (active resistance and inductance, capacitance and conductance) are distributed evenly along the line.

In the long lines, at a frequency of 50 Hz, active and reactive currents do not remain constant. This, in turn, causes the voltage to change along the line. Therefore, by writing differential equations expressing the electrical state of a line with distributed parameters, in which voltage with a non-sinusoidal law of change is applied to the terminals, and solving these equations, one can determine the patterns of voltage and current distribution along the line.

Based on Figure 1, we can write the corresponding equations. To do this, the balance equations are written down for each element dx

$$\left.\begin{aligned} i_x &= i_x + \frac{\partial i_x}{\partial x}dx + u_x dg + dc\frac{\partial x}{\partial t} \\ u_x &= u_x + \frac{\partial u_x}{\partial x}dx + \left(i_x + \frac{\partial i_x}{\partial x}dx\right)dr + dL\frac{\partial}{\partial t}\left(i_x + \frac{\partial i_x}{\partial x}dx\right) \end{aligned}\right\} \tag{3}$$

This equation can be written as follows

$$\frac{\partial i_x}{\partial x}dx + u_x dg + dc\frac{\partial u_x}{\partial t} = 0$$

Let us divide both sides of the equation by dx. Then we have

$$\frac{\partial i_x}{\partial x} + u_x\frac{dg}{dx} + \frac{dc}{dx}\cdot\frac{\partial u_x}{\partial t} = 0\,.$$

From here we arrive at

$$\frac{\partial i_x}{\partial x} + g_0 u_x + c_0\frac{\partial u_x}{\partial t} = 0 \tag{4}$$

Accordingly, the following can be written for the second equation of system (3)

$$\frac{\partial u_x}{\partial x} + i_x\frac{\partial r}{\partial x} + \frac{\partial i_x}{\partial x}dr + \frac{\partial i_x}{\partial t}\cdot\frac{dL}{dx} + dL\frac{\partial i_x}{\partial x} = 0$$
$$\frac{\partial u_x}{\partial x} + r_0 i_x + \partial i_x r_0 + L_0\frac{\partial i_x}{\partial t} + dL\frac{\partial i_x}{\partial x} = 0$$

Here, given that $\partial i_x r_0 \approx 0,\ dL\frac{\partial i_x}{\partial x} \approx 0$, we obtain

$$\frac{\partial u_x}{\partial x} + r_0 i_x + L_0\frac{\partial i_x}{\partial t} = 0 \tag{5}$$

Thus, we have the following systems of partial differential equations:

$$\left.\begin{aligned}\frac{\partial u_x}{\partial x}+r_0 i_x+L_0\frac{\partial i_x}{\partial t}=0\\ \frac{\partial i_x}{\partial x}+g_0 u_x+c_0\frac{\partial u_x}{\partial t}=0\end{aligned}\right\} \tag{6}$$

In the case of non-sinusoidal voltage $\underline{U}$ and current $\underline{I}$, they can be expanded into a Fourier series

$$\left.\begin{aligned}\underline{u}=\sum_{n=1}^{\infty}U_{n,m}\exp(jnmt)\\ \underline{i}=\sum_{n=1}^{\infty}I_{n,m}\exp(jnmt)\end{aligned}\right\} \tag{7}$$

where $U_{n,m}=f_1(x)$ and $I_{n,m}=f_2(x)$. Let us obtain partial derivatives from equations (7)

$$\left.\begin{aligned}\frac{\partial u_x}{\partial x}=\sum_{n=1}^{\infty}\frac{dU_{n,m}}{dx}\exp(jnmt)\\ \frac{\partial i_x}{\partial x}=\sum_{n=1}^{\infty}\frac{dI_{n,m}}{dx}\exp(jnmt)\end{aligned}\right\} \tag{8}$$

Substituting equations (7) and (8) into (6), we obtain

$$\left.\begin{aligned}\sum_{n=1}^{\infty}\left[\frac{dU_{n,m}}{dx}+r_0 I_{n,m}+L_0 I_{n,m}jnm\right]\cdot\exp(jnmt)=0\\ \sum_{n=1}^{\infty}\left[\frac{dI_{n,m}}{dx}+g_0 U_{n,m}+c_0 U_{n,m}jnm\right]\cdot\exp(jnmt)=0\end{aligned}\right\}$$

It is known that $\exp(jnmt)\neq 0$. The sign of the sum is not considered. In this case, for the n-th harmonic voltage and current, the following expressions can be written:

$$\left.\begin{aligned}\frac{d\underline{U}_{n,m}}{dx}=r_0\underline{I}_{n,m}+jnmL_0\underline{I}_{n,m}=0\\ \frac{d\underline{I}_{n,m}}{dx}=g_0\underline{U}_{n,m}+jnmc_0\underline{U}_{n,m}=0\end{aligned}\right\}$$

From here, we obtain:

$$\left.\begin{aligned}-\frac{dU_{n,m}}{dx}=(r_0+jn\omega L_0)I_{n,m}\\ -\frac{dI_{n,m}}{dx}=(g_0+jn\omega c_0)U_{n,m}\end{aligned}\right\} \tag{9}$$

Considering that $r_0+jn\omega L_0=Z_0$ and $g_0+jn\omega c_0=Y_0$, we have the following systems of equations:

$$\left.\begin{aligned}-\frac{dU_{n,m}}{dx}=Z_0 I_{n,m}\\ -\frac{dI_{n,m}}{dx}=Y_0 U_{n,m}\end{aligned}\right\} \tag{10}$$

Equation (10) is a telegraph equation obtained for the n-th harmonic voltage and current

according to a VT distributed-parameter circuit and an applied non-sinusoidal supply voltage. We solve equation (10). To do this, we obtain the derivative of the first equation and take it into account in the second equation. Then we write

$$-\frac{d^2\underline{U}_{n,m}}{dx^2} = Z_0 \frac{d\underline{I}_{n,m}}{dx} \tag{11}$$

or

$$-\frac{d^2\underline{U}_{n,m}}{dx^2} = Z_0 Y_0 \underline{U}_{n,m} = 0 \tag{12}$$

Denoting $d/dx = P$, we obtain the characteristic equation

$$P^2 - Z_0 Y_0 = 0 \tag{13}$$

From equation (12), we obtain the following expression

$$P = \pm\sqrt{Z_0 Y_0} = \pm P_0$$

As is known, solution to (12) will be in the form

$$\underline{U}_{n,m} = A_{1,n} \exp(P_0 x) + A_{2,n} \exp(-P_0 x), \tag{14}$$

where P_0 is the solution to equation (13); $A_{1,n}$, $A_{2,n}$ are still unknown coefficients. These unknown coefficients are determined from the initial conditions. For x=0, we have

$$U\big|_{x=0} = U_{n,m,0} \text{ and } A_{1,n} + A_{2,n} = U_{n,m,0} \tag{15}$$

The second condition is written based on the first equation of system (10) for the case $x = l$. Then we have

$$\left.\frac{d\underline{U}_{n,m}}{dx}\right|_{x=l} = -Z_0 \underline{I}_{n,m}\big|_{x=l}$$

In the case of an open circuit at the end of the transformer, we have $i_2 = 0$, and since the VT operates in close-to–no-load conditions, we obtain:

$$\left.\frac{d\underline{U}_{n,m}}{dx}\right|_{x=l} = 0 \tag{16}$$

Calculating the derivative from equation (14) and substituting it into (16) we have:

$$\left.\frac{d\underline{U}_{n,m}}{dx}\right|_{x=l} = P_0\left(A_{1,n} \exp(P_0 l) - A_{2,n} \exp(-P_0 l)\right) = 0$$

Or

$$A_{1,n} \exp(P_0 l) - A_{2,n} \exp(-P_0 l) = 0 \Rightarrow A_{2,n} = A_{1,n} \exp(2P_0 l) \tag{17}$$

Substituting (17) into (15), we get:

$$A_{1,n} + A_{1,n}\exp(2P_0 l) = U_{n,m,0}$$

$$A_{1,n} = \frac{U_{n,m,0}}{1+\exp(2P_0 l)} \tag{18}$$

and

$$A_{2,n} = \frac{U_{n,m,0}\exp(2P_0 l)}{1+\exp(2P_0 l)} = \frac{U_{n,m,0}}{1+\exp(-2P_0 l)} \tag{19}$$

Given the obtained expressions for the integral constants $A_{1,n}$ and $A_{2,n}$ in (14), we have:

$$\underline{U}_{n,m} = U_{n,m,0}\left(\frac{\exp(P_0 x)}{1+\exp(2P_0 l)} + \frac{\exp(-P_0 x)}{1+\exp(-2P_0 l)}\right)$$

At the end of the line, voltage at $x = l$ will be $U_{n,m}\big|_{x=l} = U_{n,m,l}$. Then we have:

$$\begin{aligned} U_{n,m,l} &= U_{n,m,0}\left(\frac{1}{\exp(P_0 l)+\exp(-P_0 l)} + \frac{1}{\exp(P_0 l)+\exp(-P_0 l)}\right) = \\ &= \frac{2U_{n,m,0}}{\exp(P_0 l)+\exp(-P_0 l)} = \frac{U_{n,m,0}}{ch(P_0 l)} \end{aligned} \tag{20}$$

Let us determine the VT transformation ratio for the *n*-th harmonic:

$$K_{t,n} = \frac{U_{n,m,0}}{U_{n,m,l}} = ch(P_0 l) \tag{21}$$

To obtain a more explicit formula, we simplify the expression of the hyperbolic cosine:

$$P = \sqrt{Z_0 Y_0} = \sqrt{(r_0 + jnx_0)\cdot(g_0 + jnb_0)} = \sqrt{(r_0 + jnx_0)\,jnb_0} = \sqrt{\frac{nb_0}{2}(1+j)}\cdot\sqrt{r_0 + jnx_0}$$

In almost all cases, $g_0 \approx 0$ is assumed. Then, given the following

$$\sqrt{j} = \sqrt{\exp(j90^0)} = \exp(j45^0) = \frac{\sqrt{2}}{2} + j\frac{\sqrt{2}}{2} = \frac{1+j}{\sqrt{2}}$$

and after some transformation, we obtain:

$$P_0 = \sqrt{2nb_0}\cdot\sqrt[4]{r_0^2 + (nx_0)^2}\cdot\left(\frac{r_0 - nx_0}{\sqrt{r_0^2 + (nx_0)^2}} + j\frac{r_0 + nx_0}{\sqrt{r_0^2 + (nx_0)^2}}\right)$$

Thus, we arrive at:

$$P_0 \cdot l = \sqrt{2nb_0} \cdot \sqrt[4]{r_0^2 + (nx_0)^2} \cdot \left(\frac{r_0 - nx_0}{\sqrt{r_0^2 + (nx_0)^2}} + j \frac{r_0 + nx_0}{\sqrt{r_0^2 + (nx_0)^2}} \right) \cdot l = a_n + jb_n$$

The hyperbolic cosine argument is complex and therefore we can write $ch(P_0 l) = ch(a_n + jb_n)$. After the transformation, we obtain

$$ch(a_n + jb_n) = \sqrt{(ch(a_n) \cdot \cos(b_n))^2 + (sh(a_n) \cdot \sin(b_n))^2} = \sqrt{ch^2(a_n) - \sin^2(b_n)}$$

The VT transformation ratio for the *n*-th harmonic voltage (current) is finally expressed as follows:

$$K_{t,n} = ch(P_0 l) = \sqrt{ch^2(a_n) - \sin^2(b_n)} \tag{22}$$

where

$$a_n = \sqrt{2nb_0} \cdot \sqrt[4]{r_0^2 + (nx_0)^2} \cdot \left(\frac{r_0 - nx_0}{\sqrt{r_0^2 + (nx_0)^2}} \right) \cdot l = \frac{\sqrt{2nb_0} \cdot (r_0 - nx_0)}{\sqrt[4]{r_0^2 + (nx_0)^2}} \cdot l \tag{23}$$

$$b_n = \frac{\sqrt{2nb_0} \cdot (r_0 + nx_0)}{\sqrt[4]{r_0^2 + (nx_0)^2}} \cdot l \tag{24}$$

Thus, under a non-sinusoidal supply voltage, the VT transformation ratio for the n-th harmonic voltage components can be estimated by formulas (22) – (24) expressed with distributed parameters. As can be seen from these formulas, the expressions are simple and this simplifies the computational procedures.

IV. Modeling results

Calculations were carried out to estimate the transformation ratios for the voltage transformer of type MOM-6, using the obtained formulas (22) – (24). The calculation results are shown in Figure 3. Comparative curves $K_{t,n} = f(n)$ obtained using formulas (1) and (22) show that when estimating the VT transformation ratio using formulas (22) - (24), which are obtained using a VT distributed-parameter circuit, the error of the measuring instruments is significantly reduced, and the accuracy of the electricity metering instruments is increased.

As follows from Figure 3, when using formula (1) (curve 1), significant distortions occur at $n > 15$, and when using formula (2) (curve 2), the error up to the 27th harmonic inclusive can be considered acceptable. At $n > 27$, the error is tens of percent.

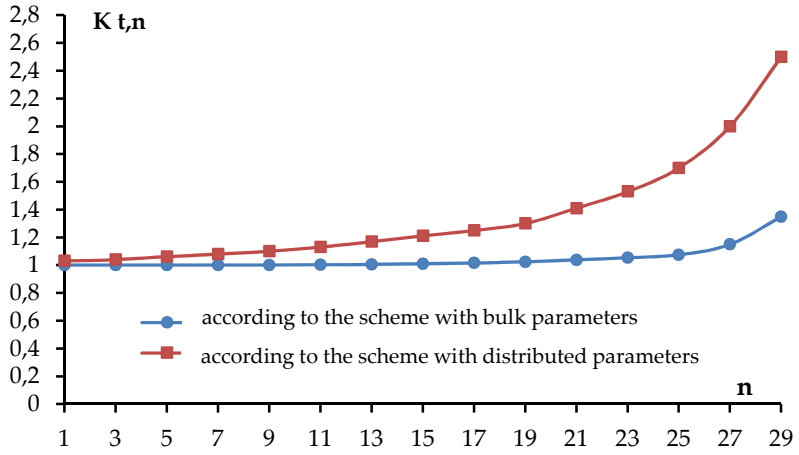


Figure 3: *VT transformation ratio curves for high harmonics*

Figure 4 shows the curve of a change in the relative error of the VT transformation ratio calculated using formulas (1) and (24) depending on the number of harmonics. The comparative analysis of the error indicates that at $n>11$, the calculation error is more than tens of percent.

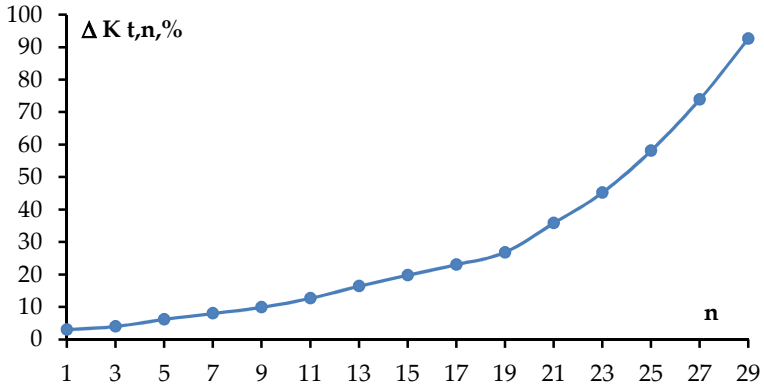


Figure 4: *The curve of a change in the relative error in determining the VT transformation ratio*

Therefore, it is essential to conduct additional research to ensure the reliability of measurements when using instrument transducers.

Conclusion

1. A mathematical model of the transformation ratio is obtained based on a distributed-parameter circuit for the voltage transformer used in the electricity metering system in the presence of harmonic distortion of the supply voltage, which allows boosting the reliability of the calculation and measurement results.

2. The obtained comparative curves suggest that when estimating transformer ratios for the harmonic components of the supply voltage using formulas obtained using a distributed-parameter circuit, the error in the measuring instruments is significantly decreased and the accuracy of the electricity metering instruments is increased. A comparative analysis of the voltage transformer error shows that at $n>11$, the error in calculation and measurement can amount to more than tens of percent.

3. These findings emphasize the importance of conducting further research to ensure the reliability of measurements using instrument transducers based on instrument transformers.

References

[1] Huanghui Zhang, Haiming Shao, Jiafu Wang. The Research on Additional Errors of Voltage Transformer Connected in Series // Published in: IEEE Access (Volume: 8), February 2020, IEEE Access PP(99):1-1, DOI: 10.1109/ACCESS.2020.2972304, Electronic ISSN: 2169 -3536, https://ieeexplore.ieee.org/document/8986609

[2] Maja Grbić, Dragutin Salamon, Jovan Mikulović. Analysis of influence of measuring voltage transformer ratio error on single-circuit overhead power line electric field calculation results// Electric Power Systems Research, Volume 166 , January 2019, Pages 232-240, https://www.sciencedirect.com/science/ article/abs/pii/S0378779618303183

[3] Artemov AI, Andreenkov ES, Errors of measuring voltage transformers with nonlinear loading//Electrical equipment: Exploitation and Repair, No. 3, 2017, https://panor.ru/articles/pogreshnosti-izmeritelnykh-transformatorov-napryazheniya-pri-nelineynoy-nagruzke/50668.html

[4] Andreenkov ES, Artemov AI Utochnenie ergorestei three-winding transformers of voltage // Information technologies, energy and economics. Proceedings of the XII International scientific and technical conference of students and postgraduates. In 3 Volumes, Smolensk: Universum, 2015, P. 6–10.

[5] GOST (State Standard) 30804.4.7–2013. Electromagnetic compatibility of technical means. General guide to measuring instruments and measurements of harmonics and interharmonics for power systems and technical means connected to them. M.: Standartinform. 2013.

[6] R. Ghosh, B. Chatterjee, and S. Chakravorti, Investigations on the effect of voltage harmonics on leakage current for condition monitoring in insulators, in 2016 IEEE 6th International Conference on Power Systems, ICPS 2016, 2016.

[7] B. Singh and V. Verma, An improved hybrid filter for compensation of current and voltage harmonics for varying rectifier loads, Int. J. Electrical Power & Energy Systems., Vol . 29 , No. 4, 2007, pp. 312-321.

[8] B. Suechoey, C. Thammarat , M. Leelajindakraireak. Estimation of core loss of transformer under non-sinusoidal voltage supply // International Conference on Power System Technology, 2004. PowerCon 2004 . Vol.1.

[9] Michal Kaczmarek and Ernest Stano. Application of the Sinusoidal Voltage for Detection of the Resonance in Inductive Voltage Transformers//*Energies* 2021, *14* (21), 7047, pp.1-16, https://doi.org/10.3390/en14217047

[10] Simkin LA, Domanov VF, Avanesov VM, Plahotnyuk AN Experience in operation and verification of measuring transformers of electric energy accounting systems // Industrial Energy, 2000, No. 10, pp. 17-21.

[11] Anders Bergman . Estimation of errors of voltage transformers on-site and the uncertainty of the estimate // Conference on Precision Electromagnetic Measurements. Conference Digest. CPEM 2000 (Cat. No.00CH37031), 14-19 May 2000, Sydney, NSW, Australia, https://ieeexplore.ieee.org/document/851052

[12] Feng Zhou, Peng Zhao, Min Lei, Changxi Yue, Jicheng Yu. Capacitive voltage transformer measurement error prediction by improved long short-term memory neural network// Energy Reports, Volume 8, Supplement 10, November 2022, pp. 1011-1021, https://www.sciencedirect.com/science/article/pii/ S235248472201109X

[13] Guliyev H.B., Sadykhov E.V., Dzhabbarova S.M. Reliability of calculation and measurements of indicators of sinusoidal distortion // Nauchnye Izvestiya, SDD, 2001, No. 4, pp. 77-81.

[14] Guliyev H.B. Evaluation of the transformation coefficient of the voltage transformer during polyharmonic supply voltages // Nauchnye Izvestiya, SDD, 2008, No. 1, pp.94-100.

[15] Andreenkov E.S., Artemov A.I. Estimation of errors of voltage transformers under non-sinusoidal load // Bulletin of the Kuzbass State Technical University, 2016, pp.97-103, https://elibrary.ru/item.asp?id=28963215

[16] Michal Kaczmarek. Measurement error of non-sinusoidal electrical power and energy caused by instrument transformers // *IET Generation, Transmission & Distribution, 2016, 10 (14),* pp.3492 -3498, https://doi.org/10.1049/iet-gtd. 2016.0131

[17] Jayanth R. Ramamurthya, Nicola Chiesa B., Hans K. Høidalenb, Bruce A. Morka, Nils M. Stenviga, Adam C. Mantya. Influence of voltage harmonics on transformer no-load loss measurements and calculation of magnetization curves // Electric Power Systems Research 146 , 2017, p.43–50, https://daneshyari.com/article/preview/5001105.pdf

[18] Jawad Faiz , Milad Ghazizadeh , Hashem Oraee . Derating of transformers under non-linear load current and non-sinusoidal voltage – an overview // IET Electric Power Applications, Volume 9, Issue 7, 2015, pp.486-495, https://doi.org/10.1049/iet- epa.2014.0377

[19] Anh -Tuan Phung, Hoang-Phuong Vu, Trinh-Tuan Nguyen, Dang- Hai Nguyen. Error analysis for inductive current transformers under non-sinusoidal waveform current// The University of Danang - Journal of Science and Technology, 3(124), 2018, pp.49-53, https://media.neliti.com/media/publications/441077-error-analysis-for-inductive-current-tra-cff611bf.pdf

[20] Seyed Reza Karimpour-Mouziraji, Iman Ahmadi-joneidi, Amir Shayegani-Akmal. Evaluation of 3 Phase Dry Transformer under Non-Sinusoidal Load Currents and Unbalanced Voltage Supply//The 23rd International Conference on Electricity Distribution, Lyon, 15-18 June 2015, Paper 0153, p.1-4, http://cired.net/ publications/cired2015/papers/CIRED2015_0153_final.pdf

ENHANCING SENSITIVITY OF MTTF AND NETWORK RELIABILITY STABILITY OF K-OUT-OF-N ACYCLIC TRANSMISSION NETWORK

Hemalatha.G[1], Vijayalakshmi.G[2*]

•

[1, 2*] Department of Mathematics, College of Engineering and Technology,
SRM Institute of Science and Technology, Kattankulathur,
Tamil Nadu- 603 203, India.
[1] hg9080@srmist.edu.in, [2*]vijayalg02@gmail.com

Abstract

Networks are used in many areas of engineering, including computers and communication, transportation and electric transmission. Component importance measures are applicable to improve the system design. The reliability characteristics of k-out-of-n acyclic transmission network have been evaluated using universal generating function (UGF). This paper explores and analyses Mean Time To Failure(MTTF), Birnbaum importance measurement, critical component importance, component risk growth factor, average risk growth factor and MTTF sensitivity for acyclic transmission network. Finally, The stability illustrates the impact of component failure on the reliability of the network system. The network reliability stability for acyclic transmission network is computed. In this work, we have evaluated reliability characteristics and of the acyclic transmission network using UGF method. The network reliability stability for k-out-of-n acyclic transmission network is computed. Furthermore, numerical illustrations are provided.

Keywords: Birnbaum importance measurement, risk growth factor, average risk growth factor, MTTF sensitivity, universal generating function, critical component importance.

I. Introduction

Reliability is a critical aspect of network systems, ensuring their ability to function under various conditions. Over the years, significant efforts have been made to improve network reliability using advanced methodologies. The universal generating function (UGF), introduced by Ushakov [15], was a pivotal development for system reliability evaluation. Levitin and Lisnianski [5] utilized UGF for significance and sensitivity analysis of multistate systems, while Levitin[6] extended its application to acyclic networks. Lisnianski and Levitin [7] further applied UGF to analyze real-world systems with distinct failure modes. Consistency in signal transmission from source to destination, known as terminal reliability *or* network reliability, has been extensively discussed in Colbourn [2]. Yeh [16] evaluated multi-state node network reliability using traditional binary-state methods, while Yeh [17] enhanced the UGF method for general multi-state networks, offering practical solutions for reliability evaluation.

Birnbaum [1] first introduced component importance measures, highlighting the role of individual components in system reliability. Building on this, Gao and Yao [3] explored the importance indices of components in uncertain systems, while Zhu et al. [21] tackled multi-type component assignment problems using Birnbaum importance heuristics. He et al. [4] employed cellular automata to measure component importance in network systems, advancing the understanding of individual component contributions.

In network reliability studies, Trivedi et al. [14] emphasized the mathematical modeling of reliability in modern communication systems, while Negi and Singh [9] analyzed non-repairable complex systems with weighted subsystems. Yeh [18] used a path-based UGF approach to examine network reliability and multistate connectivity, particularly in binary-state systems.

Complex system reliability has been studied using diverse modeling techniques. Ram and Singh [10] applied copula-based approaches to assess engineering systems, and Ram [11] reviewed various reliability methods. Ram and Manglik [12], [13] used stochastic and extra variable techniques to evaluate industrial systems, emphasizing critical components.

Graph-theoretic methods were applied by Zarghami et al. [20], who evaluated infrastructure network reliability by analyzing series and parallel structures. Yeh [19] further refined the UGF algorithm for acyclic binary-state network reliability, enabling better performance assessment.

These studies underscore the value of tools like UGF, importance measures, and advanced modeling techniques in improving network reliability. Their contributions have significantly enhanced system robustness, minimizing failures and ensuring stable operations across multiple domains.

Notations and Abbreviations:

$u(z)$	Universal Generating Function of z^{th} vertex of the network
$U(z)$	Universal Generating Function of z^{th} subnet vertex of the network
$\otimes$	Composition operator
$\mathrm{I}_i^{\mathrm{BM}}(t)$	Birnbaum component importance of i^{th} component
$\mathrm{I}_i^{\mathrm{CIM}}(t)$	Critical importance measure of i^{th} component
$RGF_i(t)$	Risk growth factor of i^{th} component
ξ	Function operator
R	Reliability of the suggested network
N	Number of sink vertices in the suggested network
L	Number of vertices in the suggested network
$\eta_{i:j}$	Failure rate of flow from vertex i to j
h(t)	Hazard rate
t	Mission time
p_n	Probability that x which equ als x_n
$\mathrm{p}_{n,\chi^{m:n}}$	Probability of a set of vertices $\chi^{m:n}$ getting a signal directly from the vertex at f_m
$\mathrm{q}_{n,\chi^{m:n}}$	Probability that the set of vertex $\chi^{m:n}$ cannot receive a signal directly from the vertex at f_m
f_m	m th vertex of the suggested network
$\chi^{m:n}$	A signal is received by a set of vertices from the vertex at f_m
UGF	Universal generating function
MTTF	Mean Time To Failure

II. Network Reliability using the UGF

Ushakov[15] first presented the UGF method for determining a system's or network's reliability. A detailed discussion of UGF, composition operator, and network reliability is provided by Levitin[8].

The UGF of a discrete random variable is defined by a polynomial as follows:

$$u(z) = \sum_{n=1}^{N} \mathrm{p}_n z^{x_n} \tag{1}$$

where p_n is the probability of x, which is equal to x_n and x has n probable values.

Let us suppose that a vertex is placed at f_m. If in-state n is the vertex available for signal transmission from f_m to a set of vertices $\chi^{m:n}$, then $\mathrm{v}_{m:n} = \begin{cases} 1 & \text{if } f_m \in \chi^{m:n} \\ 0 & \text{if } f_m \notin \chi^{m:n} \end{cases}$ (2)

The vertex-UGF of $v_{m:n}$ elements are expressed as:

$$u_i(z) = \sum_{n=1}^{N} p_{n,\chi^{m:n}} q_{n,\chi^{m:n}} z^{v_{m:n}} \quad (3)$$

The subnet UGF of $U_{i+1}(z)$ elements are expressed as:

$$U_{i+1}(z) = \{\xi[U_i(z)], u_{i+1}(z)\} \quad (4)$$

- If the path fails to pass through the regarded vertex (unit) and if the path fails to complete the connection from the source vertex to the considered vertex, remove the term from UGF from the operator η.
- Gather every term that is similar for every vertex in the generated UGF.

I. Algorithm for the Reliability of Networks

The following algorithm is created to assess network reliability for a binary state network:

Step 1: Find $v_{m:n}$ corresponding to $\chi^{m:n}$ for the vertices located at the positions f_1, …… , f_{L-N} in the network.
Step 2: Determine $u_i(z)$ of the vertices located at f_1, …….. , f_{L-N} places using equation(3).
Step 3: Set $U_1(z) = u_1(z)$.
Step 4: Determine the value of $U_{i+1}(z) = \{\xi[U_i(z)], u_{i+1}(z)\}$ where $i = 1$ to $L - N - 1$.
Step 5: Obtain the network reliability at the sink vertices by simplifying polynomial $U_{L-N}(z)$ and applying operator ξ.

II. Model Description

Moving signal flows can occur under the following conditions:

- If the signal flow from vertex i to vertex j is successful and vertex i to vertex k fails then the probability becomes $p_{i:j}(1 - p_{i:k}) = p_{i:j} q_{i:k}$ (5)
- If signal flow at a vertex k is interrupted, then the probability becomes $p_{k:\varphi}$, where k =1, 2, ……, $L - 1$.
- If the signal flows from vertex 1 to vertices 2 and 3 then probability becomes

$$p_{1:[2,3]} = p_{1:2} p_{1:3}. \quad (6)$$

III. Numerical Illustration

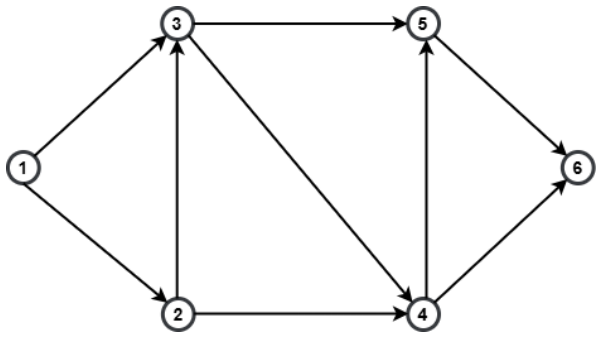


Figure 1: *k-out-of-n network*

Step 1: Consider the k-out-of-n network depicted in Figure 1, where signal flow starts at vertex 1 and ends at vertex 6. Here $L = 6$ and $N = 1$
Step 2: The Universal generating function of all the vertices are expressed as:

$$u_1(z) = p_{1:\varphi}\ q_{1:2}\ q_{1:3}\ z^{000000} + p_{1:2}\ q_{1:3}\ z^{010000} + p_{1:3}\ q_{1:2}\ z^{001000} + p_{1:[2,3]}\ z^{011000} \quad (7)$$

$$u_2(z) = p_{2:\varphi}\ q_{2:3}\ q_{2:4}\ z^{000000} + p_{2:3}\ q_{2:4}\ z^{001000} + p_{2:4}\ q_{2:3}\ z^{000100} + p_{2:[3,4]}\ z^{001100} \quad (8)$$

$$u_3(z) = p_{3:\varphi}\, q_{3:4}\, q_{3:5}\, z^{000000} + p_{3:4}\, q_{3:5}\ z^{000100} + p_{3:5}\, q_{3:4}\ z^{000010} + p_{3:[4,5]}\, z^{000110} \quad (9)$$

$$u_4(z) = p_{4:\varphi}\, q_{4:5}\, q_{4:6}\, z^{000000} + p_{4:5}\, q_{4:6}\ z^{000010} + p_{4:6}\, q_{4:5}\ z^{000001} + p_{4:[5,6]}\, z^{000011} \quad (10)$$

$$u_5(z) = p_{5:\varphi}\, q_{5:6}\, z^{000000} + p_{5:6}\, z^{000001} \quad (11)$$

Step 3: The UGF of vertex 1 is $U_1(z) = u_1(z)$

$$U_1(z) = p_{1:\varphi}\, q_{1:2}\, q_{1:3}\, z^{000000} + p_{1:2}\, q_{1:3}\ z^{010000} + p_{1:3}\, q_{1:2}\ z^{001000} + p_{1:[2,3]}\, z^{011000} \quad (12)$$

$$\xi[U_1(z)] = p_{1:2}\, q_{1:3}\ z^{010000} + p_{1:3}\, q_{1:2}\ z^{001000} + p_{1:[2,3]}\, z^{011000} \quad (13)$$

Step 4: The UGF of vertex 2 is $U_2(z) = \{\xi[U_1(z)], u_2(z)\}$

$$U_2(z) = [\, p_{1:2}\, q_{1:3}\ z^{010000} + p_{1:3}\, q_{1:2}\ z^{001000} + p_{1:[2,3]}\, z^{011000}] \otimes [p_{2:\varphi}\, q_{2:3}\, q_{2:4}\, z^{000000} + p_{2:3}\, q_{2:4}\ z^{001000} + p_{2:4}\, q_{2:3}\ z^{000100} + p_{2:[3,4]}\, z^{001100}\ ]$$

$$= p_{1:2}\, p_{2:3}\, q_{2:4}\, q_{1:3}\, z^{001000} + p_{1:[2,3]}\, p_{2:3}\, q_{2:4}\, z^{001000} + p_{1:2}\, p_{2:4}\, q_{2:3}\, q_{1:3}\ z^{000100} + p_{1:[2,3]}\, p_{2:4}\, q_{2:3}\, z^{001100} + p_{1:2}\, p_{2:[3,4]}\, q_{1:3}\ z^{001100} + p_{1:[2,3]}\, p_{2:[3,4]}\, z^{001100} \quad (14)$$

$$\xi[U_2(z)] = \{\, p_{1:2}\, p_{2:3}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, q_{2:4}\}z^{001000} + p_{1:2}\, p_{2:4}\, q_{2:3}\, q_{1:3}\ z^{000100} + \{p_{1:[2,3]}\, p_{2:4}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}\ \}z^{001100} \quad (15)$$

The UGF of vertex 3 is $U_3(z) = \{\xi[U_2(z)], u_3(z)\}$

$$U_3(z) = [\ \{p_{1:2}\, p_{2:3}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, q_{2:4}\}z^{001000} + p_{1:2}\, p_{2:4}\, q_{2:3}\, q_{1:3}\ z^{000100} + \{p_{1:[2,3]}\, p_{2:4}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}\}z^{001100}] \otimes [p_{3:\varphi}\, q_{3:4}\, q_{3:5}\, z^{000000} + p_{3:4}\, q_{3:5}\ z^{000100} + p_{3:5}\, q_{3:4}\ z^{000010} + p_{3:[4,5]}\, z^{000110}\ ]$$

$$= \{\, p_{1:2}\, p_{2:3}\, p_{3:4}\, q_{3:5}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, p_{3:4}\, q_{3:5}\, q_{2:4}\}z^{000100} + \{p_{1:[2,3]}\, p_{2:4}\, p_{3:4}\, q_{3:5}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:4}\, q_{3:5}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}\, p_{3:4}\, q_{3:5}\}z^{000100} + \{p_{1:2}\, p_{2:3}\, p_{3:5}\, q_{3:4}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, p_{3:5}\, q_{3:4}\, q_{2:4}\}z^{000010} + \{p_{1:[2,3]}\, p_{2:4}\, p_{3:5}\, q_{3:4}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:5}\, q_{3:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}p_{3:5}\, q_{3:4}\}z^{000110} + \{p_{1:2}\, p_{2:3}\, p_{3:[4,5]}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, p_{3:[4,5]}\, q_{2:4}\}z^{000110} + \{p_{1:[2,3]}\, p_{2:4}\, p_{3:[4,5]}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:[4,5]}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}\, p_{3:[4,5]}\ \}z^{000110} \quad (16)$$

$$\xi[U_3(z)] = \{\, p_{1:2}\, p_{2:3}\, p_{3:4}\, q_{3:5}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, p_{3:4}\, q_{3:5}\, q_{2:4} + p_{1:[2,3]}\, p_{2:4}\, p_{3:4}\, q_{3:5}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:4}\, q_{3:5}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}\, p_{3:4}\, q_{3:5}\}z^{000100} + \{p_{1:2}\, p_{2:3}\, p_{3:5}\, q_{3:4}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, p_{3:5}\, q_{3:4}\, q_{2:4}\}z^{000010} + \{p_{1:[2,3]}\, p_{2:4}\, p_{3:5}\, q_{3:4}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:5}\, q_{3:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}p_{3:5}\, q_{3:4} + p_{1:2}\, p_{2:3}\, p_{3:[4,5]}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, p_{3:[4,5]}\, q_{2:4} + p_{1:[2,3]}\, p_{2:4}\, p_{3:[4,5]}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:[4,5]}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}\, p_{3:[4,5]}\ \}z^{000110} \quad (17)$$

The UGF of vertex 4 is $U_4(z) = \{\xi[U_3(z)], u_4(z)\}$

$$\xi[U_4(z)] = \{p_{1:2}\, p_{2:3}\, p_{3:4}\, p_{4:5}\, q_{4:6}\, q_{3:5}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, p_{3:4}\, p_{4:5}\, q_{4:6}\, q_{3:5}\, q_{2:4} + p_{1:[2,3]}\, p_{2:4}\, p_{3:4}\, p_{4:5}\, q_{4:6}\, q_{3:5}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:4}\, p_{4:5}\, q_{4:6}\, q_{3:5}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}\, p_{3:4}\, p_{4:5}\, q_{4:6}\, q_{3:5} + p_{1:[2,3]}\, p_{2:4}\, p_{3:5}\, p_{4:5}\, q_{4:6}\, q_{3:4}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:5}\, p_{4:5}\, q_{4:6}\, q_{3:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}p_{3:5}\, p_{4:5}\, q_{4:6}\, q_{3:4} + p_{1:2}\, p_{2:3}\, p_{3:[4,5]}\, p_{4:5}\, q_{4:6}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, p_{3:[4,5]}\, p_{4:5}\, q_{4:6}\, q_{2:4} + p_{1:[2,3]}\, p_{2:4}\, p_{3:[4,5]}\, p_{4:5}\, q_{4:6}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:[4,5]}\, p_{4:5}\, q_{4:6}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}\, p_{3:[4,5]}\, p_{4:5}\, q_{4:6}\}z^{000010} + \{p_{1:2}\, p_{2:3}\, p_{3:4}\, p_{4:6}\, q_{4:5}\, q_{3:5}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, p_{3:4}\, p_{4:6}\, q_{4:5}\, q_{3:5}\, q_{2:4} + p_{1:[2,3]}\, p_{2:4}\, p_{3:4}\, p_{4:6}\, q_{4:5}\, q_{3:5}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:4}\, p_{4:6}\, q_{4:5}\, q_{3:5}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}\, p_{3:4}\, p_{4:6}\, q_{4:5}\, q_{3:5}\}z^{000001} + \{p_{1:[2,3]}\, p_{2:4}\, p_{3:5}\, p_{4:6}\, q_{4:5}\, q_{3:4}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:5}\, p_{4:6}\, q_{4:5}\, q_{3:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}p_{3:5}\, p_{4:6}\, q_{4:5}\, q_{3:4} + p_{1:2}\, p_{2:3}\, p_{3:[4,5]}\, p_{4:6}\, q_{4:5}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, p_{3:[4,5]}\, p_{4:6}\, q_{4:5}\, q_{2:4} + p_{1:[2,3]}\, p_{2:4}\, p_{3:[4,5]}p_{4:6}\, q_{4:5}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:[4,5]}\, p_{4:6}\, q_{4:5}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}\, p_{3:[4,5]}\, p_{4:6}\, q_{4:5} + p_{1:2}\, p_{2:3}\, p_{3:4}\, p_{4:[5,6]}\, q_{3:5}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, p_{3:4}\, p_{4:[5,6]}\, q_{3:5}\, q_{2:4} + p_{1:[2,3]}\, p_{2:4}\, p_{3:4}\, p_{4:[5,6]}\, q_{3:5}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:4}\, p_{4:[5,6]}\, q_{3:5}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}\, p_{3:4}\, p_{4:[5,6]}\, q_{3:5} + p_{1:[2,3]}\, p_{2:4}\, p_{3:5}\, p_{4:[5,6]}\, q_{3:4}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:5}\, p_{4:[5,6]}\, q_{3:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}p_{3:5}\, p_{4:[5,6]}\, q_{3:4} + p_{1:2}\, p_{2:3}\, p_{3:[4,5]}\, p_{4:[5,6]}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, p_{3:[4,5]}\, p_{4:[5,6]}\, q_{2:4} + p_{1:[2,3]}\, p_{2:4}\, p_{3:[4,5]}\, p_{4:[5,6]}\, q_{2:3} + p_{1:2}\, p_{2:[3,4]}\, p_{3:[4,5]}\, p_{4:[5,6]}\, q_{1:3} + p_{1:[2,3]}\, p_{2:[3,4]}\, p_{3:[4,5]}\, p_{4:[5,6]}\}\ z^{000011} \quad (18)$$

The UGF of vertex 5 is $U_5(z) = \{\xi[U_4(z)], u_5(z)\}$

$$\xi[U_5(z)] = \{\, p_{1:2}\, p_{2:3}\, p_{3:4}\, p_{4:5}\, q_{4:6}\, q_{3:5}\, q_{2:4}\, q_{1:3} + p_{1:[2,3]}\, p_{2:3}\, p_{3:4}\, p_{4:5}\, q_{4:6}\, q_{3:5}\, q_{2:4} + p_{1:[2,3]}\, p_{2:4}\, p_{3:4}\, p_{4:5}$$

$q_{4:6}\ q_{3:5}\ q_{2:3} + p_{1:2}\ p_{2:[3,4]}\ p_{3:4}\ p_{4:5}\ q_{4:6}\ q_{3:5}\ q_{1:3} + p_{1:[2,3]}\ p_{2:[3,4]}\ p_{3:4}\ p_{4:5}\ q_{4:6}\ q_{3:5} + p_{1:[2,3]}\ p_{2:4}\ p_{3:5}\ p_{4:5}\ q_{4:6}\ q_{3:4}\ q_{2:3} + p_{1:2}\ p_{2:[3,4]}\ p_{3:5}\ p_{4:5}\ q_{4:6}\ q_{3:4}\ q_{1:3} + p_{1:[2,3]}\ p_{2:[3,4]}p_{3:5}\ p_{4:5}\ q_{4:6}\ q_{3:4} + p_{1:2}\ p_{2:3}\ p_{3:[4,5]}\ p_{4:5}\ q_{4:6}\ q_{2:4}\ q_{1:3} + p_{1:[2,3]}\ p_{2:3}\ p_{3:[4,5]}\ p_{4:5}\ q_{4:6}\ q_{2:4} + p_{1:[2,3]}\ p_{2:4}\ p_{3:[4,5]}\ p_{4:5}\ q_{4:6}\ q_{2:3} + p_{1:2}\ p_{2:[3,4]}\ p_{3:[4,5]}\ p_{4:5}\ q_{4:6}\ q_{1:3} + p_{1:[2,3]}\ p_{2:[3,4]}\ p_{3:[4,5]}\ p_{4:5}\ q_{4:6} + p_{1:[2,3]}\ p_{2:4}\ p_{3:5}\ p_{4:6}\ q_{4:5}\ q_{3:4}\ q_{2:3} + p_{1:2}\ p_{2:[3,4]}\ p_{3:5}\ p_{4:6}\ q_{4:5}\ q_{3:4}\ q_{1:3} + p_{1:[2,3]}\ p_{2:[3,4]}p_{3:5}\ p_{4:6}\ q_{4:5}\ q_{3:4} + p_{1:2}\ p_{2:3}\ p_{3:[4,5]}\ p_{4:6}\ q_{4:5}\ q_{2:4}\ q_{1:3} + p_{1:[2,3]}\ p_{2:3}\ p_{3:[4,5]}\ p_{4:6}\ q_{4:5}\ q_{2:4} + p_{1:[2,3]}\ p_{2:4}\ p_{3:[4,5]}p_{4:6}\ q_{4:5}\ q_{2:3} + p_{1:2}\ p_{2:[3,4]}\ p_{3:[4,5]}\ p_{4:6}\ q_{4:5}\ q_{1:3} + p_{1:[2,3]}\ p_{2:[3,4]}\ p_{3:[4,5]}\ p_{4:6}\ q_{4:5} + p_{1:2}\ p_{2:3}\ p_{3:4}\ p_{4:[5,6]}\ q_{3:5}\ q_{2:4}\ q_{1:3} + p_{1:[2,3]}\ p_{2:3}\ p_{3:4}\ p_{4:[5,6]}\ q_{3:5}\ q_{2:4} + p_{1:[2,3]}\ p_{2:4}\ p_{3:4}\ p_{4:[5,6]}\ q_{3:5}\ q_{2:3} + p_{1:2}\ p_{2:[3,4]}\ p_{3:4}\ p_{4:[5,6]}\ q_{3:5}\ q_{1:3} + p_{1:[2,3]}\ p_{2:[3,4]}\ p_{3:4}\ p_{4:[5,6]}\ q_{3:5} + p_{1:[2,3]}\ p_{2:4}\ p_{3:5}\ p_{4:[5,6]}\ q_{3:4}\ q_{2:3} + p_{1:2}\ p_{2:[3,4]}\ p_{3:5}\ p_{4:[5,6]}\ q_{3:4}\ q_{1:3} + p_{1:[2,3]}\ p_{2:[3,4]}p_{3:5}\ p_{4:[5,6]}\ q_{3:4} + p_{1:2}\ p_{2:3}\ p_{3:[4,5]}\ p_{4:[5,6]}\ q_{2:4}\ q_{1:3} + p_{1:[2,3]}\ p_{2:3}\ p_{3:[4,5]}\ p_{4:[5,6]}\ q_{2:4} + p_{1:[2,3]}\ p_{2:4}\ p_{3:[4,5]}\ p_{4:[5,6]}\ q_{2:3} + p_{1:2}\ p_{2:[3,4]}\ p_{3:[4,5]}\ p_{4:[5,6]}\ q_{1:3} + p_{1:[2,3]}\ p_{2:[3,4]}\ p_{3:[4,5]}\ p_{4:[5,6]}\ \}p_{5:6}\ z^{000001}$ (19)

Step 5:

$R =$ $p_{1:2}\ p_{2:3}\ p_{3:4}\ p_{4:5}\ p_{5:6} - p_{1:2}\ p_{1:3}\ p_{2:3}\ p_{2:4}\ p_{4:5}\ p_{5:6} + p_{1:2}\ p_{1:3}\ p_{2:4}\ p_{3:4}p_{4:5}p_{5:6} + p_{1:2}\ p_{1:3}\ p_{2:4}\ p_{3:5}\ p_{4:6}\ p_{5:6} + p_{1:2}\ p_{2:3}\ p_{2:4}\ p_{3:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{1:3}\ p_{2:3}\ p_{2:4}\ p_{3:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{2:3}\ p_{2:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} + p_{1:2}\ p_{2:3}\ p_{3:4}\ p_{3:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{2:3}\ p_{2:4}\ p_{3:4}\ p_{3:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{2:3}\ p_{3:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} + p_{1:2}\ p_{2:3}\ p_{2:4}\ p_{3:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{1:3}\ p_{2:4}\ p_{3:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} + p_{1:2}\ p_{1:3}\ p_{2:3}\ p_{2:4}\ p_{3:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6}$ (20)

III. Weibull Distribution

$P_{i:j}(t) = e^{-\eta_{i:j}t^{\beta}}$ where η is the scale parameter and β is the shape parameter.

I. Mean Time To Failure (MTTF)

Mean Time To Failure in network reliability refers to the average time elapsed between the start of network operation and the occurrence of a failure in one of its components or nodes. It's a crucial metric used to assess the reliability and availability of a network infrastructure over time. MTTF helps in predicting the frequency of failures within a network and aids in planning maintenance activities to minimize downtime.

When considering the failure rate for different network components, the MTTF is calculated as:

$$\text{MTTF} = M(t) = \int_0^{\infty} R(t)\,dt = \frac{1}{\eta}\Gamma\left(1+\frac{1}{\beta}\right) \quad (21)$$

II. Sensitivity of MTTF

The partial derivative of the function with respect to that factor is the definition of sensitivity analysis. The sensitivity analysis of MTTF can be obtained by differentiating the MTTF equation with respect to $\eta_{1:2}, \eta_{1:3}, \eta_{2:3}, \eta_{2:4}, \eta_{3:4}, \eta_{3:5}, \eta_{4:5}, \eta_{4:6}$ and $\eta_{5:6}$ the failure rates, respectively [10].

$$\text{Sensitivity of MTTF} = \frac{\partial MTTF}{\partial \eta_{i:j}} \quad (22)$$

We have the following three cases.

Case 1: $\beta < 1 \Rightarrow$ h(t) is a decreasing function of time.

If $\beta = 0.5$ then $P_{i:j}(t) = e^{-\eta_{i:j}t^{0.5}}$ substitute in equation (20),

$R(t) =$ $e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6})t^{0.5}} - e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6})t^{0.5}} + e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6})t^{0.5}} + e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t^{0.5}} + e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t^{0.5}} - e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t^{0.5}} - e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t^{0.5}} + e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t^{0.5}} - e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t^{0.5}} + e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t^{0.5}} - e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t^{0.5}} + e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t^{0.5}}$ (23)

We consider that the failure rates are as follows: $\eta_{1:2} = 0.1,\ \eta_{1:3} = 0.1, \eta_{2:3} = 0.2, \eta_{2:4} = 0.2, \eta_{3:4} = 0.3, \eta_{3:5} = 0.3, \eta_{4:5} = 0.4, \eta_{4:6} = 0.4, \eta_{5:6} = 0.5$. The reliability variation over time for the proposed network, derived from equation (23), is presented in Table 1.

Table 1: *Reliability of k-out-of-n network*

t	R
1	0.348257
2	0.208999
3	0.138394
4	0.096688
5	0.070000
6	0.052018
7	0.039447
8	0.030410
9	0.023765
10	0.018790

$$\text{MTTF}=\frac{1}{\eta}\Gamma\left(1+\frac{1}{0.5}\right)=\frac{1}{\eta}\Gamma(3)=\frac{2}{\eta} \tag{24}$$

$$\begin{aligned}\text{MTTF} = {} & \frac{2}{\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6}} - \frac{2}{\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6}} + \frac{2}{\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6}} \\ & + \frac{2}{\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} + \frac{2}{\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} - \frac{2}{\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} \\ & - \frac{2}{\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} + \frac{2}{\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} - \frac{2}{\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} \\ & - \frac{2}{\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} + \frac{2}{\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} \\ & - \frac{2}{\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} + \frac{2}{\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}}\end{aligned} \tag{25}$$

The variation in MTTF for parameters $\eta_{1:2}$, $\eta_{1:3}$, $\eta_{2:3}$, $\eta_{2:4}$, $\eta_{3:4}$, $\eta_{3:5}$, $\eta_{4:5}$, $\eta_{4:6}$ and $\eta_{5:6}$ is derived from equation (25). The values are presented in Table 2, with a graphical representation in Figure 2.

Table 2: *MTTF of k-out-of-n network*

Failure rate	MTTF $\eta_{1:2}$	MTTF $\eta_{1:3}$	MTTF $\eta_{2:3}$	MTTF $\eta_{2:4}$	MTTF $\eta_{3:4}$	MTTF $\eta_{3:5}$	MTTF $\eta_{4:5}$	MTTF $\eta_{4:6}$	MTTF $\eta_{5:6}$
0.1	5.2698	5.2698	5.2698	5.2698	5.2698	5.2698	5.2698	5.2698	5.2698
0.2	4.2619	5.0873	4.9365	5.0476	4.5000	5.0476	4.9762	5.0476	4.2619
0.3	3.5642	4.9690	4.7309	4.9055	3.9610	4.9055	4.7944	4.9055	3.5642
0.4	3.0556	4.8882	4.5960	4.8095	3.5635	4.8095	4.6746	4.8095	3.0556
0.5	2.6698	4.8307	4.5031	4.7419	3.2585	4.7419	4.5919	4.7419	2.6698
0.6	2.3680	4.7884	4.4367	4.6926	3.0173	4.6926	4.5325	4.6926	2.3680
0.7	2.1259	4.7564	4.3878	4.6557	2.8218	4.6557	4.4885	4.6557	2.1259
0.8	1.9277	4.7317	4.3507	4.6273	2.6603	4.6273	4.4551	4.6273	1.9277
0.9	1.7626	4.7122	4.3221	4.6050	2.5245	4.6050	4.4292	4.6050	1.7626
1.0	1.6230	4.6965	4.2995	4.5873	2.4087	4.5873	4.4087	4.5873	1.6230

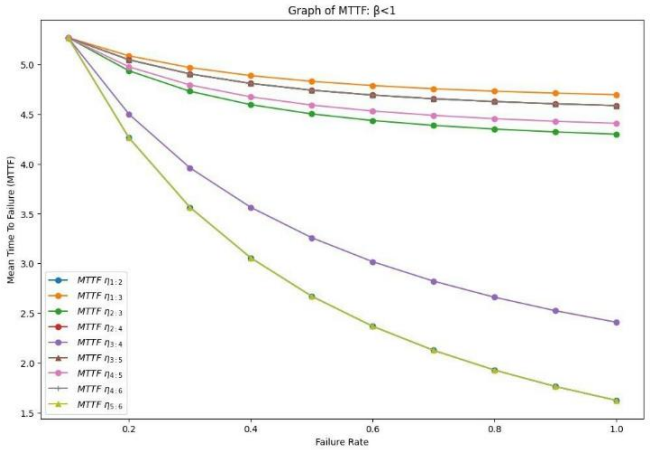


Figure 2: *MTTF of k-out-of-n network*

The sensitivity of MTTF for each edge will then be determined. By varying the failure rates to 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1 in the partial derivatives of MTTF, we can determine the sensitivity of MTTF and Table 3 have been generated. Table on the sensitivity of MTTF shows that the system is very sensitive with respect to the failure.

Table 3: *Sensitivity of MTTF*

Failure rate	∂MTTF/ $\partial\eta_{1:2}$	∂MTTF/ $\partial\eta_{1:3}$	∂MTTF/ $\partial\eta_{2:3}$	∂MTTF/ $\partial\eta_{2:4}$	∂MTTF/ $\partial\eta_{3:4}$	∂MTTF/ $\partial\eta_{3:5}$	∂MTTF/ $\partial\eta_{4:5}$	∂MTTF/ $\partial\eta_{4:6}$	∂MTTF/ $\partial\eta_{5:6}$
0.1	-12.2832	-2.2920	-4.2971	-2.8093	-9.3353	-2.8093	-9.3353	2.7463	-12.2832
0.2	-8.2571	-1.4441	-2.5630	-1.7449	-6.3438	-1.7449	-6.3438	2.3367	-8.2571
0.3	-5.8888	-0.9646	-1.6388	-1.1513	-4.5771	-1.1513	-4.5771	1.9737	-5.8888
0.4	-4.3904	-0.6743	-1.1051	-0.7963	-3.4522	-0.7963	-3.4522	1.6728	-4.3904
0.5	-3.3881	-0.4888	-0.7770	-0.5719	-2.6939	-0.5719	-2.6939	1.4281	-3.3881
0.6	-2.6875	-0.3650	-0.5652	-0.4235	-2.1595	-0.4235	-2.1595	1.2294	-2.6875
0.7	-2.1802	-0.2794	-0.4228	-0.3218	-1.7693	-0.3218	-1.7693	1.0671	-2.1802
0.8	-1.8018	-0.2184	-0.3239	-0.2499	-1.4758	-0.2499	-1.4758	0.9335	-1.8018
0.9	-1.5127	-0.1738	-0.2531	-0.1977	-1.2496	-0.1977	-1.2496	0.8227	-1.5127
1.0	-1.2870	-0.1405	-0.2013	-0.1590	-1.0717	-0.1590	-1.0717	0.7299	-1.2870

Case 2: $\beta = 1 \Rightarrow$ h(t) is constant.

If $\beta = 1$ then $P_{i:j}(t) = e^{-\eta_{i:j}(t)}$, substitute in equation (20)

$$\begin{aligned} R(t) = \; & e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6})t} - e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6})t} + \\ & e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6})t} + e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t} + \\ & e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t} - e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t} - \\ & e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t} + e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t} - \\ & e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t} - e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t} + \\ & e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t} - e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t} + \\ & e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t} \end{aligned} \tag{26}$$

We consider that the failure rates are as follows: $\eta_{1:2} = 0.1,\ \eta_{1:3} = 0.1, \eta_{2:3} = 0.2, \eta_{2:4} = 0.2, \eta_{3:4} = 0.3, \eta_{3:5} = 0.3, \eta_{4:5} = 0.4, \eta_{4:6} = 0.4, \eta_{5:6} = 0.5$. The reliability variation over time for the proposed network, derived from equation (26), is presented in Table 4.

Table 4: *Reliability of k-out-of-n network*

t	R
1	0.348257
2	0.096688
3	0.023765
4	0.005470
5	0.001214
6	0.000264
7	0.000057
8	0.000012
9	0.000003
10	0.018790

$$\begin{aligned} \text{MTTF} = \; & \frac{1}{\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6}} - \frac{1}{\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6}} + \frac{1}{\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6}} \\ & + \frac{1}{\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} + \frac{1}{\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} - \frac{1}{\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} \\ & - \frac{1}{\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} + \frac{1}{\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} - \frac{1}{\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} \\ & - \frac{1}{\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} + \frac{1}{\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} \\ & - \frac{1}{\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} + \frac{1}{\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} \end{aligned} \tag{27}$$

From equation (27), we can obtain variation in MTTF for the parameters $\eta_{1:2}$, $\eta_{1:3}$, $\eta_{2:3}$, $\eta_{2:4}$, $\eta_{3:4}$, $\eta_{3:5}$, $\eta_{4:5}$, $\eta_{4:6}$ and $\eta_{5:6}$. All values are listed in Table 5, and their corresponding graph is shown in Figure 3.

Table 5: *MTTF of the k-out-of-n network*

Failure rate	MTTF $\eta_{1:2}$	MTTF $\eta_{1:3}$	MTTF $\eta_{2:3}$	MTTF $\eta_{2:4}$	MTTF $\eta_{3:4}$	MTTF $\eta_{3:5}$	MTTF $\eta_{4:5}$	MTTF $\eta_{4:6}$	MTTF $\eta_{5:6}$
0.1	2.6349	2.6349	2.6349	2.6349	2.6349	2.6349	2.6349	2.6349	2.6349
0.2	2.1310	2.5437	2.4683	2.5238	2.25	2.5238	2.4881	2.5238	2.1310
0.3	1.7821	2.4845	2.3654	2.4527	1.9805	2.4527	2.3972	2.4527	1.7821
0.4	1.5278	2.4441	2.2980	2.4048	1.7817	2.4048	2.3373	2.4048	1.5278
0.5	1.3349	2.4153	2.2516	2.3710	1.6293	2.3710	2.2959	2.3710	1.3349
0.6	1.1840	2.3942	2.2184	2.3463	1.5087	2.3463	2.2662	2.3463	1.1840
0.7	1.0629	2.3782	2.1939	2.3278	1.4109	2.3278	2.2442	2.3278	1.0629
0.8	0.9638	2.3658	2.1754	2.3136	1.3301	2.3136	2.2275	2.3136	0.9638
0.9	0.8813	2.3561	2.1610	2.3025	1.2622	2.3025	2.2146	2.3025	0.8813
1.0	0.8115	2.3483	2.1497	2.2937	1.2044	2.2937	2.2043	2.2937	0.8115

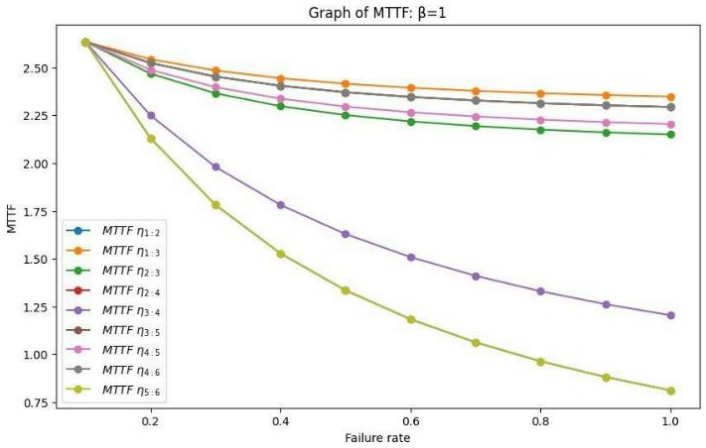


Figure 3: *MTTF of k-out-of-n network*

By varying the failure rates to 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1 in the partial derivatives of MTTF, we can determine the sensitivity of MTTF and Table 6 have been generated. Table on the sensitivity of MTTF shows that the system is very sensitive with respect to the failure.

Table 6: *Sensitivity of MTTF*

Failure rate	∂MTTF/ $\partial\eta_{1:2}$	∂MTTF/ $\partial\eta_{1:3}$	∂MTTF/ $\partial\eta_{2:3}$	∂MTTF/ $\partial\eta_{2:4}$	∂MTTF/ $\partial\eta_{3:4}$	∂MTTF/ $\partial\eta_{3:5}$	∂MTTF/ $\partial\eta_{4:5}$	∂MTTF/ $\partial\eta_{4:6}$	∂MTTF/ $\partial\eta_{5:6}$
0.1	-6.1416	-1.1460	-2.1485	-1.4046	-4.6677	-1.4046	-4.6677	1.3731	-6.1416
0.2	-4.1285	-0.7221	-1.2815	-0.8724	-3.1719	-0.8724	-3.1719	1.1684	-4.1285
0.3	-2.9444	-0.4823	-0.8194	-0.5757	-2.2886	-0.5757	-2.2886	0.9868	-2.9444
0.4	-2.1952	-0.3371	-0.5525	-0.3981	-1.7261	-0.3981	-1.7261	0.8364	-2.1952
0.5	-1.6941	-0.2444	-0.3885	-0.2859	-1.3469	-0.2859	-1.3469	0.7141	-1.6941
0.6	-1.3438	-0.1825	-0.2826	-0.2118	-1.0798	-0.2118	-1.0798	0.6147	-1.3438
0.7	-1.0901	-0.1397	-0.2114	-0.1609	-0.8846	-0.1609	-0.8846	0.5335	-1.0901
0.8	-0.9009	-0.1092	-0.1619	-0.1250	-0.7379	-0.1250	-0.7379	0.4668	-0.9009
0.9	-0.7563	-0.0869	-0.1265	-0.0989	-0.6248	-0.0989	-0.6248	0.4113	-0.7563
1.0	-0.6435	-0.0703	-0.1006	-0.0795	-0.5359	-0.0795	-0.5359	0.3650	-0.6435

Case 3: $\beta > 1 \Rightarrow$ h(t) is an increasing function of time.

If $\beta = 2$ then $P_{i:j}(t) = e^{-\eta_{i:j}t^2}$, substitute in equation (20)

$$\begin{aligned} R(t) = \; & e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6})t^2} - e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6})t^2} + \\ & e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6})t^2} + e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t^2} + \\ & e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t^2} - e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t^2} - \\ & e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t^2} + e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t^2} - \\ & e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6})t^2} - e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t^2} + \\ & e^{-(\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t^2} - e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t^2} + \\ & e^{-(\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6})t^2} \end{aligned} \tag{28}$$

We consider that the failure rates are as follows: $\eta_{1:2} = 0.1,\ \eta_{1:3} = 0.1, \eta_{2:3} = 0.2, \eta_{2:4} = 0.2, \eta_{3:4} = 0.3, \eta_{3:5} = 0.3, \eta_{4:5} = 0.4, \eta_{4:6} = 0.4, \eta_{5:6} = 0.5$. The reliability variation over time for the proposed network, derived from equation (28), is presented in Table 7.

Table 7: *Reliability of k-out-of-n network*

t	R
1	0.348257
2	0.083151
3	0.044262
4	0.027324
5	0.017865
6	0.012166
7	0.008545
8	0.006151
9	0.004517
10	0.003373

$$\text{MTTF} = \frac{1}{\eta}\Gamma\left(1+\frac{1}{2}\right) = \frac{1}{\eta}\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{\eta} \tag{29}$$

$$\begin{aligned} \text{MTTF} = \; & \frac{\sqrt{\pi}}{\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6}} - \frac{\sqrt{\pi}}{\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6}} + \frac{\sqrt{\pi}}{\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:4}+\eta_{4:5}+\eta_{5:6}} \\ & + \frac{\sqrt{\pi}}{\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} + \frac{\sqrt{\pi}}{\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} - \frac{\sqrt{\pi}}{\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} \\ & - \frac{\sqrt{\pi}}{\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} + \frac{\sqrt{\pi}}{\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} - \frac{\sqrt{\pi}}{\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:6}+\eta_{5:6}} \\ & - \frac{\sqrt{\pi}}{\eta_{1:2}+\eta_{2:3}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} + \frac{\sqrt{\pi}}{\eta_{1:2}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} \\ & - \frac{\sqrt{\pi}}{\eta_{1:2}+\eta_{1:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} + \frac{\sqrt{\pi}}{\eta_{1:2}+\eta_{1:3}+\eta_{2:3}+\eta_{2:4}+\eta_{3:4}+\eta_{3:5}+\eta_{4:5}+\eta_{4:6}+\eta_{5:6}} \end{aligned} \tag{30}$$

From equation (30), we can obtain variation in MTTF for the parameters $\eta_{1:2}, \eta_{1:3}, \eta_{2:3}, \eta_{2:4}, \eta_{3:4}, \eta_{3:5}, \eta_{4:5}, \eta_{4:6}$ and $\eta_{5:6}$. All values are listed in Table 8, and their corresponding graph is shown in Figure 4.

By varying the failure rates to 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1 in the partial derivatives of MTTF, we can determine the sensitivity of MTTF and Table 9 have been generated. Table on the sensitivity of MTTF shows that the system is very sensitive with respect to the failure.

Table 8: *MTTF of k-out-of-n network*

Failure rate	MTTF$\eta_{1:2}$	MTTF$\eta_{1:3}$	MTTF$\eta_{2:3}$	MTTF$\eta_{2:4}$	MTTF$\eta_{3:4}$	MTTF$\eta_{3:5}$	MTTF$\eta_{4:5}$	MTTF$\eta_{4:6}$	MTTF$\eta_{5:6}$
0.1	4.6701	4.6701	4.6701	4.6701	4.6701	4.6701	4.6701	4.6701	4.6701
0.2	3.7769	4.5084	4.3747	4.4732	3.9879	4.4732	4.4099	4.4732	3.7769
0.3	3.1586	4.4035	4.1925	4.3472	3.5103	4.3472	4.2488	4.3472	3.1586
0.4	2.7078	4.3319	4.0729	4.2622	3.1580	4.2622	4.1426	4.2622	2.7078
0.5	2.3660	4.2809	3.9907	4.2023	2.8877	4.2023	4.0693	4.2023	2.3660
0.6	2.0985	4.2435	3.9318	4.1586	2.6739	4.1586	4.0167	4.1586	2.0985
0.7	1.8839	4.2151	3.8884	4.1259	2.5007	4.1259	3.9777	4.1259	1.8839
0.8	1.7083	4.1932	3.8556	4.1007	2.3575	4.1007	3.9481	4.1007	1.7083
0.9	1.5620	4.1759	3.8302	4.0810	2.2372	4.0810	3.9252	4.0810	1.5620
1.0	1.4383	4.1621	3.8102	4.0653	2.1346	4.0653	3.9070	4.0653	1.4383

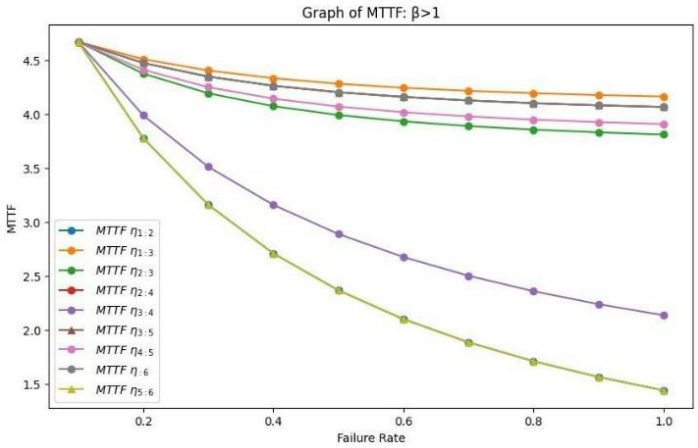


Figure 4: *MTTF of k-out-of-n network*

Table 9: *Sensitivity of MTTF*

Failure rate	∂MTTF/ $\partial\eta_{1:2}$	∂MTTF/ $\partial\eta_{1:3}$	∂MTTF/ $\partial\eta_{2:3}$	∂MTTF/ $\partial\eta_{2:4}$	∂MTTF/ $\partial\eta_{3:4}$	∂MTTF/ $\partial\eta_{3:5}$	∂MTTF/ $\partial\eta_{4:5}$	∂MTTF/ $\partial\eta_{4:6}$	∂MTTF/ $\partial\eta_{5:6}$
0.1	-10.8854	-2.0312	-3.8081	-2.4896	-8.2730	-2.4896	-8.2730	2.4338	-10.8854
0.2	-7.3174	-1.2798	-2.2713	-1.5463	-5.6219	-1.5463	-5.6219	2.0708	-7.3174
0.3	-5.2187	-0.8548	-1.4523	-1.0203	-4.0562	-1.0203	-4.0562	1.7490	-5.2187
0.4	-3.8908	-0.5975	-0.9793	-0.7057	-3.0593	-0.7057	-3.0593	1.4825	-3.8908
0.5	-3.0025	-0.4331	-0.6886	-0.5068	-2.3873	-0.5068	-2.3873	1.2656	-3.0025
0.6	-2.3817	-0.3234	-0.5009	-0.3753	-1.9138	-0.3753	-1.9138	1.0895	-2.3817
0.7	-1.9321	-0.2476	-0.3747	-0.2852	-1.5679	-0.2852	-1.5679	0.9456	-1.9321
0.8	-1.5968	-0.1936	-0.2870	-0.2215	-1.3078	-0.2215	-1.3078	0.8273	-1.5968
0.9	-1.3405	-0.1540	-0.2243	-0.1752	-1.1074	-0.1752	-1.1074	0.7291	-1.3405
1.0	-1.1405	-0.1245	-0.1784	-0.1409	-0.9498	-0.1409	-0.9498	0.6469	-1.1405

Reliability comparison results from tables 1, 4 and 7 indicate that the system's reliability generally decreases as the failure rate increases is presented in Figure 5. For all time points (t), the reliability is highest when $\beta < 1$, indicating a decreasing failure rate over time. When $\beta = 1$, representing a constant failure rate, the reliability is significantly lower than when $\beta < 1$. The lowest reliability is observed when $\beta > 1$, indicating an increasing failure rate over time. Specifically, the reliability drops sharply as β transitions from 1 to values greater than 1. This trend suggests that

systems with a higher failure rate experience more rapid degradation and reduced reliability over time. The reliability values decrease by orders of magnitude as β increases, indicating a substantial impact on the system's ability to maintain functionality over an extended period. These results highlight the importance of managing and reducing failure rates to ensure in critical systems.

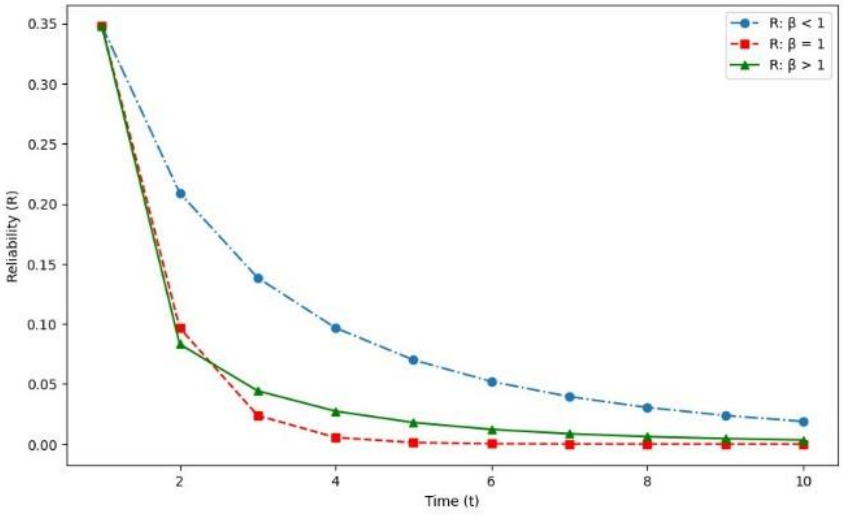


Figure 5: *Reliability comparison analysis*

IV. Birnbaum Measure

The significance of network system components importance is determined by the reliability of the network system and components can be expressed by the Birnbaum measure[1] which indicates the influence degree of network system components on network system connectivity. Due to the mathematical structure of this definition, the reliability of the network system is partially derived from the reliability of each of its components; in other words, changes in the reliability of its components will also affect the reliability of the network system. The Birnbaum measure of a network system with N components is defined as:

$$\mathrm{I}_i^{\mathrm{BM}}(t) = \frac{\partial R_N(t)}{\partial R_i(t)} \tag{31}$$

Here, $R_N(t)$ is the reliability function of the network and $R_i(t)$ is the working function of the network component i. The Birnbaum component importance of the components in the network is discussed below:

$$\mathrm{I}_{1:2}^{\mathrm{BM}}(t) = \mathrm{p}_{2:3}\,\mathrm{p}_{3:4}\,\mathrm{p}_{4:5}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:3}\,\mathrm{p}_{2:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{4:5}\,\mathrm{p}_{5:6} + \mathrm{p}_{1:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:4}\mathrm{p}_{4:5}\mathrm{p}_{5:6} + \mathrm{p}_{1:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} + \mathrm{p}_{2:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:3}\,\mathrm{p}_{2:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{2:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} + \mathrm{p}_{2:3}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{2:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{2:3}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} + \mathrm{p}_{2:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} + \mathrm{p}_{1:3}\,\mathrm{p}_{2:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} \tag{32}$$

$$\mathrm{I}_{1:3}^{\mathrm{BM}}(t) = -\mathrm{p}_{1:2}\,\mathrm{p}_{2:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{4:5}\,\mathrm{p}_{5:6} + \mathrm{p}_{1:2}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:4}\mathrm{p}_{4:5}\mathrm{p}_{5:6} + \mathrm{p}_{1:2}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:2}\,\mathrm{p}_{2:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:2}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} + \mathrm{p}_{1:2}\,\mathrm{p}_{2:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} \tag{33}$$

$$\mathrm{I}_{2:3}^{\mathrm{BM}}(t) = \mathrm{p}_{1:2}\,\mathrm{p}_{3:4}\,\mathrm{p}_{4:5}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:2}\,\mathrm{p}_{1:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{4:5}\,\mathrm{p}_{5:6} + \mathrm{p}_{1:2}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:2}\,\mathrm{p}_{1:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:2}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} + \mathrm{p}_{1:2}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:2}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:2}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} + \mathrm{p}_{1:2}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} + \mathrm{p}_{1:2}\,\mathrm{p}_{1:3}\,\mathrm{p}_{2:4}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} \tag{34}$$

$$\mathrm{I}_{2:4}^{\mathrm{BM}}(t) = -\mathrm{p}_{1:2}\,\mathrm{p}_{1:3}\,\mathrm{p}_{2:3}\,\mathrm{p}_{4:5}\,\mathrm{p}_{5:6} + \mathrm{p}_{1:2}\,\mathrm{p}_{1:3}\,\mathrm{p}_{3:4}\mathrm{p}_{4:5}\mathrm{p}_{5:6} + \mathrm{p}_{1:2}\,\mathrm{p}_{1:3}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} + \mathrm{p}_{1:2}\,\mathrm{p}_{2:3}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:2}\,\mathrm{p}_{1:3}\,\mathrm{p}_{2:3}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:2}\,\mathrm{p}_{2:3}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:2}\,\mathrm{p}_{2:3}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} + \mathrm{p}_{1:2}\,\mathrm{p}_{2:3}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} - \mathrm{p}_{1:2}\,\mathrm{p}_{1:3}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} + \mathrm{p}_{1:2}\,\mathrm{p}_{1:3}\,\mathrm{p}_{2:3}\,\mathrm{p}_{3:4}\,\mathrm{p}_{3:5}\,\mathrm{p}_{4:5}\,\mathrm{p}_{4:6}\,\mathrm{p}_{5:6} \tag{35}$$

$$I_{3:4}^{BM}(t) = p_{1:2}\,p_{2:3}\,p_{4:5}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:4}p_{4:5}p_{5:6} + p_{1:2}\,p_{2:3}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} \quad (36)$$

$$I_{3:5}^{BM}(t) = p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{4:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{4:5}\,p_{4:6}\,p_{5:6} \quad (37)$$

$$I_{4:5}^{BM}(t) = p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} \quad (38)$$

$$I_{4:6}^{BM}(t) = p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:5}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:5}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{5:6} \quad (39)$$

$$I_{5:6}^{BM}(t) = p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{4:5} - p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{4:5} + p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}p_{4:5} + p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:5}\,p_{4:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6} - p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:5}\,p_{4:6} + p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:6} - p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6} - p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6} + p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6} \quad (40)$$

We consider that the various probabilities for the various components are as follows: $p_{1:2} = 0.4$, $p_{1:3} = 0.45$, $p_{2:3} = 0.5$, $p_{2:4} = 0.55$, $p_{3:4} = 0.6$, $p_{3:5} = 0.65$, $p_{4:5} = 0.7$, $p_{4:6} = 0.75$, $p_{5:6} = 0.8$.
Different values of Birnbaum component importance for different edges are obtained by entering different probabilities for various components, as shown in Table 10.

Table 10: *Birnbaum component importance measure*

	BCIM
$I_{1:2}^{BM}(t)$	0.1184
$I_{1:3}^{BM}(t)$	0.0372
$I_{2:3}^{BM}(t)$	0.0787
$I_{2:4}^{BM}(t)$	0.1406
$I_{3:4}^{BM}(t)$	0.1645
$I_{3:5}^{BM}(t)$	0.0468
$I_{4:5}^{BM}(t)$	0.0284
$I_{4:6}^{BM}(t)$	0.0405
$I_{5:6}^{BM}(t)$	0.1289

V. Critical Important Measure

The concept of critical importance in network reliability describes the significance of specific network components that significantly affect the overall performance and stability of the network. Determining the critical components is essential to maintaining network stability.

The critical importance of components can be computed based on the Birnbaum measurement.

$$\text{Since } I_i^{CIM}(t) = \frac{I_i^{BM}(t)(1-p_i(t))}{(1-R(t))} \quad (41)$$

Entering different probabilities for different components yields varied values of the critical importance of components for different edges, as Table 11 illustrates.

Table 11: *Critical importance measure*

CIM	
$I_{1:2}^{CIM}(t)$	0.0799
$I_{1:3}^{CIM}(t)$	0.0230
$I_{2:3}^{CIM}(t)$	0.0443
$I_{2:4}^{CIM}(t)$	0.0712
$I_{3:4}^{CIM}(t)$	0.0740
$I_{3:5}^{CIM}(t)$	0.0184
$I_{4:5}^{CIM}(t)$	0.0096
$I_{4:6}^{CIM}(t)$	0.0114
$I_{5:6}^{CIM}(t)$	0.0290

VI. Risk Growth Factor

Network system average risk growth factor has a significant impact on network system stability and reliability, according to the definition of network system stability. When $R(t) \to 1$, the network system component failure has little impact on the network system reliability, and vice versa.
The reliability stability depends on the risk growth fact of the component, so the reliability stability can be calculated:

$$RGF_{i:j}(t) = R(t) - R_{e_{i:j}=0}(t) \tag{42}$$

$$RGF_{1:2}(t) = R(t) - R_{e_{1:2}=0}(t) = p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{4:5}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{4:5}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}p_{4:5}p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} \tag{43}$$

$$RGF_{1:3}(t) = R(t) - R_{e_{1:3}=0}(t) = -p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{4:5}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}p_{4:5}p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} \tag{44}$$

$$RGF_{2:3}(t) = R(t) - R_{e_{2:3}=0}(t) = p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{4:5}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{4:5}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} \tag{45}$$

$$RGF_{2:4}(t) = R(t) - R_{e_{2:4}=0}(t) = -p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{4:5}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}p_{4:5}p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} \tag{46}$$

$$RGF_{3:4}(t) = R(t) - R_{e_{3:4}=0}(t) = p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{4:5}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}p_{4:5}p_{5:6} + p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} \tag{47}$$

$$RGF_{3:5}(t) = R(t) - R_{e_{3:5}=0}(t) = p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} \tag{48}$$

$$RGF_{4:5}(t) = R(t) - R_{e_{4:5}=0}(t) = p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{4:5}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{4:5}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}p_{4:5}p_{5:6} - p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{2:3}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{1:3}\,p_{2:3}\,p_{2:4}\,p_{3:4}\,p_{3:5}\,p_{4:5}\,p_{4:6}\,p_{5:6} \tag{49}$$

$$RGF_{4:6}(t) = R(t) - R_{e_{4:6}=0}(t) = p_{1:2}\,p_{1:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} + p_{1:2}\,p_{2:3}\,p_{2:4}\,p_{3:5}\,p_{4:6}\,p_{5:6} - p_{1:2}\,p_{1:3}\,p_{2:3}$$

$$p_{2:4}\ p_{3:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{2:3}\ p_{2:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} + p_{1:2}\ p_{2:3}\ p_{3:4}\ p_{3:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{2:3}\ p_{2:4}\ p_{3:4}\ p_{3:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{2:3}\ p_{3:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} + p_{1:2}\ p_{2:3}\ p_{2:4}\ p_{3:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{1:3}\ p_{2:4}\ p_{3:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} + p_{1:2}\ p_{1:3}\ p_{2:3}\ p_{2:4}\ p_{3:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} \quad (50)$$

$$RGF_{5:6}(t) = R(t) - R_{e_{5:6}=0}(t) = p_{1:2}\ p_{2:3}\ p_{3:4}\ p_{4:5}\ p_{5:6} - p_{1:2}\ p_{1:3}\ p_{2:3}\ p_{2:4}\ p_{4:5}\ p_{5:6} + p_{1:2}\ p_{1:3}\ p_{2:4}\ p_{3:4}p_{4:5}p_{5:6} + p_{1:2}\ p_{1:3}\ p_{2:4}\ p_{3:5}\ p_{4:6}\ p_{5:6} + p_{1:2}\ p_{2:3}\ p_{2:4}\ p_{3:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{1:3}\ p_{2:3}\ p_{2:4}\ p_{3:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{2:3}\ p_{2:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} + p_{1:2}\ p_{2:3}\ p_{3:4}\ p_{3:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{2:3}\ p_{2:4}\ p_{3:4}\ p_{3:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{2:3}\ p_{3:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} + p_{1:2}\ p_{2:3}\ p_{2:4}\ p_{3:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} - p_{1:2}\ p_{1:3}\ p_{2:4}\ p_{3:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} + p_{1:2}\ p_{1:3}\ p_{2:3}\ p_{2:4}\ p_{3:4}\ p_{3:5}\ p_{4:5}\ p_{4:6}\ p_{5:6} \quad (51)$$

Table 12: *Risk growth factor*

RGF	
$RGF_{1:2}(t)$	0.1031
$RGF_{1:3}(t)$	0.0635
$RGF_{2:3}(t)$	0.0475
$RGF_{2:4}(t)$	0.2189
$RGF_{3:4}(t)$	0.0987
$RGF_{3:5}(t)$	0.0304
$RGF_{4:5}(t)$	0.0199
$RGF_{4:6}(t)$	0.0304
$RGF_{5:6}(t)$	0.1031

Average risk growth factor: $ARGF(t) = \frac{\sum_{i=1}^{d} RGF_i(t)}{n} = 0.0795$ (52)

Network reliability stability: $NR_S(t) = \frac{R(t) - ARGF(t)}{R(t)} = \frac{0.1112 - 0.0795}{0.1112} = 0.2851$ (53)

VII. Conclusion

The network, structured as a directed acyclic graph, was evaluated based on performance metrics including MTTF, sensitivity analysis, and reliability stability indices. Our findings indicate a noticeable decline in system reliability over time, as illustrated in Tables 1, 4, and 7. The relationship between increasing failure rates and the decreasing MTTF, depicted in Figures 2, 3, and 4, emphasizes the significant influence that variations in failure rates exert on system lifespan. Node 1 and its connections are identified as the most critical, while node 3 also plays a vital role in sustaining reliability. A reliability stability index of 0.2851 highlights moderate operational deterioration, necessitating proactive measures to address vulnerabilities. By applying MTTF sensitivity, organizations can prioritize maintenance on critical components, particularly nodes 1, 2 and their edges, to enhance system stability and resilience.

Declarations

Conflict of interest: The authors declare that they have no conflict of interest.
Funding: No funding received for this research.
Consent for publication: This article do not contain studies with human participants or animals by authors.
Data Availability: The no data used to in this article.
The authors declare that they have read and approved the final manuscript.

References

[1] Birnbaum, Z. W. (1968). On the importance of different components in a multicomponent system, 591–592.

[2] Colbourn, C. J. (1987). *The combinatorics of network reliability*. Oxford University Press, Inc.

[3] Gao, R., & Yao, K. (2016). Importance index of components in uncertain random systems. *Knowledge-Based Systems, 109*, 208–217.

[4] He, L., Cao, Q., & Shang, F. (2019). Measuring component importance for network system using cellular automata. *Complexity*, 397–597.

[5] Levitin, G., & Lisnianski, A. (1999). Importance and sensitivity analysis of multi-state systems using the universal generating function method. *Reliability Engineering & System Safety, 65*(3), 271–282.

[6] Levitin, G. (2001). Reliability evaluation for acyclic consecutively connected networks with multistate elements. *Reliability Engineering & System Safety, 73*(2), 137–143.

[7] Lisnianski, A., & Levitin, G. (2003). *Multi-state system reliability: Assessment, optimization and applications*. World Scientific.

[8] Levitin, G. (2005). *The universal generating function in reliability analysis and optimization* (Vol. 6). London: Springer.

[9] Negi, S., & Singh, S. B. (2015). Reliability analysis of non-repairable complex system with weighted subsystems connected in series. *Applied Mathematics and Computation, 262*, 79–89.

[10] Ram, M., & Singh, S. B. (2009). Analysis of reliability characteristics of a complex engineering system under copula. *Journal of Reliability and Statistical Studies*, 91–102.

[11] Ram, M. (2013). On system reliability approaches: A brief survey. *International Journal of System Assurance Engineering and Management, 4*, 101–117.

[12] Ram, M., & Manglik, M. (2016). Reliability measures analysis of an industrial system under standby modes and catastrophic failure. *International Journal of Operations Research and Information Systems (IJORIS, 7*(3), 36–56.

[13] Ram, M., & Manglik, M. (2016). Stochastic biometric system modelling with rework strategy. *International Journal of Mathematical, Engineering and Management Sciences, 1*(1), 1–7.

[14] Trivedi, S., & Andrea Bobbio. (2017). *Reliability and availability engineering: Modeling, analysis, and applications*. Cambridge University Press.

[15] Ushakov, I. A. (1986). A universal generating function. *Soviet Journal of Computer and Systems Sciences, 24*(5), 118–129.

[16] Yeh, W. C. (2003). An evaluation of the multi-state node networks reliability using the traditional binary-state networks reliability algorithm. *Reliability Engineering & System Safety*, *81*(1), 1-7.

[17] Yeh, W. C. (2008). A simple universal generating function method for estimating the reliability of general multi-state node networks. *Iie Transactions*, *41*(1), 3-11.

[18] Yeh, W. C. (2009). A simple universal generating function method to search for all minimal paths in networks. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *39*(6), 1247-1254.

[19] Yeh, W. C. (2012). A modified universal generating function algorithm for the acyclic binary-state network reliability. *IEEE Transactions on Reliability*, *61*(3), 702-709.

[20] Zarghami, S. A., Gunawan, I., & Schultmann, F. (2020). Exact reliability evaluation of infrastructure networks using graph theory. *Quality and Reliability Engineering International, 36*(2), 498–510.

[21] Zhu, X., Fu, Y., Yuan, T., & Wu, X. (2017). Birnbaum importance-based heuristics for multi-type component assignment problems. *Reliability Engineering & System Safety, 165*, 209–221.

TWO PARAMETRIC GENERALIZED FUZZY ENTROPY MEASURE WITH SOURCE CODING

Tabasum Fatima*[1], M. A. K. Baig[2]

•

[1,2] Department of Statistics, University of Kashmir, Srinagar, India
tabasumf11@gmail.com*, baigmak@gmail.com

Abstract

In this manuscript, we introduce a new generalized fuzzy information measure for a fuzzy set. We establish its validity as a fuzzy entropy measure. Also, we define a new generalized fuzzy average code-word length for a fuzzy set, and explore its relationship with the fuzzy information measure. We prove coding theorems for discrete noiseless channels. The measures outlined in this communication are not only novel, but also encompass certain well-established measures from the existing literature on fuzzy information theory. We also show the proposed measure through a dataset and and observe its monotonicity through tabular and graphical forms.

Keywords: Fuzzy set, Membership function, Shannon's entropy, Fuzzy entropy, Code-word length, Kraft inequality, Coding theorem, Holder's inequality.

1. Introduction

Fuzzy logic and probability both deal with uncertainty, but they do so in fundamentally different ways; Probability Theory deals with randomness and stochastic uncertainty. It describes how likely an event is to occur, assuming a well-defined sample space while Fuzzy Logic deals with vagueness and imprecision. Instead of assigning probabilities, it assigns degrees of truth to statements. Fuzziness and uncertainty are inherent aspects of human thinking and many real-world problems. Uncertainty is present in our decisions, language, and the way we process information. The primary goal of information is to reduce this uncertainty and fuzziness. In fact, the amount of information gained from an experiment is measured by the reduction of probabilistic uncertainty, and this reduction is often referred to as a measure of information. On the other hand, the measure of fuzziness quantifies the vagueness and ambiguity within uncertainties. The concept of entropy, which is central to understanding uncertainty, has been widely applied across various fields, including communication theory, statistical mechanics, finance, pattern recognition, and neural networks. Fuzzy set theory, developed by Lotfi A. Zadeh [20], has also found extensive applications in areas such as clustering, image processing, and decision-making due to its ability to model non-statistical imprecision and vague concepts. The significance of fuzzy sets lies in their ability to handle imprecise and uncertain information. Numerous fuzzy measures have been explored and developed by Kapur [11], Lowen [13], Nguyen and Walker [15], and Parkash [17], Pal and Bezdek [16], Zadeh [20], Hooda [8], Javid and Baig [2], Bhat and Baig [3], [4] etc. Application of fuzzy measures to engineering, fuzzy traffic control, fuzzy aircraft control, medicines, computer science and decision making etc, have already been established. The basic noiseless coding theorems see for instance; the papers, Aczel [1], Kapur [10], Renyi [18], Campbell [5] give the lower bound for the mean code-word length of a uniquely decipherable code in terms of Shannon's [19] entropy etc. Kapur has developed connections

between probability, entropy, and coding. However, there are instances where probabilistic entropy measures are not effective. To address these situations, rather than relying on probability, the concept of fuzziness can be considered.
Suppose there be a universe of discourse $X = \{x_1, x_2, x_3, \ldots, x_n\}$, then a fuzzy subset of universe X is defined as:

$$A = \{(x_i, \mu_A(x_i)) : x_i \in X, \mu_A(x_i) \in [0,1]\}$$

Where $\mu_A(x_i) \in [0,1]$ is a membership function and represents the degree to which the element x_i belongs to the set A and is defined as follows:

$$\mu_A(x_i) = \begin{cases} 0 & \text{if } x_i \notin A \text{ and there is no ambiguity,} \\ 1 & \text{if } x_i \in A \text{ and there is no ambiguity,} \\ 0.5 & \text{if } x_i \in A \text{ or } x_i \notin A \text{ and there is maximum ambiguity.} \end{cases}$$

The membership function $\mu_A(x_i)$ associates each element $x_i \in X$ with a grade of membership in the set A. When $\mu_A(x_i)$ takes only the values 0 or 1, there is no uncertainty about the membership, and the set is considered crisp (i.e., non-fuzzy).

2. Basic Concepts

Let X be a discrete random variable that can take the values $x_1, x_2, x_3, \ldots, x_n$ with corresponding probabilities $p_1, p_2, p_3, \ldots, p_n$ where $p_i \geq 0 \quad \forall\ i = 1, 2, 3, \ldots, n$ and $\sum_{i=1}^{n} p_i = 1$. Shannon [19] defines the entropy of this random variable as a measure of information which is given as:

$$H(A) = -\sum_{i=1}^{n} p_i \log_D p_i \tag{1}$$

The measure (1) acts as an appropriate way to quantify entropy. Let $p_1, p_2, p_3, \ldots, p_n$ represent the probabilities of n code-words to be transmitted, and let their corresponding lengths $l_1, l_2, l_3, \ldots, l_n$ adhere to the Kraft's [12] inequality,

$$\sum_{i=1}^{n} D^{-l_i} \leq 1 \tag{2}$$

Shannon [19] demonstrated that for uniquely decipherable codes, the mean code-word length,

$$L = \sum_{i=1}^{n} p_i l_i \tag{3}$$

has a lower bound for all codes that satisfy (2) and lies between H(P) and H(P)+1, where D is the size of code alphabet.
Let $x_1, x_2, x_3, \ldots, x_n$ be elements of the universe of discourse, with corresponding membership functions $\mu_A(x_1)$, $\mu_A(x_2)$, $\mu_A(x_3)$, ..., $\mu_A(x_n)$. Each of these membership values lies between 0 and 1, but they are not probabilities because their sum does not equals unity. The $\mu_A(x_i)$ represents the degree to which the element x_i belongs to the set A. This function $\mu_A(x_i)$ assigns a grade of membership to each $x_i \in \mathbb{R}^n$ and is known as the membership function.
Define,

$$\text{Fuzzy Set} = \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ \mu_A(x_1) & \mu_A(x_2) & \ldots & \mu_A(x_n) \end{pmatrix}, \quad 0 \leq \mu_A(x_i) \leq 1, \forall\ x_i \in X \tag{4}$$

We refer to the scheme (4) as a finite fuzzy information scheme. Each finite scheme represents a state of uncertainty, which is analogous to Shannon's probabilistic entropy. For Havrda and Charvat's [7] information measure, Kapur [11] suggests the following fuzzy information measure:

$$H^{\beta}(A) = \frac{1}{1-\beta}\left[\frac{1}{n}\sum_{i=1}^{n}\left[\mu_A^{\beta}(x_i) + (1-\mu_A(x_i))^{\beta}\right] - 1\right], \quad \beta \neq 1,\ \beta \geq 0 \tag{5}$$

The measure (5) serves as a very suitable measure of fuzzy entropy for the finite fuzzy information scheme given in (4).
De-Luca and Termini [6] introduced a set of four properties and these properties are widely accepted as for defining new fuzzy entropy. In fuzzy set theory, the entropy is a measure of fuzziness which expresses the amount of average ambiguity in making a decision whether an element belongs to a set or not. So, a measure of average fuzziness H(A) in a fuzzy set A should have the following properties to be valid fuzzy entropy:

1. Sharpness: $H(A)$ is minimum if and only if A is a crisp set, i,e,.

$$\mu_A(x_i) = 0 \text{ or } 1, \quad \forall\, x_i,\, i = 1, 2, \ldots, n$$

2. Maximality: $H(A)$ is maximum if and only if A is most fuzzy set, i.e.,

$$\mu_A(x_i) = \frac{1}{2}, \quad \forall\, x_i,\, i = 1, 2, \ldots, n$$

3. Resolution: $H(A^*) \leq H(A)$, where A^* is sharpened version of A

4. Symmetry: $H(A) = H(A^c)$, where A^c is the complement of A, i.e.,

$$\mu_{A^c}(x_i) = 1 - \mu_A(x_i), \quad \forall\, x_i,\, i = 1, 2, \ldots, n$$

2.1. Example of Fuzzy Sets

Example 1 Let $X = \{0, 2, 4, 6, 8, 10, 12\}$ represent the number of working hours in a day. Define the fuzzy set A as "desirable working hours" with the following membership grades:

- $\mu_A(0) = 0.1$ (not working at all is not desirable)
- $\mu_A(2) = 0.3$
- $\mu_A(4) = 0.7$
- $\mu_A(6) = 1.0$ (ideal working hours)
- $\mu_A(8) = 0.7$
- $\mu_A(10) = 0.3$
- $\mu_A(12) = 0.1$ (working too long is not desirable)

We have plotted the graphs using R software and from the below two graphs the ambiguity can easily be seen and it is maximum at $\mu_A(6) = 1.0$, which shows the ideal working hours. These graphs define a fuzzy set like "Moderate Working Hours" or "Optimal Working Time". This concept is widely used in decision-making, AI, control systems, and human behavior modeling. Here 6 hours is ideal (membership = 1) suggesting that it is the most optimal or representative number of working hours; 4 or 8 hours is still acceptable but less ideal (membership = 0.7) and 0 or 12 hours is not ideal (membership = 0.1) indicating these extremes are least preferred. Here the discrete membership function highlights how specific working hours"ranging from 0 to 12"align with varying degrees of membership, showing a peak at 6 hours where the membership reaches 1 and the continuous membership function graph provides a smooth, gradual transition between different working hours, rather than assigning membership values only to specific points. So, such a function can be used in decision-making processes where strict classification is impractical, such as determining an ideal workload for employees based on efficiency and fatigue considerations.



Figure 1: *Membership Function on a Discrete Universe*

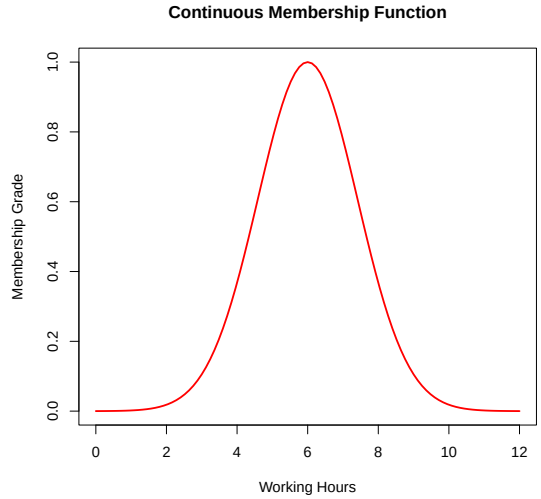


Figure 2: *Membership Function on a Continuous Universe*

3. New Generalized Fuzzy Information Measure and Its Properties

Suppose we have a fuzzy set A defined on a discrete universe of discourse consisting of the elements $x_1, x_2, x_3, \ldots, x_n$, with corresponding membership values $\mu_A(x_1), \mu_A(x_2), \mu_A(x_3), \ldots, \mu_A(x_n)$. We introduce a new generalized fuzzy information measure of order α and type β as:

$$H_\alpha^\beta(A) = \frac{\alpha}{\alpha - \beta}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\mu_A^{\frac{\beta}{\alpha}}(x_i) + (1-\mu_A(x_i))^{\frac{\beta}{\alpha}}\right) - 1\right], \quad \alpha > \beta,\ \beta > 0,\ \alpha \neq \beta \tag{6}$$

To demonstrate that $H_\alpha^\beta(A)$ (as stated in equation (6) is a valid fuzzy information measure, we will prove that it satisfies its key properties: sharpness, maximality, resolution, and symmetry (denoted as P1-P4).

- **P1**. **Sharpness**: $H_\alpha^\beta(A)$ is minimum if and only if A is a crisp set, i.e,

$$H_\alpha^\beta(A) = 0 \text{ if and only if } \mu_A(x_i) = 0 \text{ or } 1, \forall\ x_i,\ i = 1, 2, \ldots, n$$

Proof: Let $H_\alpha^\beta(A)$=0

$$\frac{\alpha}{\alpha - \beta}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\mu_A^{\frac{\beta}{\alpha}}(x_i) + (1 - \mu_A(x_i))^{\frac{\beta}{\alpha}}\right) - 1\right] = 0$$

After some mathematical calculations, we have

$$\sum_{i=1}^{n}\left(\mu_A^{\frac{\beta}{\alpha}}(x_i) + (1 - \mu_A(x_i))^{\frac{\beta}{\alpha}}\right) = n$$

Conversely Suppose

$$\left[\frac{1}{n}\sum_{i=1}^{n}\left(\mu_A^{\frac{\beta}{\alpha}}(x_i) + (1 - \mu_A(x_i))^{\frac{\beta}{\alpha}}\right)\right] = 1$$

Or, Multiplying both sides of the above equation by $\frac{\alpha}{\alpha-\beta}$, we have

$$\frac{\alpha}{\alpha - \beta}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\mu_A^{\frac{\beta}{\alpha}}(x_i) + (1 - \mu_A(x_i))^{\frac{\beta}{\alpha}}\right) - 1\right] = 0$$

Which implies,

$$H_\alpha^\beta(P) = 0$$

Hence, $H_\alpha^\beta(A) = 0$ if and only if A is non fuzzy set or crisp set.

- **P2**. **Maximality**: $H_\alpha^\beta(A)$ is maximum if and only if A is most fuzzy set.
Proof: We have

$$\frac{\alpha}{\alpha - \beta}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\mu_A^{\frac{\beta}{\alpha}}(x_i) + (1 - \mu_A(x_i))^{\frac{\beta}{\alpha}}\right) - 1\right] \tag{7}$$

Differentiating equation (7) with respect to $\mu_A(x_i)$, we have

$$\frac{\partial H_\sigma^\beta(A)}{\partial \mu_A(x_i)} = \frac{\beta}{n(\alpha - \beta)}\left[\mu_A^{\frac{\beta}{\alpha}-1}(x_i) + (1 - \mu_A(x_i))^{\frac{\beta}{\alpha}-1}\right]$$

  - When $0 \leq \mu_A(x_i) < 0.5$

$$\frac{\partial H_\sigma^\beta(A)}{\partial \mu_A(x_i)} > 0, \quad \forall\ 0 < \alpha < 1,\ 0 \leq \beta < 1.$$

    Hence, $H_\alpha^\beta(A)$ is an increasing function of $\mu_A(x_i)$, whenever, $0 \leq \mu_A(x_i) < 0.5$
  - Similarly, for $0.5 < \mu_A(x_i) < 1$, we have

$$\frac{\partial H_\sigma^\beta(A)}{\partial \mu_A(x_i)} < 0, \quad \forall\ 0 < \alpha < 1,\ 0 \leq \beta < 1.$$

    Hence, $H_\alpha^\beta(A)$ is decreasing function of $\mu_A(x_i)$, whenever, $0 \leq \mu_A(x_i) < 0.5$
  - And for $\mu_A(x_i) = 0.5$

$$\frac{\partial H_\sigma^\beta(A)}{\partial \mu_A(x_i)} = 0, \quad \forall\ 0 < \alpha < 1,\ 0 \leq \beta < 1.$$

    Thus $H_\alpha^\beta(A)$ is a concave function which has a global maximum at $\mu_A(x_i) = 0.5$. Hence $H_\alpha^\beta(A)$ is maximum if and only if A is the most fuzzy set, i.e. $\mu_A(x_i) = 0.5, i = 1, 2, 3, \ldots, n$.

- **P3**. **Resolution**: $H_\alpha^\beta(A) \geq H_\alpha^\beta(A^*)$ where A^* is sharped version of A.
  Proof: Since $H_\alpha^\beta(A)$
  is increasing function of $\mu_A(x_i)$ in the interval [0,0.5) and is decreasing function of $\mu_A(x_i)$ in the interval (0.5,1) therefore
  $$\mu_{A^*}(x_i) \leq \mu_A(x_i)$$
  Which implies,
  $$H_\alpha^\beta(A^*) \leq H_\alpha^\beta(A), \quad \text{in } [0, 0.5) \tag{8}$$
  and,
  $$\mu_{A^*}(x_i) \geq \mu_A(x_i)$$
  Which implies,
  $$H_\alpha^\beta(A^*) \geq H_\alpha^\beta(A), \quad \text{in } (0.5, 1) \tag{9}$$
  and
  $$\mu_{A^*}(x_i) \geq \mu_A(x_i)$$
  Then from above two equations, we conclude
  $$H_\alpha^\beta(A) \geq H_\alpha^\beta(A^*)$$

- **P4**. **Symmetry**: $H_\alpha^\beta(A) = H_\alpha^\beta(A^c)$, where A^c is the complement of A.
  Proof: It may be noted that from the definition of $H_\alpha^\beta(A)$ and $\mu_{A^c}(x_i) = 1 - \mu_A(x_i)$, we conclude that
  $$H_\alpha^\beta(A) = H_\alpha^\beta(A^c)$$
  Hence, It may be noted that from the definition of $H_\alpha^\beta(A)$ satisfies all the properties of fuzzy entropy, therefore, $H_\alpha^\beta(A)$ is a valid measure of fuzzy entropy.

4. Noiseless Coding Theorem of Two Parametric Harvda and Charvat's Generalized Fuzzy Information Measure

Here we introduced a new corresponding code-word length according to equation (6) as:

$$L_\alpha^\beta(A) = \frac{\alpha}{\alpha - \beta} \left\{ \left[\left[\sum_{i=1}^{n} \left(\mu_A^{\frac{\beta}{\alpha}}(x_i) + (1 - \mu_A(x_i))^{\frac{\beta}{\alpha}} \right) \right] D^{-l_i\left(\frac{\beta-1}{\beta}\right)} \right]^\beta - 1 \right\} \tag{10}$$

$$; \alpha, \beta > 0,\ \alpha \neq \beta,\ \alpha > \beta$$

Now, we have determined the bounds of (10) in terms of equation (6) under the condition

$$\sum_{i=1}^{n} \left(\mu_A^{\frac{\beta}{\alpha}}(x_i) + (1 - \mu_A(x_i))^{\frac{\beta}{\alpha}} \right) D^{-l_i} \leq 1 \tag{11}$$

Or we can write

$$\sum_{i=1}^{n} \left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right) \right] D^{-l_i} \leq 1 \tag{12}$$

Where,

$$f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right) = \left(\mu_A^{\frac{\beta}{\alpha}}(x_i) + (1 - \mu_A(x_i))^{\frac{\beta}{\alpha}} \right)$$

Which is generalized fuzzy Kraft's inequality, where 'D' is the size of code alphabet.

Theorem 1. For all uniquely decipherable codes $L_\alpha^\beta(A) \geq H_\alpha^\beta(A)$ where,

$$L_\alpha^\beta(A) = \frac{\alpha}{\alpha-\beta}\left\{\left[\left[\sum_{i=1}^{n}\left(\mu_A^{\frac{\beta}{\alpha}}(x_i) + (1-\mu_A(x_i))^{\frac{\beta}{\alpha}}\right)\right] D^{-l_i\left(\frac{\beta-1}{\beta}\right)}\right]^{\beta} - 1\right\}; \alpha,\beta > 0,\ \alpha \neq \beta,\ \alpha > \beta$$

The inequality holds if and only if

$$l_i = -\log_D\left[\frac{1}{\sum_{i=1}^{n}\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]}\right] \tag{13}$$

Proof: From Holder's Inequality, we have

$$\left(\sum_{i=1}^{n} a_i b_i\right) \geq \left(\sum_{i=1}^{n} a_i^\theta\right)^{\frac{1}{\theta}}\left(\sum_{i=1}^{n} b_i^\delta\right)^{\frac{1}{\delta}} \tag{14}$$

$\forall,\ a_i, b_i > 0,\ i = 1,2,3,\ldots,n$ and $\frac{1}{\theta} + \frac{1}{\delta} = 1;\ \theta < 1\,(\neq 0),\ \delta < 0$ or $\delta < 1\,(\neq 0),\ \theta < 0$

The equality holds if and only if there exists a positive constant 'c' such that

$$a_i^\theta = c b_i^\delta$$

Make substitution

$$a_i = \left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]^{\frac{\beta}{\beta-1}} D^{-l_i} \qquad b_i = \left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]^{\frac{1}{1-\beta}}$$

$$\theta_i = \frac{\beta-1}{\beta} \qquad \delta_i = 1-\beta$$

Now using the above values in (14), we have

$$\sum_{i=1}^{n}\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right) D^{-l_i}\right] \geq \left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right) D^{-l_i}\right]^{\frac{\beta-1}{\beta}}\right]^{\frac{\beta}{\beta-1}}\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]\right]^{\frac{1}{1-\beta}} \tag{15}$$

After some mathematical calculations, we have

$$\left\{\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right) D^{-l_i\left(\frac{\beta-1}{\beta}\right)}\right]\right]^{\beta} - 1\right\} \geq \left\{\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]\right] - 1\right\} \tag{16}$$

Multiplying $\frac{\alpha}{\alpha-\beta}$ both sides to (16) and the subtracting 1, we have

$$\frac{\alpha}{\alpha-\beta}\left\{\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right) D^{-l_i\left(\frac{\beta-1}{\beta}\right)}\right]\right]^{\beta} - 1\right\} \geq \frac{\alpha}{\alpha-\beta}\left\{\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]\right] - 1\right\} \tag{17}$$

Which implies,

$$L_\alpha^\beta(A) \geq H_\alpha^\beta(A)$$

Now we have

$$l_i = -\log_D\left[\frac{1}{\sum_{i=1}^{n}\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]}\right]$$

Or,

$$D^{-l_i\left(\frac{\beta-1}{\beta}\right)} = \left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]\right]^{\frac{1-\beta}{\beta}}$$

After some mathematical calculations, we have

$$\frac{\alpha}{\alpha-\beta}\left\{\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)D^{-l_i\left(\frac{\beta-1}{\beta}\right)}\right]\right]^{\beta}-1\right\}=\frac{\alpha}{\alpha-\beta}\left\{\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]\right]-1\right\}$$

Or, we can write that

$$L_{\alpha}^{\beta}(A)=H_{\alpha}^{\beta}(A)$$

Theorem 2. For all uniquely decipherable codes, the following inequality holds:

$$L_{\alpha}^{\beta}(A)<H_{\alpha}^{\beta}(A)D^{(1-\beta)}+\frac{\alpha}{\alpha-\beta}\left(D^{(1-\beta)}-1\right),\ \alpha>\beta,\ 0\leq\beta<1,\ 0<\beta<1 \tag{18}$$

Proof: From theorem (1)

$$L_{\alpha}^{\beta}(A)=H_{\alpha}^{\beta}(A) \tag{19}$$

Equality holds if and only if,

$$D^{-l_i}=\frac{1}{\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]}$$

Or, we can write the above equation as

$$l_i=-\log_D\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]\right]$$

Code-word lengths $l_i,\ (i=1,2,3,,n)$ are chosen in such a way that they satisfy the inequality

$$\log_D\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]\right]\leq l_i<\log_D\left\{\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]\right]+1\right\} \tag{20}$$

Consider the interval

$$\lambda_i=\left\{\log_D\left[\frac{1}{n}\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]\right],\left\{\log_D\left[\frac{1}{n}\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]\right]+1\right\}\right\}$$

of length unity. In every λ_i, there lies exactly one positive integer l_i, such that

$$0<\log_D\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]\right]\leq l_i<\log_D\left\{\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]\right]+1\right\} \tag{21}$$

We will first show the sequence $l_1,l_2,,l_n$, thus defined satisfies the inequality given in (21) which is generalized fuzzy Kraft's Inequality of equation (21), we have

$$\log_D\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]\right]\leq l_i$$

Again after some mathematical manipulation, we have

$$\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]\right]D^{-l_i\left(\frac{1-\beta}{\beta}\right)}<\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]\right]^{\frac{1}{\beta}}D^{\left(\frac{1-\beta}{\beta}\right)} \tag{22}$$

Again after simplifying, we have

$$\frac{\alpha}{\alpha-\beta}\left\{\left[\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]\right]D^{-l_i\left(\frac{1-\beta}{\beta}\right)}\right]^{\beta}-1\right\}<\frac{\alpha}{\alpha-\beta}\left\{\left[\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]\right]D^{(1-\beta)}-1\right\}$$

Or, we can write that,

$$L_{\alpha}^{\beta}(A)<H_{\alpha}^{\beta}(A)D^{(1-\beta)}+\frac{\alpha}{\alpha-\beta}\left(D^{(1-\beta)}-1\right)$$

5. Real World Application

This section showcases the real-world applicability of the Proposed Fuzzy measure using an actual dataset. The dataset pertains to daily snowfall amounts recorded in inches of water, from 30 observations taken near Climax [14], and is also used by [9]. We evaluate the monotonic behavior by holding one parameter constant and changing the other, and similarly, by fixing the second parameter and adjusting the first.

Table 1: *The behaviour of our proposed fuzzy measure When α is fixed and β varies in snowfall dataset.*

	α	$H_\alpha^\beta(A)$		α	$H_\alpha^\beta(A)$
	0.10	0.6856531		0.25	0.8532246
	0.11	0.7077139		0.26	0.8582822
	0.12	0.7269247		0.27	0.8630038
	0.13	0.7437958		0.28	0.8674217
	0.14	0.7587248		0.29	0.8715642
	0.15	0.7720250		0.30	0.8754563
	0.16	0.7839468		0.31	0.8791199
$\beta = 0.02$	0.17	0.7946922	$\beta = 0.02$	0.32	0.8825745
	0.18	0.8044259		0.33	0.8858374
	0.19	0.8132833		0.34	0.8889243
	0.20	0.8213772		0.35	0.8918489
	0.21	0.8288016		0.36	0.8946236
	0.22	0.8356359		0.37	0.8972597
	0.23	0.8419475		0.38	0.8997674
	0.24	0.8477939		0.39	0.9021557

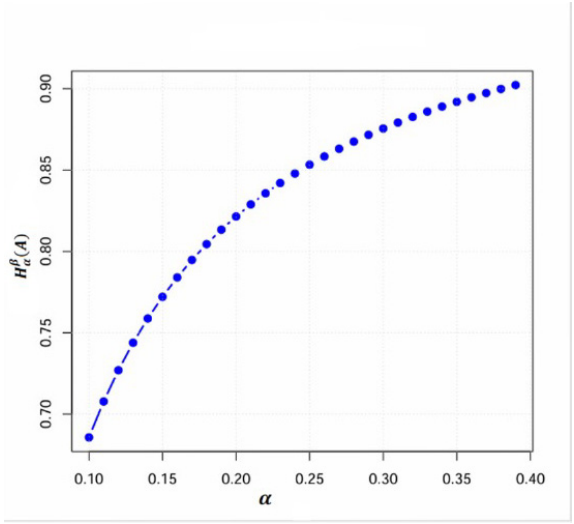


Figure 3: *Graphical representation of $H_\alpha^\beta(A)$ when α is fixed and β varies*

As observed in the table 1 above, when the value of α is held constant and β is varied, $H_\alpha^\beta(A)$ demonstrates an increasing trend. This behavior is illustrated graphically in Figure3. So, we can say that when α is held constant, there is an increase in snowfall level which could be attributed to changing climatic patterns or regional weather dynamics.

Table 2: *The behaviour our proposed measure, when β is fixed and α varies in the snowfall data set.*

	α	$H_\alpha^\beta(A)$		α	$H_\alpha^\beta(A)$
	0.10	0.8213772		0.25	0.6301113
	0.11	0.8060979		0.26	0.6198356
	0.12	0.7912303		0.27	0.6098182
	0.13	0.7767615		0.28	0.6000514
	0.14	0.7626791		0.29	0.5905279
	0.15	0.7489710		0.30	0.5812403
	0.16	0.7356256		0.31	0.5721818
$\beta = 1$	0.17	0.7226318	$\beta = 1$	0.32	0.5633456
	0.18	0.7099787		0.33	0.5547253
	0.19	0.6976558		0.34	0.5463145
	0.20	0.6856531		0.35	0.5381071
	0.21	0.6739607		0.36	0.5300972
	0.22	0.6625694		0.37	0.5222792
	0.23	0.6514698		0.38	0.5146476
	0.24	0.6406533		0.39	0.5071969

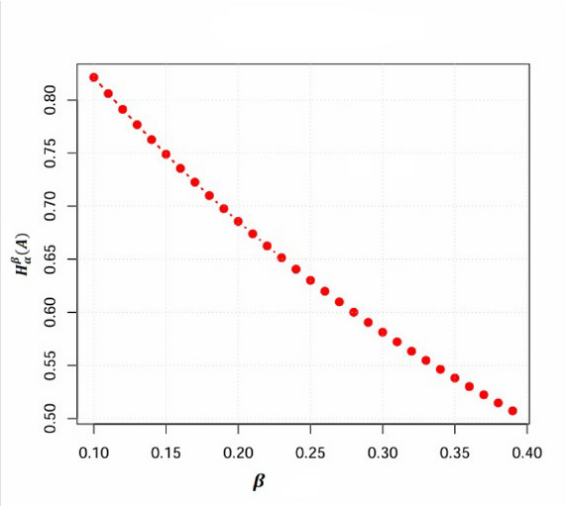


Figure 4: *Graphical representation of $H_\alpha^\beta(A)$ when α is fixed and β varies*

As observed in the table 2 above, when the value of β is held constant and α is varied, $H_\alpha^\beta(A)$ demonstrates an decreasing trend. This behavior is illustrated graphically in Figure4. So, we can say that when α is held constant, there is a decrease in snowfall level which could be attributed to changing climatic patterns or regional weather dynamics. This notable decrease may reflect broader climatic changes, with profound implications for ecosystems, water management, and local economies. By unraveling these factors, we can better anticipate future changes and develop adaptive strategies to mitigate potential impacts.

6. Conclusion

In this manuscript, we introduced a new generalized fuzzy entropy measure of order α and type β, and demonstrated that it is a valid measure of fuzzy entropy. This new measure generalizes several well-known fuzzy entropy measures found in existing literature. Additionally, we considered the

generalized fuzzy average code-word length and derived bounds in terms of the newly defined generalized fuzzy entropy measure of order α and type β. And finally, We have also shown the monotonicity of the proposed measure through a snowfall dataset and presented it in both tabular and graphical form.

References

[1] J. Aczél. *On Shannon's Inequality, Optimal Coding, and Characterizations of Shannon's and Rényi's Entropies*. University of Waterloo, 1973.

[2] M. Baig and M. Javid. Some new generalization of fuzzy average codeword length and their bounds. *American Journal of Applied Mathematics and Statistics*, 2(2):73–76, 2014.

[3] A. H. Bhat and M. Baig. Coding theorems on generalized useful fuzzy inaccuracy measure. *International Journal of Modern Mathematical Science*, 14(1):54–62, 2016.

[4] A. H. Bhat and M. Baig. Some coding theorems on generalized reyni™s entropy of order α and type β. *International Journal of Applied Mathematics and Information Sciences Letters*, 5(1):13–19, 2016.

[5] L. L. Campbell. A coding theorem and rényi's entropy. *Information and control*, 8(4):423–429, 1965.

[6] A. Deluca and S. Termini. A definition of a non-probabilistic entropy in the setting of fuzzy sets, int, 1972.

[7] J. Havrda and F. Charvát. Quantification method of classification processes. concept of structural *a*-entropy. *Kybernetika*, 3(1):30–35, 1967.

[8] D. Hooda. On generalized measures of fuzzy entropy. *Mathematica slovaca*, 54(3):315–325, 2004.

[9] M. Jan and S. Ahmad. A new extension of kumaraswamy distribution for improved data modeling: Properties and applications. *Reliability: Theory & Applications*, 19(4 (80)):654–665, 2024.

[10] J. Kapur. A generalization of campbell's noiseless coding theorem. *Jour. Bihar Math, Society*, 10:1–10, 1986.

[11] J. N. Kapur. *Measures of fuzzy information*. Mathematical Sciences Trust Society, 1997.

[12] L. G. Kraft. *A device for quantizing, grouping, and coding amplitude-modulated pulses*. PhD thesis, Massachusetts Institute of Technology, 1949.

[13] R. Lowen. *Fuzzy set theory: basic concepts, techniques and bibliography*. Springer Science & Business Media, 2012.

[14] P. W. Mielke Jr, L. O. Grant, and C. F. Chappell. An independent replication of the climax wintertime orographic cloud seeding experiment. *Journal of Applied Meteorology (1962-1982)*, pages 1198–1212, 1971.

[15] H. T. Nguyen, C. Walker, and E. A. Walker. *A first course in fuzzy logic*. Chapman and Hall/CRC, 2018.

[16] N. R. Pal and J. C. Bezdek. Measuring fuzzy uncertainty. *IEEE Transactions on Fuzzy Systems*, 2(2):107–118, 1994.

[17] O. Parkash and P. Sharma. Noiseless coding theorems corresponding to fuzzy entropies. In *10th IEEE International Conference on Fuzzy Systems.(Cat. No. 01CH37297)*, volume 1, pages 176–179. IEEE, 2001.

[18] A. Rényi. On measures of entropy and information. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, volume 1: contributions to the theory of statistics*, volume 4, pages 547–562. University of California Press, 1961.

[19] C. E. Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[20] L. A. Zadeh. Fuzzy sets. *Information and Control*, 1965.

RT&A, No 2 (84)
Volume 20, June 2025

SHREEKANT DISTRIBUTION WITH PROPERTIES AND APPLICATIONS TO MODEL STRESS AND STRENGTH DATA FROM ENGINEERING

Rama Shanker
•
Department of Statistics, Assam University, Silchar, India
shankerrama2009@gmail.com

Abstract

Modeling and analyzing real lifetime data that are stochastic in nature with the existing lifetime distributions is really a challenging task for researchers. During the last decade several one parameter lifetime distributions have been proposed in statistics but almost each of them has been found having some problems regarding their goodness of fit on certain datasets. This might have been due to their distributional properties or due to stochastic nature of datasets. In this paper, a new lifetime distribution named Shreekant distribution has been proposed which has been found to give improved fits to lifetime datasets than those by other existing ones such as exponential, Lindley, Shanker, Akash, Sujatha and Uma distributions. This has been shown by presenting its fittings to two datasets relating to stress and strength data from engineering. Its various structural and other properties and estimation of parameters by different methods of estimation have also been studied.

Keywords: Lifetime distributions, statistical properties, estimation of parameter, applications.

I. Introduction

Due to stochastic nature of lifetime data, search for lifetime distribution in the field of lifetime data analysis seems to expand fast and getting popularity among researchers and policy makers to model data of various natures. In past decades several one parameter lifetime distributions have been suggested in statistics literature. For example, Lindley distribution by Lindley [1], Shanker distribution by Shanker [2], Akash distribution by Shanker [3] and Sujatha distribution by Shanker [4], is some among others. Shanker et al [5] discussed the modeling of lifetime data from various fields of knowledge using exponential and Lindley distributions and observed that there are some datasets to which these two distributions do not provide satisfactory fit. Further, Shanker et al [6] made an effort to have comparative study on modeling of lifetime data using exponential, Lindley and Akash distributions and found that Akash distribution gives much better fit than both exponential and Lindley distributions but still there are some datasets in which these three distributions do not provide satisfactory fit. Then, Shanker and Hagos [7] tried to model the real

lifetime datasets arising from different fields of knowledge using exponential, Lindley, Shanker and Akash distributions and observed that still there are some data to which these distributions do not provide satisfactory fits. Shanker [8] proposed another lifetime distribution named Uma distribution which provides better fit than exponential, Lindley, Shanker, Akash and Sujatha distributions but there were some datasets where Uma distribution failed to provide better fit. Flexibility and tractability are the two important distributional characteristics of lifetime distributions and if the existing distributions are not flexible or tractable for given datasets, then search for a new distribution starts. It has been observed that sometimes data are transformed to satisfy some assumptions of the distribution so that distribution fits well. But this is not useful and should not be a preferred practice because the original nature of the dataset is lost. Also, many researchers try to modify the distributions by introducing some additional shape parameters to the existing distribution so that the distribution provides satisfactory fit. But again, this is not very much preferable because of the simple reason that with the introduction of additional parameters may create problems in finding estimates of parameters and also other computational complexities. Therefore, the most practical and preferable way is to search a lifetime distribution which fits the given data well than to modify the existing distributions or to transform the data.

In this paper, in course of search for a new distribution, an attempt has been made to propose a new one parameter lifetime distribution named Shreekant distribution which fits the data arising from different fields of knowledge well over the existing one parameter lifetime distributions. Its statistical properties, characterizations, estimation of parameter and applications to model stress and strength data from engineering have been presented systematically. It is hoped that Shreekant distribution would draw attention of researchers to model lifetime data from engineering and biomedical sciences and have preference over the existing one parameter lifetime distributions.

II. Shreekant Distribution

The Shreekant distribution is defined by its probability density function (pdf) and cumulative distribution function (cdf)

$$f(x;\theta)=\frac{\theta^5}{\theta^4+\theta^3+24}\left(1+x+x^4\right)e^{-\theta x};x>0,\theta>0$$

$$F(x;\theta)=1-\left[1+\frac{\theta x\left\{\theta^3x^3+4\theta^2x^2+12\theta x+\left(\theta^3+24\right)\right\}}{\theta^4+\theta^3+24}\right]e^{-\theta x};x>0,\theta>0\,.$$

It can be easily seen that the Shreekant distribution is a convex combination of exponential (θ), gamma $(2,\theta)$ and gamma $(5,\theta)$ distributions with respective mixing proportions $\frac{\theta^4}{\theta^4+\theta^3+24},\frac{\theta^3}{\theta^4+\theta^3+24}$ and $\frac{24}{\theta^4+\theta^3+24}$. Since it is a convex combination of exponential and gamma distributions, it is hoped to provide better fit over other one parameter lifetime distributions proposed in statistics literature using convex combinations of exponential distribution and gamma distribution with different shape parameter.

The pdf and the cdf of Shreekant distribution for varying values of parameter θ have been presented in figures 1 and 2 respectively. It is quite obvious that the distribution is positively skewed and its skewness depends upon the values of the parameter.

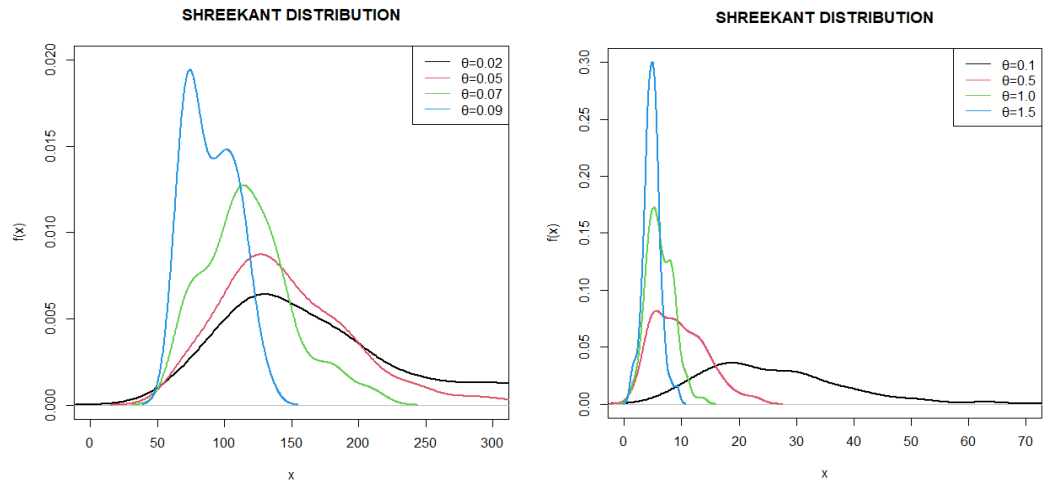


Figure1: *The pdf of Shreekant distribution*

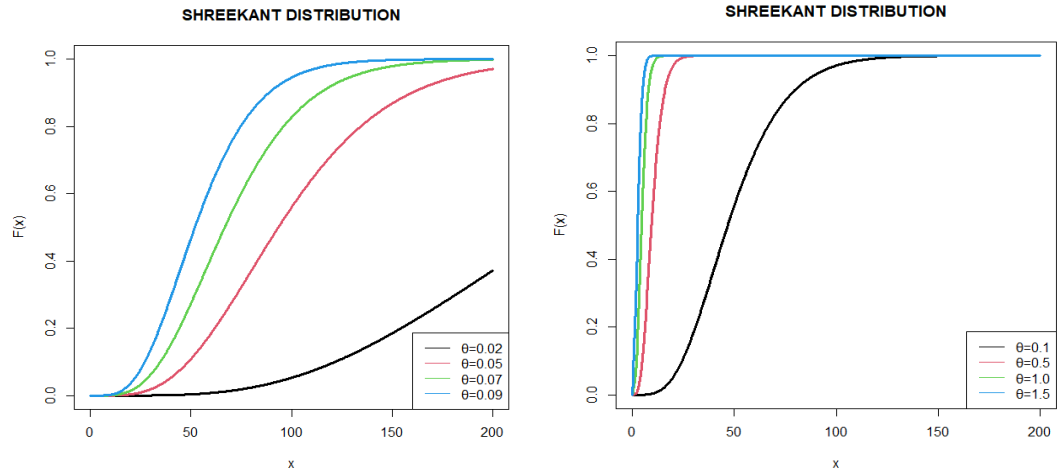


Figure2: *The cdf of Shreekant distribution*

III. Reliability Properties

I. Hazard function

The hazard function of a random variable X having pdf $f(x;\theta)$ and cdf $F(x;\theta)$ is defined as

$$h(x)=\lim_{\Delta x\to 0}\frac{P\left(X<x+\Delta x\,|\,X>x\right)}{\Delta x}=\frac{f(x;\theta)}{1-F(x;\theta)}.$$

Thus, the hazard function of the Shreekant distribution can be obtained as

$$h(x,\theta)=\frac{\theta^5\left(1+x+x^4\right)}{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+\left(\theta^3+24\right)\theta x+\left(\theta^4+\theta^3+24\right)}.$$

This gives $h(0,\theta)=\dfrac{\theta^5}{\theta^4+\theta^3+24}=f(0,\theta)$. The behaviour of the hazard function of the distribution for various values of parameter θ is shown in the following figure 3. It is quite obvious that for increasing values of x the hazard function is increasing and for increasing values of the parameter the hazard function scaled up.

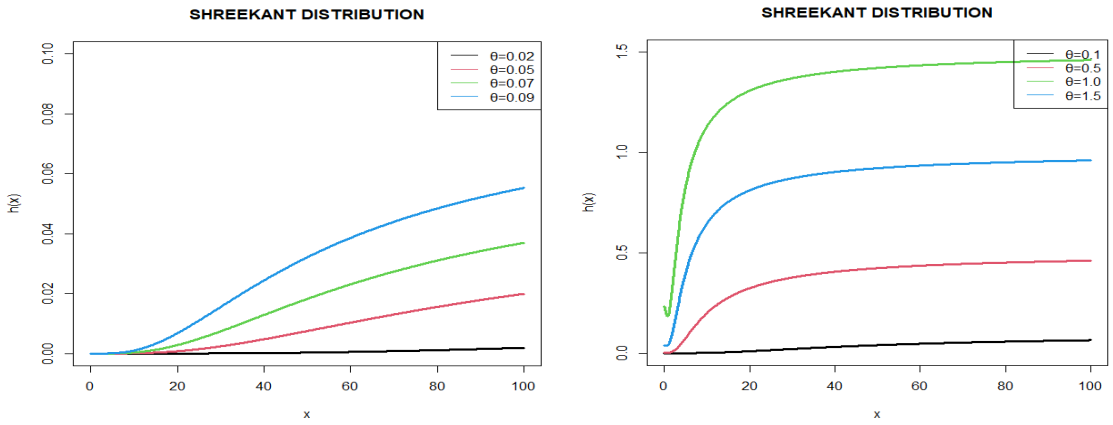


Figure 3: *The hazarad function of the Shreekant distribution.*

II. Mean Residual Life Function

Let X be a random variable over the support $(0,\infty)$ representing the lifetime of a system. Mean residual life (MRL) function measures the expected value of the remaining lifetime of the system, provided it has survived up to time x. Let us consider the conditional random variable $X_x = (X - x \mid X > x); x > 0$. Then, the MRL function, denoted by $m(x)$, is defined as

$$m(x) = E(X_x) = \frac{1}{S(x)} \int_x^\infty [1 - F(t)]\, dt \quad .$$

The MRL function of the Shreekant distribution can thus be obtained as

$$m(x,\theta) = \frac{1}{\left\{\theta^4 x^4 + 4\theta^3 x^3 + 12\theta^2 x^2 + (\theta^3 + 24)\theta x + (\theta^4 + \theta^3 + 24)\right\} e^{-\theta x}} \int_x^\infty t\left(1 + t + t^4\right) e^{-\theta t}\, dt - x$$

$$= \frac{\theta^4 x^4 + 8\theta^3 x^3 + 36\theta^2 x^2 + (\theta^4 + 96\theta)x + (\theta^4 + 2\theta^3 + 120)}{\theta\left\{\theta^4 x^4 + 4\theta^3 x^3 + 12\theta^2 x^2 + (\theta^3 + 24)\theta x + (\theta^4 + \theta^3 + 24)\right\}} .$$

This gives $m(0,\theta) = \dfrac{\theta^4 + 2\theta^3 + 120}{\theta(\theta^4 + \theta^3 + 24)} = \mu_1'$. The behaviour of the mean residual life function of Shreekant distribution for various values of parameter θ is shown in the following figure 4. It is quite obvious that MRL function is decreasing for increasing values of x and it scaled down for increasing values of parameter.

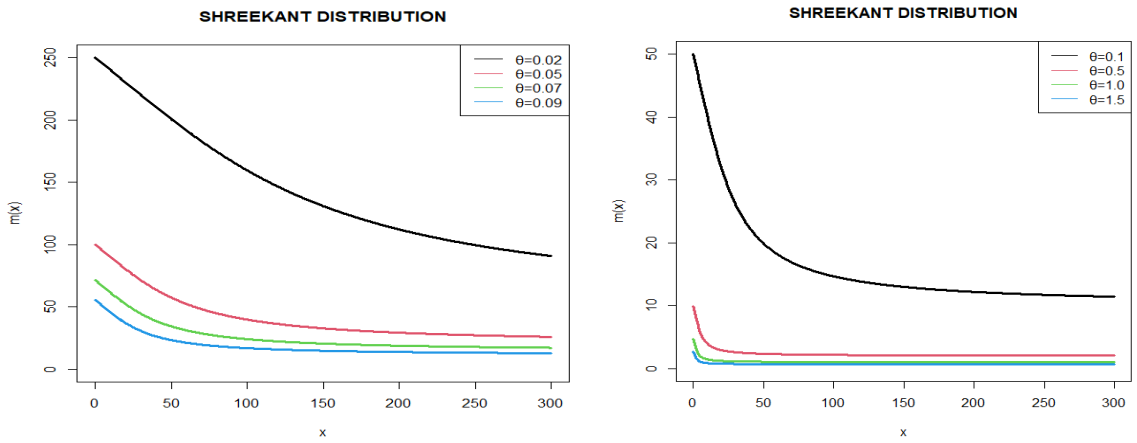


Figure 4: *The mean residual life function of the Shreekant distribution*

III. Reverse hazard function and Mill's ratio

The reverse hazard function of a random variable X having pdf $f(x;\theta)$ and cdf $F(x;\theta)$ is defined as

$$h_r(x,\theta)=\frac{f(x;\theta)}{F(x;\theta)}.$$

Thus, the reverse hazard rate function of the Shreekant distribution can be obtained as

$$h_r(x,\theta)=\frac{\theta^5\left(1+x+x^4\right)e^{-\theta x}}{\left(\theta^4+\theta^3+24\right)-\left[\theta^4x^4+4\theta^3x^3+12\theta^2x^2+\left(\theta^3+24\right)\theta x+\left(\theta^4+\theta^3+24\right)\right]e^{-\theta x}}.$$

Mill's ratio of a random variable X having pdf $f(x;\theta)$ and cdf $F(x;\theta)$ is defined as

$$\text{Mill's ratio}=\frac{1}{h(x)}=\frac{1-F(x;\theta)}{f(x;\theta)}.$$

Thus, the Mill's ratio of the Shreekant distribution can be obtained as

$$\frac{1}{h(x,\theta)}=\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+\left(\theta^3+24\right)\theta x+\left(\theta^4+\theta^3+24\right)}{\theta^5\left(1+x+x^4\right)}.$$

IV. Stochastic Ordering

The stochastic order, in probability theory and statistics, quantifies the concept of one random variable being bigger than another. A random variable X is said to be smaller than a random variable Y in the

(i) Stochastic order $\left(X\le_{st}Y\right)$ if $F_X(x)\ge F_Y(y)$ for all x

(ii) Hazard rate order $\left(X\le_{hr}Y\right)$ if $h_X(x)\ge h_Y(y)$ for all x

(iii) Mean residual life order $\left(X\le_{mrl}Y\right)$ if $m_X(x)\ge m_Y(y)$ for all x

(iv) Likelihood ratio order $\left(X\le_{lr}Y\right)$ if $\frac{f_X(x)}{f_Y(y)}$ decrease in x

The following results due to Shaked and Shantikumar [9] are well known for establishing stochastic ordering of distributions

$$X<_{lr}Y\Rightarrow X<_{hr}Y\Rightarrow X<_{mrl}Y$$
$$\Downarrow$$
$$X<_{st}Y$$

Theorem: Let X and Y have Shreekant distributions with their respective parameters θ_1 and θ_2. If $\theta_1>\theta_2$, then $X<_{lr}Y$ hence $X<_{hr}Y$, $X<_{mrl}Y$ and $X<_{st}Y$.

Proof: We have

$$\frac{f_X(x;\theta_1)}{f_Y(x;\theta_2)}==\frac{\theta_1^{\,5}\left(\theta_2^{\,4}+\theta_2^{\,3}+24\right)}{\theta_2^{\,5}\left(\theta_1^{\,4}+\theta_1^{\,3}+24\right)}e^{-(\theta_1-\theta_2)x}.$$

This gives

$$\log\left[\frac{f_X(x;\theta_1)}{f_Y(x;\theta_2)}\right]=\log\left[\frac{\theta_1^{\,5}\left(\theta_2^{\,4}+\theta_2^{\,3}+24\right)}{\theta_2^{\,5}\left(\theta_1^{\,4}+\theta_1^{\,3}+24\right)}\right]-(\theta_1-\theta_2)x.$$

Therefore, $\frac{d}{dx}\log\left[\frac{f_X(x;\theta_1)}{f_Y(x;\theta_2)}\right] = -(\theta_1 - \theta_2)$

Thus, for $\theta_1 > \theta_2$, $\frac{d}{dx}\log\left[\frac{f_X(x)}{f_Y(x)}\right] < 0$. This means that $X <_{lr} Y$ hence $X <_{hr} Y$, $X <_{mrl} Y$ and $X <_{st} Y$. This proves the theorem.

V. Stress-Strength Reliability

Suppose X and Y be independent strength and stress random variables having Shreekant distribution with parameter θ_1 and θ_2 respectively. Then $R = P(Y < X)$ is known as stress-strength parameter and is a measure of the component reliability. Thus

$$R = P(Y < X) = \int_0^\infty P(Y < X \mid X = x) f_X(x)\,dx = \int_0^\infty f(x;\theta_1)\, F(x;\theta_2)\,dx$$

$$= 1 - \frac{\theta_1^5 \left[\begin{array}{l} 40320\theta_2^4 + 20160(\theta_1+\theta_2)\theta_2^3 + 8640(\theta_1+\theta_2)^2\theta_2^2 + 240(\theta_2^4+12\theta_2)(\theta_1+\theta_2)^3 \\ +24(2\theta_2^4+5\theta_2^3+24)(\theta_1+\theta_2)^4 + 24(\theta_2^3+3\theta_2^2)(\theta_1+\theta_2)^5 + 2(\theta_2^4+12\theta_2^2+24\theta_2)(\theta_1+\theta_2)^6 \\ +(2\theta_2^4+\theta_2^3+24\theta_2+24)(\theta_1+\theta_2)^7 + (\theta_2^4+\theta_2^3+24)(\theta_1+\theta_2)^8 \end{array}\right]}{(\theta_1^4+\theta_1^3+24)(\theta_2^4+\theta_2^3+24)(\theta_1+\theta_2)^9}$$

IV. Moments based descriptive measures

The r th moment about origin μ_r' of the Shreekant distribution can be obtained as

$$\mu_r' = E(X^r) = \frac{\theta^5}{\theta^4+\theta^3+24}\int_0^\infty x^r\left(1+x+x^4\right)e^{-\theta x}dx$$

$$= \frac{r!\left\{\theta^4+(r+1)\theta^3+(r+1)(r+2)(r+3)(r+4)\right\}}{\theta^r\left(\theta^4+\theta^3+24\right)}; r = 1,2,3,\cdots.$$

Substituting $r = 1,2,3,4$ in the above equation, the first four moments about origin of the Shreekant distribution can be obtained as

$$\mu_1' = \frac{\theta^4+2\theta^3+120}{\theta\left(\theta^4+\theta^3+24\right)}, \qquad \mu_2' = \frac{2\left(\theta^4+3\theta^3+360\right)}{\theta^2\left(\theta^4+\theta^3+24\right)}$$

$$\mu_3' = \frac{6\left(\theta^4+4\theta^3+840\right)}{\theta^3\left(\theta^4+\theta^3+24\right)}, \qquad \mu_4' = \frac{24\left(\theta^4+5\theta^3+1680\right)}{\theta^4\left(\theta^4+\theta^3+24\right)}.$$

The moments about the mean of the Shreekant distribution, using relationship between moments about the mean and the moments about the origin, can thus be obtained as

$$\mu_2 = \frac{\theta^8+4\theta^7+2\theta^6+528\theta^4+384\theta^3+2880}{\theta^2\left(\theta^4+\theta^3+24\right)^2}$$

$$\mu_3 = \frac{2\left(\theta^{12}+6\theta^{11}+6\theta^{10}+2\theta^9+1512\theta^8+2160\theta^7+864\theta^6+1728\theta^4+6912\theta^3+69120\right)}{\theta^3\left(\theta^4+\theta^3+24\right)^3}$$

$$\mu_4 = \frac{3\left(\begin{array}{l}3\theta^{16}+24\theta^{15}+44\theta^{14}+32\theta^{13}+7976\theta^{12}+19776\theta^{11}+16896\theta^{10}+4800\theta^{9}\\+1762256\theta^{8}+354816\theta^{7}+158976\theta^{6}+3704832\theta^{4}+2764800\theta^{3}+11612160\end{array}\right)}{\theta^4\left(\theta^4+\theta^3+24\right)^4}.$$

The descriptive constants including coefficient of variation (CV), coefficient of skewness (CS), coefficient of kurtosis (CK) and the index of dispersion (ID) of the Shreekant distribution are thus obtained as

$$CV = \frac{\sqrt{\mu_2}}{\mu_1'} = \frac{\sqrt{\theta^8+4\theta^7+2\theta^6+528\theta^4+384\theta^3+2880}}{\theta^4+2\theta^3+120}$$

$$CS = \frac{\mu_3^{\,2}}{\mu_2^{\,3}} = \frac{4\left(\theta^{12}+6\theta^{11}+6\theta^{10}+2\theta^9+1512\theta^8+2160\theta^7+864\theta^6+1728\theta^4+6912\theta^3+69120\right)^2}{\left(\theta^8+4\theta^7+2\theta^6+528\theta^4+384\theta^3+2880\right)^3}$$

$$CK = \frac{\mu_4}{\mu_2^{\,2}} = \frac{3\left(\begin{array}{l}3\theta^{16}+24\theta^{15}+44\theta^{14}+32\theta^{13}+7976\theta^{12}+19776\theta^{11}+16896\theta^{10}+4800\theta^{9}\\+1762256\theta^{8}+354816\theta^{7}+158976\theta^{6}+3704832\theta^{4}+2764800\theta^{3}+11612160\end{array}\right)}{\left(\theta^8+4\theta^7+2\theta^6+528\theta^4+384\theta^3+2880\right)^2}$$

$$ID = \frac{\mu_2}{\mu_1'} = \frac{\theta^8+4\theta^7+2\theta^6+528\theta^4+384\theta^3+2880}{\theta\left(\theta^4+\theta^3+24\right)\left(\theta^4+2\theta^3+120\right)}.$$

Behaviour of coefficient of variation (CV), coefficient of skewness (CS), coefficient of kurtosis (CK) and index of dispersion (ID) of the Shreekant distribution for various values of parameter are shown in the figure 5. CV is increasing for increasing values of parameter. CS and CK are decreasing for increasing values of parameter whereas ID is decreasing for increasing values of parameter. The over-dispersion (mean less than the variance), under-dispersion (mean more than the variance) and equi-dispersion (mean equal to the variance) of the Shreekant, Uma, Sujatha, Akash, Shanker, Lindley and exponential distributions are presented in table 1

Table 1: *Over-dispersion, equi-dispersion and under-dispersion of distributions*

Distributions	**Over-dispersion** $\left(\mu<\sigma^2\right)$	**Equi-dispersion** $\left(\mu=\sigma^2\right)$	**Under-dispersion** $\left(\mu>\sigma^2\right)$
Shreekant	$\theta<2.21145$	$\theta=2.21145$	$\theta>2.21145$
Uma	$\theta<1.73845$	$\theta=1.73845$	$\theta>1.73845$
Sujatha	$\theta<1.36427$	$\theta=1.36427$	$\theta>1.36427$
Akash	$\theta<1.51540$	$\theta=1.51540$	$\theta>1.51540$
Shanker	$\theta<1.17153$	$\theta=1.17153$	$\theta>1.17153$
Lindley	$\theta<1.17008$	$\theta=1.17008$	$\theta>1.17008$
	$\theta<1$	$\theta=1$	$\theta>1$

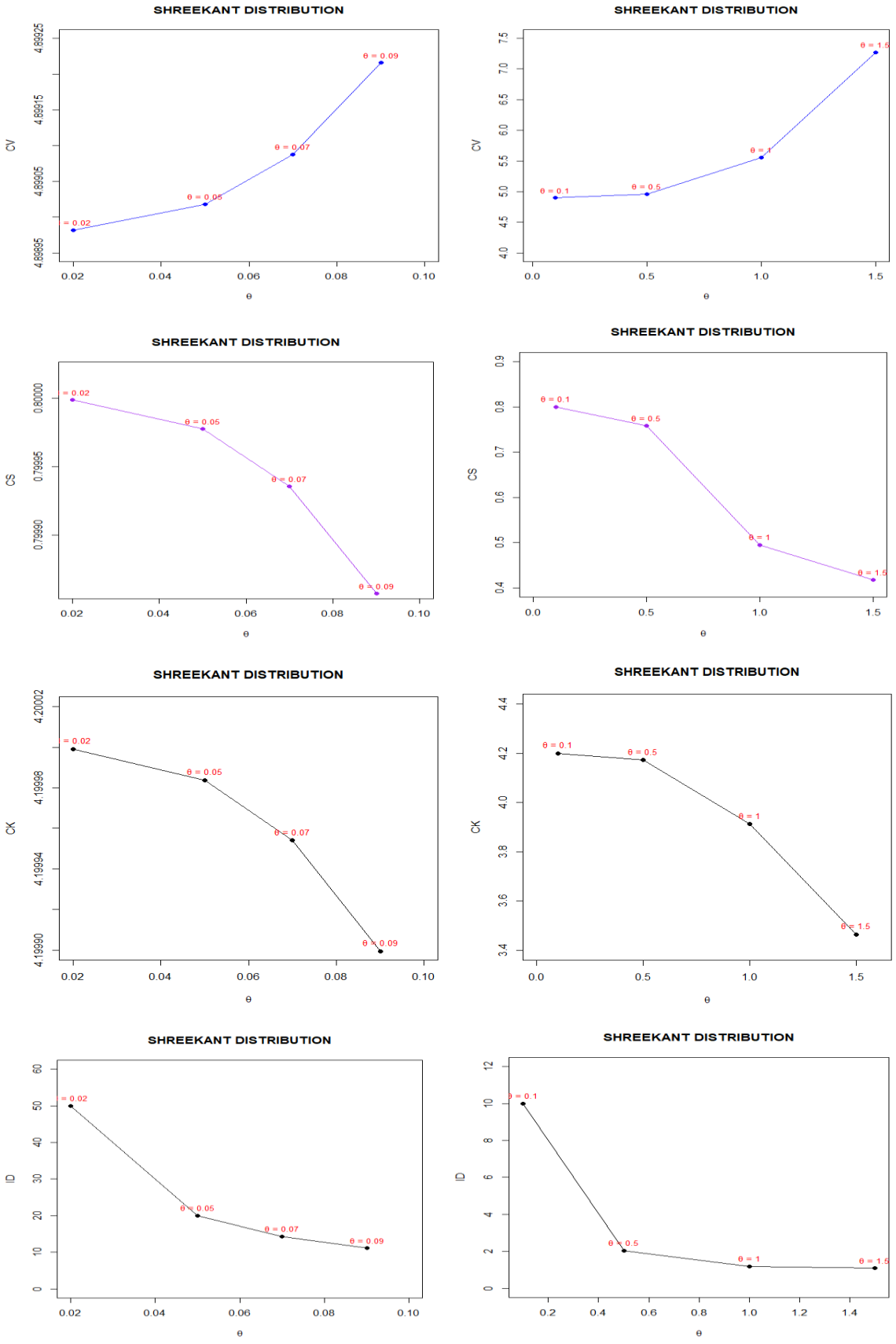


Figure 5. *The CV, CS, CK and ID of the Shreekant distribution*

V. Deviation from Mean and Median

Mean deviation about the mean and the mean deviation about median of a random variable X having pdf $f(x)$ and cdf $F(x)$ are defined by

$$\delta_1(x) = \int_0^{\infty} |x-\mu|\, f(x)dx = 2\mu F(\mu) - 2\int_0^{\mu} x\, f(x)dx$$

and $\delta_2(x) = \int_0^{\infty} |x-M|\, f(x)dx = -\mu + 2\int_M^{\infty} x\, f(x)dx$ respectively, where $\mu = E(X)$ and $M = Median(X)$.

Using the pdf and the expression for the mean of the Shreekant distribution, we get

$$\int_0^{\mu} x\, f(x;\theta)dx = \mu - \frac{\left|\begin{matrix}\left(\mu^5+\mu^2+\mu\right)\theta^5+\left(5\mu^4+2\mu+1\right)\theta^4+\left(20\mu^3+2\right)\theta^3 \\ +60\mu^2\theta^2+120\mu\theta+120\end{matrix}\right| e^{-\theta\mu}}{\theta\left(\theta^4+\theta^3+24\right)}$$

$$\int_0^{M} x\, f(x;\theta)dx = \mu - \frac{\left|\begin{matrix}\left(M^5+M^2+M\right)\theta^5+\left(5M^4+2M+1\right)\theta^4+\left(20M^3+2\right)\theta^3 \\ +60M^2\theta^2+120M\theta+120\end{matrix}\right| e^{-\theta M}}{\theta\left(\theta^4+\theta^3+24\right)}.$$

Using above expressions and after some algebraic simplifications, the mean deviation about the mean $\delta_1(x)$, and the mean deviation about the median $\delta_2(x)$ of the Shreekant distribution are obtained as

$$\delta_1(X) = \frac{2\left\{\left(\mu^4+\mu+1\right)\theta^4+\left(8\mu^3+2\right)\theta^3+36\mu^2\theta^2+96\mu\theta+120\right\}e^{-\theta\mu}}{\theta\left(\theta^4+\theta^3+24\right)}.$$

$$\delta_2(X) = \frac{2\left\{\begin{matrix}\left(M^5+M^2+M\right)\theta^5+\left(5M^4+2M+1\right)\theta^4+\left(20M^3+2\right)\theta^3 \\ +60M^2\theta^2+120M\theta+120\end{matrix}\right\}e^{-\theta M}}{\theta\left(\theta^4+\theta^3+24\right)} - \mu.$$

VI. Distribution of Order Statistics

Let $X_i \sim$ Shreekant distribution(θ), $i = 1,2,3,...,n$ be a *random* sample of size n. Let the corresponding order statistics of these n samples observations be $Y_1, Y_2, ..., Y_n$. The pdf of the r th order statistic Y_r of the Shreekant distribution can be obtained as

$$f_{Y_r}(y) = \frac{1}{B(r,n-r+1)}\left[F(y)\right]^{r-1}\left[1-F(y)\right]^{n-r} f(y)$$

$$= \frac{\theta^5}{\left(\theta^4+\theta^3+24\right)B(r,n-r+1)}\left[1-\left[1+\frac{\theta y\left\{\theta^3y^3+4\theta^2y^2+12\theta y+\left(\theta^3+24\right)\right\}}{\theta^4+\theta^3+24}\right]e^{-\theta y}\right]^{r-1}$$

$$\left[1-\left[1-\left[1+\frac{\theta y\left\{\theta^3y^3+4\theta^2y^2+12\theta y+\left(\theta^3+24\right)\right\}}{\theta^4+\theta^3+24}\right]e^{-\theta y}\right]\right]^{n-r}\left(1+y+y^4\right)e^{-\theta y}$$

The pdf of the nth order statistic Y_n and that of the 1st order statistics Y_1 can be obtained by putting $r = n$ and $r = 1$ respectively in the above equation.

The pdf and cdf of the nth order statistic Y_n can be obtained as

$$f_{Y_n}(y) = \frac{n\theta^5}{\theta^4+\theta^3+24}\left(1+y+y^4\right)e^{-\theta y}\left[1-\left[1+\frac{\theta y\left\{\theta^3y^3+4\theta^2y^2+12\theta y+\left(\theta^3+24\right)\right\}}{\theta^4+\theta^3+24}\right]e^{-\theta y}\right]^{n-1}$$

$$F_{Y_n}(y) = \left[1-\left[1+\frac{\theta y\left\{\theta^3y^3+4\theta^2y^2+12\theta y+\left(\theta^3+24\right)\right\}}{\theta^4+\theta^3+24}\right]e^{-\theta y}\right]^{n}.$$

VII. Estimation of Parameter of the Shreekant Distribution

In this section, following five estimation techniques for estimating the parameter of the Shreekant distribution have been discussed.

I. Maximum Likelihood Estimation

Let $X_i \sim$ Shreekant distribution(θ), $i = 1,2,3,...,n$ be n random samples. The log-likelihood function, L of the Shreekant distribution is given by

$$\log L = \sum_{i=1}^{n}\log f(x_i;\theta) = n\left\{5\log\theta - \log\left(\theta^4+\theta^3+24\right)\right\} + \sum_{i=1}^{n}\log\left(1+x_i+x_i^4\right) - n\theta\bar{x}$$

The maximum likelihood estimate (MLE) $\left(\hat{\theta}\right)$ of the parameters (θ) of the Shreekant distribution is the solution of the following log likelihood equation

$$\frac{d\log L}{d\theta} = \frac{5n}{\theta} - \frac{n\left(4\theta^3+3\theta^2\right)}{\theta^4+\theta^3+24} - n\bar{x} = 0$$

This gives

$$\bar{x}\theta^5 + (\bar{x}-1)\theta^4 - 2\theta^3 + 24\bar{x}\theta - 120 = 0.$$

This is a fifth-degree polynomial equation in θ. It should be noted that the method of moment estimate is also the same as that of the MLE. The above equation can easily be solved using Newton-Raphson method, taking the initial value of the parameter $\theta = 0.5$.

II. Maximum product spacing estimation

The maximum product spacing estimates (MPSE) $\left(\hat{\theta}\right)$ of the parameter (θ) can be obtained numerically by maximizing the following function in relation to θ.

$$MPS = \frac{1}{n+1}\sum_{i=1}^{n+1}\log\left[F(x_i,\theta) - F(x_{i-1},\theta)\right]$$

III. Least Squares Estimation

Let $(x_1, x_2, ..., x_n)$ be a random sample from Shreekant (θ) distribution. The least square function of the parameter based on the pdf of Shreekant distribution can be given as

$$LSE = \sum_{i}^{n}\left(F\left(x_{(i)}\right) - \frac{i}{n+1}\right)^2$$

The least square estimates (LSE) $\hat{\theta}$ of θ can be obtained by minimizing the above equation with respect to θ

IV. Weighted Least Squares Estimation

Let $(x_1, x_2, ..., x_n)$ be a random sample from Shreekant (θ) distribution. The weighted least square function of the parameter based on the pdf of Shreekant distribution can be given as

$$WLSE = \sum_{i}^{n} w_i\left(F\left(x_{(i)}\right) - \frac{i}{n+1}\right)^2, \text{ where } w_i = \frac{(n+1)^2(n+2)}{i(n-i+1)}$$

The weighted least square estimates (WLSE) $\hat{\theta}$ of θ can be obtained by minimizing the above equation with respect to θ .

V. Cramer-Von Mises Estimation

Let $(x_1, x_2, ..., x_n)$ be a random sample from Shreekant (θ) distribution. The Cramer-Von Mises function of the parameter based on the pdf of the Shreekant distribution can be given as

$$CVM = \frac{1}{12n} + \sum_{i=1}^{n}\left(F\left(x_{(i)}\right) - \frac{2i-1}{2n}\right)^2$$

The Cramer-Von Mises estimates (CVME) $\hat{\theta}$ of θ can be obtained by minimizing the above equation with respect to θ .

VIII. Numerical Simulation Study

A simulation study has been conducted to assess the consistency of the estimation of the parameter by different methods of estimation for Shreekant distribution. The study involved examining biases (B) and mean square errors (MSEs) of the different estimation methods viz. MLE, MPSE, LSE. WLSE and CVME discussed in section 7 for the Shreekant distribution using the formulas

$$B = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\theta} - \theta\right), \text{ and } \text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\theta} - \theta\right)^2$$

The acceptance-rejection method of simulation study has been used to generate random samples from the Shreekant distribution.

The B and MSE of the parameter for the five estimation methods are presented in table 2. The values of B and MSE are decreasing for increasing values of sample size.

Table 2: *The Biases and MSE of Shreekant distribution for* $\theta = 0.07$ *and* $\theta = 0.16$

Parameter $(\hat{\theta})$	Sample size	MLE		MPSE		LSE		WLSE		CVME	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
0.07	**20**	-0.00762	0.00008	0.00168	0.00005	0.00216	0.00007	0.00147	0.00007	-0.00289	0.00001
	40	-0.00409	0.00005	0.00141	0.00004	0.00199	0.00007	0.00119	0.00007	-0.00259	0.00001
	60	-0.00238	0.00002	0.00115	0.00003	0.00156	0.00005	0.00104	0.00007	-0.00207	0.00000
	80	-0.00177	0.00001	0.00099	0.00002	0.00133	0.00004	0.00096	0.00007	-0.00183	0.00000
	100	-0.00139	0.00001	0.00089	0.00002	0.00119	0.00003	0.00075	0.00006	-0.00168	0.00000
0.16	**20**	0.01630	0.00036	-0.00031	0.00003	0.00036	0.00002	-0.00259	0.00041	-0.00110	0.00013
	40	0.01580	0.00033	-0.00024	0.00002	0.00028	0.00002	-0.00156	0.00031	-0.00083	0.00012
	60	0.01115	0.00019	-0.00017	0.00002	0.00018	0.00002	-0.00057	0.00027	-0.00985	0.00011
	80	0.00427	0.00008	-0.00008	0.00002	0.00008	0.00002	-0.00928	0.00023	-0.00898	0.00009
	100	0.00113	0.00003	-0.00006	0.00002	0.00005	0.00001	-0.00775	0.00020	-0.00767	0.00007

IX. Real Data Analysis

The applications and the goodness of fit of Shreekant distribution have been discussed with two datasets relating to strength and stress data from engineering. Keeping in mind the flexibility and tractability of the distribution with the dataset following two datasets have been considered. The total time on test (TTT) plot of dataset 1 and 2 and the simulated datasets of the Shreekant distribution are shown in figure 6 and 7 respectively. The descriptive summary of the datasets 1 and 2 are presented in the table 3.

Dataset 1: This dataset is the strength data of glass of the aircraft window reported by Fuller et al [10]:

18.83, 20.80, 21.657, 23.03, 23.23, 24.05, 24.321, 25.5, 25.52, 25.80, 26.69, 26.77, 26.78, 27.05, 27.67, 29.90, 31.11, 33.2, 33.73, 33.76, 33.89, 34.76, 35.75, 35.91, 36.98, 37.08, 37.09, 39.58, 44.045, 45.29, 45.381.

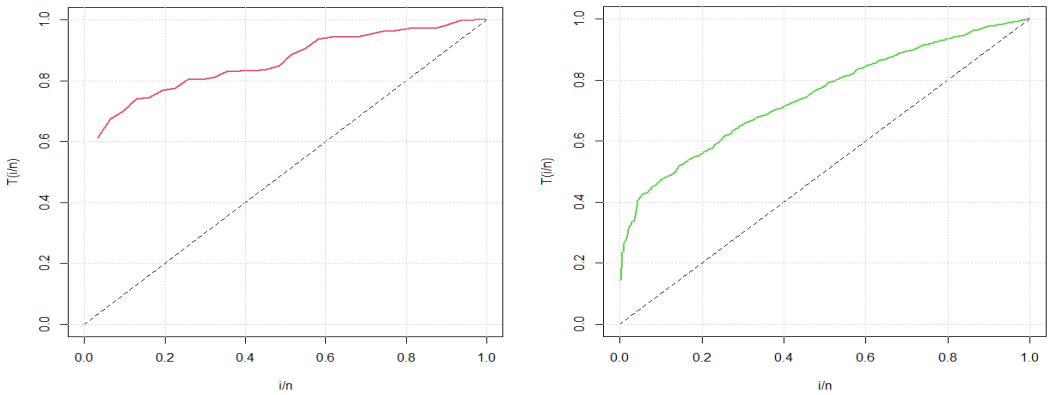


Figure 6: TTT plot of dataset 1 and simulated dataset of the Shreekant distribution.

Dataset 2: The data is given by Birnbaum & Saunders [11] on the fatigue life of 6061 – T6 aluminum coupons cut parallel to the direction of rolling and oscillated at 18 cycles per second. The data set consists of 101 observations with maximum stress per cycle 31,000 psi. The data (× 3 10−) are presented below (after subtracting 65).

5, 25, 31, 32, 34, 35, 38, 39, 39, 40, 42, 43, 43, 43, 44, 44, 47, 47, 48, 49, 49, 49, 51, 54, 55, 55, 55, 56, 56, 56, 58, 59, 59, 59, 59, 59, 63, 63, 64, 64, 65, 65, 65, 66, 66, 66, 66, 66, 67, 67, 67, 68, 69, 69, 69, 69, 71, 71, 72, 73, 73, 73, 74, 74, 76, 76, 77, 77, 77, 77, 77, 77, 79, 79, 80, 81, 83, 83, 84, 86, 86, 87, 90, 91, 92, 92, 92, 92, 93, 94, 97, 98, 98, 99, 101, 103, 105 ,109 ,136, 147

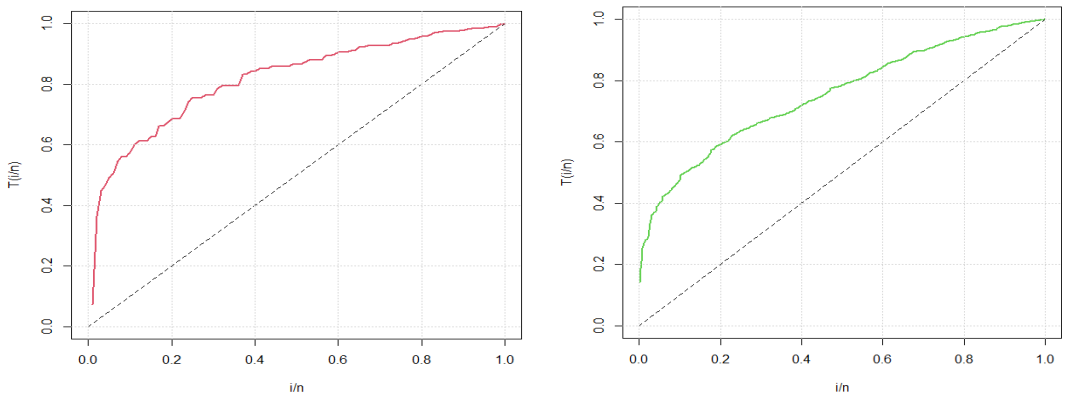


Figure 7: *TTT-plot of dataset 2 and simulated dataset of the Shreekant distribution.*

In order to compare the lifetime distributions, values of $-2\log L$, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Consistent Akaike Information Criterion (CAIC), Hannan-Quinn Information Criterion (HQIC), Kolmogorov-Smirnov (K-S) Statistics and the corresponding probability value (p-value) for the above two datasets have been computed using the formulae

$$AIC = -2\log L + 2p\,,\ BIC = -2\log L + p\log(n)\,,\ CAIC = -2\log L + \frac{2pn}{n-p-1}$$

$$HQIC = -2\log L + 2p\log\lfloor \log(n)\rfloor,\quad K\text{-}S = \underset{x}{Sup}\,|F_m(x) - F_o(x)|$$

where, p = number of parameters, n = sample size, $F_m(x)$ = empirical cdf of considered distribution and $F_o(x)$ = cdf of considered distribution

The estimated values of parameter with standard error (SE) of the considered distributions for the given datasets are presented in the table 4 and 6 respectively. The values of $-2\log L$, AIC, BIC, CAIC, HQIC, K-S and p-values of the considered distributions for the given datasets are presented in the table 5 and 7 respectively. The confidence intervals of the parameter of the proposed distribution are presented in table 8.

Table 3: *Summary of the datasets 1 and 2*

Datasets	Minimum	1st Quartile	Median	Mean	Standard deviation	3rd Quartile	Maximum
1	18.83	25.51	29.90	30.81	7.25	30.81	45.38
2	5.00	55.00	67.00	68.33	22.41	80.25	147.00

Table 4: *Estimated values by different methods of estimation and standard error (S.E) of the parameter for the considered distributions of the dataset-1*

Distributions	MLE $\hat{\theta}$(S.E)	MPSE $\hat{\theta}$(S.E)	LSE $\hat{\theta}$(S.E)	WLSE $\hat{\theta}$(S.E)	CVME $\hat{\theta}$(S.E)
Shreekant	0.1622	0.1611	0.1545	0.1555	0.1557
	(0.0130)	(0.0128)	(0.0046)	(0.0015)	(0.0239)
Uma	0.1296	0.1285	0.1212	0.1225	0.1224
	(0.0116)	(0.0114)	(0.0039)	(0.0012)	(0.0206)
Sujatha	0.0956	0.0946	0.0868	0.0881	0.0879
	(0.0099)	(0.0097)	(0.0032)	(0.0010)	(0.0168)
Akash	0.0970	0.0961	0.0881	0.0896	0.0893
	(0.0100)	(0.0099)	(0.0032)	(0.0010)	(0.0171)
Shanker	0.0647	0.0638	0.0553	0.0568	0.0565
	(0.0082)	(0.0080)	(0.0025)	(0.0008)	(0.0133)
Lindley	0.0629	0.0621	0.0537	0.0551	0.0548
	(0.0080)	(0.0078)	(0.0024)	(0.0007)	(0.0129)
Exponential	0.0324	0.0317	0.0229	0.0242	0.0238
	(0.0058)	(0.0057)	(0.0015)	(0.0005)	(0.0083)

Table 5: *Goodness of fit for the dataset-1*

Sl. No	Distributions	$-2\log L$	AIC	BIC	CAIC	HQIC	K-S	P-value
1	Shreekant	227.25	229.25	231.86	229.30	230.31	0.2120	0.1460
2	Uma	232.88	234.88	237.49	234.93	235.94	0.4502	0.0000
3	Sujatha	241.50	243.50	246.10	243.54	244.55	0.4501	0.0000
4	Akash	240.68	242.68	245.28	242.72	243.73	0.3026	0.0000
5	Shanker	252.35	254.35	256.95	254.39	255.40	0.3341	0.0000
6	Lindley	253.98	255.98	258.59	256.02	257.04	0.3444	0.0000
7	Exponential	274.52	276.52	279.13	276.56	277.58	0.4777	0.0000

Table 6: *Estimate values with standard error of the considered distributions for the dataset-2*

Distributions	MLE $\hat{\theta}$(S.E)	MPSE $\hat{\theta}$(S.E)	LSE $\hat{\theta}$(S.E)	WLSE $\hat{\theta}$(S.E)	CVME $\hat{\theta}$(S.E)
Shreekant	0.0731	0.0728	0.0703	0.0717	0.0705
	(0.0032)	(0.0032)	(0.0007)	(0.0003)	(0.0064)
Uma	0.0585	0.0582	0.0553	0.0568	0.0556
	(0.0029)	(0.0029)	(0.0006)	(0.0002)	(0.0055)
Sujatha	0.0435	0.0433	0.0400	0.0416	0.0405
	(0.0025)	(0.0025)	(0.0004)	(0.0001)	(0.0045)
Akash	0.0438	0.0436	0.0403	0.0419	0.0408
	(0.0025)	(0.0025)	(0.0004)	(0.0001)	(0.0045)
Shanker	0.0292	0.0290	0.0254	0.0270	0.0259
	(0.0020)	(0.0020)	(0.0003)	(0.0001)	(0.0035
Lindley	0.0288	0.0286	0.0250	0.0266	0.0255
	(0.0020)	(0.0286)	(0.0003)	(0.0001)	(0.0034)
Exponential	0.0146	0.0144	0.0105	0.0119	0.0109
	(0.0014)	(0.0014)	(0.0002)	(0.0007)	(0.0021)

Table 7: *Goodness of fit for the dataset-2*

Sl. No	Distributions	$-2\log L$	AIC	BIC	CAIC	HQIC	K-S	P-value
1	Shreekant	924.19	926.19	928.80	926.23	927.25	0.1271	0.1390
2	Uma	934.03	936.03	938.63	936.07	937.08	0.1891	0.0050
3	Sujatha	951.78	953.78	956.38	953.82	954.83	0.2149	0.0000
4	Akash	950.97	952.97	955.57	953.01	954.02	0.1944	0.0010
5	Shanker	980.97	982.97	985.57	983.01	984.02	0.2333	0.0000
6	Lindley	983.10	985.10	987.70	985.14	986.15	0.2326	0.0000
7	Exponential	1044.87	1046.87	1049.47	1046.91	1047.92	0.3700	0.0000

Table 8: *Confidence interval of the parameter θ of the Shreekant distribution*

Dataset	90%Confidence interval		95%Confidence interval		99%Confidence interval	
	Lower	Upper	Lower	Upper	Lower	Upper
1	0.0731	0.0732	0.0730	0.0732	0.0730	0.0732
2	0.0679	0.0786	0.0669	0.0797	0.0650	0.0819

The fitted plots of the considered distributions, the Q-Q plot and the P-P plot of Shreekant distribution for the datasets 1 and 2 are presented in the figure 8 and 9 respectively. The profile plots of the parameters of Shreekant distribution are presented in the figure 10.

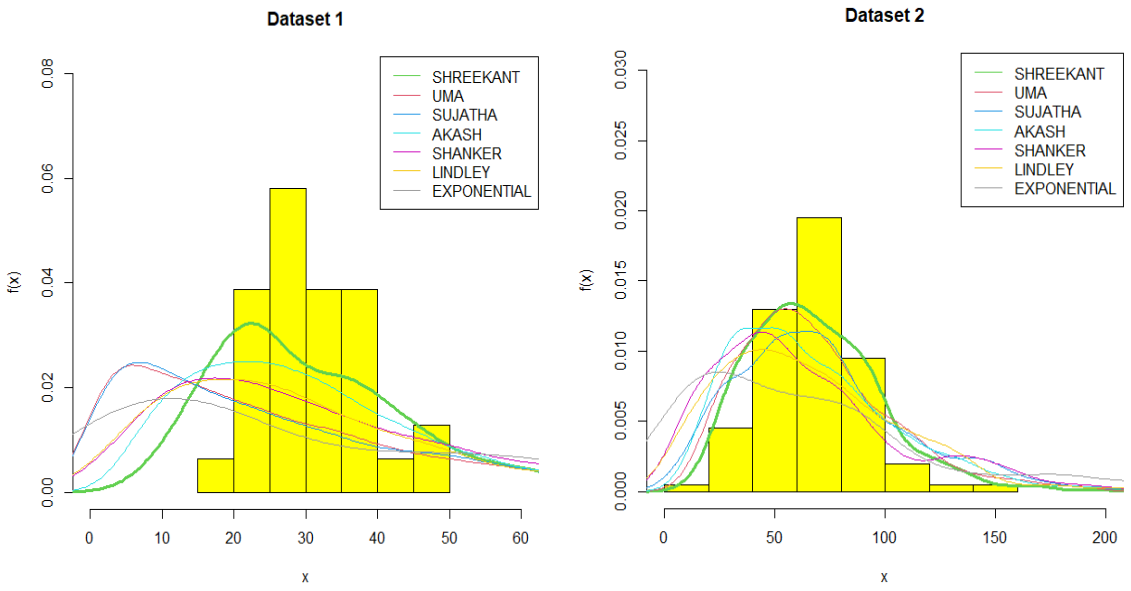


Figure 8. *Fitted plots of the considered distributions for the datasets 1 and 2.*

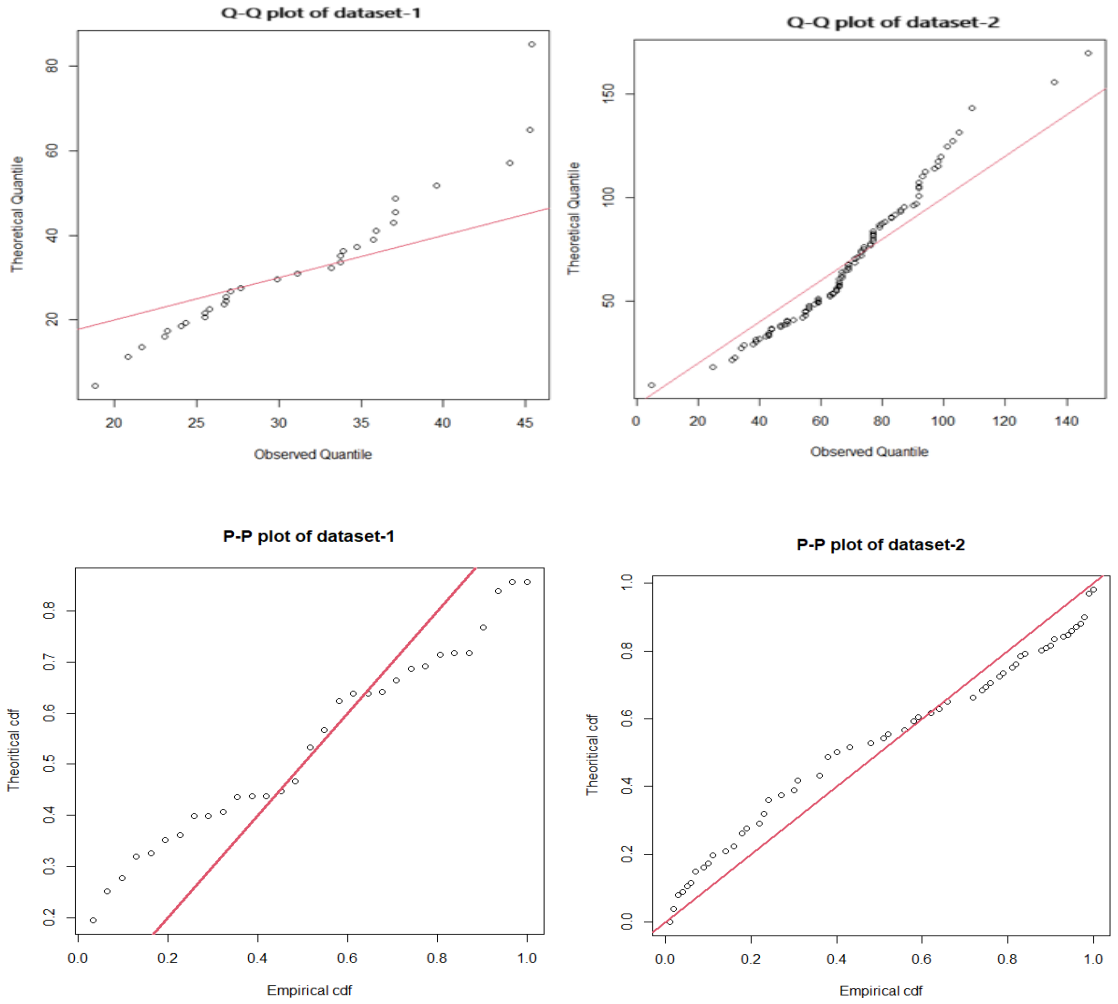


Figure 9: *Q-Q plot and P-P plot of the Shreekant distribution for the datasets 1 and 2.*

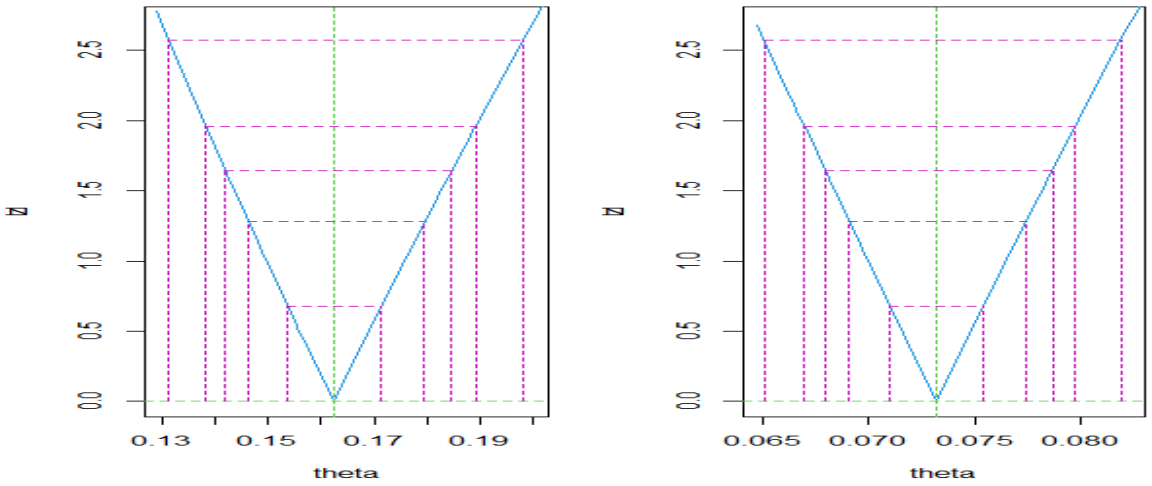


Figure 10: *Profile plot of the parameter of Shreekant distribution for the dataset 1 and 2*

It is clear from the goodness of fit in tables 5 and 7 and from the fitted plots of distributions in figure 8 that the Shreekant distribution gives much better fit over exponential, Lindley, Shanker, Akash, Sujatha and Uma distributions.

X. Conclusion and Future Work

A new lifetime distribution named Shreekant distribution has been suggested to model strength and stress data from engineering. Statistical properties, estimation of parameter and applications of the distribution have been presented. As the distribution is new one, it is expected and hoped that it will be of great use to statisticians and policy makers working in the field of modeling lifetime data from different fields of knowledge.

Being a new lifetime distribution having flexibility, tractability and practicability, the distribution is expected to generate much scope for future works on it.

References

[1] Lindley, D.V. (1958). Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society, Series B*, 20: 102- 107.

[2] Shanker, R. (2015a). Shanker distribution and its applications. *International Journal of Statistics and Applications*, 5(6): 338 – 348.

[3] Shanker, R. (2015b). Akash distribution and its applications. *International Journal of Probability and Statistics*, 4 (3): 65 – 75.

[4] Shanker, R. (2016). Sujatha distribution and its applications. *Statistics in Transition-New Series*, 17(3): 391- 410.

[5] Shanker, R., Hagos, F. and Sujatha, S. (2015). On Modelling of Lifetimes data using exponential and Lindley distributions. *Biometrics & Biostatistics International Journal*, 2(5): 140 - 147.

[6] Shanker, R., Hagos, F. and Sujatha, S. (2016). On modelling of Lifetime data using one parameter Akash, Lindley and exponential distributions. *Biometrics & Biostatistics International Journal*, 3(2): 54 – 62.

[7] Shanker, R. and Hagos, F. (2016). On modelling of Lifetime data using Akash, Shanker, Lindley and exponential distributions. *Biometrics & Biostatistics International Journal,* 3(6): 214 – 224

[8] Shanker, R. (2022): Uma distribution with properties and applications, *Biometrics & Biostatistics International Journal*, 11 (5): 165 – 169.

[9] Shaked, M. and Shanthikumar, J.G. (1994). Stochastic Orders and Their Applications. Academic Press, New York.

[10] Fuller, E.R., Frieman, S.W., Quinn, J.B., Quinn, G.D., and Carter, W.C. (1994). Fracture mechanics approach to the design of glass aircraft windows: A case study, *Proceedings of SPIE - The International Society for Optical Engineering*, 2286: 419-430.

[11] Birnbaum, Z.W. and Saunders, S.C. (1969). Estimation for a family of life distributions with applications to fatigue, *Journal of Applied Probability*, 6(2): 328-347.

ON THE FAILURE RATE FUNCTION OF A SYSTEM UNDER THE δ-SHOCK MODEL

Reza Farhadian[1], Habib Jafari[2] and Hamed Lorvand[3]

•

[1,2]Department of Statistics, Razi University, Kermanshah, Iran
[3]Department of Mathematical Sciences, Isfahan University of Technology, Isfahan, Iran
[1]farhadian.reza@yahoo.com
[2]h.jafari@razi.ac.ir
[3]lorvandhamed@iut.ac.ir

Abstract

An interesting type of shock model in reliability theory is the delta-shock model, which is defined based on the length of time between successive shocks. Whenever this length is less than or equal to a pre-fixed threshold called delta, the system will fail. In this paper, the failure rate function of a system subjected to random shocks under the delta-shock model is discussed. To be more precise, a simple approximate formula for the failure rate function is established. Numerical studies are performed under the assumption that the arrival of shocks follows a Poisson process. Based on the numerical results, the accuracy of the approximation is evaluated, and it is also shown that reducing the mean of inter-arrival times between successive shocks leads to an increase in the failure rate of the system. Conclusions are presented at the end.

Keywords: δ-shock model, failure rate function, approximation

1. Introduction

Shock models are useful tools for analyzing systems that are subjected to random shocks from external sources. It is important to know how a system will behave when it gets shocks. Therefore, shock models have attracted much attention in applied probability, reliability theory, and engineering. Some basic types of shock models are cumulative shock models, extreme shock models, and run shock models. In a cumulative shock model, a system breaks down due to a cumulative effect; in an extreme shock model, a system breaks down due to a single large shock; in a run shock model, the range of a certain number of consecutive shocks is considered a failure criterion. For more on these, see, e.g., [1], [7], [13], and [15]. There are other shock models that have been developed in recent years. The so-called δ-shock model is one of them that has attracted more attention. According to the classical δ-shock model, if the inter-arrival time between two consecutive shocks falls below a critical threshold δ, the system fails. The δ-shock model was first introduced by Li et al. [8] in 1999, after which it was widely studied by many researchers; see, e.g., [9], [10], [11], [12], [18], and [19]. Many extensions and generalizations have been proposed for the δ-shock model. We briefly review some of them below. A generalized δ-shock model is studied in [2], in which the system fails when k consecutive inter-arrival times are less than a threshold δ. A discrete-time version of the δ-shock model is studied in [3], where the shocks occur according to a binomial process, i.e., the inter-arrival times between successive shocks follow a geometric distribution. In [4], the lifetime behavior of the δ-shock model and the censored δ-shock model is studied under the assumption that shocks arrive according to a

renewal process with uniformly distributed inter-arrival times. In [16], the survival function and the mean lifetime of the system failure under the δ-shock model is investigated by considering the proportional hazard rate model. In [14], the reliability analysis of an extension of the discrete time version of the δ-shock model is studied by considering two different critical thresholds and a probable failure region. A general δ-shock model is introduced in [4], in which the recovery time depends on both the inter-arrival time and the magnitude of the shocks. The work by [5] revisited the classical δ-shock model and generalized it to the case of renewal processes of external shocks with arbitrary inter-arrival times and arbitrary distribution of the threshold δ.

In this paper, we will study the failure rate function of a system subjected to random shocks under the δ-shock model. In fact, it is not possible to provide an exact formula for this failure rate function, but here we are trying to establish a simple approximation. To the best of our knowledge, the failure rate function in δ-shock modeling has not been discussed in the literature. Therefore, the results of this paper constitute a significant advancement in this area.

The remainder of the paper is organized as follows. The description of the δ-shock model is derived in Section 2. The failure rate function for a system subjected to random shocks under the δ-shock model is discussed with a simple approximation in Section 3. Numerical results are presented in Section 4. Section 5 concludes the paper.

2. The δ-shock model

Assume that a system is subject to a sequence of external random shocks. Let X_i denote the ith intershock time, i.e., the time between ith and $(i+1)$th shocks for $i = 1, 2, \ldots$. Let also intershock times $X_1, X_2, \ldots$ are independent and identical (i.i.d.) with the cumulative distribution function (cdf) $F(x) = \Pr(X \leq x)$ (with $F(0) = 0$). Under the δ-shock model, the performance of the system is such that if $X_n \leq \delta$ for a given threshold δ, the system fails. Otherwise, the system continues to work without any performance problems under the influence of shocks. Thus, the lifetime of the system is defined as

$$T_\delta = \sum_{i=1}^{N} X_i,$$

where the stopping random variable N is

$$\{N = n\} \Leftrightarrow \{X_1 > \delta, X_2 > \delta, \ldots, X_{n-1} > \delta, X_n \leq \delta\},$$

for $n = 1, 2, \ldots$ and $\delta > 0$.

Obviously, the probability mass function of N is

$$\Pr(N = n) = F(\delta)(1 - F(\delta))^{n-1}, \quad for\ n = 1, 2, \ldots,$$

that is, N follows a geometric distribution with the mean $\frac{1}{F(\delta)}$.

Moreover, the reliability probability function (or survival function) of the system's lifetime T_δ is given by the following formula (see [4]):

$$\Pr(T_\delta > t) = \sum_{n=1}^{\infty} (1 - F(\delta))^{n-1} \int_0^\delta \Pr\left(S_{n-1}^* > t - x\right) dF(x), \tag{1}$$

where S_n^* is the nth arrival time of a renewal process, whose inter-arrival times have the cdf,

$$F^*(x) = \frac{F(x) - F(\delta)}{1 - F(\delta)}, \quad for\ x > \delta.$$

Furthermore, the mean lifetime of the system, that is, the system's mean time to failure (MTTF), is calculated as $E(T_\delta) = \dfrac{E(X)}{F(\delta)}$.

3. Approximation of the failure rate function

Suppose the density function, cumulative distribution function, and survival function of the lifetime T_δ are denoted by $g(t)$, $G(t)$, and $\bar{G}(t)$, respectively. The survival function $\bar{G}(t)$ is the same in Eq. (1). The failure rate (or hazard rate) function is defined as

$$h(t) = \frac{g(t)}{\bar{G}(t)}. \tag{2}$$

Since it is difficult to obtain $g(t)$ from formula (1), it is not possible to obtain an exact formula for the failure rate function of a system subject to random shocks under the δ-shock model. Therefore, approximate formulas can be useful in this case. In the following theorem, we present an approximate formula for $h(t)$ by using Eq. (1).

Theorem 1. We have

$$h(t) \approx \frac{\sum_{n=1}^{\infty} (1-F(\delta))^{n-1} \int_0^\delta \Pr\left(t-x < S_{n-1}^* \leq t+\epsilon-x\right) dF(x)}{\epsilon \left(\sum_{n=1}^{\infty} (1-F(\delta))^{n-1} \int_0^\delta \Pr\left(S_{n-1}^* > t-x\right) dF(x)\right)},$$

where $\epsilon > 0$ is a small number and S_n^* is the nth arrival time of a renewal process, whose inter-arrival times have the cdf, $F^*(x) = \dfrac{F(x)-F(\delta)}{1-F(\delta)}$ for $x > \delta$.

Proof. We have $g(t) = \frac{d}{dt}G(t)$. By using the definition of the derivative of a function, we can write

$$g(t) = \frac{d}{dt}G(t) = \lim_{\epsilon\to 0} \frac{G(t+\epsilon)-G(t)}{\epsilon} = \lim_{\epsilon\to 0} \frac{\bar{G}(t)-\bar{G}(t+\epsilon)}{\epsilon}$$

This implies that when ϵ is small (or $t \approx t+\epsilon$), then

$$g(t) \approx \frac{\bar{G}(t)-\bar{G}(t+\epsilon)}{\epsilon}. \tag{3}$$

Using Eq. (3) in Eq. (2), we have

$$h(t) \approx \frac{\frac{\bar{G}(t)-\bar{G}(t+\epsilon)}{\epsilon}}{\bar{G}(t)} = \frac{\bar{G}(t)-\bar{G}(t+\epsilon)}{\epsilon\bar{G}(t)} \tag{4}$$

Now, by using Eq. (1) in Eq. (4), we obtain

$$\begin{aligned} h(t) &\approx \frac{\sum_{n=1}^{\infty} (1-F(\delta))^{n-1} \int_0^\delta \left(\Pr\left(S_{n-1}^* > t-x\right) - \Pr\left(S_{n-1}^* > t+\epsilon-x\right)\right) dF(x)}{\epsilon \left(\sum_{n=1}^{\infty} (1-F(\delta))^{n-1} \int_0^\delta \Pr\left(S_{n-1}^* > t-x\right) dF(x)\right)} \\ &= \frac{\sum_{n=1}^{\infty} (1-F(\delta))^{n-1} \int_0^\delta \Pr\left(t-x < S_{n-1}^* \leq t+\epsilon-x\right) dF(x)}{\epsilon \left(\sum_{n=1}^{\infty} (1-F(\delta))^{n-1} \int_0^\delta \Pr\left(S_{n-1}^* > t-x\right) dF(x)\right)}, \end{aligned}$$

where S_n^* is the nth arrival time of a renewal process, whose inter-arrival times have the cdf,

$$F^*(x) = \frac{F(x)-F(\delta)}{1-F(\delta)}, \quad for\ x > \delta.$$

The proof is complete. ■

Remark 1. The approximate formula proven in Theorem 1 can also be written in the following simpler form:

$$h(t) \approx \frac{1}{\epsilon}\left(1 - \frac{\sum_{n=1}^{\infty} (1-F(\delta))^{n-1} \int_0^\delta \Pr\left(S_{n-1}^* > t+\epsilon-x\right) dF(x)}{\sum_{n=1}^{\infty} (1-F(\delta))^{n-1} \int_0^\delta \Pr\left(S_{n-1}^* > t-x\right) dF(x)}\right),$$

where ϵ and S_n^* hold under the assumptions of Theorem 1.

4. Numerical results

Assume that the inter-arrival times $X_1, X_2, \ldots$ have exponential distribution with mean $\frac{1}{\lambda}$ for $\lambda > 0$. Such a situation occurs when the arrival of shocks follows a Poisson process. Using Theorem 1, we obtain numerical results on the approximation of the failure rate function $h(t)$. The calculations have been performed considering different values of the parameter λ and the critical threshold δ at some different time points t. For the approximation accuracy trend in terms of ϵ, values $0.01, 0.001$, and 0.0001 have been considered for ϵ. All numerical results are summarized in Table 1. It can be seen from Table 1 that as the parameter λ becomes larger and ϵ becomes smaller, the changes in the approximate values of the failure rate function $h(t)$ gradually decrease until the changes stop at least to four decimal places for $\lambda = 10$ and $\epsilon = 0.0001$. Regardless of the structural features of the approximation formula used, the numerical values show that the values of the failure rate function are increasing as the parameter λ increases. This seems reasonable because an increase in the parameter λ means a decrease in the mean of inter-arrival times between successive shocks, and this leads to the system being stressed by shocks and increasing the probability of system failure.

Table 1: Approximate values for the failure rate function $h(t)$.

λ	δ	t	ϵ	$h(t)$	ϵ	$h(t)$	ϵ	$h(t)$
5	3	4	0.01	4.7580	0.001	4.9973	0.0001	4.8517
		7		4.8587		4.9800		4.9166
		9		4.7457		4.9956		4.8908
		11		4.8784		5.0015		5.3862
		13		4.8778		4.9998		5.0845
	6	4	0.01	4.8768	0.001	4.9973	0.0001	5.3369
		7		4.8770		4.9959		5.0752
		9		4.8768		4.9956		5.2401
		11		4.8711		5.0019		4.6171
		13		4.8762		4.9830		4.9152
10	3	4	0.01	9.5167	0.001	9.9569	0.0001	9.8863
		7		9.5160		9.9612		10.0618
		9		9.5167		9.9585		10.0073
		11		9.5032		9.9473		10.0657
		13		9.5171		10.0830		10.0542
	6	4	0.01	9.5167	0.001	9.9569	0.0001	9.8863
		7		9.5184		9.9612		10.0618
		9		9.5167		9.9585		10.0073
		11		9.5209		9.9473		10.0657
		13		9.5171		9.9681		10.0542

5. Conclusions

In this paper, we establish an approximate formula to calculate the failure rate function of a system subjected to random shocks under the δ-shock model. The accuracy of the approximation was evaluated by numerical studies for the case where the arrival of the shocks follows a Poisson process. Future studies could focus on establishing more approximate formulas and comparing the accuracy of the formulas in approximating the system failure rate function.

References

[1] Anderson, K.K., (1988). A note on cumulative shock models. *J. Appl. Probab*, **25**, 220–223.
[2] Eryilmaz, S., Generalized δ-shock model via runs, *Statistics and Probability Letters*, **82**, 326Ű331, 2012.
[3] Eryilmaz, S. (2013). On the lifetime behavior of a discrete time shock model. *Journal of Computational and Applied Mathematics*, 237, 384–388.
[4] Eryilmaz, S and Bayramoglu, K. (2014). Life behavior of δ-shock models for uniformly distributed interarrival times, *Stat Papers*. **55**, 841–852.
[5] Finkelstein, M. and Cha, J.H. (2024). Discussing some approaches to delta-shock modeling, *TOP* **32**, 245–262.
[6] Goyal, D., Hazra, N.K., and Finkelstein, M. (2022). On the general δ-shock model. *TEST*, **31**, 994–1029.
[7] Gut, A. (1990). Cumulative shock models. *Advances in Applied Probability*, **22**, 504–507.
[8] Li, Z.H., Chan, L.Y. and Yuan, Z.X. (1999). Failure time distribution under a δ-shock model and its application to economic design of system. *International Journal of Reliability, Quality and Safety Engineering*, **6**, 237–247.
[9] Li, Z.H., Huang, B.S. and Wang, G.J. (1999). Life distribution and its properties of shock models under random shocks. *Journal of Lanzhou University*, **35**, 1–7.
[10] Li, Z.H. and Kong, X.B. (2007). Life behavior of δ-shock model. *Statistics and Probability Letters*, **77**, 577–587.
[11] Li, Z.H. and Zhao, P. (2007). Reliability analysis on the δ-shock model of complex systems. *IEEE Transactions on Reliability*, **56**, 340–348.
[12] Ma, M. and Li, Z. (2010). Life behavior of censored δ-shock model. *Indian J Pure Appl Math*, **41**,401–420.
[13] Mallor, F. and Omey E. (2001). Shocks, runs and random sums. *Journal of Applied Probability*, **38**, 438–448.
[14] Poursaeed, M.H. (2019). A generalization of δ-shock model. *Journal of Advanced in Mathematical Modeling*, **9**, 58–70.
[15] Sumita, U. and Shanthikumar, J.G. (1985). A class of correlated cumulative shock models. *Advances in Applied Probability*, **17**, 347–366.
[16] Tuncel, A. and Eryilmaz, S. (2018). System reliability under δ-shock model, *Commun Stat Theory Methods*, **47**, 4872–4880.
[17] Wang, G.J., and Zhang, Y.L. (2001). δ-shock model and its optimal replacement policy. *Journal of Southeast University*, **31**, 121–124.
[18] Wang, G.J. and Zhang, Y.L. (2007). A shock model with two-type failures and optimal replacement policy. *International Journal of Systems Science*, **36**, 209–214.
[19] Xu, Z.Y. and Li, Z.H. (2004). Statistical inference on d shock model with censored date. *Chinese J. Appl. Probab. Statist*. **20**, 147–153.

RT&A, No 2 (84)
Volume 20, June 2025

COMPARATIVE ANALYSIS OF BAYESIAN ESTIMATION TECHNIQUES FOR GAMMA DISTRIBUTIONS

Vijay Kumar Lingutla[1], Nagamani Nadiminti[2], R. K. Davala[3]

•

[1,2,3]Department of Mathematics, School of Advanced Sciences, VIT-AP University, Inavolu, Beside AP Secretariat, Amaravati AP-522237, India

[1]vijay.lingutla@gmail.com, [2]maniinadiminti999@gmail.com, [3]davalaravikumar@gmail.com

Abstract

This paper explores the estimation of the common scale parameter for two Gamma populations using three Bayesian methods: Approximate Bayesian Computation (ABC), Hamiltonian Monte Carlo (HMC), and Metropolis-Hastings (MH). Unlike traditional methods like Maximum Likelihood Estimation (MLE), which can be sensitive to model misspecification, these Bayesian methods are more flexible in handling parameter uncertainty and complex data. ABC bypasses explicit likelihood calculations, HMC improves sampling efficiency through gradient information, and MH offers simplicity and adaptability. A simulation study compares these methods based on bias, mean squared error (MSE), and computational efficiency. Additionally, we apply these methods to real glucose level data from male and female individuals with diabetes, highlighting their strengths and limitations. The results offer practical recommendations for real-world applications involving Gamma distributions, contributing to the advancement of Bayesian estimation techniques in complex parameter estimation problems.

Keywords: Gamma Distribution, Bayesian estimation, Approximate Bayesian Computation, Hamiltonian Monte Carlo, Metropolis-Hastings.

1. Introduction

The gamma distribution is widely used in reliability analysis, survival studies, and queuing theory due to its flexibility in modeling positive continuous data Lawless [1]. A key parameter in these applications is the scale parameter, which determines the distribution's spread. Estimating this parameter is relatively straightforward when the shape parameters of the gamma populations are known. However, when these shape parameters are unknown and may differ across populations, the estimation problem becomes significantly more complex.

Traditional methods, such as Maximum Likelihood Estimation (MLE), are commonly used for parameter estimation in gamma distributions. Although MLE is asymptotically efficient, it is highly sensitive to deviations from model assumptions, such as outliers or misestimated shape parameters Huber [2]. These limitations highlight the need for more flexible and robust estimation techniques, especially in real-world applications.

Computational Bayesian methods provide a promising alternative. This study explores three advanced Bayesian techniques: Approximate Bayesian Computation (ABC), Hamiltonian Monte Carlo (HMC), and the Metropolis-Hastings (MH) algorithm. ABC is particularly useful when likelihood functions are intractable, as it relies on summary statistics for inference Beaumont [3]. HMC efficiently samples high-dimensional spaces by leveraging gradient information, leading to faster convergence and lower autocorrelation compared to traditional Markov Chain Monte Carlo

(MCMC) methods Neal [4]. Robert and Casella [5] further examined its efficiency in improving Markov chain mixing. Betancourt [6] provided a detailed explanation of why HMC excels in high-dimensional settings, reinforcing its inclusion in this study. Meanwhile, MH remains widely used for posterior sampling due to its simplicity and flexibility Hastings [7]. Chib and Greenberg [8] later explored its practical applications in Bayesian inference, demonstrating its adaptability across statistical models.

A common assumption in analyzing multiple gamma populations is that their scale parameters are equal, simplifying model formulation and improving statistical efficiency Gelman et al. [9]. Rubin [10] further explored the advantages of this assumption, particularly in Bayesian inference, where hierarchical structures enable borrowing strength across related populations. Ibrahim, Chen, and Sinha [11] provided a comprehensive discussion on Bayesian methods for survival data, highlighting their relevance to this study. Recent advancements in adaptive MCMC techniques have further improved estimation robustness Andrieu et al. [12]. Additionally, Andrieu et al. [13] introduced particle MCMC methods, which integrate sequential Monte Carlo with MCMC strategies to handle complex posterior distributions.

While Bayesian methods have been studied in various contexts, few have directly compared their performance in estimating the common scale parameter of two gamma populations with unknown shape parameters. This research addresses that gap by systematically evaluating ABC, HMC, and MH based on bias, mean squared error (MSE), and computational efficiency. The findings will highlight each method's strengths and limitations, offering practical guidance for their application.

This study builds on the work of Nagamani and Tripathy [14], who estimated the common scale parameter for two gamma populations using simulation-based methods. We extend their approach by incorporating advanced Bayesian techniques for a more comprehensive evaluation. Through a simulation-based study, we contribute to the growing literature on Bayesian estimation Gilks et al. [15]. Spiegelhalter et al. [16] refined these methods, particularly in model checking and predictive inference. Christensen et al. [17] provided insights into Bayesian estimation for various statistical models, including gamma distributions, reinforcing the theoretical foundation of this study. Prior research has also examined Bayesian inference for complex distributions [18], with Liu et al. [19] exploring Bayesian approaches for intricate distributional structures. Kamal et al. [20] further contributed by studying Bayesian estimation for parametric models with unknown distributions, highlighting the need for robust Bayesian methods. Additionally, Sisson, Fan, and Beaumont [21] provided an authoritative discussion on ABC methods, detailing their advantages and limitations, further justifying their use in complex estimation scenarios.

The paper is structured as follows: Section 2 discusses the methodology, covering the gamma distribution model, simulation setup, and parameter estimation methods—ABC, HMC, and MH—along with their assumptions and hyperparameters. Section 3 presents the simulation results, evaluating and comparing the methods based on key performance metrics. Section 4 applies these methods to real-life glucose level data from male and female individuals with diabetes, demonstrating the practical use of Bayesian estimation techniques in analyzing real-world complex data. Finally, Section 5 concludes the paper by summarizing the findings, offering recommendations for choosing the appropriate estimation method, and suggesting directions for future research in Bayesian parameter estimation.

2. Methodology

Consider two independent random samples drawn from Gamma distributions: $X = (x_1, x_2, \ldots, x_m)$ and $Y = (y_1, y_2, \ldots, y_n)$ with sizes m and n, respectively. These samples are assumed to follow Gamma distributions $G(\mu_1, \beta)$ and $G(\mu_2, \beta)$ with a common scale parameter β and different shape parameters μ_1 and μ_2. The probability density functions (PDFs) for these Gamma distributions are:

$$f(x; \mu_1, \beta) = \frac{x^{\mu_1 - 1} e^{-x/\beta}}{\Gamma(\mu_1)\beta^{\mu_1}}, \quad x > 0,\ \mu_1 > 0,\ \beta > 0, \tag{1}$$

$$f(y;\mu_2,\beta)=\frac{y^{\mu_2-1}e^{-y/\beta}}{\Gamma(\mu_2)\beta^{\mu_2}},\quad y>0,\ \mu_2>0,\ \beta>0. \tag{2}$$

The Bayesian framework provides a robust approach to parameter estimation, leveraging prior knowledge to improve inference, particularly in complex models such as Gamma distributions. In this study, we consider Bayesian estimation of the common scale parameter β and the shape parameters μ_1 and μ_2 for two Gamma populations. Let $\pi_1(\beta)$, $\pi_2(\mu_1)$, and $\pi_3(\mu_2)$ represent the prior density functions for β, μ_1, and μ_2, respectively.

The likelihood function of (β,μ_1,μ_2) for the observed data (X,Y) is expressed as:

$$L(\beta,\mu_1,\mu_2|X,Y)=\prod_{i=1}^{m}\frac{1}{\Gamma(\mu_1)\beta^{\mu_1}}x_i^{\mu_1-1}e^{-x_i/\beta}\prod_{j=1}^{n}\frac{1}{\Gamma(\mu_2)\beta^{\mu_2}}y_j^{\mu_2-1}e^{-y_j/\beta},$$

where $X=\{x_1,x_2,\ldots,x_m\}$ and $Y=\{y_1,y_2,\ldots,y_n\}$ are samples from the two Gamma populations. The posterior density function of (β,μ_1,μ_2) given (X,Y) is:

$$f(\beta,\mu_1,\mu_2|X,Y)=\frac{f(\beta,\mu_1,\mu_2,X,Y)}{\int_0^\infty\int_0^\infty\int_0^\infty f(\beta,\mu_1,\mu_2,X,Y)\,d\beta\,d\mu_1\,d\mu_2}.$$

The posterior expectation of any function $g(\beta,\mu_1,\mu_2)$ is:

$$E[g(\beta,\mu_1,\mu_2)|X,Y]=\frac{\int_0^\infty\int_0^\infty\int_0^\infty g(\beta,\mu_1,\mu_2)f(\beta,\mu_1,\mu_2,X,Y)\,d\beta\,d\mu_1\,d\mu_2}{\int_0^\infty\int_0^\infty\int_0^\infty f(\beta,\mu_1,\mu_2,X,Y)\,d\beta\,d\mu_1\,d\mu_2}. \tag{3}$$

The integral ratio in equation (3) is analytically intractable due to the complexity of the posterior distribution. Therefore, numerical techniques such as Hamiltonian Monte Carlo (HMC), Metropolis-Hastings (MH), or Approximate Bayesian Computation (ABC) are employed to approximate the posterior distributions. These methods enable efficient computation of Bayesian estimators for the parameters under different prior distributions and loss functions.

This Bayesian framework provides a comprehensive basis for comparing computational methods in terms of their accuracy, efficiency, and applicability to real-world data involving Gamma populations.

2.1. Approximate Bayesian Computation (ABC)

Approximate Bayesian Computation (ABC) is a powerful Bayesian inference method used when the likelihood function is difficult or impossible to compute directly. Instead of using the likelihood function, ABC uses a set of summary statistics to approximate the posterior distribution [3]. The ABC method is particularly useful when the model is complex and simulating data from the model is easier than evaluating the likelihood function.

ABC Procedure

- Two populations are modeled as Gamma distributions, with the following relationships for the survival times:
$$X\sim\text{Gamma}(\mu_1,\beta),\quad Y\sim\text{Gamma}(\mu_2,\beta).$$
Here, X and Y represent the survival times for the two populations, with β being the common scale parameter, and μ_1 and μ_2 as the shape parameters.

- The prior distributions for the parameters are chosen as:
$$\beta\sim\text{Gamma}(b_1,c_1),\quad \mu_1\sim\text{Gamma}(b_2,c_2),\quad \mu_2\sim\text{Gamma}(b_3,c_3).$$

- Synthetic data are generated by drawing samples from the prior distributions and simulating data based on the Gamma model for both populations.

- The summary statistics (such as the mean and variance) of both the observed data (S_{obs}) and the simulated data (S_{sim}) are computed. A distance metric, typically the Euclidean distance, is then used to quantify the difference:

$$d = \|S_{\text{obs}} - S_{\text{sim}}\|.$$

- A tolerance level ϵ is defined, and parameter sets are accepted if the distance satisfies $d \leq \epsilon$.
- The procedure is repeated iteratively, generating synthetic data and comparing it to the observed data. Accepted parameter sets help build an approximate posterior distribution.
- From the approximate posterior distribution, point estimates such as the posterior mean or median are derived, along with credible intervals for the parameters.

This method provides a flexible and robust framework for Bayesian inference under complex data scenarios by approximating posterior distributions without requiring explicit likelihood functions. It efficiently handles parameter uncertainty and supports reliability analysis for two gamma populations.

2.2. Hamiltonian Monte Carlo (HMC)

Hamiltonian Monte Carlo (HMC) is a Markov Chain Monte Carlo (MCMC) method that utilizes Hamiltonian dynamics to propose new states in the parameter space. This method is known for its efficiency in exploring high-dimensional parameter spaces by using gradient information from the posterior distribution [4]. Unlike traditional random walk-based MCMC algorithms, HMC uses the gradient of the target distribution to guide the chain's movements, leading to faster mixing and reduced autocorrelation between samples. In our study, HMC is applied to sample from the posterior distribution of the common scale parameter, allowing for efficient estimation even in higher-dimensional settings.

HMC Procedure

- Two populations are modeled as Gamma distributions with common scale parameter β and shape parameters μ_1 and μ_2:

$$X \sim \text{Gamma}(\mu_1, \beta), \quad Y \sim \text{Gamma}(\mu_2, \beta).$$

- The prior distributions for the parameters are chosen as:

$$\beta \sim \text{Gamma}(b_1, c_1), \quad \mu_1 \sim \text{Gamma}(b_2, c_2), \quad \mu_2 \sim \text{Gamma}(b_3, c_3).$$

- The Hamiltonian Monte Carlo method involves simulating the joint distribution of parameters, leveraging their gradients to propose new values for the parameters. This is done by introducing auxiliary momentum variables and using Hamiltonian dynamics to simulate the movement of the system.
- The trajectory of the parameters is simulated using the Hamiltonian dynamics, and the positions are updated using the leapfrog integration method to minimize the energy (or maximize the posterior distribution).
- A Metropolis acceptance step is applied to the proposed new position, where the new state is accepted with a probability given by the ratio of the posterior densities at the new and old positions.
- The process is repeated for multiple iterations, and a sequence of samples is generated from the posterior distribution of the parameters.

- From the posterior samples, point estimates such as the posterior mean or median are derived, along with credible intervals for the parameters.

HMC provides an efficient method for exploring complex high-dimensional posterior distributions by leveraging gradient information, which improves convergence and reduces autocorrelation compared to traditional MCMC methods.

2.3. Metropolis-Hastings (MH)

Metropolis-Hastings (MH) is one of the most widely used MCMC algorithms for sampling from posterior distributions. It relies on a proposal distribution to generate candidate samples and accepts or rejects these candidates based on a specified acceptance probability [7]. The MH algorithm is a versatile method that can be adapted to a wide variety of problems. The general procedure for MH involves:

MH Procedure

- Two populations are modeled as Gamma distributions with the following relationships for the survival times:
$$X \sim \text{Gamma}(\mu_1, \beta), \quad Y \sim \text{Gamma}(\mu_2, \beta).$$
Here, X and Y represent the survival times for the two populations, with β being the common scale parameter, and μ_1 and μ_2 as the shape parameters.
- The prior distributions for the parameters are chosen as:
$$\beta \sim \text{Gamma}(b_1, c_1), \quad \mu_1 \sim \text{Gamma}(b_2, c_2), \quad \mu_2 \sim \text{Gamma}(b_3, c_3).$$
- The Metropolis-Hastings algorithm starts with an initial value for the parameters and proposes new values based on a proposal distribution. The proposal distribution is typically chosen to be symmetric, such as a normal distribution.
- The ratio of the posterior densities at the new and old parameter values is computed, and the new parameters are accepted with a probability given by the Metropolis criterion:
$$\mu = \min\left(1, \frac{p(\theta_{\text{new}}|X, Y)}{p(\theta_{\text{old}}|X, Y)}\right).$$
If the new parameters are accepted, they replace the old ones; otherwise, the old parameters are retained.
- This process is repeated iteratively for a set number of steps or until convergence is reached, generating a sequence of samples from the posterior distribution of the parameters.
- From the posterior samples, point estimates such as the posterior mean or median are derived, along with credible intervals for the parameters.

The Metropolis-Hastings algorithm is widely used for sampling from complex posterior distributions due to its simplicity and flexibility. It is particularly effective in situations where the likelihood is difficult to compute directly.

3. Simulation Study

To assess the performance of the proposed methodology for estimating the parameters of two Gamma populations, we conducted an extensive simulation study. The study aimed to evaluate the accuracy, robustness, and efficiency of the Metropolis-Hastings (MH) algorithm in recovering the true parameter values across various experimental conditions, while also including the

Hamiltonian Monte Carlo (HMC) and Approximate Bayesian Computation (ABC) methods for comparison.

In this section, we present experimental results evaluating the performance of different estimation methods across varying sample sizes. We estimate the common scale parameter β and the potentially distinct shape parameters μ_1 and μ_2 using the ABC, HMC, and MH algorithms.

Two sets of parameter values are considered:

- Set 1: $(\mu_1, \mu_2, \beta) = (1, 2, 1.5)$
- Set 2: $(\mu_1, \mu_2, \beta) = (1, 2, 2.5)$

We evaluate the performance of the estimators using two metrics:

- Average Value (Estimate)
- Mean Square Error (MSE)

To facilitate comparison, we generate 10,000 random samples from the two Gamma populations across various sample sizes and parameter combinations. Specific hyperparameters ($b_1 = b_2 = b_3 = 1$ and $c_1 = c_2 = c_3 = 2$) are employed to calculate biases and MSE for all estimators, with results summarized in the tables below.

Table 1 & 2 presents the Estimate and MSE of MLE and Bayesian estimators under a SEL (squared error loss) function. The first column indicates sample sizes and the second column details the parameters μ_1, μ_2, and β. Columns 3 and 4 show the Estimate and MSE of the ABC estimates, while subsequent columns reflect the estimates of HMC and MH algorithm.

Tables 3 & 4 display the 95% HPD intervals derived from HMC and MH for the parameters. The first two columns indicate sample sizes and parameters, the third column shows asymptotic confidence intervals calculated using the HMC, and the fourth column presents the HPD intervals estimated through MH algorithm.

The observations from our simulation study yield several insights into the performance of the estimators:

- Increasing sample sizes consistently reduce MSE across all estimators, underscoring the importance of data size for estimation accuracy.
- ABC, HMC and MH converge to stable estimates as sample sizes increase, ensuring consistency and reliability.
- ABC, HMC and MH algorithms performed well, but HMC consistently produced slightly more precise estimates with narrower posterior distributions. ABC & MH occasionally exhibited higher variability, especially for smaller sample sizes.
- The accuracy and precision of parameter estimates improved significantly as sample sizes increased. Smaller sample sizes resulted in wider 95% Highest Posterior Density (HPD) intervals, indicating greater uncertainty, while larger sample sizes yielded narrower intervals.
- Compared to traditional methods like Gibbs sampling and MLE, our study shows that HMC offers smoother sampling and faster convergence, while ABC demonstrated strong robustness, especially in noisy data scenarios. The Metropolis-Hastings algorithm performed competitively with Gibbs sampling for larger sample sizes but showed slightly higher MSE in smaller samples.

4. Real Data analysis

To validate our model, we analyze glucose levels among male and female individuals with diabetes who have never smoked. The objective is to identify potential gender-based differences in blood sugar regulation. The dataset, sourced from [22] and openly available, serves as the

Table 1: *For various sample sizes, we compare biases and mean square errors of multiple estimators under squared error loss for $\theta = (\mu_1, \mu_2, \beta)$.*

		ABC		HMC		MH	
(m, n)	Observed	Avg	Error	Avg	Error	Avg	Error
(10,10)	μ_1=1	1.7905	0.6628	1.1535	0.1273	1.3618	0.2541
	μ_2=2	1.6857	0.5596	1.2527	0.6027	1.8651	0.4216
	β=1.5	1.2389	0.3975	1.6196	0.0678	1.4224	0.1889
(20,10)	μ_1=1	1.6831	0.5541	0.8695	0.04	0.9881	0.2450
	μ_2=2	2.6094	0.4185	2.0003	0.1467	2.1232	0.3532
	β=1.5	1.04	0.2177	1.5748	0.0362	1.4375	0.1432
(10,20)	μ_1=1	1.7955	0.5234	1.1233	0.0563	1.4720	0.2424
	μ_2=2	2.5821	0.6474	1.9726	0.1209	2.2212	0.2847
	β=1.5	1.038	0.2586	1.4731	0.0301	1.2743	0.1611
(30,30)	μ_1=1	1.2239	0.1997	0.9195	0.0218	0.9970	0.0301
	μ_2=2	1.7548	0.2896	1.71	0.1552	1.7605	0.1887
	β=1.5	1.4231	0.1077	1.4259	0.0295	1.3532	0.0941
(50,50)	μ_1=1	1.1262	0.1623	0.9873	0.0201	0.9643	0.0191
	μ_2=2	1.9885	0.2107	2.0923	0.0886	2.3846	0.1020
	β=1.5	1.4474	0.0963	1.4652	0.0198	1.3364	0.0718

Table 2: *For various sample sizes, we compare biases and mean square errors of multiple estimators under squared error loss for $\theta = (\mu_1, \mu_2, \beta)$.*

		ABC		HMC		MH	
(m, n)	Observed	Avg	Error	Avg	Error	Avg	Error
(10,10)	μ_1=1	1.7383	0.9155	1.291	0.2019	1.4806	0.3656
	μ_2=2	1.4756	0.4434	1.4266	0.3709	2.0112	0.4304
	β=2.5	2.5085	0.714	2.1637	0.266	2.1064	0.4615
(20,10)	μ_1=1	1.0409	0.0017	0.9642	0.1341	1.0525	0.2491
	μ_2=2	2.5289	0.2797	2.238	0.2216	2.2882	0.3655
	β=2.5	2.3103	0.036	2.2247	0.1688	2.1455	0.3645
(10,20)	μ_1=1	1.5312	0.289	1.2376	0.1073	1.4720	0.2424
	μ_2=2	1.5578	0.2059	2.2075	0.1531	2.2212	0.2847
	β=2.5	2.2135	0.0906	2.1395	0.1662	1.2743	0.3611
(30,30)	μ_1=1	0.801	0.1975	0.9858	0.0195	1.0349	0.0307
	μ_2=2	1.4272	0.7557	1.8763	0.1037	1.8547	0.2535
	β=2.5	2.6424	0.6259	2.1338	0.1073	2.1037	0.2996
(50,50)	μ_1=1	1.7408	0.5488	0.9148	0.0155	0.9937	0.0194
	μ_2=2	2.2353	0.0554	2.1885	0.0943	2.4751	0.1760
	β=2.5	1.7755	0.5248	2.3097	0.0947	2.1290	0.2338

Table 3: *95% HPD intervals using HMC and MH algorithams for the parameters (μ_1, μ_2, β)=(1, 1.5, 2) at various sample sizes*

(m,n)	Observed	HMC	MH
(10,10)	μ_1=1	[0.5781, 1.6895]	[0.7294, 2.0582]
	μ_2=2	[0.8307, 2.6554]	[0.7122, 2.8704]
	β=1.5	[1.1448, 2.087]	[0.857, 2.3381]
(20,10)	μ_1=1	[0.6122, 1.1922]	[0.6248, 1.4402]
	μ_2=2	[1.2255, 2.4953]	[0.8936, 3.4056]
	β=1.5	[1.2714, 1.9175]	[0.871, 2.2512]
(10,20)	μ_1=1	[0.7751, 1.5938]	[0.7907, 2.127]
	μ_2=2	[1.2922, 2.4633]	[1.1446, 3.1959]
	β=1.5	[1.1423, 1.8007]	[0.7812, 2.0972]
(30,30)	μ_1=1	[0.6976, 1.1776]	[0.6866, 1.3578]
	μ_2=2	[1.2081, 2.108]	[1.0913, 2.3701]
	β=1.5	[1.1481, 1.7068]	[0.9123, 1.9463]
(50,50)	μ_1=1	[0.7833, 1.2777]	[0.7247, 1.2439]
	μ_2=2	[1.5996, 2.5279]	[1.7072, 3.102]
	β=1.5	[1.1658 1.8268]	[0.9585, 1.781]

Table 4: *95% HPD intervals using HMC and MH algorithams for the parameters (μ_1, μ_2, β)=(1, 1.5, 2) at various sample sizes*

(m,n)	Observed	HMC	MH
(10,10)	μ_1=1	[0.6585, 1.8312]	[0.8111, 2.1997]
	μ_2=2	[0.9728, 2.7928]	[0.8452, 3.2114]
	β=2.5	[1.7432, 2.627]	[1.3433, 3.3201]
(20,10)	μ_1=1	[0.6758, 1.3285]	[0.6798, 1.481]
	μ_2=2	[1.4236, 2.7487]	[1.0481, 3.6334]
	β=2.5	[1.8816, 2.6857]	[1.4003, 3.1843]
(10,20)	μ_1=1	[0.8573, 1.7827]	[0.8826, 2.2122]
	μ_2=2	[1.5443, 2.6648]	[1.2453, 3.3589]
	β=2.5	[1.7822 2.9561]	[1.2655, 2.965]
(30,30)	μ_1=1	[0.735, 1.2751]	[0.7218, 1.3871]
	μ_2=2	[1.4345, 2.2498]	[1.22, 2.5561]
	β=2.5	[1.8064, 2.8706]	[1.4851, 2.8742]
(50,50)	μ_1=1	[0.7341, 1.0959]	[0.7384, 1.2774]
	μ_2=2	[1.7131, 2.6098]	[1.7623, 3.1623]
	β=2.5	[1.833, 2.7369]	[1.9046, 2.8637]

foundation for this study. The primary goal is to assess glucose level distributions across genders and evaluate whether they follow a Gamma distribution. Both visual and statistical techniques are employed to examine the goodness-of-fit of the Gamma model, ensuring a robust assessment of its applicability to the data.

A collection of 50 sample observations related to glucose levels among male and female individuals are as follows:
Glucose Levels Among Male: 124.3, 180.5, 72.9, 90.8, 196.2, 75.1, 131.5, 128.5, 117.8, 75, 118, 124, 184.7, 152.3, 149.9, 162.8, 70.7, 101.7, 99.4, 125.7, 120.3, 73.6, 161.1, 92.1, 119.7, 163.5, 166.8, 187.9, 103.1, 198.9, 89.3, 160.7, 169, 195.9, 91.2, 194.1, 109.4, 84.8, 122.9, 101, 162.7, 138.3, 106.9, 152.8, 161.2, 88.9, 80.7, 163.7, 104.6, 177.4.

Glucose Levels Among Females: 157.2, 125.9, 127.7, 147.9, 145.8, 190.2, 151.4, 105.7, 85.8, 179.4, 154.1, 86.8, 186.2, 94.6, 94.8, 170, 95.8, 108.1, 139.5, 180.4, 165.9, 165.7, 81.9, 173.1, 92.4, 118.9, 165.9, 137.1, 92.3, 75.8, 181.1, 70.5, 163.4, 156.3, 85.8, 81.6, 177.8, 104.1, 79.8, 144, 174.8, 94.3, 90.8, 72.8, 123.8, 190.8, 140.2, 134.9, 71, 100.3.

In this study, we employ both visual and statistical techniques to assess the goodness-of-fit of the Gamma distribution for glucose levels across genders.

The Q-Q plot shown in Figure 1 demonstrates that the Gamma distribution provides a strong fit for the observed glucose levels in both male and female groups. The majority of data points align well with the reference line, particularly in the middle range, indicating that the Gamma model effectively captures the central trend of glucose levels. While some deviations are observed in the upper tail, suggesting slight underestimation of higher glucose values, the overall fit remains robust. This suggests that the Gamma distribution is a suitable choice for modeling glucose levels, as it accurately represents the underlying data structure while capturing key distributional characteristics.

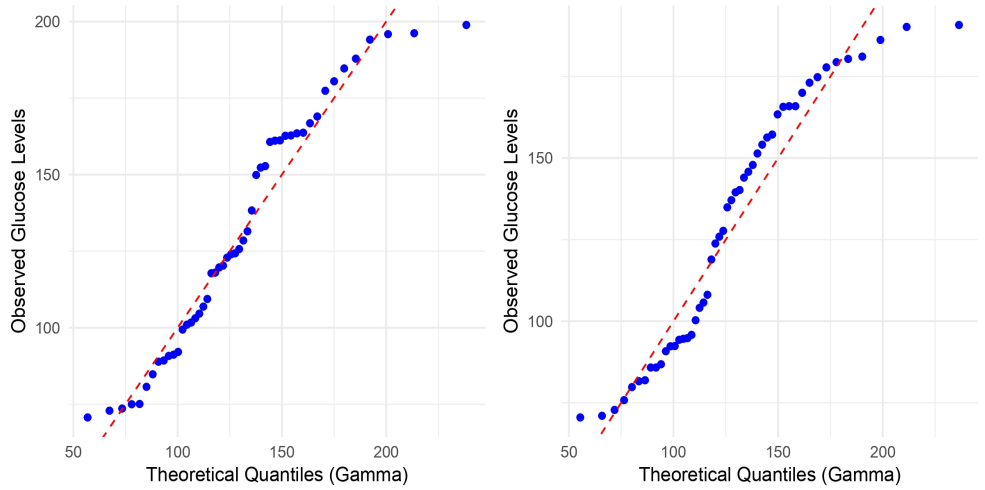


Figure 1: *Q-Q Plot curve compared with the theoretical Gamma model.*

To statistically test whether the scale parameters of the two populations are equal, we performed the Wald test. The null hypothesis for the Wald test is that the scale parameters of the two populations are equal ($H_0 : \beta_1 = \beta_2 = \beta$). The test results are presented in Table 5, with the computed Wald statistic of 0.000832 and a corresponding p-value of 0.97698. Since the p-value is greater than the standard significance level ($\alpha = 0.05$), we fail to reject the null hypothesis. This implies that there is insufficient evidence to conclude that the scale parameters are significantly different between the two populations, suggesting that the scale parameters may be equal.

The results of parameter estimation for the Weibull distribution, using Bayesian estimation methods, are summarized in Tables 6 and 7 . Table 6 presents the point estimates for the scale

Table 5: *Wald Test for Equality of Shape Parameters*

Data	$\hat{\beta}$	**V(β)**	**Statistic**	**P-value**
Set 1	11.82508	2.378837	0.000832	0.05726
Set 2	11.92254	2.398886		

Table 6: *Maximum Likelihood estimators and Bayes estimators of the combined model*

Method	μ_1	μ_2	β
ABC	9.1273	9.4199	12.3821
HMC	9.4844	9.9056	12.7042
MH	9.3260	9.1939	12.8135

parameter (β) and the two shape parameters (μ_1 and μ_2) obtained through various Bayesian estimation techniques: Approximate Bayesian Computation(ABC), Hamiltonian Monte Carlo (HMC), and Metropolis-Hastings (MH).

The point estimates for the Weibull distribution parameters, including the shape parameters (μ_1, μ_2) and the scale parameter (β), were obtained using three different Bayesian estimation methods: Approximate Bayesian Computation (ABC), Hamiltonian Monte Carlo (HMC), and Metropolis-Hastings (MH). Among these, HMC provided the most accurate and consistent estimates, with values closest to each other for all parameters. ABC produced estimates that were slightly lower than those from HMC, especially for the shape parameters, while MH showed slightly more variability in its estimates. Overall, HMC yielded the best point estimates, indicating its efficiency in parameter estimation compared to the other methods.

The 95% Highest Posterior Density (HPD) intervals for the parameters revealed the precision of each Bayesian method. HMC produced narrower and more precise HPD intervals, indicating higher confidence in the parameter estimates, with values tightly clustered around the point estimates. In contrast, MH generated wider HPD intervals, reflecting greater uncertainty in the parameter estimates. ABC's HPD intervals were not provided here, but based on the point estimates, it is expected that the intervals would be less precise than those from HMC. The narrow HPD intervals from HMC suggest it as the most reliable method for quantifying uncertainty in the parameters.

5. Conclusion

This study compared three Bayesian parameter estimation methods—Approximate Bayesian Computation (ABC), Hamiltonian Monte Carlo (HMC), and Metropolis-Hastings (MH)—for estimating the parameters of two Gamma populations in the context of analyzing glucose levels in male and female individuals with diabetes who have never smoked. The simulation results revealed that increasing the sample size consistently improved the accuracy of all methods. HMC generally provided the most precise estimates, with narrower posterior distributions and lower Mean Squared Error (MSE), particularly for larger sample sizes. ABC, while robust and effective in

Table 7: *95% HPD intervals using HMC and MH Algorithms for the parameters* (μ_1, μ_2, β)

Method	μ_1	μ_2	β
HMC	[8.6691, 10.0549]	[9.2609, 10.8336]	[12.2959, 13.0811]
MH	[4.9786, 11.4956]	[5.7222, 11.7434]	[10.3235, 15.5177]

handling noisy data, showed higher variability with smaller samples. MH produced results similar to Gibbs sampling, with competitive outcomes for larger sample sizes but slightly higher MSE for smaller ones. These findings suggest that HMC was the most efficient method, offering faster convergence and better mixing, making it particularly suitable for high-dimensional estimation tasks, such as analyzing complex glucose level data. On the other hand, ABC demonstrated robustness in more complex models, especially when the data showed significant noise. Overall, while HMC emerged as the preferred choice for accurate and efficient estimation, ABC and MH each showcased their strengths in particular contexts, emphasizing the trade-offs between precision and computational efficiency in real-world applications like diabetes research.

References

[1] Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*. Wiley-Interscience.
[2] Huber, P. J. (2009). *Robust Statistics*. Wiley.
[3] Beaumont, M., Zhang, W., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, **162**(4), 2025–2035.
[4] Neal, R. M. (2011). MCMC using Hamiltonian dynamics. In *Handbook of Markov Chain Monte Carlo* (pp. 113–162). CRC Press.
[5] Robert, C. P., Casella, G., & Chopin, N. (2013). *Monte Carlo Statistical Methods* (3rd ed.). Springer.
[6] Betancourt, M. (2017). A Conceptual Introduction to Hamiltonian Monte Carlo. *arXiv preprint arXiv:1701.02434*.
[7] Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**(1), 97–109.
[8] Chib, S. (1995). Marginal likelihood from the Metropolis-Hastings output. *Journal of the American Statistical Association*, **90**(432), 1313–1321.
[9] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd ed.). CRC Press.
[10] Rubin, D. B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *The Annals of Statistics*, **12**(4), 1151–1172.
[11] Ibrahim, J. G., Chen, M. H., & Sinha, D. (2001). Bayesian Survival Analysis. *Springer*.
[12] Andrieu, C., Thoms, J., & Doucet, A. (2010). Adaptive MCMC: Theory and applications. *Springer*.
[13] Andrieu, C., Doucet, A., & Holenstein, R. (2008). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **70**(2), 269–292.
[14] Nagamani, N., & Tripathy, M. R. (2017). Estimating common scale parameter of two gamma populations: A simulation study. *American Journal of Mathematical and Management Sciences*, **36**(4), 346–362.
[15] Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (1995). *Markov Chain Monte Carlo in Practice*. CRC Press.
[16] Spiegelhalter, D. J., Abrams, K. R., & Myles, J. P. (2002). *Bayesian Methods in Health Technology Assessment*. Wiley.
[17] Christensen, R., Johnson, W., Branscum, A., & Hanson, T. (2011). Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians. *CRC Press*.
[18] Bernardo, J. M. (2000). *Bayesian Theory*. Wiley.
[19] Liu, F., Li, H., & Wang, Y. (2023). Efficient parameter estimation for hierarchical models using ABC and Hamiltonian Monte Carlo. *Bayesian Analysis*, **18**(1), 1–21.
[20] Kamal, M. I., Singh, G., & Chatterjee, S. (2022). Bayesian analysis of hierarchical models using approximate Bayesian computation. *Statistical Modelling*, **22**(3), 221–246.
[21] Sisson, S. A., Fan, Y., & Beaumont, M. (2018). Handbook of Approximate Bayesian Computation. *CRC Press*.
[22] Ankit. (2025). Diabetes Prediction in India [Data set]. Kaggle. DOI: 10.34740/KAGGLE/DSV/10594461.

STOCHASTIC ANALYSIS THROUGH MATLAB OF TWO-UNIT PARALLEL SYSTEM

Sonia[1*], Shakuntla Singla[2]

•

[1*]Department of mathematics, Dr. B. R. Ambedkar Govt. College Kaithal (Jagdishpura), Haryana, India
[1*,2]Departent of mathematics and humanities, MMEC, Maharishi Markandeshwar(Deemed to be University), Mullana, Ambala, Haryana, India
[1*]Sonia_garg99@yahoo.com, [2]shakus25@gmail.com
Corresponding Author Email: sonia_garg99@yahoo.com

Abstract

This paper presents a reliability modelling of a two parallel unit of manufacturing Industry. Original conservation figures of the manufacturing industry have been applied for this determination. Four kinds of failure were observed: developmental, electronical, motorized and infrastructural failures. Both units of manufacturing plant work in parallel and independently. Various reliability parameters of the plant such as availability, busy period for repair, expected number of repairs and profitability for each type of failure are examined. Graphical Technique of Regenerative Point, process of Markov and Genetic Algorithm are applied for analysis. Profit analysis for the plant is examined with respect to returns of the plant and repair rate along with a graphical representation.

Keywords: Manufacturing, Markov process, regenerative point graphical techniques, repairs, failures

I. Introduction

Many investigators have studied reliability modelling and analysis of huge complex industrial systems under various operating conditions and presumptions for contributing to the discipline. Al Rahbi et al. [1,2] analyzed the reliability of single unit/ multiple units of a rodding anode plant in the aluminum industry with single/numerous repairs and optimized maintenance strategies, improved system reliability, and reduced downtime, leading to increased productivity and cost savings. Rizwan et al. [3] presented a Sensitivity analysis of reliability of Fuel cell. Sonia et al [4] discussed in her paper analysis of reliability through multiple way like normal distribution, exponential distribution, weibull distribution and observe that the normal distribution is the best fit for reliability analyze. Shakuntla Singla, Sonia [5] analysis the reliability of mushroom plant by the use of regenerative point graphical technique. Shakuntla Singla et al [6] represent a accurate model for explainng the availability through the use of chapman kolomogorov approach under the reduced capacity. Sunita Kumari et al [7] examining the performance and profit of thresher plant under steady state condition. Sunita Kumari et al [8] discussed the dependability of rubber plant with the help of scrutiny of results with and COGA with the limitation of cost and RAP illustration. In this paper non-poison problem of system were analyzed. Shakuntla Singla et al [9] discussed single unit subdivision with complete failure and in working condition with two types of malfunction one is direct and the second is in partial mode.

Sara Salim Al Oraimi et al [10] discussed in their paper sensitivity of Ammonia/Urea plant by taking parallel system. Thus, the literature has widely discussed for analysis of reliability and modelling of complex industrial systems in various failure/maintenance circumstances. But the concept of reliability analysis for manufacturing plants has yet to be discussed.

The demand for manufacturing items is growing day by day around the world to meet the daily requirements. The most used or consumed items of manufacturing plant is called modernization or fashion in daily life which is manufactured from manufacturing plant, from different industries. To meet the growing demand of manufacturing items in the current market, manufacturer must keep production continuously with maximum capacity to meet the market's growing demand. For continuous operation, these facilities must function the plant and equipment efficiently throughout the year without significant technical or maintenance issues. Any unexpected operational failure, breakdown, or downtime may affect plant productivity and efficiency. For that, the operation and maintenance strategies are critical as they help maintain the life and smooth operation of the equipment. These strategies also help reduce plant downtime. Also, further analysis and research techniques for plant performance, productivity, reliability, availability, maintainability, sensitivity, etc., may be carried out to ensure continuous and smooth plant operations. This paper provides profitability analysis, along with reliability analysis of parallel units of plant worldwide that have operated for more than 15 years. The research is based on the actual plant data of the probability of various failures and repair rates with some assumptions.

II. Assumptions and Notation

These are some assumptions which are as follows:

- Initially, we have an operative manufacturing plant composed of two parallel units: Unit 1 and Unit 2.
- The four types of failures are observed in both units, i.e., developmental, electronical, motorized, and infrastructural.
- Both units work independently.
- By repairing of units after failure as new one.
- Failure rate and repair rates all are taken as general.

The following are the notations used in the analysis:

λ_u : failure rate of manufacturing plant;

$p_1/\ p_2/\ p_3/\ p_4$: probability of developmental/ electronical/motorized/infrastructural in unit 1;

$p_5/\ p_6/\ p_7/\ p_8$: probability of developmental/ electronical/motorized/ infrastructural failure in unit 2;

$\alpha_1/\ \alpha_2/\ \alpha_3/\ \alpha_4$: repair rate of developmental/ electronical/motorized/ infrastructural failure in unit 1;

$\alpha_5/\ \alpha_6/\ \alpha_7/\ \alpha_8$: repair rate of developmental/ electronical/motorized/ infrastructural failure in unit 2;

$f^{1,2}(t)$: p.d.f. of failure time;

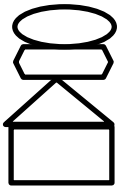
$g_1(t)/\ g_2(t)/\ g_3(t)/\ g_4(t)$: p.d.f. of repair time due to developmental/electronical/ motorized/ infrastructural failure in unit 1;

$g_5(t)/g_6(t)/\ g_7(t)/\ g_8(t)$: p.d.f. of repair time due to developmental/electronical/ motorized/ infrastructural failure in unit 2;

A_0^u : Availability of the System;

$DB^{u1}/\ EB^{u1}/\ MB^{u1}/IB^{u1}$: Busy period of the server due to developmental/ electronical/ motorized/ infrastructural failure in unit 1;

$DB^{u2}/\ EB^{u2}/\ MB^{u2}/IB^{u2}$: Busy period of the server due to developmental/ electronical/ motorized/ infrastructural failure in unit 2;



: Developmental Failure of unit '1' and unit '2';
: Electronical Failure of unit '1' and unit '2';
: Motorized Failure of unit '1' and unit '2';
: Infrastructural Failure of unit '1' and unit '2';

III. Data Summary

The real data from a manufacturing company is summarized as follows:

- Probability of developmental failure in unit 1, $p_1 = 0.3044$;
- Probability of electronical failure in unit 1, $p_2 = 0.02$;
- Probability of motorized failure in unit 1, $p_3=0.21$;
- Probability of infrastructural failure in unit 1, $p_4 =0.0124$;
- Probability of developmental failure in unit 2, $p_5 =0.1648$;
- Probability of electronical failure in unit 2, $p_6 = 0.152$;
- Probability of motorized failure in unit 2, $p_7 = 0.0112$;
- Probability of infrastructural failure in unit 2, $p_8 = 0.1252$;
- Failure rate of manufacturing plant, $\lambda_u = 0.00034$ per hour;
- Repair rate of developmental failure in unit 1, $\alpha_1 = 0.0149$ per hour;
- Repair rate of electronical failure in unit 1, $\alpha_2 = 0.12$ per hour;
- Repair rate of motorized failure in unit 1, $\alpha_3 = 0.0072$ per hour;
- Repair rate of infrastructural failure in unit 1, $\alpha_4 = 0.0253$ per hour;
- Repair rate of developmental failure in unit 2, $\alpha_5 = 0.006$ per hour;
- Repair rate of electronical failure in unit 2, $\alpha_6 = 0.0275$ per hour;
- Repair rate of motorized failure in unit 2, $\alpha_7 = 0.0047$ per hour;
- Repair rate of infrastructural failure in unit 2, $\alpha_8 = 0.0192$ per hour;

IV. Stochastic Model and Model Description

There are different transitions from different stages to different stages which shows in Figure 1. Figure 1 shows the transitions from stage i (S_i) to stage j (S_j). The set of stages {0,1,2,3...,8} all are functioning and regenerative.

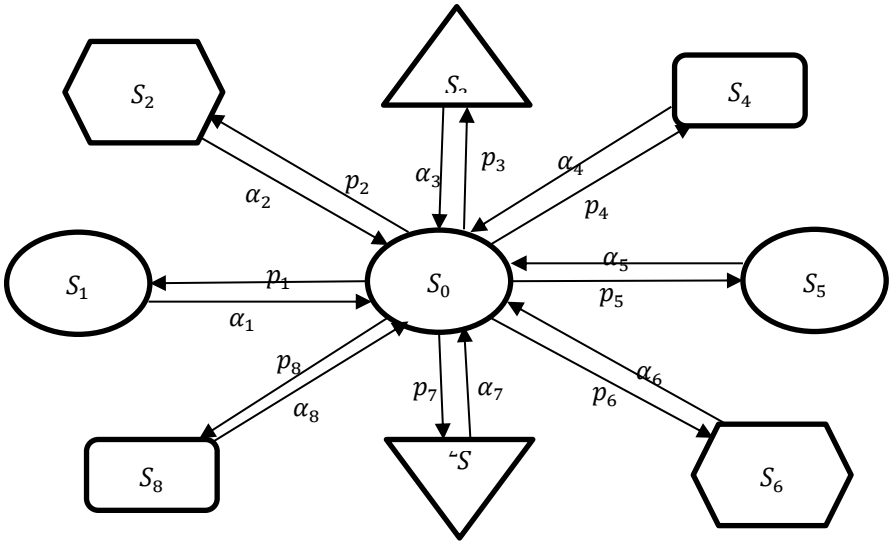


Figure 1: *State Transition Diagram*

where,
Stage 0 (S_0) – Both machines of manufacturing Industry unit 1 and unit 2 are functioning.
Stage 1 (S_1) - Unit 1 failed due to developmental failure, and Unit 2 is still functioning.
Stage 2 (S_2) - Unit 1 failed due to electronical failure, and Unit 2 is still functioning.
Stage 3 (S_3) - Unit 1 failed due to motorized failure, and Unit 2 is still functioning.
Stage 4 (S_4) - Unit 1 failed due to infrastructural failure, and Unit 2 is still functioning.
Stage 5 (S_5)- Unit 2 failed due to developmental failure and Unit 1 is still functioning.
Stage 6 (S_6) - Unit 2 failed due to electronical failure and Unit 1 is still functioning.
Stage 7 (S_7) - Unit 2 failed due to motorized failure and Unit 1 is still functioning.
Stage 8 (S_8) - Unit 2 failed due to infrastructural failure and Unit 1 is still functioning.

I. Transition Probabilities

The transition probabilities Table 1 from stage i (S_i) to stage j (S_j), $q_{ij}(t)$ where (i, j=1 to 8) are given by:

Table 1: *Transition Probabilities*

$q_{0i}(t)$	$qi_0(t)$
$q_{01}(t) = p_1 f^{1,2}(t)$	$q_{10}(t) = g_1(t)$
$q_{02}(t) = p_2 f^{1,2}(t)$	$q_{20}(t) = g_2(t)$
$q_{03}(t) = p_3 f^{1,2}(t)$	$q_{30}(t) = g_3(t)$
$q_{04}(t) = p_4 f^{1,2}(t)$	$q_{40}(t) = g_4(t)$
$q_{05}(t) = p_5 f^{1,2}(t)$	$q_{50}(t) = g_5(t)$
$q_{06}(t) = p_6 f^{1,2}(t)$	$q_{60}(t) = g_6(t)$
$q_{07}(t) = p_7 f^{1,2}(t)$	$q_{70}(t) = g_7(t)$
$q_{08}(t) = p_8 f^{1,2}(t)$	$q_{80}(t) = g_8(t)$

The steady-stage probability is p_{ij} as $p_{0j} = p_j$ and $p_{i0} = 1$ where i and j vary from 1 to 8
and $\sum_{j=1}^{8} p_j$=1

II. Mean Sojourn time

(μ_i), i.e., mean stay time in particular stage i, is given in Table 2 as

Table 2: *Mean sojourn time*

$\mu_i = \int_0^\infty tp.d.fdt$	
$\mu_0 = \int_0^\infty t f^{1,2}(t)dt$	$\mu_5 = \int_0^\infty t g_5(t)dt$
$\mu_1 = \int_0^\infty t g_1(t)dt$	$\mu_6 = \int_0^\infty t g_6(t)dt$
$\mu_2 = \int_0^\infty t g_2(t)dt$	$\mu_7 = \int_0^\infty t g_7(t)dt$
$\mu_3 = \int_0^\infty t g_3(t)dt$	$\mu_8 = \int_0^\infty t g_8(t)dt$
$\mu_4 = \int_0^\infty t g_4(t)dt$	

V. System Performance Measures

I. Availability of the System

Define $A_0^u(t)$ is the possibility of the system at time t, when the structure is in stage i at time t = 0. By using the state transitions, we get the following equation:

$$A_0^{u*}(s) = \frac{N_1^u(s)}{D_1^u(s)} \quad (1)$$

Where

$$N_1^u(s) = M_0^*(s) + q_{01}^*(s)M_1^*(s) + q_{02}^*(s)M_2^*(s) + q_{03}^*(s)M_3^*(s) + q_{04}^*(s)M_4^*(s) + q_{05}^*(s)M_5^*(s) + q_{06}^*(s)M_6^*(s) + q_{07}^*(s)M_7^*(s) + q_{08}^*(s)M_8^*(s)$$

$$D_1^u(s) = 1 - q_{01}^*(s)q_{10}^*(s) - q_{02}^*(s)q_{20}^*(s) - q_{03}^*(s)q_{30}^*(s) - q_{04}^*(s)q_{40}^*(s) - q_{05}^*(s)q_{50}^*(s) - q_{06}^*(s)q_{60}^*(s) - q_{07}^*(s)q_{70}^*(s) - q_{08}^*(s)q_{80}^*(s)$$

where

$M_i(t)$= probability that the system stays in stage i while operating rather than transferring to any other stage.

And (*) denotes the transformation of laplace of the respective function.

For solving $A_0^{u*}(s)$, we get the steady-stage accessibility of the structure is given by

$$A_0^u = \lim_{s\to 0} s.A_0^{u*}(s) = \lim_{s\to 0} s.\frac{N_1^u(s)}{D_1^u(s)} = \frac{N_1^{u'}(0)}{D_1^u(0)} = \frac{N_1^u}{D_1^u}(say) \quad (2)$$

Where

$$N_1^u = \mu_0 + p_1\mu_1 + p_2\mu_2 + p_3\mu_3 + p_4\mu_4 + p_5\mu_5 + p_6\mu_6 + p_7\mu_7 + p_8\mu_8$$

$$D_1^u = p_1\mu_1 + p_2\mu_2 + p_3\mu_3 + p_4\mu_4 + p_5\mu_5 + p_6\mu_6 + p_7\mu_7 + p_8\mu_8 + \mu_0$$

II. Busy Period for Repair

The expected time for which the repair man is in demanding for the reoperative of unit 1 and unit 2 due to developmental failure in steady state is given by:

$$DB^{u1} = DN^{u1}/D_1^u \qquad DB^{\mu 2} = DN^{u2}/D_1^u \quad (3)$$

where

$DN^{u1} = p_1\mu_1$ $\qquad$ $DN^{u2} = p_5\mu_5$

The expected time for which the repair man is in demanding for the reoperative of unit 1 and unit 2 due to electronical failure in steady state is given by:

$$EB^{u1} = EN^{u1}/D_1^\mu \qquad EB^{u2} = EN^{u2}/D_1^\mu \quad (4)$$

Where

$EN^{u1} = p_2\mu_2$ $\qquad$ $EN^{u2} = p_6\mu_6$

The expected time for which the repair man is in demanding for the reoperative of unit 1 and unit 2 due to motorized failure in steady state is given by:

$$MB^{u1} = MN^{u1}/D_1^\mu \qquad MB^{u2} = MN^{u2}/D_1^\mu \quad (5)$$

Where

$MN^{u1} = p_3\mu_3$ $\qquad$ $MN^{u2} = p_7\mu_7$

The expected time for which the repair man is in demanding for the reoperative of unit 1 and unit 2 due to infrastructural failure in steady state is given by:

$$IB^{u1} = IN^{u1}/D_1^\mu \qquad IB^{u2} = IN^{u2}/D_1^\mu \quad (6)$$

Where
$IN^{u1} = p_4\mu_4$ $IN^{u2} = p_8\mu_8$

III. Expected Number of Repairs

The expected number of repairs in unit 1 and unit 2 due to developmental failure in steady stage is as follow:

$$DR^{u1} = DS^{u1}/D_1^u \qquad DR^{\mu 2} = DS^{u2}/D_1^u \tag{7}$$

where
$DS^{u1} = p_1$ $DS^{u2} = p_5$
The expected number of repairs in unit 1 and unit 2 due to electronical failure in steady state is given by:

$$ER^{u1} = ES^{u1}/D_1^u \qquad ER^{\mu 2} = ES^{u2}/D_1^u \tag{8}$$

where
$ES^{u1} = p_2$ $ES^{u2} = p_6$
The expected number of repairs in unit 1 and unit 2 due to motorized failure in steady state is given by:

$$MR^{u1} = MS^{u1}/D_1^u \qquad MR^{\mu 2} = MS^{u2}/D_1^u \tag{9}$$

Where
$MS^{u1} = p_3$ $MS^{u2} = p_7$
The expected number of repairs in unit 1 and unit 2 due to infrastructural failur e in steady state is given by:

$$IR^{u1} = IS^{u1}/D_1^u \qquad IR^{\mu 2} = IS^{u2}/D_1^u \tag{10}$$

Where
$IS^{u1} = p_4$ $IS^{u2} = p_8$

IV. Profit Function of the System

The profit function of the system is as follows:

$$P^u = C_0A_0^u - C_1(DB^{u1} + DR^{u1}) - C_2(DB^{u2} + DR^{u2}) - C_3(EB^{u1} + ER^{u1}) - C_4(EB^{u2} + ER^{u2}) - C_5(MB^{u1} + MR^{u1}) - C_6(MB^{u2} + MR^{u2}) - C_7(IB^{u1} + IR^{u1}) - C_8(IB^{u2} + IR^{u2}) \tag{11}$$

Where
C_0= Revenue generated by the system and
$C_1(C_2)$/ $C_3(C_4)$/ $C_5(C_6)$/ $C_7(C_8)$: - Cost per unit time for engaging the repair man and cost for repair due to developmental/electronical/motorized/infrastructural failure in unit 1 (unit 2).

VI. Result and Discussion

In this section, we numerically solve all the system parameters in equation (1), (2), (3), (4), (5), (6), (7), (8), (9), (10) and (11) by using all the equations of table 1, table 2 and given as below. Let us adopt all the failures and repair times follow the distribution of exponential along with their p.d.f. as:

$$f^{1,2}(t) = \lambda_u e^{-\lambda_u t}$$

$$g_i(t) = \alpha_i e^{-\alpha_i t}, \text{ i = 1 to 8}$$

Using the values as written in equation (1), (2), (3), (4), (5), (6), (7), (8), (9), (10) and (11) that is calculated from real data of a manufacturing company, we get system effectiveness measures in Table 3 as:

Table 3: *System Parameters*

Numerical Analysis of system parameters		
System Parameters	Unit-1	Unit-II
Availability of the manufacturing plant	1	1
Busy period of Repair	Developmental Failure $DB^{u1} = 0.0067$	Developmental Failure $DB^{u2} = 0.009$
	Electronical Failure $EB^{u1} = 5.495E - 5$	Electronical Failure $EB^{u2} = 0.0018$
	Motorized Failure $MB^{u1} = 0.0096$	Motorized Failure $MB^{u2} = 0.0008$
	Infrastructural Failure $IB^{u1} = 0.0002$	Infrastructural Failure $IB^{u2} = 0.0021$
Expected number of Repair	Developmental Failure $DR^{u1} = 0.0001$	Developmental Failure $DR^{u2} = 5.433E - 5$
	Electronical Failure $ER^{u1} = 6.5934 - 6$	Electronical Failure $ER^{u2} = 5.011E - 5$
	Motorized Failure $MR^{u1} = 6.9230 - 5$	Motorized Failure $MR^{u2} = 3.6923 - 6$
	Infrastructural Failure $IR^{u1} = 4.0879E - 6$	Infrastructural Failure $IR^{u2} = 4.1274E - 5$

I. Profit Analysis:

The graph of profit function (P^u) w.r.t. returns (C_0) for different kind of values of regenerative rate (α_i) has been shown in Figure 1.

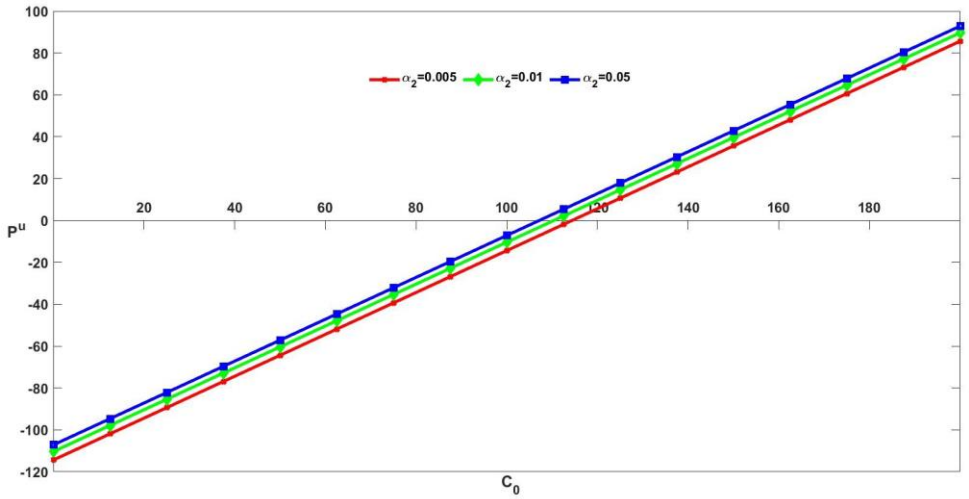


Figure 2: *Change in Profit w.r.t. Returns and Repair Rate*

Figure 2 shows that the increase in returns and repair rate increases profit Also, the cut-off points for the system to be profitable can be observed in Fig. 2.:

- For $C_0 > 108.2567$ and $\alpha_2 = 0.05$, $P^u > 0$.
- For $C_0 > 112.6834$ and $\alpha_2 = 0.01$, $P^u > 0$.
- For $C_0 > 115.8273$ and $\alpha_2 = 0.005$, $P^u > 0$

Similarly, we can draw graphs of the profit function with other parameters to see its effect and cut-off points when the system is profitable.

VII. Conclusion

In this paper, reliability of the parallel system of two units manufacturing plant has been examined and all these parameters such as accessibility of the plant, the busy period for repairs and the anticipated number of repairs for each type of failure are calculated with the help of regenerative technique of graphical point. Also, Profit analysis for the plant is also carried out along with the graphical representation with respect to various parameters. Profit increases when returns and repair rates both increases. The cut-off point is also drawn to determine when a system is profitable. It demonstrates that, in comparison to other factors, returns and system failure rate have the most significant impact on the profit function. The model forecasts the failure and repair conditions based on the optimized reliability and profitability results.

References

[1] Al Rahbi, Y., Rizwan, S., Alkali, B., Cowell, A., & Taneja, G. (2019). Reliability analysis of multiple units with multiple repairmen of rodding anode plant in aluminum industry. *Advances and Applications in Statistics*, *54*(1), 151-178.

[2] Taj, S. Z., Rizwan, S., Alkali, B., Harrison, D., & Taneja, G. (2018). Reliability analysis of a 3-unit subsystem of a cable plant. *Advances and Applications in Statistics*, *52*(6), 413-429.

[3] Rizwan, S. M., Sachdeva, K., Al Balushi, N., Al Rashdi, S., & Taj, S. Z. (2023). Reliability and sensitivity analysis of membrane biofilm fuel cell. *International Journal of Engineering Trends and Technology*, *71*(3), 73-80.

[4] Sonia, Singla, S. (2024). EXPLORE THE DYNAMICS OF MANUFACTURING INDUSTRIES: RELIABILITY ANALYSIS THROUGH STOCHASTIC PROCESS MODELING. *Reliability: Theory & Applications*, *19*(2 (78)), 467-471.

[5] Singla, S., Sonia & Panwar, P. (2024). STOCHASTIC OPTIMIZATION AND RELIABILITY ANALYSIS OF MUSHROOM PLANT. *Reliability: Theory & Applications*, *19*(1 (77)), 729-743.

[6] Singla, S., Modibbo, U. M., Mijinyawa, M., Malik, S., Verma, S., & Khurana, P. (2022). Mathematical model for analysing availability of threshing combine machine under reduced capacity. *Yugoslav Journal of Operations Research*, *32*(4), 425-437.

[7] Kumari, S., Khurana, P., & Singla, S. (2022). Behavior and profit analysis of a thresher plant under steady state. *International Journal of System Assurance Engineering and Management*, 1-6.

[8] Kumari, S., Khurana, P., & Singla, S. (2021). RAP via constraint optimization genetic algorithm. *Life Cycle Reliability and Safety Engineering*, *10*(4), 341-345.

[9] Singla, S., & Dhawan, P. (2022). Mathematical analysis of regenerative point graphical technique (RPGT). *Mathematical Analysis and its Contemporary Applications*, *4*(4), 49-56.

[10] Al Oraimi, S. S., Rizwan, S. M., & Sachdeva, K. (2024). SENSITIVITY AND PROFITABILITY ANALYSIS OF TWO-UNITS AMMONIA/UREA PLANT. *Reliability: Theory & Applications*, *19*(1 (77)), 376-386.

PERFORMANCE ANALYSIS OF SOFTWARE SYSTEM SUBJECT TO PREVENTIVE MAINTENANCE AND SOFTWARE UP-GRADATION FACILITY

Kavita Thukral[1] and Shiv Kumar Sharma[2]

•

[1, 2] Department of Mathematics, Chandigarh University, Mohali, Punjab, India
ktkavita540@gmail.com, shivksharma2005@gmail.com

Abstract

Every industrialist as well as scientist try to satisfy the customers requirement by providing reliable software systems and a software system is reliable and profitable if it is easily manageable. This study analyzes the criteria under which software performs correctly with software upgrade and preventive maintenance facility. It is observed that researcher threw light on negligible, major, and load faults cause software to fail, after which load recovery, major, and negligible upgrades are carried out, but there are no discussion on preventive maintenance concept. Sometimes, to enhance the software reliability and performance, preventive maintenance is utilized and it is better to upgrade the software before completely failure. The weibull distributions are followed by all of these failures and upgrade rates. The malfunctioning program can always be fixed by a skilled technician. The software performance is analyzed by applying the regenerative point technique and the semi markov process. Software reliability metrics including mean time to software system failure, software availability, and software profit values are calculated using tables.

Keywords: Software reliability, negligible fault, major fault, load failure, preventive maintenance, regenerative point technique.

I. Introduction

Software is a key component of the majority of software systems. In the rapidly expanding era of digitization, most people now rely significantly on software tools to complete tasks accurately and on schedule. Therefore, in order to satisfy the clients, the software system must perform well. However, it is truth that no system can function indefinitely since software systems and machinery malfunction due to little fault, major fault, and high loads. Software upgrades and preventive maintenance can be employed in certain circumstances to enhance the availability and profit values of software systems.

Over the years, researchers have suggested some reliability improvement techniques using stochastic process with semi markov approach. Smith [20] described the performance parameters of software system under different conditions. Coleman et al. [4] examined the application of software maintainability models in industrial software systems. Huang [5] evaluated the performance of software reliability growth models with testing-effort and change-point. Misra [16] threw light on modeling design and coding factors that analyzed the maintainability of software systems. Kumar and Malik [7] discussed reliability measures of computer systems with priority in repair activities subject to maximum operation and repair times. Malhotra and Chug [14] evaluated the software maintainability with systematic literature review and current trends. The profit analysis of a computer system has been discussed by Kumar and Saini [8] regarding preventive maintenance. Malik et al. [20] examined the reliability of a non series parallel system with weibull failure laws. Barak et al. [1] examined the reliability, availability and profit values of

milk industry subject to repair facility. Kavita and Sharma [6] analyzed the estimation of software reliability and performance using lindley distribution. Barak et al. [2] described the effects of refreshment for the technician in a stochastic cold standby system. Kumar and Sharma [9] descried the availability and profit values of stochastic repairable cold standby system under refreshments. Kumar et al. [11] described the reliability measures of hardware and software system using weibull distribution under load failure facility. Kumar et al. [13] analyzed the mean time to system failure, availability of the system, busy period of the system, estimated number of visits made by the technician and profit function under inspection and refreshment facility. Kumar et al. [12] analyzed the reliability measures of a computer system under hardware repair and software upgrade facility. Kumar and Sharma [10] analyzed the reliability measures of repairable cold standby system subject to reboot facility under refreshments. Behera and Agarwal [3] explored the software reliability prediction and management incorporating change points based on testing effort. Rahul et al. [17] analyzed the reliability characteristics of juice industry using stochastic process and regenerative point graphical technique. Reddy et al. [18] examined the reliability characteristics of utensil industry subject to repair facility. Sharma and Rana [19] described the advancing software reliability with time series approach under identification of defects with artificial neural network.

It has been determined that software upgradation improve the machines as well as software systems availability and profitability. Regenerative point technique with weibull distribution has been used to analyze the reliability characteristics such as mean sojourn times, mean time to software system failure, availability, technician busy period due to software upgrades, expected number of technician visits, and profit function. Tables explore these reliability metrics.

II. System Assumptions

There are following software system assumptions

- Initially, it is assumed that the software work properly.
- The software system is timely up-graded to avoid the negligible fault/major fault/ load failure of software.
- Preventive maintenance is performed to enhance the reliability of the software.
- An expert technician is available to up-grade the software.
- After up-grade, the failed software work like a new software.
- Parameters used to describe the software system are all statistically independent.
- Software failure time and up-gradation time follow the weibull distribution.
- Preventive maintenance request time and preventive maintenance time follow weibull distribution.

III. System Notations

To explain the software system, there are following notations

R	Set of regenerative state (S_0, S_1, S_2, S_3).
S_{Op}	Software in operative mode.
S_{Nf}	Software with negligible fault/failure.
S_{Mf}	Software with major fault/failure.
S_{Lf}	Software with load failure.
$q_{ij}(t)$	PDF of the first passage time from a regenerative state i to a regenerative state j or to a failed state j without visiting any other regenerative state in (0,1].
μ_i	Mean sojourn time spent in the state 'i' before visiting any other states.

μ_i'	Total unconditional time spent before transiting to any other regenerative state while the system entered regenerative state '*i*' at t=0.
$M_i(t)$	Probability that the system is originally up in the regenerative state $S_i \in R$ up to at the time (*t*) without passing through any other $S_i \in R$.
$W_i(t)$	The probability that the technician is busy in the state S_i up to time (t) without making any transition to any other $S_i \in R$ or returning to the same via one or more non- regenerative states.
$f_1(t) = \alpha\eta t^{\eta-1} e^{-\alpha t^{\eta}}$	PDF of the negligible failure time.
$f_2(t) = \beta\eta t^{\eta-1} e^{-\beta t^{\eta}}$	PDF of the major failure time.
$f_3(t) = \gamma\eta t^{\eta-1} e^{-\gamma t^{\eta}}$	PDF of the load failure time.
$f_4(t) = \psi\eta t^{\eta-1} e^{-\psi t^{\eta}}$	PDF of the negligible failure time to major failure time.
$f_4(t) = \xi\eta t^{\eta-1} e^{-\xi t^{\eta}}$	PDF of the preventive maintenance request time.
$g_1(t) = k\eta t^{\eta-1} e^{-kt^{\eta}}$	PDF of the negligible up-gradation time.
$g_2(t) = l\eta t^{\eta-1} e^{-lt^{\eta}}$	PDF of the major repair time.
$g_3(t) = m\eta t^{\eta-1} e^{-mt^{\eta}}$	PDF of the load repair time.
$g_5(t) = h\eta t^{\eta-1} e^{-ht^{\eta}}$	PDF of the preventive maintenance time.
○ / ⬠ / ▭ / ●	Upstate/ negligible failed state/ failed state/ regenerative state.

IV. Circuits Descriptions

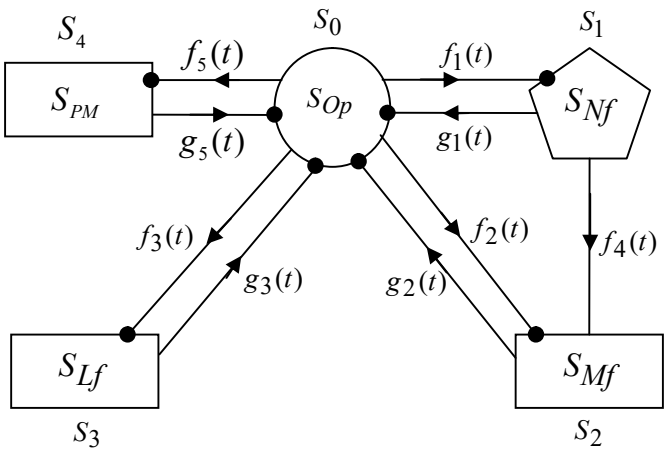


Figure 1 *State Transition Diagram*

V. Transition Probabilities

There are following transition probabilities

$p_{0,1} = \alpha/(\alpha+\beta+\gamma+\xi)$, $p_{0,2} = \beta/(\alpha+\beta+\gamma+\xi)$

$p_{0,3} = \gamma/(\alpha+\beta+\gamma+\xi)$, $p_{0,4} = \xi/(\alpha+\beta+\gamma+\xi)$,

$p_{1,0} = k/(k+\psi)$, $p_{1,0;2} = \psi/(k+\psi)$, $p_{1,2} = \psi/(k+\psi)$, $p_{2,0} = p_{3,0} = 1$ (1)

It is also simplified that

$p_{0,1} + p_{0,2} + p_{0,3} + p_{0,4} = p_{1,0} + p_{1,2} = 1$, $p_{1,0} + p_{1,0;2} = 1$ (2)

VI. Mean Sojourn Time

For the particular state, it becomes

$$\mu_0 = \frac{\Gamma(1+\frac{1}{\eta})}{(\alpha+\beta+\gamma+\xi)^{1/\eta}},\ \mu_1 = \frac{\Gamma(1+\frac{1}{\eta})}{(k+\psi)^{1/\eta}},\ \mu_2 = \frac{\Gamma(1+\frac{1}{\eta})}{(l)^{1/\eta}},\ \mu_3 = \frac{\Gamma(1+\frac{1}{\eta})}{(m)^{1/\eta}}$$

$$\mu_1' = \Gamma(1+\frac{1}{\eta})\left[\frac{1}{(k+\psi)^{1/\eta}} + \frac{\psi}{l(k+\psi)^{1/\eta}}\right] \tag{3}$$

VII. Evaluation of Parameters

I. Mean Time to System Failure

Let $\varphi_i(t)$ is the cumulative distribution function of the first passage time from $S_i \in R$ to a failed state and treating the failed states as an absorbing states then upcoming recursive relation for $\varphi_i(t)$ becomes

$$\varphi_0(t) = Q_{0,1}(t) \otimes \varphi_1(t) + Q_{0,2}(t) + Q_{0,3}(t) + Q_{0,4}(t)$$
$$\varphi_1(t) = Q_{1,0}(t) \otimes \varphi_0(t) + Q_{1,2}(t)$$
$$\varphi_2(t) = Q_{2,0}(t)$$
$$\varphi_3(t) = Q_{3,0}(t)$$
$$\varphi_4(t) = Q_{4,0}(t) \tag{4}$$

Now taking LST on above equation (4) and solving for $\varphi_0^{**}(s)$, obtained

$$R^*(s) = \frac{1-\varphi_0^{**}(s)}{s} \tag{5}$$

The system reliability is obtained by using inverse LT on the above equation (5), get

$$MTSF = \lim_{s\to 0} \frac{1-\varphi_0^{**}(s)}{s} = \frac{\mu_0 + p_{0,1}\mu_1}{1-p_{0,1}p_{1,0}}$$

$$MTSF = \frac{\Gamma(1+\frac{1}{\eta})\left[\frac{\alpha}{(\alpha+\beta+\gamma+\xi)(k+\psi)^{1/\eta}} + \frac{1}{(\alpha+\beta+\gamma+\xi)^{1/\eta}}\right]}{\left[1-\frac{\alpha k}{(\alpha+\beta+\gamma+\xi)(k+\psi)}\right]} \tag{6}$$

II. Software Availability

Let $A_i(t)$ is the probability that software system is in up-state at the time 't' specified that the system arrives at the regenerative-state S_i at t = 0 and then upcoming recursive relation for $A_i(t)$ becomes

$$A_0(t) = M_0(t) + q_{0,1}(t) \oplus A_1(t) + q_{0,2}(t) \oplus A_2(t) + q_{0,3}(t) \oplus A_3(t) + q_{0,4}(t) \oplus A_4(t)$$
$$A_1(t) = M_1(t) + [q_{1,0}(t) + q_{1,0;2}(t)] \oplus A_0(t)$$
$$A_2(t) = q_{2,0}(t) \oplus A_0(t)$$
$$A_3(t) = q_{3,0}(t) \oplus A_0(t)$$
$$A_4(t) = q_{4,0}(t) \oplus A_0(t) \tag{7}$$

where, $M_0(t) = e^{-(\alpha+\beta+\gamma+\xi)t^n}$, $M_1(t) = e^{-(k+\psi)t^n}$

Using LT on equation (7) and solving for $A_0^*(s)$, then steady-state availability is given by

$$A_0 = \lim_{s\to 0} sA_0^*(s) = \frac{N_A}{D'} = \frac{\mu_0 + p_{0,1}\mu_1}{\mu_0 + p_{0,1}\mu_1' + p_{0,2}\mu_2 + p_{0,3}\mu_3 + p_{0,4}\mu_4} \tag{8}$$

where, $$N_A = \Gamma(1+\frac{1}{\eta})\left[\frac{1}{(\alpha+\beta+\gamma+\xi)^{1/\eta}} + \frac{\psi}{(\alpha+\beta+\gamma+\xi)(k+\psi)^{1/\eta}}\right] \tag{9}$$

$$D' = \Gamma(1+\frac{1}{\eta})\left[\begin{array}{l}\left\{\begin{array}{l}\frac{1}{(\alpha+\beta+\gamma+\xi)^{1/\eta}} + \frac{\beta}{(\alpha+\beta+\gamma+\xi)(l)^{1/\eta}} \\ + \frac{\gamma}{(\alpha+\beta+\gamma+\xi)(m)^{1/\eta}} + \frac{\xi}{(\alpha+\beta+\gamma+\xi)(h)^{1/\eta}}\end{array}\right\} + \\ \frac{\alpha}{(\alpha+\beta+\gamma+\xi)}\left\{\frac{1}{(k+\xi)^{1/\eta}} + \frac{\psi}{l(k+\psi)^{1/\eta}}\right\}\end{array}\right] \tag{10}$$

III. Busy Period of the Technician

Let $B_i(t)$ is the probability that the technician is busy at time '*t*' specified that the system arrives at the S_i at *t* = 0. Then the upcoming recursive relation for $B_i(t)$ obtained as

$$\begin{aligned}
&B_0(t) = q_{0,1}(t) \oplus B_1(t) + q_{0,2}(t) \oplus B_2(t) + q_{0,3}(t) \oplus B_3(t) + q_{0,4}(t) \oplus B_4(t) \\
&B_1(t) = W_1(t) + [q_{1,0}(t) + q_{1,0;2}(t)] \oplus B_0(t) \\
&B_2(t) = W_2(t) + q_{2,0}(t) \oplus B_0(t) \\
&B_3(t) = W_3(t) + q_{3,0}(t) \oplus B_0(t) \\
&B_4(t) = W_4(t) + q_{4,0}(t) \oplus B_0(t)
\end{aligned} \tag{11}$$

where, $W_1(t) = \overline{F}_4(t) \oplus \overline{G}_1(t)$, $W_2(t) = \overline{G}_2(t)$, $W_3(t) = \overline{G}_3(t)$ and $W_4(t) = \overline{G}_5(t)$

Using LT on equation (11) and solving for $B_0^*(s)$ then

$$B_0 = \lim_{s\to 0} sB_0^*(s) = \frac{N_B}{D'} = \frac{p_{0,1}\mu_1' + p_{0,2}\mu_2 + p_{0,3}\mu_3}{\mu_0 + p_{0,1}\mu_1' + p_{0,2}\mu_2 + p_{0,3}\mu_3 + p_{0,4}\mu_4} \tag{12}$$

where, $$N_B = \Gamma(1+\frac{1}{\eta})\left[\begin{array}{l}\left\{\frac{\beta}{(\alpha+\beta+\gamma+\xi)(l)^{1/\eta}} + \frac{\gamma}{(\alpha+\beta+\gamma+\xi)(m)^{1/\eta}}\right\} + \\ \frac{\alpha}{(\alpha+\beta+\gamma+\xi)}\left\{\frac{1}{(\alpha+\beta+\gamma+\xi)^{1/\eta}} + \frac{\psi}{(\alpha+\beta+\gamma+\xi)(k+\psi)^{1/\eta}}\right\}\end{array}\right] \tag{13}$$

and D' is defined earlier in equation (10).

IV. Estimated number of visits made by the Technician

Let $N_i(t)$ is the estimated number of visits by the technician for software up-gradation in (0, t] specified that technician arrives $S_i \in R$ at *t*=0. The upcoming recursive relation for $N_i(t)$ becomes

$$\begin{aligned}
&N_0(t) = Q_{0,1}(t) \otimes [1 + N_1(t)] + Q_{0,2}(t) \otimes [1 + N_2(t)] + Q_{0,3}(t) \otimes [1 + N_3(t)] + Q_{0,4}(t) \otimes [1 + N_4(t)] \\
&N_1(t) = [Q_{1,0}(t) + Q_{1,0;2}(t)] \otimes N_0(t) \\
&N_2(t) = Q_{2,0}(t) \otimes N_0(t) \\
&N_3(t) = Q_{3,0}(t) \otimes N_0(t) \\
&N_3(t) = Q_{4,0}(t) \otimes N_0(t)
\end{aligned} \tag{14}$$

Using LST on equation (14) and solving for $N_0^{**}(s)$. Then

$N_0 = \lim_{s\to 0} sN_0^{**}(s) = \frac{N_r}{D'}$, where $N_r = 1$

and D' is defined earlier in equation (10).

V. Profit Analysis

The profit analysis of the system can be evaluated by using the profit function

$$P = E_0 A_0 - E_1 B_0 - E_2 N_0 \quad (15)$$

Where, $E_0 = 5000$ (Revenue per unit uptime of the system is available)

$E_1 = 300$ (Charge per unit time for which technician is busy in software up-grade)

$E_2 = 100$ (Charge per unit visit made by the technician for the defined purpose)

VIII. Discussion

Here, preventive maintenance rate play an important role in software reliability and performance evaluation. To calculate the reliability measures such that mean time to software system failure, software availability, and profit values are calculated by using particular cases of failure time and up-gradation time such as $\alpha = \beta = \gamma = \psi = \xi = \alpha_1$ and $k = l = m = h = k_1$ respectively. Here, we have used the shape parameter η=0.5 for the reliability measures.

From table-1 it is clarified that mean time to software system failure values decreases with the increment of software failure rate (α_1) but it is enhance with increment in software up-gradation rate (k_1) from 2 to 2.5, 3, 3.5, & 4 respectively.

Table 1: *MTSF vs. Software Failure Rate* (α_1)

(α_1) ↓	k_1=2, η=0.5	k_1=2.5, η=0.5	k_1=3, η=0.5	k_1=3.5, η=0.5	k_1=4, η=0.5
0.03	51.72014	62.18437	69.09374	79.42516	86.55808
0.035	48.62661	57.35737	63.73042	73.25431	79.84645
0.04	45.87798	53.21386	59.12651	67.95793	74.08405
0.045	43.41996	49.61988	55.13325	63.36451	69.08506
0.05	41.20925	46.47405	51.63784	59.34428	64.70879
0.055	39.21056	43.69831	48.55367	55.79732	60.84685
0.06	37.39508	41.23157	45.81285	52.64549	57.41445
0.065	35.73288	39.02541	43.36157	49.82682	54.34432
0.07	34.22656	37.04096	41.15662	47.29159	51.58246
0.075	32.82775	35.24665	39.16294	44.99854	49.08502

Table 2: *Availability vs. Software Failure Rate* (α_1)

(α_1) ↓	k_1=2, η=0.5	k_1=2.5, η=0.5	k_1=3, η=0.5	k_1=3.5, η=0.5	k_1=4, η=0.5
0.03	0.942288	0.958545	0.966458	0.970756	0.980568
0.035	0.939973	0.956109	0.964075	0.968365	0.978178
0.04	0.937661	0.953672	0.961704	0.965976	0.976241
0.045	0.935358	0.951233	0.959333	0.963594	0.974239
0.05	0.933036	0.948788	0.956965	0.961216	0.972568
0.055	0.930717	0.946336	0.954582	0.958827	0.978622
0.06	0.928392	0.943874	0.952197	0.956437	0.967798
0.065	0.926059	0.941402	0.949804	0.954046	0.965492
0.07	0.923716	0.938917	0.947401	0.951634	0.963101
0.075	0.921362	0.936419	0.944987	0.949217	0.960628

Table 4: *Profit vs. Software Failure Rate* (α_1)

(α_1) ↓	k_1=2, η=0.5	k_1=2.5, η=0.5	k_1=3, η=0.5	k_1=3.5, η=0.5	k_1=4, η=0.5
0.03	4710.33	4791.29	4831.37	4852.87	4898.38
0.035	4698.61	4778.95	4819.76	4840.85	4889.88
0.04	4686.89	4766.68	4807.39	4828.74	4880.07
0.045	4675.16	4754.23	4795.46	4816.78	4869.92
0.05	4663.43	4741.82	4783.38	4804.64	4859.42
0.055	4651.65	4729.37	4771.34	4792.56	4848.54
0.06	4639.83	4716.86	4759.26	4780.44	4837.25
0.065	4627.96	4704.28	4747.18	4768.28	4825.56
0.07	4616.58	4691.64	4734.93	4756.77	4816.28
0.075	4604.56	4678.92	4722.67	4743.86	4802.83

Table-2 threw light on the availability of the software and it is decreases with increment of software failure rate (α_1). Also, it is analyzed that availability of software enhance with increment in software up-gradation rate (k_1) from 2 to 2.5, 3, 3.5, & 4 respectively.

Profit is the essential parameter of reliability that inform customers or industrialist whether the software system is in their favour or not. Software are used in the computing machine and well as to handle the robotic systems in particular manners. Sometimes, it is observed that preventive maintenance enhance these reliability measures.

Profit values are calculated from table-3, and it is observed that it decreases with increment of software failure rate (α_1) but it is enhances with increment in software up-gradation rate(k_1) from 2 to 2.5, 3, 3.5, & 4 respectively.

IX. Conclusion

Mean time to software system failure, availability, and profit function are essential parameters of reliability measure that inform customers or industrialist whether the software system is in their favour or not. Software are used in the computing machine and well as to handle the robotic systems in particular manners. Sometimes, it is observed that preventive maintenance enhance these reliability measures. The performance of the software is discussed using the regenerative point technique. The above tables explore that when the upgradation rate increases then the MTSF, availability and profit values also increase but when the failure rate increases then the MTSF, availability and profit values decrease.

X. Future Scope

It is observed that the role of the regenerative point technique with preventive maintenance for the software system will be beneficial and also used by the management, manufacturers and the persons engaged in reliability engineering and working on analyzing the nature and performance of the computing system like robot and automatic electric machine.

References

[1] Barak, M. S., Garg, R., and Kumar, A. (2021). Reliability measures analysis of a milk plant using RPGT. *Life Cycle Reliability and Safety Engineering*, *10*(3), 295-302.

[2] Barak, M. S., Kumar, A., and Garg, R. (2023). Contribution of Refreshment Provided to the Server During His Job in the Repairable Cold Standby System. *Computational Intelligence in Sustainable Reliability Engineering*, 251-267.

[3] Behera, A. K., and Agarwal, P. (2024). A software reliability prediction and management incorporating change points based on testing effort. *Reliability: Theory and Applications*, *19*(2(78)), 91-100.

[4] Coleman, D., Lowther, B., and Oman, P. (1995). The application of software maintainability models in industrial software systems. *Journal of Systems and Software*, *29*(1), 3-16.

[5] Huang, C. Y. (2005). Performance analysis of software reliability growth models with testing-effort and change-point. *Journal of Systems and Software*, *76*(2), 181-194.

[6] Kavita, and Sharma, S. K. (2022, August). Estimation of software reliability using Lindley distribution based on MLE and UMVUE. In *International workshop of Mathematical Modelling, Applied Analysis and Computation* (pp. 299-316). Cham: Springer Nature Switzerland.

[7] Kumar, A., and Malik, S. C. (2011). Profit analysis of a computer system with priority to so- ftware replacement over hardware repair subject to maximum operation and repair time. *International Journal of Engineering Science and Technology (IJEST)*, *3*(10), 7452-7468.

[8] Kumar, A., and Saini, M. (2018). Profit analysis of a computer system with preventive mai- ntenance and priority subject to maximum operation and repair times. Iran *Journal of Computer Science*, *1*(3), 147-153.

[9] Kumar, A., and Sharma, A. (2023). Profit analysis of repairable cold standby system under refreshment. *Reliability Theory and Applications*, *18*(4), 604-612.

[10] Kumar, A., and Sharma, A. (2024). Profit analysis of repairable cold standby system subject to reboot facility under. *Reliability: Theory and Applications*, *19*(2 (78)), 594-603.

[11] Kumar, A., Garg, R., and Barak, M. S. (2023). Performance analysis of computer systems with Weibull distribution subject to software upgrade and load recovery. *Life Cycle Reliability and Safety Engineering*, *12*(1), 51-63.

[12] Kumar, A., Garg, R., and Barak, M. S. (2023). Performance of computer system with three types of failure using weibull distribution subject to hardware repair and software up gradation. *International Journal of System Assurance Engineering and Management*, *14*(1), 483-491.

[13] Kumar, A., Garg, R., and Barak, M. S. (2023). Reliability measures of a cold standby system subject to refreshment. *International Journal of System Assurance Engineering and Management*, *14*(1), 147-155.

[14] Malhotra, R., and Chug, A. (2016). Software maintainability: Systematic literature review and current trends. *International Journal of Software Engineering and Knowledge Engineering*, *26*(08), 1221-1253.

[15] Malik, S. C., Chauhan, S. K., and Ahlawat, N. (2020). Reliability analysis of a non series parallel system of seven components with Weibull failure laws. *International Journal of System Assurance Engineering and Management*, *11*(3), 577-582.

[16] Misra, S. C. (2005). Modeling design/coding factors that drive maintainability of software systems. *Software Quality Journal*, *13*, 297-320.

[17] Rahul, Yadav, M., and Kumar, H. (2024). Profit analysis of repairable juice plant. *Reliability: Theory and Applications*, *19*(1 (77)), 688-695.

[18] Reddy, H. I., Yadav, M., and Kumar, H. (2024). Stochastic analysis of the utensil industry subject to repair facility. *Reliability: Theory and Applications*, *19*(2 (78)), 170-177.

[19] Sharma, S. K., and Rana, R. K. (2024). Advancing software reliability with time series insights: A non-autoregressive ANN approach. *Quality and Reliability Engineering International*, *40*(8), 4166-4186.

[20] Smith, J. (1991). Performance engineering of software systems. *Journal of the Operation Research Society*, *42*(10), 903-904.

PERFORMANCE MEASURES AND PARAMETRIC ESTIMATION OF A PARALLEL-SERIES SYSTEM MODEL WITH WEIBULL FAILURE AND REPPAIR TIME DISTRIBUTIONS

Shallu Sharma[1], Pawan Kumar*[2], Poonam Sharma[2]

•

[1,*2,2]Department of Statistics, University of Jammu, J&K.
shallusharma28udh@gmail.com[1], pkk_skumar@yahoo.co.in*[2], sharmapoonam1038@gmail.com[2]

Abstract

The present paper aims at the study of three non-identical unit parallel-series system model with Weibull failure and repair time distributions. The main focus is on the analysis of reliability characteristics and also estimation of parameters in Classical and Bayesian paradigms. The units in the system are named as A, B and C. In order that system works successfully, unit A and anyone of the units B or C should work.Upon failure a unit will immediately be taken up to repair facility for necessary remedial actions. A single repair facility is available to deal with any kind of failure/fault detected with any unit. First come first served service discipline is followed. The failure and repair time distributions for all the units are taken Weibull with varying parameters. Regenerative point technique is used to study various measure of effectiveness. MLE and Bayes estimates of various failure and repair parameters involved in the study have been obtained. A simulation study has also been undertaken to exhibit the behaviour of obtained characteristics in Classical and Bayesian setup and a comparison is made thereupon. Various conclusions have been drawn from the tables and graphs plotted for numerous performance measures for varying values of repair and failure parameters. Having obtained all the reliability characteristics, the profit incurred by the system has been obtained and studied graphically too.

Keywords: Reliability, Availability, Mean Time to System Failure, Regenerative Point Technique, Bayesian Estimation, Maximum Likelihood Estimation.

1. Introduction

In today's context, the configuration and design of industrial systems—such as transportation system, aircraft, telecommunication, automobiles, textile manufacturing systems, computer systems, and satellite systems—are becoming increasingly complex. The reliability, availability, and efficiency of any system largely depends on its design. As a result, system designers are constantly striving to create high-reliability systems through various techniques. Reliability denotes the capacity of a system, process, product, or service to consistently fulfill its intended purpose without failure over time or under particular conditions. It embodies dependability and trust, guaranteeing that results are predictable and align with expectations. In fields like engineering, manufacturing, and service industries, reliability is crucial for ensuring safety, customer satisfaction, and overall effectiveness. In the recent past, extensive research has been carried out by several authors on reliability characteristics and optimization of cost-benefit analysis for varieties of stochastic models under different sets of assumptions and different failure and repair time distributions. Sharma and Kumar[1] analyzed reliability measures of two identical unit system with one switching device and imperfect coverage. Sharma and Gupta[9] and Kumar

and Saini[6] respectively undertaken analytical study on cost-benefit analysis of three-unit system model with correlated repair and failure time distributions and a single unit system model having Weibull repair and failure laws along with preventive maintenance and obtained various performance measures. Malik[7] also worked on cost benefit evaluation of a repairable system with alternative repair and Weibull distribution. Guilani et al.[3] discussed a simulation-based optimization approach for the redundancy allocation in a system with increasing failure rates of components based on Weibull distribution. Classical and Bayesian analysis of reliability characteristics was done by Kishan and Jain[4] for two-unit parallel system with Weibull failure and repair laws for each unit with common shape parameters but different scale parameters and derived important measures of system effectiveness using Regenerative point techniques. Devi et al.[2] and Saini at el. [8] demonstrated the performance of a non-identical unit system with priority and Weibull repair and failure laws by obtaining multiple measures of system effectiveness using regenerative point technique and semi-markov process. They also worked on stochastic modeling and profit evaluation of a redundant system with priority subject to Weibull densities for failure and repair. Casteren[10] made use of Weibull Markov model for the assessment of interruption costs in electric power systems. Kumar at el.[5] analysed the performance of computer systems with Weibull distribution subject to software upgrade and load recovery. Based on the literature reviewed above, it is evident that the research to a great extent has been conducted so far and still in progress on cost benefit analysis for stochastic system models under different configuration and different sets of assumptions. In view of above, an attempt has been made to analyze the performance of three unit redundant system model with Weibull repair and failure time distributions. Keeping in view the above ideas, this paper deals with the estimation of parameters and deriving important performance measures of three non-identical unit system model. Initially, all units are in operative state. Single repairman is always available to repair the failed units with FCFS service discipline. Repair and failure times for all the units are assumed to follow Weibull distributions. Regenerative point technique is used to analyze system model and derive various performance measures such as mean time to system failure, reliability, availability, and expected number of repairs. Furthermore, some characteristics of interest of the considered system model are examined and compared in both the Bayesian and Classical sets up using a simulation study and significant inferences are presented in the form of tables and graphs.

2. System Description and Assumptions

• The system consists of three non-identical units A,B and C. For successful operation of the system unit A and atleast one of the units B and C should function.
• The system failure occurs when either unit A or both the units B and C fail.
• The failure and repair times of units are assumed to have Weibull distributions with different scale parameters but common shape parameter.
• Only a single repair facility is available to repair failed units.
• Service discipline is FCFS.
• A repaired units work as good as new.

3. Notations and Symbols

$f_1(.)/f_2(.)/f_3(.)$: Failure rate function for Unit A/B/C.
$h_1(.)/h_2(.)/h_3(.)$: Repair rate function for unit A/B/C
$\alpha_i{}'s$& $\beta_i{}'s, 1 \leq i \leq 3$: Scale parameter for failure and repair time distributions for Units A/B/C respectively.
γ: Shape parameter for failure and repair time distributions of units
$A_o/A_g/A_r/A_R/A_w$:Unit A is operative/ good/ under repair/ repair continued from earlier

state and waiting for repair respectively. Similar notations hold for units B and C.

3.1. SYMBOLS FOR THE STATES OF THE SYSTEM

UpStates	*DownStates*
$S_0 = [A_o, B_o, C_o]$	$S_3 = [A_r, B_g, C_g]$
$S_1 = [A_o, B_r, C_o]$	$S_4 = [A_g, B_r, C_w]$
$S_2 = [A_o, B_o, C_r]$	$S_5 = [A_g, B_w, C_r]$
$S_6 = [A_w, B_g, C_r]$	$S_7 = [A_w, B_r, C_g]$

The transition diagram of the model is shown in Figure 1.

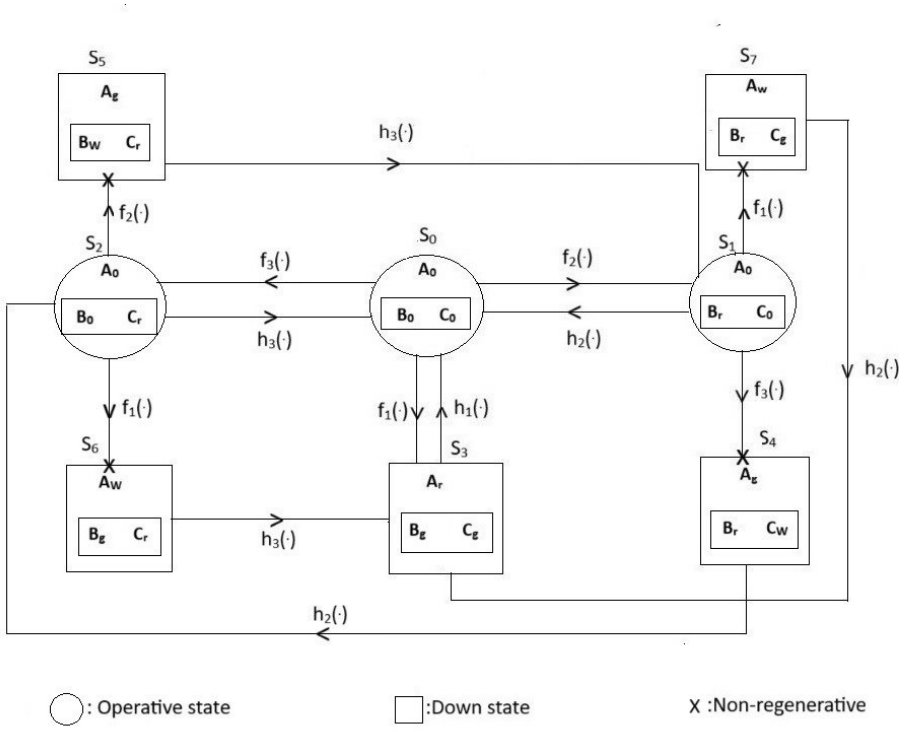


Figure 1: *Transition Diagram*

4. Transition Probabilities and Sojourn Times

The PDF of Weibull distribution with scale parameter α and shape parameter γ is given as:

$$f_T(t) = \begin{cases} \alpha\gamma t^{\gamma-1}\exp^{-}\alpha t^{\gamma}, & \text{if } \alpha, \gamma > 0 \ \ and \ \ t \geq 0 \\ 0 \qquad\qquad\qquad\quad , & \text{if } otherwise \end{cases}$$

and CDF is:

$$F_T(t) = \begin{cases} 1 - \exp^{-}\alpha t^{\gamma}, & \text{if } \alpha, \gamma > 0 \ \ and \ \ t \geq 0 \\ 0 \qquad\qquad\quad , & \text{if } otherwise \end{cases}$$

Also, the failure and repair rate functions respectively are:

$$f_i(t) = \alpha_i \gamma t^{\gamma-1}$$

and

$$h_i(t) = \beta_i \gamma t^{\gamma-1}; \ \ where, 1 \leq i \leq 3, \ \alpha, \gamma > 0 \ \ and \ \ t \geq 0$$

The reliability function is:

$$R_i(t) = \exp^{-}\alpha_i t^{\gamma}, 1 \leq i \leq 3,\ \alpha, \gamma > 0 \ \ and \ \ t \geq 0$$

The long-run or the steady state probabilities are obtained as under,

$p_{ij} = \lim_{t\to\infty} Q_{ij}(t) = \int q_{ij}(t)dt, \qquad p_{ij}^{(k)} = \lim_{t\to\infty} Q_{ij}^{(k)}(t) \qquad \text{and} \qquad p_{ij}^{(k,l)} = \lim_{t\to\infty} Q_{ij}^{k,l}(t).$

In particular we have

$$p_{01} = \int \alpha_2 \gamma t^{\gamma-1} e^{-\alpha_2 t^{\gamma}} e^{-\alpha_3 t^{\gamma}} e^{-\alpha_1 t^{\gamma}} dt = \frac{\alpha_2}{(\alpha_2+\alpha_3+\alpha_1)}$$

Similarly,

$p_{02} = \frac{\alpha_3}{\alpha_1+\alpha_2+\alpha_3}$ $\qquad$ $p_{13}^{(7)} = \frac{\alpha_1}{\beta_2+\alpha_1+\alpha_3}$

$p_{03} = \frac{\alpha_1}{\alpha_1+\alpha_2+\alpha_3}$ $\qquad$ $p_{20} = \frac{\beta_3}{\beta_3+\alpha_1+\alpha_2}$

$p_{10} = \frac{\beta_2}{\beta_2+\alpha_1+\alpha_3}$ $\qquad$ $p_{21}^{(5)} = \frac{\alpha_2}{\beta_3+\alpha_1+\alpha_2}$

$p_{12}^{(4)} = \frac{\alpha_3}{\beta_2+\alpha_1+\alpha_3}$ $\qquad$ $p_{23}^{(6)} = \frac{\alpha_1}{\beta_3+\alpha_1+\alpha_2}$

Following relationships hold

$p_{01} + p_{02} + p_{03} = 1$
$p_{10} + p_{12}^{(4)} + p_{13}^{(7)} = 1$
$p_{20} + p_{21}^{(5)} + p_{23}^{(6)} = 1$
$p_{30} = p_{42} = p_{51} = p_{63} = p_{73} = 1$

4.1. Mean Sojourn times

The expected duration of time for which system remains in a particular state S_i before making transition to another state is known as the Mean Sojourn time ψ_i i.e. as long as the system is in state S_i, there is no transition from one state to another. With this information, we calculate ψ_i for state S_i. If T_i denotes sojourn time in state S_i, the mean sojourn time ψ_i can be calculated as follows.

$$\psi_i = E[T_i] = \int P(T_i > t)dt$$

Hence, using the above formula following values for mean sojourn time are obtained:

$\psi_0 = \frac{\sqrt{(\frac{1}{\gamma}+1)}}{(\alpha_1+\alpha_2+\alpha_3)^{\frac{1}{\gamma}}}$ $\qquad$ $\psi_1 = \frac{\sqrt{(\frac{1}{\gamma}+1)}}{(\alpha_1+\alpha_3+\beta_2)^{\frac{1}{\gamma}}}$

$\psi_2 = \frac{\sqrt{(\frac{1}{\gamma}+1)}}{(\alpha_1+\alpha_2+\beta_3)^{\frac{1}{\gamma}}}$ $\qquad$ $\psi_3 = \frac{\sqrt{(\frac{1}{\gamma}+1)}}{\beta_1^{\frac{1}{\gamma}}}$

$\psi_4 = \psi_7 = \frac{\sqrt{(\frac{1}{\gamma}+1)}}{\beta_2^{\frac{1}{\gamma}}}$ $\qquad$ $\psi_5 = \psi_6 = \frac{\sqrt{(\frac{1}{\gamma}+1)}}{\beta_3^{\frac{1}{\gamma}}}$

5. Analysis of Reliability and MTSF

If the random variable T_i represents the system's life time when it starts from state $S_i \in E_i$, then the system's reliability is:

$$R_i(t) = P[T_i > t]$$

Here, the failed states are treated as absorbing states.
The state transition diagram and probabilistic arguments can be used to establish the recursive

relations among $R_i(t)$. After identifying the set of equations for $R_i(t)$ and applying the Laplace transform, we obtain

$$R_0^*(s) = \frac{N_1(s)}{D_1(s)} \tag{1}$$

[1] where,

$$N_1(s) = Z_0^* + Z_1^* q_{01}^* + Z_2^* q_{02}^*$$

and

$$D_1(s) = 1 - q_{01}^* q_{10}^* - q_{02}^* q_{20}^*$$

We obtain the system's reliability by taking the inverse Laplace transform of (1).
To obtain MTSF, we use well-known formula

$$E(T_0) = \int R_0(t)dt = \lim_{s\to 0} R_0^*(s) = \frac{N_1(0)}{D_1(0)} \tag{2}$$

where,

$$N_1(0) = \psi_0 + \psi_1 p_{01} + \psi_2 p_{02}$$

and

$$D_1(0) = 1 - p_{01}p_{10} - p_{02}p_{20}$$

Here we have used the results, $q_{ij}^*(0) = p_{ij}$ and $Z_i^*(0)$= ψ_i

6. Availability Analysis

Availability is the likelihood that a system, starting from $S_i \in E_i$, accomplishes its specified purpose at time 't'. The system's availability at a given time is referred to as point wise availability. It is a system performance metric that indicates if a system is potentially functional and capable of delivering the desired service at a specific moment. Establishing recursive relations among various pointwise availabilities and the applications of stochastic reasoning and Laplace transforms gives

$$A_0^*(s) = \frac{N_2(s)}{D_2(s)}$$

where,

$$N_2(s) = Z_0^*(1 - q_{12}^{*(4)} q_{21}^{*(5)}) + Z_1^*(q_{01}^* + q_{02}^* q_{21}^{*(5)}) + Z_2^*(q_{02}^* + q_{01}^* q_{12}^{*(4)})$$

and,

$$\begin{aligned} D_2(s) =& (1 - q_{03}^* q_{30}^*)(1 - q_{12}^{*(4)} q_{21}^{*(5)}) - q_{01}^*[q_{10}^* + q_{12}^{*(4)} q_{20}^* + q_{30}^*(q_{13}^{*(7)} + q_{12}^{*(4)} q_{23}^{*(6)})] \\ & - q_{02}^*[q_{20}^* + q_{10}^* q_{21}^{*(5)} + q_{30}^*(q_{23}^{*(6)} + q_{21}^{*(5)} q_{13}^{*(7)})] \end{aligned} \tag{3}$$

[1] Limits of integration whenever they are 0 to ∞ are not mentioned.

The steady state availability that the system will be up in long run is given as under

$$A_0 = \lim_{t\to\infty} A_0(t) = \lim_{s\to 0}\ sA_0^*(s) = \frac{N_2(0)}{D_2(0)}$$

Now using the results

$$\lim_{s\to 0} Z_i^*(s) = \int Z_i(t)dt = \psi_i \ \ and \ \ \lim_{s\to 0} q_{ij}^*(s) = p_{ij}$$

we get

$$N_2(0) = \psi_0(1 - \overset{(4)}{p_{12}}\overset{(5)}{p_{21}}) + \psi_1(p_{01} + p_{02}\overset{(5)}{p_{21}}) + \psi_2(p_{02} + p_{01}\overset{(4)}{p_{12}})$$

$$D_2(0) = (1 - p_{03})(1 - \overset{(4)}{p_{12}}\overset{(5)}{p_{21}}) - p_{01}[p_{10} + \overset{(4)}{p_{12}}p_{20} + \overset{(7)}{p_{13}} + \overset{(4)}{p_{12}}\overset{(6)}{p_{23}}] - p_{02}[p_{20} + p_{10}\overset{(5)}{p_{21}} + \overset{(6)}{p_{23}} + \overset{(5)}{p_{21}}\overset{(7)}{p_{13}}]$$

For a given system, the steady-state probability of its long-term operation is given by

$$A_0 = \lim_{t\to\infty} A_0(t) = \lim_{s\to 0}\ sA_0^*(s)$$

$$\lim_{s\to 0}\frac{sN_2(s)}{D_2(s)} = \lim_{s\to 0} N_2(s)\lim_{s\to 0}\frac{s}{D_2(s)}$$

Since as s $\to$ 0, $D_2(s)$ becomes zero. Therefore, applying L'Hospital's rule, A_0 becomes

$$A_0 = \frac{N_2(0)}{D_2'(0)} \tag{4}$$

where,

$$D_2'(0) = (1 - \overset{(4)}{p_{12}}\overset{(5)}{p_{21}})(\psi_0 + \psi_3) + \psi_1(p_{01} + p_{02}\overset{(5)}{p_{21}}) + \psi_2(p_{02} + p_{01}\overset{(4)}{p_{12}}) \tag{5}$$

Using $N_2(0)$ and $D_2'(0)$ in equation[4], the expression for A_0 can be determined.
The system's expected uptime for (0,t] is provided by

$$\mu_{up}(t) = \int_0^t A_0(u)du$$

So that,

$$\mu_{up}^*(s) = \frac{A_0^*(s)}{s}$$

7. Busy Period Analysis

$B_i(t)$ is defined as the probability that the system, which initially starts from the regenerative state Si∈ E, is under repair at time t due to failure of the unit. By applying basic probabilistic logics and the Laplace transformations and solving the resulting set of equations for $B_0^*(s)$, we get

$$B^*{}_0(s) = \frac{N_3(s)}{D_2(s)}$$

$$N_3(s) = (q_{01}^* + q_{02}^*\overset{*(5)}{q_{21}})(Z_1^* + q_{14}^*Z_4^* + q_{17}^*Z_7^*) + (q_{02}^* + q_{01}^*\overset{*(4)}{q_{12}})(Z_2^* + q_{23}^*Z_3^* + q_{26}^*Z_6^*) + Z_3^*[q_{01}^*(\overset{*(7)}{q_{13}} + \overset{*(4)}{q_{12}}\overset{*(6)}{q_{23}}) + q_{02}^*(\overset{*(6)}{q_{23}} + \overset{*(7)}{q_{13}}\overset{*(5)}{q_{21}}) + q_{03}^*(1 - \overset{*(4)}{q_{12}}\overset{*(5)}{q_{21}})]$$

and, D_2(s) is same as given in equation [3].
The probability that the repairman will be busy in the long run is as follows:

$$B_0 = \lim_{t\to\infty} B_0(t) = \lim_{s\to 0}\ sB_0^*(s) = \frac{N_3(0)}{D_2^{'}(0)}$$

where,

$$N_3(0) = (p_{01} + p_{02}p_{21}^{(5)})(\psi_1 - p_{14}\psi_4 + p_{17}\psi_7) + (p_{02} + p_{01}p_{12}^{(4)})(\psi_2 + p_{25}\psi_5 + p_{26}\psi_6) + \psi_3[p_{01}(p_{13}^{(7)} + p_{12}^{(4)}p_{23}^{(6)}) + p_{02}(p_{23}^{(6)} + p_{13}^{(7)}p_{21}^{(5)}) + p_{03}(1 - p_{12}^{(4)}p_{21}^{(5)})$$

and $D_2^{'}(0)$ is same as obtained in [5].
During (0,t], the repairman's expected busy time is given by

$$\mu_b(t) = \int_0^t B_0(u)du$$

So that,

$$\mu_b^*(s) = \frac{B_0^*(s)}{s}$$

8. Expected Number of Repairs

Define $V_i(t)$ the expected number of repairs of the failed units during the time range (0,t] when the system starts from the regenerative state S_i. Furthermore, the recurrence relations may simply be framed using the definition of $V_i(t)$. Then, by making use of Laplace-Stieltjes transformations the solution of resulting set of equations for $\widetilde{V}_0(s)$, is given as
$\widetilde{V}_0(s) = N_4(s)/D_3(s)$
where, and $D_3(s)$ is written on replacing q_{ij}^* and $q_{ij}^{(k)*}$ by $\widetilde{Q}_{ij}$ and $\widetilde{Q}_{ij}^{(k)}$ respectively in equation[3].

$$V_0 = \lim_{t\to\infty} V_0(t) = \lim_{s\to 0} s\ \widetilde{V}_0(s) = \frac{N_4(0)}{D_2^{'}(0)}$$

where,

$$N_4(s) = (\widetilde{Q}_{01} + \widetilde{Q}_{02} + \widetilde{Q}_{21}^{(5)})(\widetilde{Q}_{12}^{(4)} + \widetilde{Q}_{13}^{(7)}) + (\widetilde{Q}_{01}\widetilde{Q}_{12}^{(4)} + \widetilde{Q}_{02})(\widetilde{Q}_{20} + \widetilde{Q}_{21}^{(5)} + \widetilde{Q}_{23}^{(6)}) + \widetilde{Q}_{30}[\widetilde{Q}_{01}(\widetilde{Q}_{13}^{(7)} + \widetilde{Q}_{12}^{(4)}\widetilde{Q}_{23}^{(6)}) + \widetilde{Q}_{02}(\widetilde{Q}_{23}^{(6)} + \widetilde{Q}_{13}^{(7)}\widetilde{Q}_{21}^{(5)}) + \widetilde{Q}_{03}(1 - \widetilde{Q}_{12}^{(4)}\widetilde{Q}_{21}^{(5)})]$$

In steady state the expected number of repairs per unit time is given by
$V_0 = \lim_{t\to\infty} V_0(t) = \lim_{s\to 0} s\ \widetilde{V}_0(s) = \frac{N_4(0)}{D_2^{'}(0)}$
where,

$$N_4(0) = (p_{01} + p_{02} + p_{21}^{(5)})(p_{12}^{(4)} + p_{13}^{(7)}) + (p_{01}p_{12}^{(4)} + p_{02})(p_{20} + p_{21}^{(5)} + p_{23}^{(6)}) + p_{30}[p_{01}(p_{13}^{(7)} + p_{12}^{(4)}p_{23}^{(6)}) + p_{02}(p_{23}^{(6)} + p_{13}^{(7)}p_{21}^{(5)}) + p_{03}(1 - p_{12}^{(4)}p_{21}^{(5)})]$$

9. Profit Function Analysis

With the help of the characteristics obtained two profit functions $P_1(t)$ and $P_2(t)$ can be obtained for the system model under study. The expected total profit incurred during (0,t) are:
$P_1(t)$ = Expected total revenue in(0,t] - Expected total expenditure in(0,t]

$$= K_0\mu_{up}(t) - K_1\mu_b(t)$$

and

$$P_2(t) = K_0\mu_{up}(t) - K_2V_0(t)$$

where,
K_0 = revenue per unit up time of the system.
K_1 and K_2 = Amounts paid to the repairmen per unit of time and per unit repair while repairing the failed units.
The expected total gain per unit of time in steady state is provided by:

$$P_1 = \lim_{t\to\infty}\frac{P_1(t)}{t} = \lim_{s\to 0} s^2\overset{*}{P}(s)$$

Therefore, we have

$$P_1 = K_0A_0 - K_1B_0 \tag{6}$$

Similarly,

$$P_2 = \lim_{t\to\infty}\frac{P_2(t)}{t} = K_0A_0 - K_1V_0 \tag{7}$$

10. Estimation of the Parameters, MTSF, And Profit Function

10.1. Classical Estimation

10.1.1 ML Estimation

Let us take

$\underset{\sim}{X_1} = (x_{11}, x_{12}, ..., x_{1n_1})$, $\underset{\sim}{X_2} = (x_{21}, x_{22}, ...x_{2n_2})$, $\underset{\sim}{X_3} = (x_{31}, x_{32}, ..., x_{3n_3})$,
$\underset{\sim}{X_4} = (x_{4_1}, x_{42}, ..., x_{4n_4})$, $\underset{\sim}{X_5} = (x_{51}, x_{52}, ..., x_{5n_5})$, $\underset{\sim}{X_6} = and\ \ (x_{61}, x_{62}, ..., x_{6n_6})$

Therefore, Likelihood function of combined sample is :

$$L = (X_1, X_2, X_3, X_4, X_5, X_6|\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3)$$

The pdf of Weibull distribution is $f(x) = \alpha\gamma t^{\gamma-1}\exp^{-}\alpha t^{\gamma}$,$\alpha, \gamma\ \ and\ \ t \geq 0$

$$L = \alpha_1^{n_1}\alpha_2^{n_2}\alpha_3^{n_3}\beta_1^{n_4}\beta_2^{n_5}\beta_3^{n_6}\gamma^{n_1+n_2+n_3+n_4+n_5+n_6}Z_1Z_2Z_3Z_4Z_5Z_6\exp^{-(\alpha_1W_1+\alpha_2W_2+\alpha_3W_3+\beta_1W_4+\beta_2W_5+\beta_3W_6)} \tag{8}$$

Here, $W_i = \sum_{j=1}^{n_i} x_{ij}^{\gamma}$ and $Z_i = \prod_{j=1}^{n_i} x_{ij}^{(\gamma-1)}$; i = 1,2,3,4,5,6
On solving, we get

$$logL = n_1log\alpha_1 + n_2log\alpha_2 + n_3log\alpha_3 + n_4log\beta_1 + n_5log\beta_2 + n_6log\beta_3 + logZ_1 + logZ_2 + logZ_3 + logZ_4 + logZ_5 + logZ_6 - (\alpha_1W_1 + \alpha_2W_2 + \alpha_3W_3 + \beta_1W_4 + \beta_2W_5 + \beta_3W_6) + (n_1 + n_2 + n_3 + n_4 + n_5 + n_6)log\gamma \tag{9}$$

The, MLE $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3\ , \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)$ of the parameters
$(\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3)$ are as under

$\hat{\alpha}_1 = \frac{n_1}{W_1}$, $\hat{\alpha}_2 = \frac{n_2}{W_2}$
$\hat{\alpha}_3 = \frac{n_3}{W_3}$, $\hat{\beta}_1 = \frac{n_4}{W_4}$
$\hat{\beta}_2 = \frac{n_5}{W_5}$, $\hat{\beta}_3 = \frac{n_6}{W_6}$

The asymptotic distribution of$(\hat{\alpha_1} - \alpha_1, \hat{\alpha_2} - \alpha_2, \hat{\alpha_3} - \alpha_3, \hat{\beta}_1 - \beta_1, \hat{\beta}_2 - \beta_2, \hat{\beta}_3 - \beta_3,) \sim N_6(0, I^{-1})$, where I is the Fisher Information matrix with diagonal elements as

$I_{11} = \frac{n_1}{\alpha_1^2}, \quad I_{22} = \frac{n_2}{\alpha_2^2}, \quad I_{33} = \frac{n_3}{\alpha_3^2}, \quad I_{44} = \frac{n_4}{\beta_1^2}, \quad I_{55} = \frac{n_5}{\beta_2^2}, \quad I_{66} = \frac{n_6}{\beta_3^2}$

and all non-diagonal elements are zero. Using invariance property of MLE's, we can extract The MLE $\hat{M}$ & $\hat{P}$ of MTSF and Profit function. Also, asymptotic distribution of $(\hat{M} - M) is N(0, A' I^{-1} A)$ & that of $(\hat{P} - P) is N(0, B' I^{-1} B)$, where

$A' = (\frac{\delta M}{\delta \alpha_1}, \frac{\delta M}{\delta \alpha_2}, \frac{\delta M}{\delta \alpha_3}, \frac{\delta M}{\delta \beta_1}, \frac{\delta M}{\delta \beta_2}, \frac{\delta M}{\delta \beta_3})$

$B' = (\frac{\delta P}{\delta \alpha_1}, \frac{\delta P}{\delta \alpha_2}, \frac{\delta P}{\delta \alpha_3}, \frac{\delta P}{\delta \beta_1}, \frac{\delta P}{\delta \beta_2}, \frac{\delta P}{\delta \beta_3})$

10.2. Bayesian Estimation

A statistical method called Bayesian estimation is used to ascertain how sample information and past knowledge affect the prior distributions of the parameters being studied. It is assumed that the model's parameters are random variables with independent Gamma prior distributions. In this case, we use the gamma prior distribution and the associated PDFs to estimate the unknown parameters as

$$\alpha_1 \sim Gamma(a_1, b_1) \qquad (\alpha_1, a_1, b_1) > 0, \tag{10}$$

$$\alpha_2 \sim Gamma(a_2, b_2) \qquad (\alpha_2, a_2, b_2) > 0, \tag{11}$$

$$\alpha_3 \sim Gamma(a_3, b_3) \qquad (\alpha_3, a_3, b_3) > 0, \tag{12}$$

$$\beta_1 \sim Gamma(a_4, b_4) \qquad (\beta_1, a_4, b_4) > 0, \tag{13}$$

$$\beta_2 \sim Gamma(a_5, b_5) \qquad (\beta_2, a_5, b_5) > 0, \tag{14}$$

$$\beta_3 \sim Gamma(a_6, b_6) \qquad (\beta_3, a_6, b_6) > 0, \tag{15}$$

Here, a_i *and* b_i (i = 1,2,3,4,5,6) denotes the shape and scale parameters
The posterior distributions of these parameters are computed, taking into account the likelihood function and the prior distributions:

$$\alpha_1 | \underset{\sim}{X_1} \sim Gamma(n_1 + a_1, b_1 + W_1) \tag{16}$$

$$\alpha_2 | \underset{\sim}{X_2} \sim Gamma(n_2 + a_2, b_2 + W_2) \tag{17}$$

$$\alpha_3 | \underset{\sim}{X_3} \sim Gamma(n_3 + a_3, b_3 + W_3) \tag{18}$$

$$\beta_1 | \underset{\sim}{X_4} \sim Gamma(n_4 + a_4, b_4 + W_4) \tag{19}$$

$$\beta_2 | \underset{\sim}{X_5} \sim Gamma(n_5 + a_5, b_5 + W_5) \tag{20}$$

$$\beta_3 | \underset{\sim}{X_6} \sim Gamma(n_6 + a_6, b_6 + W_6) \tag{21}$$

Observations are created from the above-mentioned posterior distributions in order to obtain the width of HPD intervals and Bayes estimates for parameters. The aforementioned draws are directly entered into equations [2], [6], and [7] to obtain Bayesian estimates of the MTSF and profit function. From the sample averages of the relevant drawings, Bayesian estimates of parameters and reliability characteristics are obtained using a squared error loss function.

11. Simulation Study

A simulation study is conducted to investigate the behavior of parameters, estimates, and reliability features. Table 1-6 displays the values of the Standard Error (SE)/Posterior Standard Error (PSE) and the width of confidence/HPD intervals. Samples of sizes $n_1 = n_2 = n_3 = n_4 = n_5 = n_6 = 100$ were taken from the six investigated distributions while presuming various parameter values as shown in Tables 1-6. R software is used to perform the computation with ten thousand iterations.

Table 1: *MTSF values for fixed $\beta_2 = 0.5$ and varying α_2*

α_2	True MTSF	MLE.MTSF	SE	C.I	Bayes MTSF	PSE	HPD Interval
0.1	3.59	3.55	0.00155	0.0059	4.79	0.0326	0.01173
0.2	2.87	2.71	0.00078	0.0036	4.71	0.0324	0.01160
0.3	2.41	2.22	0.00049	0.0027	4.53	0.0318	0.01131
0.4	2.10	1.86	0.00038	0.0022	4.50	0.0316	0.01129
0.5	1.86	1.65	0.00031	0.0020	4.38	0.0312	0.01124
0.6	1.68	1.44	0.00028	0.0018	4.21	0.0306	0.01113
0.7	1.54	1.32	0.00026	0.0016	4.20	0.0302	0.01111
0.8	1.42	1.19	0.00024	0.0015	4.19	0.0292	0.01086
0.9	1.32	1.11	0.00023	0.0014	4.06	0.0302	0.01082
1	1.24	1.01	0.00021	0.0013	3.99	0.0295	0.01077

Table 2: *MTSF values for fixed β_2=1 and varying α_2*

α_2	True MTSF	MLE.MTSF	SE	C.I	Bayes MTSF	PSE	HPD Interval
0.1	3.95	4.08	0.00244	0.0077	5.86	0.0330	0.01186
0.2	3.09	2.97	0.00113	0.0044	5.63	0.0332	0.01163
0.3	2.56	2.35	0.00068	0.0031	5.19	0.0318	0.02613
0.4	2.20	1.95	0.00049	0.0025	5.06	0.0316	0.02420
0.5	1.94	1.69	0.00039	0.0022	4.75	0.0318	0.02423
0.6	1.75	1.51	0.00034	0.0020	4.70	0.0309	0.02326
0.7	1.59	1.36	0.00030	0.0018	4.69	0.0308	0.01909
0.8	1.46	1.24	0.00028	0.0017	4.53	0.0305	0.02901
0.9	1.36	1.11	0.00027	0.0016	4.26	0.0301	0.02423
1	1.27	1.03	0.00025	0.0015	4.11	0.0297	0.02136

Table 3: *MTSF values for fixed β_2=1.5 and varying α_2*

α_2	True MTSF	MLE.MTSF	SE	C.I	Bayes MTSF	PSE	HPD Interval
0.1	4.25	4.47	0.00333	0.0093	5.90	0.0330	0.02450
0.2	3.27	3.15	0.00149	0.0050	5.83	0.0325	0.02118
0.3	2.68	2.49	0.00086	0.0035	5.67	0.0325	0.03697
0.4	2.29	2.08	0.00061	0.0027	5.59	0.0322	0.02498
0.5	2.01	1.76	0.00048	0.0023	5.23	0.0318	0.02648
0.6	1.79	1.55	0.00041	0.0021	5.01	0.0319	0.02250
0.7	1.63	1.37	0.00036	0.0019	4.99	0.0308	0.02110
0.8	1.50	1.26	0.00033	0.0018	4.91	0.0306	0.02223
0.9	1.39	1.15	0.00031	0.0017	4.70	0.0306	0.06198
1	1.29	1.05	0.00029	0.0016	4.69	0.0300	0.02095

Table 4: *Profit values for fixed β_2=0.5 and varying α_2*

α_2	True profit	MLE.Profit	SE	C.I	Bayes Profit	PSE	HPD Interval
0.1	93.76	89.02	0.097	0.05770	180.56	0.160	0.2386
0.2	86.61	81.54	0.091	0.04630	180.52	0.163	0.2422
0.3	80.59	75.04	0.085	0.06095	180.33	0.167	0.2488
0.4	75.41	69.38	0.086	0.06123	180.23	0.170	0.2516
0.5	70.90	64.29	0.088	0.06591	180.00	0.172	0.2529
0.6	66.91	59.74	0.077	0.05611	179.98	0.171	0.2571
0.7	63.34	55.67	0.082	0.06343	179.56	0.172	0.2563
0.8	60.13	51.97	0.071	0.06374	179.21	0.173	0.2569
0.9	57.22	48.56	0.075	0.06991	179.01	0.175	0.2565
1	54.57	45.46	0.089	0.077320	178.88	0.173	0.2548

Table 5: *Profit values for fixed β_2=1 and varying α_2*

α_2	True profit	MLE.Profit	SE	C.I	Bayes Profit	PSE	HPD Interval
0.1	101.49	107.01	0.084	0.06113	180.73	0.117	0.1734
0.2	94.31	99.72	0.086	0.07476	180.71	0.120	0.1779
0.3	88.19	93.34	0.094	0.07369	180.53	0.125	0.1832
0.4	82.88	87.67	0.093	0.06006	180.46	0.127	0.1877
0.5	78.20	82.53	0.092	0.06135	180.37	0.130	0.1903
0.6	74.03	77.87	0.094	0.06475	180.33	0.131	0.1929
0.7	70.30	73.61	0.085	0.06013	180.21	0.133	0.1954
0.8	66.91	69.70	0.089	0.07130	180.09	0.136	0.1956
0.9	66.83	66.07	0.080	0.07032	180.00	0.135	0.2023
1	61.00	62.70	0.077	0.06212	179.99	0.134	0.1995

Table 6: *Profit values for fixed $\beta_2 = 1.5$ and varying α_2*

α_2	True profit	MLE.Profit	SE	C.I	Bayes Profit	PSE	HPD Interval
0.1	103.26	110.57	0.098	0.06532	181.82	0.110	0.1630
0.2	96.26	103.61	0.091	0.06521	181.62	0.114	0.1690
0.3	90.25	97.46	0.087	0.06222	181.53	0.118	0.1741
0.4	85.02	91.95	0.086	0.06509	181.36	0.120	0.1771
0.5	80.38	87.00	0.081	0.06011	181.09	0.122	0.1827
0.6	76.24	82.44	0.082	0.06529	181.00	0.124	0.1816
0.7	72.51	78.26	0.085	0.06246	180.76	0.124	0.1846
0.8	69.12	74.42	0.092	0.06479	180.53	0.125	0.1863
0.9	66.03	70.83	0.097	0.07351	180.19	0.128	0.1877
1	63.18	67.51	0.082	0.06274	180.00	0.128	0.1862

12. Graphical Study

A more detailed and vivid depiction of the behavior of the system can be obtained by performing a graphical analysis of the system model. For more extensive analysis, we plot MTSF and Profit function wrt α_2 failure rate of unit B for different values of β_2 repair rate of unit B as 0.5, 1 and 1.5 while keeping other parameters as fixed α_1= 0.9, α_3= 1.5, β_1=0.5, β_3=1.

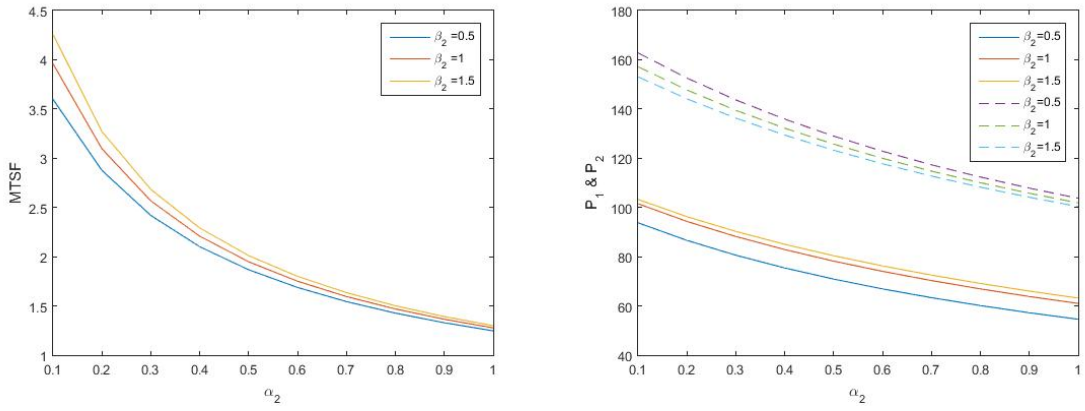


Figure 2: *(a) Behaviour of MTSF wrt to α_2 for different values of β_2 and (b) Behaviour of P_1&P_2 wrt to α_2 for different values of β_2*

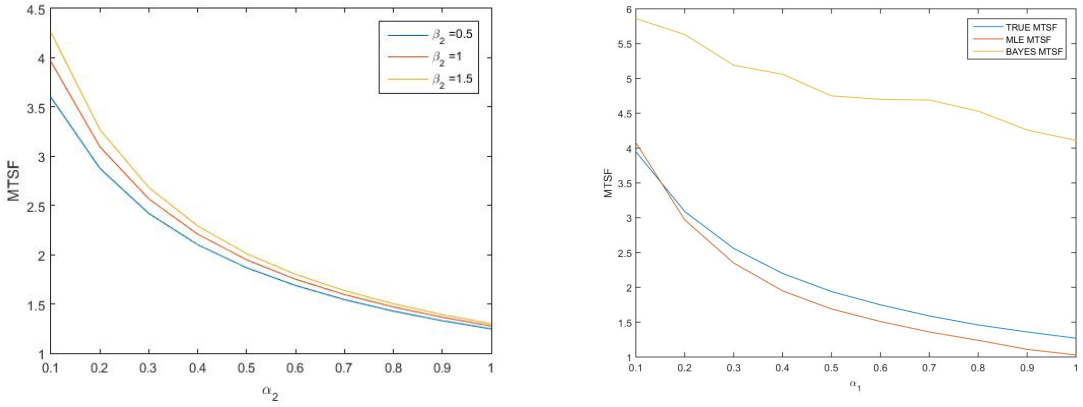


Figure 3: *(a) Behaviour of True MTSF, MLE MTSF & Bayes MTSF wrt to α_2 for β_2= 0.5 and (b) Behaviour of True MTSF, MLE MTSF & Bayes MTSF wrt to α_2 for β_2= 1*

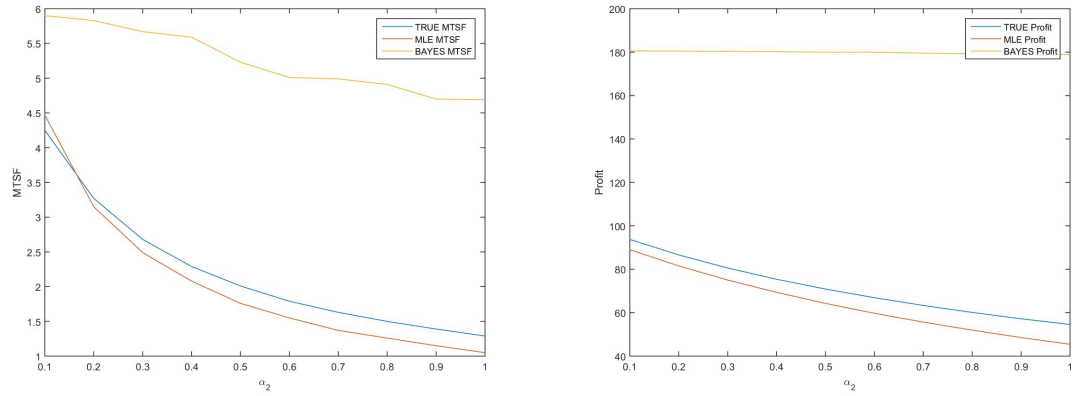


Figure 4: *(a) Behaviour of True MTSF, MLE MTSF & Bayes MTSF wrt to α_2 for β_2= 1.5 and (b) Behaviour of True Profit, MLE Profit & Bayes Profit wrt to α_2 for β_2= 0.5*

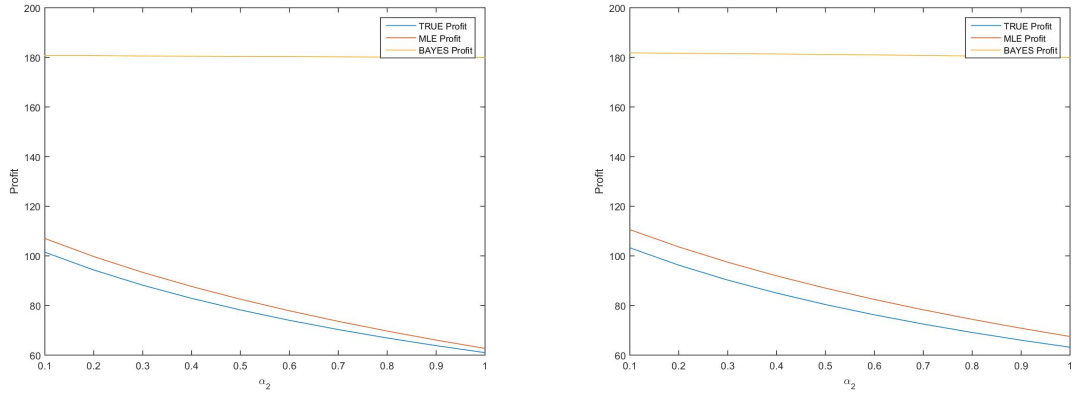


Figure 5: *(a) Behaviour of True Profit, MLE Profit & Bayes Profit wrt to α_2 for β_2= 1 and (b) Behaviour of True Profit, MLE Profit & Bayes Profit wrt to α_2 for β_2= 1.5*

13. Discussion and Conclusion

1. Tables and figures exhibit that MTSF decreases as the failure rate α_2 increases, but increases as the repair rate β_2 increases same trend is observed for profit function. Profit function decreases as the failure rate α_2 increases, and increases when value of repair rate β_2 increases.
2. Tables 1-6 indicate that for fixed and variable parameters, MLEs estimates of the profit function perform better than Bayes in terms of SE as well as in terms of the width of the confidence intervals as Bayes estimates have higher PSE and the width of HPD intervals as compared to the MLE's counterpart.
3. Based on the above discussions, we conclude that for estimating the MTSF and Profit function of the analyzed model, the Classical approach outperforms the Bayes approach.

References

[1] Sharma, A. and Kumar, P. (2019). Analysis of Reliability Measures of Two Identical Unit System with One Switching Device and Imperfect Coverage *Reliability: Theory & Applications*, 14(1):44–52.

[2] Kuntal D., Ashish K., and Monika S.(2017). Performance analysis of a non-identical unit system with priority and weibull repair and failure laws *International Journal of Computer Applications*, 975:8887.

[3] Pedram P. G., Parham A., Seyed T. A. N., and Seyed A. A. N. (2016). Redundancy allocation problem of a system with increasing failure rates of components based on weibull distribution: A simulation-based optimization approach *Reliability Engineering & System Safety*, 152:187–196.

[4] Kishan, R. and Jain, D. (2014). Classical and Bayesian analysis of reliability characteristics of a two-unit parallel system with Weibull failure and repair laws *International Journal of System Assurance Engineering and Management*, 5:252–261.

[5] Kumar, A., Garg, R. and Barak, MS (2023). Performance analysis of computer systems with Weibull distribution subject to software upgrade and load recovery *Life Cycle Reliability and Safety Engineering*, 12(1):51–63.

[6] Kumar, A. and Saini, M. (2014). Cost-benefit analysis of a single-unit system with preventive maintenance and Weibull distribution for failure and repair activities *Journal of Applied Mathematics, Statistics and Informatics*, 10(2):5–19.

[7] Malik, SC and others (2019). Cost-benefit evaluation of a repairable system with alternate repair and weibull distribution *International Journal of Agricultural & Statistical Sciences*, 15(2).

[8] Saini, M., Devi, K. and Kumar, A. (2021). Stochastic modeling and profit evaluation of a redundant system with priority subject to Weibull densities for failure and repair *Information and Communication Technology for Intelligent Systems: Proceedings of ICTIS 2020, Volume 1*, pages 11–20.

[9] Sharma, VK and Gupta, Surendra P. (2012). Cost Benefit Analysis of Three Unit Redundant System Model with Correlated Failures and Repairs *Journal of Pure & Applied Science & Technology*, 2(1).

[10] Van ,C., Jasper (2003). Assessment of interruption costs in electric power systems using the Weibull-Markov model *Department of Electric Power Engineering, Chalmers University of Technology*.

[11] Sharma, P. and Kumar, P. (2024) Analysis of two non-identical unit system having safe and unsafe failures with rebooting and parametric estimation in classical and bayesian paradigms *Reliability: Theory & Applications*, 19, 3(79):364–379.

RT&A, No 2 (84)
Volume 20, June 2025

EXPONENTIAL TYPE ESTIMATORS OF POPULATION MEAN - A REVIEW

Sajad Hussain*[1], Vilayat Ali Bhat[2]

•

[1] Department of Mathematics, School of Engineering (SOE), Presidency University, Bengaluru, Karnataka, India - 560064.
[2] Department of Mathematics, School of Computer Science and Artificila Intelegence, SR University, Telengana, India - 506371
sajad.stat321@gmail.com*[1]
vilayat.stat@gmail.com[2]

Abstract

Survey sampling is a crucial statistical approach for making inferences about population characteristics based on the analysis of a representative sample. This method offers a practical solution for saving time and resources while ensuring reliable outcomes. Auxiliary information plays a pivotal role in enhancing the accuracy of estimators used in finite population sampling. Traditional estimators, such as ratio, product, and regression types, have proven effective in improving the precision of population parameter estimation. The introduction of exponential-type transcendental functions marked a significant advancement in this field, further increasing the efficiency of ratio and product estimators. This study provides an in-depth review of the literature on exponential-type ratio and product estimators, emphasizing their relevance in survey sampling. It examines theoretical progress and practical implementations, showcasing their potential to reduce bias and improve estimation accuracy by leveraging complex relationships between study variables and auxiliary information. The insights from this review contribute to advancing research in finite population sampling and offer valuable directions for future studies in the field.

Keywords: Exponential Estimator; Auxiliary information; Study variable; Mean square error; Auxiliary attribute

1. Introduction

The concept of sampling theory is as old as human civilization. The main aim of sampling theory is to know about the population or universe by studying only a part called the sample drawn from it. This process of inspection is very widespread and commonly used on various occasions. However, this job is never done on a very large scale.

The sampling methods are extensively used by government bodies throughout the world for assessing different characteristics of national economy as they are required for taking decisions regarding imposition of taxes, fixation of prices, minimum wages etc. for planning and projection of future economic policies. Surveys are conducted for estimation of yield rates and acreages under different crops, value added by manufacturer in the industries sector, number of unemployed persons in the labor force, construction of cost of living indices for persons in different professions and so on. Sample survey techniques are widely used in market research for assessing the preferential pattern of consumers towards the products of different types, the future possible demand for a new product which a company wishes to introduce, scope for any diversification in the production schedule and so on. Bowley [1] and Neyman [2] laid the foundations of modern sampling theory. The works of R.A. Fisher during twenties provided a scientific basis for selecting

a random sample from the population. Many research papers relating to sampling theory and methods have been written for enhancing the knowledge and concepts in practice.

In sampling theory, a variable called auxiliary variable works as an instrumental variable in the estimation of the population parameters of the study variable. An auxiliary variable is a variable that is correlated with the study variable and about which information is available prior to sampling. A judicious use of this information may provide us with estimators which are more precise than those which do not make use of auxiliary information. The concept of using auxiliary information is well known in sampling theory. Its use is of paramount importance in sample surveys as it leads to increased precision of estimators for estimating the population parameters. The origin of utilizing, auxiliary information in sample surveys can be traced back to the origin of sampling theory itself. It has been a general view of the survey statisticians for last five decades that the usual methods for estimating population mean (or total) of a variable of interest (study variable), say Y may lead to much improvement in the precision by using auxiliary information. In most of the survey situations, the auxiliary information is always available in one form or the other or can be made available by diverting for this purpose a part of the survey resources at moderate cost. The method of utilizing auxiliary information depends on the form in which it is available. In whatever form it is available, one may always utilize it to devise sampling strategies which are better than those in which no auxiliary information is used. Watson [3], Neyman [4], Cochran [5] and Hansen & Hurwitz [6] were pioneer to use auxiliary information in devising estimation procedures which lead to improvement in precision of estimation. Searls [7], Thompson [8], Singh [9], Ray & Singh [10], Hirano [11], Singh and Singh [12], Maiti and Tripathi [13] have considered the problem of estimation of population parameter when such auxiliary information is available. The use of auxiliary information in theory of sampling design has been continuously gaining the attention of survey researchers from the inception due to its wide importance in all walks of real life situations.

John Graunt in 16^{th} century was the first to estimate the ratio $\frac{y}{x}$, where y represented the total population and x the known total number of registered births in the same area during the preceding year. Later, Messance and Moheau in 17^{th} century very carefully prepared estimates based on enumeration of population in certain districts of France and on the count of births, deaths and marriages as reported for the whole country. In survey sampling, when auxiliary information about all the sampling units is available then it is convenient to estimate unknown population mean of study variable using ratio estimator provided that there is a positive correlation between study variable and auxiliary variable and $\frac{C_x}{2C_y} < \rho \leq +1$. The ratio estimate of Y i.e. ratio estimate of the population total of y_i, $\hat{t}_r = \hat{R}X = \left(\frac{y}{x}\right)X = \left(\frac{\bar{y}}{\bar{x}}\right)X$. If the quantity to be estimated is $\bar{Y}$ i.e. the population means of y_i. Then the ratio estimate envisaged by Cochran [5] is, $\hat{t}_{cr} = \hat{R}\bar{X} = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)$. If there is negative correlation between auxiliary and the study variable and $-1 \leq \rho < -\frac{C_x}{2C_y}$, then the product method of estimation proposed by Robson [14] and revisited by Murthy [15] which can be employed efficiently is $\hat{t}_{rbp} = \bar{y}\left(\frac{\bar{x}}{\bar{X}}\right)$. The method of utilizing auxiliary information depends upon the form in which it is available. When the auxiliary information is quantitative in nature, the authors such as Singh *et al.* [16], Das [17], Sisodia and Dwivedi [18], Sharma and Bhatnagar [19], Singh *et al.* [20], Subramani and Kumarapandiyan [21], Yadav and Kadilar [22], Vishwakarma *et al.* [23], Misra *et al.* [24], Shalabh and Tsai [25], Singh and Vishwakarma [26], Mehta *et al.* [27] etc. proposed modified ratio/product type estimators which are more efficient than the sample mean and the classical/existing ratio/product type estimators. There are many practical situations where the auxiliary information is qualitative in nature, in other words the auxiliary variable correlated with the study variable is an attribute. The authors Jhajj *et al.* [28] and Shabir & Gupta [29] have elaborated through the examples (a) the height of a person may depend on its sex (b) amount of milk produced by a cow may depend on its breed (c) the yield of a crop may depend on its variety. In all these situations, the estimators where the auxiliary information is quantitative in nature cannot be used as there exists a point bi-serial correlation between the study and the auxiliary variable . Therefore, Naik and Gupta

[30], Jhajj *et al*. [28], Shabir & Gupta ([29] , [31]) and Abd-Elfattah *et al*. [32], have made some attempts in this direction and proposed the modified estimators of population mean by using prior knowledge of the parameters of auxiliary attribute.

Bahl and Tuteja [33] were the pioneer to make use of an exponential type transcendental function in the estimators of population mean and proposed the ratio and product type estimators. Later, Singh *et al*. [34] proposed the exponential ratio and product type estimatorsof population mean using auxiliary variable as attribute. It was observed by using such type of function, a significant reduction in the MSE. Generally, the exponential type estimators are found to be estimating the population mean more precisely as compared to conventional/modified ratio and product type estimators and can be used even if there is not a strong/high degree of correlation between the auxiliary and study variable. The authors Singh *et al*. [20], Onyeka [35], Khan and Siddiqi [36], Yasmeen *et al*. [37], Yadav and Sharma [38], Zaman and Kadilar [39], Hussain *et al*. ([40], [41], [42], [43], [44], [45], [46]) etc. proposed some modified exponential ratio/product type estimators to be more precise than the conventional ratio/product and exponential ratio/product type estimators.

In the absence of the auxiliary variable the estimators discussed so far are not possible. However one may think of getting additional information on the study variable and may propose ratio and linear regression type estimators to improve the performance of the estimators. Therefore, the pioneer work of Subramani [47] proposed a precise ratio estimator to estimate population mean of skewed population without the use of auxiliary variable by taking the advantage of the median of main variable. The median of the study variable may easily be available without having exact information on every data point (see Subramani, [47]) and these median based estimators are robust in nature as they are least influenced by the outliers.

It is mentioned that the efficiency of estimators can be increased by using known information of the parameters of auxiliary variable correlated with variable under study. Sometimes such information on parameters of auxiliary variables may not be known in advance, in that situations then two-phase/double sampling design is generally used. In two-phase/double sampling, a part of the budget is used to collect information on the auxiliary variable. A large first phase preliminary sample of size η (say) and a second phase (small) sample of size n which is nested within the first phase sample ($n < \eta$) is drawn. In double sampling procedure, it is proposed to use the information gathered in the first phase sample as auxiliary information to increase the precision of the information gathered in the second phase sample. Neyman [4] was the first to formulate double sampling (or two-phase sampling) in connection with collecting information on the strata sizes in a stratified sampling. Later the theory of two-phase sampling technique have been studied at various situations by Bose [48], Sukhatme and Koshal [49], Tikkiwal [50], Sukhatme [51], Rao [52], Rao ([53], [54]) and many others. It is already discussed that large number of ratio/product, ratio cum product type, regression estimators have been proposed by different statisticians using known values of population mean, population proportion, population variance etc. of auxiliary variable for the estimation of the parameters of interest. In case such known values are not available, then one may use two-phase sampling technique to make the mentioned estimators for practical use. Several authors such as Singh *et al*. [55] , Ahmed [56], Hossain *et al*. [57], Jhajj *et al*. [58], Upadhaya *et al*. [59], Samiuddin and Hanif [60], Kamal and Shahbaz [61], Hanif *et al*. [62], have proposed different estimators in two-phase sampling. Singh *et al*. [63], proposed two-phase sampling version of exponential type estimators of Bahl and Tuteja [33] estimators. They found that estimator was more precise than the simple mean per unit, usual two-phase sampling ratio and product estimators.

2. The exponential Type Estimators

As already stated Bahl and Tuteja [33] were the first who proposed exponential ratio and product type estimators of population mean as t_{btr} and t_{btp} respectively. Further by adding K to all observations taken on x, they proposed ratio and product type estimators as t_{btr_1} and t_{btp_1}

respectivelly

$$t_{btr} = \bar{y}\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) \text{ and } t_{btp} = \bar{y}\exp\left(\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}}\right)$$

$$t_{btr_1} = \bar{y}\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}+2K}\right) \text{ and } t_{btp_1} = \bar{y}\exp\left(\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}+2K}\right).$$

Where K is used for minimizing the MSE. These estimators used information on single auxiliary variable and the authors found them more efficient than mean per unit, ratio type estimators of Cochran [5] and product type estimator of Robson [14] in many practical situations.

Singh *et al.* [34] proposed proportion based exponential ratio and product type estimators when the auxiliary variable is an attribute as T_{sgr} and T_{sgp} respectively. Further, they proposed a class of proportion base exponential ratio-product type estimators, if P is not known as T_{sgrp}

$$T_{sgr} = \bar{y}\exp\left[\frac{P-p}{P+p}\right],\ T_{sgp} = \bar{y}\exp\left[\frac{p-P}{p+P}\right]$$

$$T_{sgrp} = \bar{y}\left[\alpha\exp\left(\frac{P-p}{P+p}\right) + (1-\alpha)\exp\left(\frac{p-P}{p+P}\right)\right].$$

The authors have also proposed two phase sampling version of the proposed exponential type estimators T_{sgr}, T_{sgp} and T_{sgrp} as

$$T_{sgr(d)} = \bar{y}\exp\left[\frac{p'-p}{p'+p}\right], T_{sgp(d)} = \bar{y}\exp\left[\frac{p-p'}{p+p'}\right] \text{ and}$$

$$T_{sgrp(d)} = \bar{y}\left[\alpha_1\exp\left(\frac{p'-p}{p'+p}\right) + (1-\alpha_1)\exp\left(\frac{p-p'}{p+p'}\right)\right] \text{ respectively.}$$

The real constants α and α_1 were determined such that MSE of is minimum. They have derived large sample properties of all the proposed estimators to the first order of approximation and found them efficient in both the sampling schemes theoretically as well as empirically.

Singh *et al.* [20] proposed exponential ratio type estimators of population mean as t_{sgr}, where $a(\neq 0)$, b are either the real numbers or functions of known parameters such as coefficient of variation (C_x), coefficient of kurtosis $(\beta_2(x))$, correlation coefficient (ρ_{xy}) etc.

$$t_{sgr} = \bar{y}\exp\left[\frac{(a\bar{X}+b)-(a\bar{x}-b)}{(a\bar{X}+b)+(a\bar{x}-b)}\right].$$

The known parameters taken as suitable combinations, produced ten estimators t_{sgr_1} to $t_{sgr_{10}}$. Further, combining the estimators t_{sgr_1} and t_{sgr_i} $(i = 2, ..., 10)$, an estimator

$$t^*_{sgr_i} = \alpha t_{sgr_1} + (1-\alpha)t_{sgr_i}$$

is also proposed by the authors, where α is a constant such that the MSE of $t^*_{sgr_i}$ is minimum. The choice of estimators mainly depend upon the available known parameters like $\beta_2(x)$, ρ_{xy}, C_x etc. They have obtained the expressions of Bias and MSE of all the proposed estimators up to first order of approximation, the conditions under which proposed estimators work more efficiently than some existing estimators were also obtained. Finally the authors supported the theoretical findings by numerical illustration which concludes that there was gain in efficiency by the proposed estimators.

Grover and Kaur [76] proposed exponential ratio type estimators of population mean using qualitative auxiliary information under simple random sampling as

$$T_{gkr} = [\alpha\bar{y} + \beta(P-p)]\exp\left[\frac{P-p}{P+p}\right],$$

where α and β are constants whose values are suitably chosen. The authors have derived large sample properties of the proposed estimators and compared with some existing estimators under consideration both theoretically as well as empirically, finally observed the proposed estimators were more efficient.

Using auxiliary information, Solanki *et al.* [68] proposed a class of alternative estimator of

population mean in sample surveys as

$$t_{slr} = \bar{y}\Big[2 - (\frac{\bar{x}}{\bar{X}})^{\alpha} \exp\Big(\frac{\delta(\bar{x}-\bar{X})}{\bar{x}+\bar{X}}\Big)\Big],$$

where the constants (α, δ) used in the estimator are suitably chosen. They have derived the asymptotic expressions of Bias and MSE, also identified the asymptotic optimum estimator from the proposed class of estimators. Further, the authors have made a comparison of the proposed estimators with several other existing estimators and found them efficient.

Sanaullah *et al.* [81] proposed some improved exponential ratio type estimators using two auxiliary variables in case of double sampling as

$$t_{su_1(d)} = \bar{y}_2 \exp\Big[\alpha\left(\frac{\bar{Z}-\bar{z}_2}{\bar{Z}+\bar{z}_2}\right) - (1-\alpha)\left(\frac{\bar{X}-\bar{x}_1}{\bar{X}+\bar{x}_1}\right)\Big],$$

$$t_{su_2(d)} = \bar{y}_2 \exp\left[\sqrt{\rho_{xy}}\left(\frac{\bar{Z}-\bar{z}_2}{\bar{Z}+\bar{z}_2}\right) - (1-\sqrt{\rho_{xy}})\left(\frac{\bar{X}-\bar{x}_1}{\bar{X}+\bar{x}_1}\right)\right],$$

$$t_{su_3(d)} = \bar{y}_2 \exp\Big[\sqrt{\rho_{yz}}\Big(\frac{\bar{Z}-\bar{z}_2}{\bar{Z}+\bar{z}_2}\Big) - (1-\sqrt{\rho_{yz}})\Big(\frac{\bar{X}-\bar{x}_1}{\bar{X}+\bar{x}_1}\Big)\Big],$$

$$t_{su_4(d)} = \bar{y}_2 \exp\left[\sqrt{\rho_{zx}}\left(\frac{\bar{Z}-\bar{z}_2}{\bar{Z}+\bar{z}_2}\right) - (1-\sqrt{\rho_{zx}})\left(\frac{\bar{X}-\bar{x}_1}{\bar{X}+\bar{x}_1}\right)\right],$$

$$t_{su_5(d)} = \bar{y}_2 \exp\Big[\left(\frac{\bar{X}-\bar{x}_1}{\bar{X}+\bar{x}_1}\right) + \exp\left(\frac{\bar{Z}-\bar{z}_2}{\bar{Z}+\bar{z}_2} + \frac{\bar{x}_1-\bar{x}_2}{\bar{x}_1+\bar{x}_2}\right)\Big],$$

where α is a non-zero constant whose value is to be determined such that MSE of $t_{su_1(d)}$ is minimum and $\rho_{ij} > 0$, where $i \neq j$. They have derived the properties of the proposed estimators to the first order of approximation for dependent and independent sample cases and compared them with some existing estimators. They found the proposed estimators efficient both theoretically as well as empirically.

Yadav *et al.* [70] considered the problem of the estimation of finite population mean of the study variable and proposed an estimator which is generalization of the ratio and product type estimators proposed by the authors Bahl and Tuteja [33] as

$$t_{ydg} = \bar{y} \exp\left[c\left(\frac{\bar{X}-\bar{x}}{\bar{X}-\bar{x}}\right)\right],$$

where c is non zero constant. Further, they proposed an almost unbiased exponential estimator of population mean of study variable as

$$t_{ydg}^{(u)} = \bar{y}\Big[\frac{3K+1}{4K}\exp\{2K\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right)\} + \frac{K-1}{2K}\exp\{4K\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right)\} + \frac{1-K}{4K}\exp\{6K\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right)\}\Big].$$

The authors have derived theoretical expressions of Bias and MSE for the proposed estimators to the first order of approximation and found them efficient than the existing estimators under consideration.

Onyeka [35] proposed a class of product type exponential estimators of population mean as t_{onp}. Further a modified ratio-cum product type exponential estimator as t_{onrp} is also proposed by the author in the paper

$$t_{onp} = \bar{y} \exp\Big[\frac{(a\bar{x}+b)-(a\bar{X}+b)}{(a\bar{x}+b)+(a\bar{X}+b)}\Big].$$

$$t_{onrp} = \alpha_1\left(\bar{y}\exp\left[\frac{a(\bar{X}-\bar{x})}{a(\bar{X}+\bar{x})+2b}\right]\right) + \alpha_2\left(\bar{y}\exp\Big[\frac{(a\bar{x}+b)-(a\bar{X}+b)}{(a\bar{x}+b)+(a\bar{X}+b)}\Big]\right)$$

where α_1 and α_2 are weighted fractions such that $\alpha_1 + \alpha_2 = 1$. The scalars $a(\neq 0)$ and b are either real numbers or functions of known values of the auxiliary variable such as coefficient of skewness $(\beta_1(x))$, coefficient of kurtosis $(\beta_2(x))$, standard deviation (σ_x) and the correlation coefficient (ρ_{xy}). In this study, the author has derived theoretical expressions of Bias and MSE of the estimators to the first order of approximation and compared them with some estimators existing in literature both theoretically and empirically. He found both the estimators t_{onp} and t_{onrp} more efficient than the estimators considered for comparision in the study.

Solanki and Singh [73] proposed a class of exponential ratio estimators of population mean

using known population proportion of an auxiliary character highly correlated with the study variable as

$$T_{ssr} = \bar{y} \exp\left(\frac{\alpha(p-\hat{p})}{p+\hat{p}}\right).$$

This proposed class of estimators is observed to be more generall and it includes the usual unbiased estimator and the the one proposed by Singh *et al.* [34]. They proposed two phase sampling version of the proposed class of estimators as

$$T_{ssr(d)} = \bar{y} \exp\left(\frac{\alpha(\hat{p}^{'}-\hat{p})}{\hat{p}^{'}+\hat{p}}\right)$$

in the paper. Regions of preferences for the constant used (α) under which the proposed estimators work efficiently than the existing estimators have been investigated by the authors. The authors have identified asymptotic optimum estimators (AOE) along with their mean square error formula in both simple random sampling and in two phase sampling schemes.

Yadav and Kadilar [22] proposed a family of exponential ratio estimators of population mean as

$$t_{ykr} = k\bar{y} \exp\left[\frac{(a\bar{X}+b)-(a\bar{x}-b)}{(a\bar{X}+b)+(a\bar{x}-b)}\right],$$

where k is a constant chosen suitably and a $(\neq 0)$ & b are either real numbers or may be the functions of known parameters of auxiliary variable, such as coefficient of variation, coefficient of skewness, correlation coefficient etc. They have derived the large sample properties of the proposed estimators to the first order of approximation and made their efficiency comparison with some existing estimators, which shows that every member of the family of estimators t_{ykr} has an improvement over the corresponding family of estimators as proposed by Singh *et al.* [20].

Saini and Kumar [79] suggested modified exponential ratio and product type estimators using information on auxiliary attribute in simple random sampling as

$$T_{skr} = \bar{y} - (e^{A_1} - 1) \text{ and } T_{skp} = \bar{y} - (e^{A_2} - 1) \text{ respectively,}$$

$$\text{where } A_1 = P - \frac{NP-np}{N-n} \text{ and } A_2 = \frac{NP-np}{N-n} - P.$$

The authors have obtained unbiasedness and variance of θ_5 and θ_6 to the first order of approximation. They have derived the conditions under which the proposed estimators work more efficiently than some existing estimators considered. They have also carried a comparative analysis of efficiency between the proposed and some existing estimators both theoretically as well as empirically, which concludes that there is gain in efficiency by the proposed estimators.

Using the information on single auxiliary variable, Shabir *et al.* [82] proposed difference-cum-exponential type estimator for estimating population mean of the variable under study as

$$t_{shd} = \left[\frac{\bar{y}}{2}\{\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) + \exp\left(\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}}\right)\}w_1(\bar{X}-\bar{x}) + w_2\bar{y}\right] \exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right).$$

They have obtained large sample properties of the proposed estimators to the first order of approximation. The authors found that the proposed estimators always perform better than some existing estimators considered regardless of positive or negative correlation between the study and auxiliary variable both theoretically as well as empirically.

Khan *et al.* [36] proposed exponential ratio and product type estimators as

$$t_{khr} = \bar{y} \exp\left(\frac{\sqrt{\bar{X}}-\sqrt{\bar{x}}}{\sqrt{\bar{X}}+\sqrt{\bar{x}}}\right) \text{ and } t_{khp} = \bar{y} \exp\left(\frac{\sqrt{\bar{x}}-\sqrt{\bar{X}}}{\sqrt{\bar{x}}+\sqrt{\bar{X}}}\right) \text{ respectively.}$$

They have also proposed a generalized exponential type estimator for estimating population mean by introducing some real constants (α, h, a) as

$$t_{kh}^{g} = \bar{y} \exp\left[\alpha\left(1 - \frac{a\bar{x}^{\frac{1}{h}}}{\bar{X}^{\frac{1}{h}} + (a-1)\bar{x}^{\frac{1}{h}}}\right)\right],$$

where the constants $\alpha(-\infty < \alpha < +\infty)$ and $h(h > 0)$ are assumed to be known and the constant a $(a \neq 0)$ is estimated such that the proposed estimator has minimum MSE. The authors have derived theoretical expressions of mean MSE and Bias to the first order of approximation

and analyzed the properties for independent units under simple random sampling. Finally, the efficiency comparison of the proposed estimators with some existing estimators was done which clearly depicted that the proposed estimator are most efficient among all other estimators considered for comparison.

Saini and Kumar [79] proposed a new modified exponential type ratio estimator using auxiliary attribute as $T_{skr} = \left[\bar{y} - a(E_1 - 1)\right]$, where a is any constant and

$$E_1 = \exp\left[\frac{NP - np}{N - n}\right].$$

They have obtained the unbiasedness and variance properties to the first order of approximation in the study. The authors worked out the conditions under which T_{skr} will be more efficient than existing estimators. Using different population data sets they authors concluded that the there is gain in efficiency by the proposed estimator.

Haq and Shabir [66] proposed improved estimators of population mean in simple random sampling and in two phase sampling scheme using information on two auxiliary attributes as

$$T_{hs(d)} = w_1\frac{\bar{y}}{4}\left(\frac{P_1}{p_1} + \frac{p_1}{P_1}\right)\left(\frac{P_2}{p_2} + \frac{p_2}{P_2}\right) + w_2(P_1 - p_1) + w_3(P_2 - p_2) \text{ and}$$

$$T^*_{hs(d)} = d_1\frac{\bar{y}}{4}\left(\frac{p_1^*}{p_1} + \frac{p_1}{p_1^*}\right)\left(\frac{P_2}{p_2^*} + \frac{p_2^*}{P_2}\right) + d_2(p_1^* - p_1) + d_3(P_2 - p_2^*) \text{ respectively,}$$

where w_1, w_2, w_3, d_1, d_2 and d_3 are suitably chosen constants. The authors found proposed estimators more efficient than some of the estimators existing in literature both theoretically as well as empirically under under both the sampling schemes.

Khan and Siddiqi [36] proposed an exponential ratio type estimator, a general class of exponential ratio and product type estimators of population mean using variance of an auxiliary variable in single phase sampling as

$$t_{ksr_1} = \bar{y}\exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right), t_{ksr_2} = \bar{y}\exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + (a-1)s_x^2}\right) \text{ and } t_{ksr_3} = \bar{y}\exp\left(\frac{s_x^2 - S_x^2}{S_x^2 + (b-1)s_x^2}\right),$$

where a and b $(a, b \neq 0)$ are unknown constants to be determined such tha MSE of the proposed estimators is minimum. The large sample properties of the proposed estimators have been obtained and analyzed by the authors for independent units under SRSWOR ignoring finite population corrections. They carried an efficiency comparison of the proposed estimators with some existing estimators which revealed that the proposed estimators are efficient than the existing estimators under consideration.

Grover and Kaur [77] proposed a generalized class of ratio type exponential estimators of population mean in SRS under linear transformation of auxiliary variable as

$$t_{gkr} = \left[\omega_1\bar{y} + \omega_2(\bar{X} - \bar{x})\right]\exp\left[\frac{\eta(\bar{X} - \bar{x})}{\eta(\bar{X} + \bar{x}) + 2\lambda}\right],$$

where $\eta(\neq 0)$ and λ are either any known constants or the functions of known population parameters like coefficient of variation, coefficient of skewness, coefficient of kurtosis etc. and the value of constants ω_1 and ω_2 is suitably chosen. The theoretical expressions of Bias and MSE were obtained up to first order. The performance of the proposed estimators is compared with some existing estimators by the authors by conducting a simulation study and observed the proposed estimators efficient.

John and Inyang [64] proposed two exponential ratio estimators of population mean as

$$t_{jir_1} = \theta_1\bar{y} + \theta_2(\bar{X} - \bar{x})\exp\left[\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right] \text{ and } t_{jir_2} = \psi_1\bar{y} + \psi_2(\bar{X} - \bar{x})\exp\left[\frac{2(\bar{X} - \bar{x})}{\bar{X} + \bar{x}}\right],$$

where θ_1, ψ_1 and θ_1, θ_2 are suitably chosen scalars θ_1, $\psi_1 > 0$, $-\infty < \theta_2, \theta_2, \psi_2 < +\infty$. They found the estimators t_1 and t_2 at their optimal conditions always more efficient than some existing estimators including the linear regression estimator which is assumed to be the best linear unbiased estimator in presence of auxiliary variable.

Singh and Pal [83] proposed chain Ratio-Ratio exponential type estimator of population mean as

$$t_{sgcr} = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)\exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right)$$

and a generalized chain Ratio-Ratio-Type exponential estimator as

$$y_{sggcr} = \bar{y}\left(\frac{a\bar{X}+b}{a\bar{x}+b}\right)\exp\left(\frac{a(\bar{X}-\bar{x})}{a(\bar{X}+\bar{x})+2b}\right),$$

where a and b are the real constants or the functions of the parameters such as coefficient of skewness $(\beta_1(x))$, coefficient of variation (C_x), standard deviation (S_x), coefficient of kurtosis $(\beta_2(x))$, quartiles Q_i $(i = 1,2,3)$, deciles (D_i) $(i = 1,2,...,10)$ and $\Delta = \beta_2(x)\ \beta_1(x) - 1$ of the auxiliary variable x, coefficient of kurtosis $(\beta_2(y))$, coefficient of variation (C_y) of the study variable and the correlation coefficient between the auxiliary and study variable (ρ_{xy}). The Bias and MSE expressions of the proposed estimators were derived to the first order of approximation by the authors. They observed proposed estimators were efficient as concluded through the efficiency comparison with some existing estimators.

Saini and Kumar [80] proposed a new modified exponential product type estimator of population mean using bi-serial negative correlation between the study and auxiliary variable as

$$T_{skp} = \Big[\bar{y} - k(m-1)\Big], \text{ where k is a constant and } m = \exp\Big[\frac{NP-np}{N-n} - P\Big].$$

They have obtained expression of Bias and MSE to the first order of approximation and also derived the conditions under which the proposed modified estimator works efficiently than existing estimators. The authors supported theoretical findings of the study by numerical illustrations.

Yasmeen *et al.* [37] proposed the dual to ratio-cum-exponential ratio and dual to product-cum-exponential product type estimators of population mean as

$$t_{ysrr} = \bar{y}\Big[\frac{\bar{x}_t}{\bar{X}}\Big] + \exp\Big[\frac{\bar{z}_t - \bar{Z}}{\bar{z}_t + \bar{Z}}\Big] \text{ and } t_{yspp} = \bar{y}\Big[\frac{\bar{X}}{\bar{x}_t}\Big] + \exp\Big[\frac{\bar{Z} - \bar{z}_t}{\bar{Z} + \bar{z}_t}\Big] \text{ respectively.}$$

These estimators are based on the information from two transformed auxiliary variables using simple random sampling without replacement. The large sample properties of the proposed estimators were obtained to the first order of approximation and the conditions under which the proposed estimators work efficiently than some existing estimators were also obtained in the paper. Further they carried an empirical study for efficiency comparison using different population data sets which supported the claim of gain in efficiency by the proposed class of estimators.

Etuk *et al.* [75] proposed a modified class of ratio estimators for estimating the population mean using a combination of scalars and known population mean of auxiliary variable as

$$t_{etr} = \bar{y}\left[\lambda - r\left(\frac{\bar{x}}{\bar{X}}\right)^{\alpha}\exp\left(\frac{\alpha(\bar{x}-\bar{X})}{\bar{x}+\bar{X}}\right)\right],$$

where λ, r, α and δ are suitably chosen scalars such that λ and r satisfy the condition $\lambda = 1 + r$; $-\infty < r < +\infty$. The large sample properties of the proposed class of estimators were derived up to first order of approximation and some existing members were also identified by them. They found the proposed estimator at their optimal condition superior than some existing estimators both theoretically as well as empirically.

In the absence of auxiliary variable, Subramani [47] proposed median based ratio type estimators of finite population mean as

$$t_{sbr} = \bar{y}\left(\frac{M}{m}\right).$$

The large sample properties and the conditions for which the proposed estimators are more efficient than the considered estimators were derived by the author. He found proposed estimator more efficient than some existing estimators making use of auxiliary variable and the linear regression estimator which is considered as the best linear unbiased estimator in the presence of auxiliary variable.

Vishwakarma *et al.* [23] proposed exponential ratio and product type estimators as

$$t_{vkr} = \alpha\bar{y} + (1-\alpha)\bar{y}\exp\left[\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right] \text{ and } t_{vkp} = \beta\bar{y} + (1-\beta)\bar{y}\exp\left[\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}}\right] \text{ respectively,}$$

where α and β are the constants to be obtained so that the MSE of t_3 and t_4 are minimized. The expressions for MSE and Bias of the proposed estimators were obtained to the first order of approximation. On assigning suitable values to the constants α and β, they found the proposed

estimators reducing to the estimators $\bar{y}$, ratio and product type estimators of Bahl and Tuteja [33].

Kadilar [65] proposed a new exponential type estimator in SRS for population mean as

$$t_{kdr} = \bar{y}\Big[\frac{\bar{x}}{\bar{X}}\Big]^{\alpha} \exp\Big(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\Big).$$

This estimator at its optimal condition is equal efficient as that of usual linear regression estimator. The expressions of Bias, MSE and the optimal value of α were obtained. They found the proposed estimator better than the mentioned existing estimators both theoretically as well as empirically and therefore the author recommended the proposed estimator for its practical use when the auxiliary information is available.

Enang *et al.* [74] proposed a class of exponential ratio-product type estimator of population mean as

$$t_{egpr} = \bar{y}\Big[\alpha_1 \frac{\bar{x}}{\bar{X}} \exp\Big(\frac{\delta_1(\bar{X}-\bar{x})}{\bar{X}+\bar{x}}\Big) + \alpha_2 \frac{\bar{X}}{\bar{x}} \exp\Big(\frac{\delta_2(\bar{x}-\bar{X})}{\bar{X}+\bar{x}}\Big)\Big],$$

where $\alpha_1 + \alpha_2 = 1$ while δ_1 and δ_2 are suitably chosen scalars. They found the proposed class of estimators at its optimal condition equal efficient as that of linear regression estimator.

Yadav *et al.* [38] proposed median based exponential estimator of population mean using population and sample meadian of study variable as

$$t_{ydr} = \bar{y} \exp\left(\frac{M-m}{M+(a-1)m}\right),$$

where the constant a is is determined such that MSE of proposed estimator is minimum. The expressions of Bias and MSE were obtained. It was found that the proposed estimator have minimum MSE among all other competing estimators considered and therefore is most efficient.

Enang *et al.* [74] proposed a modified class of exponential ratio-product type estimators for estimating the population mean in simple random sampling (SRS) as

$$t_{enrp} = \bar{y}\big[\alpha_1 \frac{\bar{x}}{\bar{X}} \exp\left(\frac{\delta_1(\bar{x}-\bar{X})}{\bar{x}+\bar{X}}\right) + \alpha_2 \frac{\bar{X}}{\bar{x}} \exp\left(\frac{\delta_2(\bar{x}-\bar{X})}{\bar{x}+\bar{X}}\right)\big],$$

where $\alpha_1 + \alpha_2 = 1$ and δ_1 & δ_2 are suitably chosen constants. The authors have derived the bias and MSE expressions for the proposed estimators to the first order of approximation. The proposed class of estimators were found efficient than the estimators considered for comparison.

Yadav *et al.* [38] proposed two new exponential ratio type estimators for estimating the finite population mean using sample and population median of study variable as

$$t_{ydr_1} = \bar{y}\left(\frac{m}{M}\right) \exp\big(\frac{a(M-m)}{M+m}\big) \text{ and } t_{ydr_2} = \bar{y}\left(\frac{m}{M}\right) \exp\left(\frac{M-m}{a(M+m)}\right),$$

where the arbitrary constant a is determined such that the MSE of the proposed estimators is minimum. The large sample properties of the proposed estimators were derived to the first order of approximation and compared with some existing estimators. The said comparison concluded that the proposed estimators are most efficient among all other estimators considered.

Panigrahi and Mishra [78] proposed some modified exponential ratio estimators of population mean as

$$t_{pmr_1} = \frac{\bar{y}}{1+\theta_1 C_y^2} \exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right),$$

if C_y^2 is known. Further, if the coefficient of variation of the auxiliary variable (C_y^2) is not known, they proposed the estimators

$$t_{pmr_2} = \frac{\bar{y}}{1+\theta_1 \hat{C}_y^2} \exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) \text{ and } t_{pmr_3} = (1+\theta_1 \hat{C}_y^2) \exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right),$$

where $\hat{C}_y^2 = \frac{S_y^2}{\bar{y}^2}$. The large sample properties of all the proposed estimators were derived to the first order of approximation and compared with the estimators considered in the study. They observed from both theoretical and empirical study that the proposed estimators are efficient.

Panigrahi and Mishra [78] proposed modified exponential product type estimators of population mean when the population coefficient of variation of y is known in advance as t_{pmp_1}.

Further in the absence of coefficient of variation, they proposed a product type estimator as t_{pmp_2}

$$t_{pmp_1} = \frac{\bar{y}}{1+\theta_1 C_y^2} \exp\left[\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}}\right].$$

$$t_{pmp_2} = \frac{\bar{y}}{1+\theta_1 \hat{C}_y^2} \exp\left[\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}}\right].$$

The theoretical expressions of Bias and MSE were obtained by the authors to the first order of approximation. The numerical efficiency comparison conducted by the authors concluded that the proposed estimators are most efficient among all other estimators considered.

Singh *et al.* [67] noted that nearly sixty five existing estimators are the members of the proposed estimator

$$t_{sg} = \left[\bar{y} + \hat{\beta}(\bar{X}-\bar{x})\right]\left(\frac{a\bar{X}+b}{a\bar{x}+b}\right)$$

and identified some one hundred sixty three unknown members of t. They have also proposed a general class of exponential ratio type estimators of population mean

$$t_{sggr} = \left[\bar{y} + \hat{\beta}(\bar{X}-\bar{x})\right]\left(\frac{a(\bar{X}-\bar{x})}{a(\bar{X}+\bar{x})+2b}\right)$$

and identified some members of t_{sggr}, where $\hat{\beta}$ is the regression coefficient of y on x, a and b are real constants or the functions of population parameters such as population standard deviation (S_x), total (X), variance (S_x^2), coefficient of skewness $(\beta_1(x))$, coefficient of kurtosis $(\beta_2(x))$, qartiles, median, midrange, mode, deciles, tri-mean, Δ, correlation coefficient (ρ), coefficient of variation (C_x) etc. Further, the authors have derived the large sample properties of the proposed estimators and observed the proposed class of exponential ratio type estimators t_{sggr} provides efficient estimators which are less biased than the some estimators considered in the paper.

Kumar *et al.* [69] proposed a class of generalized exponential ratio type estimators for estimating the population mean using known median (M) of study variable as

$$t_{kmr} = \bar{y}\left[\alpha + (1-\alpha)\exp\left(\frac{M-m}{M+m}\right)\right],$$

where α is a constant obtained such that the MSE of the proposed estimator is minimum. The Bias and MSE expressions of the proposed estimators are obtained by the authors to the first order of approximation. The proposed estimators were found efficient than the estimators making use of auxiliary information which is collected on additional cost of the survey.

Singh *et al.* [26] proposed generalized exponential ratio estimator under simple random sampling using auxiliary variable as

$$t_{sggr} = \bar{y}\, A^{\left[\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right]} \text{ for A } > 0.$$

They found the estimator of Bahl and Tuteja [33] a particular case of t_{sggr} for A = e = 2.71 ... . The authors have derived theoretical expressions of Bias and MSE to the first order of approximation and proved that the proposed estimators have highest efficiency among the other estimators considered in the study both theoretically as well as empirically.

Using the information on auxiliary variable as attribute, Zaman & Kadilar (2019) [39] proposed a family of exponential ratio type estimators as

$$T_{zkr} = \bar{y}\exp\left[\frac{(kP+l)-(kp+l)}{(kP+l)-(kp+l)}\right],$$

where the constants $(k \neq 0)$ are either real numbers or some functions of the known auxiliary attribute such as C_p, $\beta_2(\phi)$ and ρ_{pb}. They used the suitable combinations of the constants l and k which have generated various types of estimators. The authors have obtained large sample properties of the proposed estimators to the first order of approximation and compared these properties with the existing estimators considered concluding the proposed estimators are most efficient.

Zaman [72] proposed a class of exponential ratio type estimators using information on population proportion which possesses a certain attribute as

$$T_{zmr} = \bar{y}\left(\frac{p}{P}\right)^{\alpha}\exp\left[\frac{(kP+l)-(kp+l)}{(kP+l)+(kp+l)}\right]; i = 1,2,...,10,$$

where k and l are the constants like $0, 1,\ C_p,\ \beta_2(\phi)$ & ρ_{pb} taken in suitable combinations. The author has derived theoretical expressions of Bias and MSE up to first order for the proposed estimators and found the proposed estimators efficient than the estimators under consideration through numerical illustration.

Hussain *et al.* [71] proposed dual type exponential estimators of population mean using auxiliary information as

$$t_{de1} = \bar{y}\left[\alpha \exp\left(\frac{\bar{X}-\bar{x}}{p\bar{X}}\right) + (1-\alpha)\exp\left(\frac{\bar{x}-\bar{X}}{p\bar{X}}\right)\right] \text{ and}$$

$$t_{de2} = \bar{y}\left[\beta \exp\left(\frac{\bar{X}-\bar{x}}{q\bar{x}}\right) + (1-\beta)\exp\left(\frac{\bar{x}-\bar{X}}{q\bar{x}}\right)\right].$$

Where p and q are non zero constants whose value is chosen such that the estimators t_{de1} and t_{de2} should be unbiased. The value of constants α and β are chosen such that the MSE of t_{de1} and t_{de2} is minimum.

Hussain [46] proposed proportion based unbiased exponential ratio-cum-product estimators in double sampling plan as

$$T_{rp1} = \bar{y}\left[\omega_1 \exp\left(\frac{p_1-p}{lp_1}\right) + (1-\omega_1)\exp\left(\frac{p-p_1}{lp_1}\right)\right] \text{ and}$$

$$T_{rp2} = \bar{y}\left[\omega_2 \exp\left(\frac{p_1-p}{gp}\right) + (1-\omega_2)\exp\left(\frac{p-p_1}{gp}\right)\right].$$

Where $l(\neq 0)$, $g(\neq 0)$, ω_1 and ω_2 are constants. The values of l and g are chosen such that the proposed estimators are unbiased and the values of ω_1 & ω_2 are chosen such that MSE (T_{rpi}) $(i = 1, 2)$ is minimum.

Conclusion

In conclusion, exponential type estimators play a pivotal role in the accurate estimation of the population mean. The utilization of exponential type transcendental functions in these estimators proves to be an effective strategy for achieving precise and reliable estimates. By harnessing the mathematical properties of exponential functions, these estimators significantly reduce bias and variance, enhancing the robustness of the population mean estimates. This approach not only outperforms traditional estimation methods but also adapts well to various data distributions, making it a valuable tool in statistical analysis. The findings of this research underscore the importance and efficacy of exponential type estimators in achieving precise population mean estimates, highlighting their potential for broad application in statistical practices.

References

[1] Bowely, A. L. (1926). Measurement of precision attained in sampling. *Bulletin of the International Statistical Institute*, 22: 1–6.

[2] Neyman, J. (1934). On the two different aspects of the representative method : the method of stratified sampling and the method of purposive selection. *Journal of Royal Statistical Society*, 97: 558–606.

[3] Watson, D. J. (1937). The estimation of leaf areas. *Journal of the Agricultural Science*, 27: 474–483.

[4] Neyman, J. (1938). Contribution in the theory of sampling human population. *Journal of the American Statistical Association*, 33: 111–116.

[5] Cochran, W. G. (1940). The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce. *The Journal of Agricultural Science*, 30: 262–275.

[6] Hansen, M. H. and Hurwitz, W. N. (1943). On the theory of sampling from finite populations. *The Annals of Mathematical Statistics*, 14 (4): 333–362.

[7] Searls, D. T. (1964). The utilization of known coefficient of variation in the estimation procedure. *Journal of American Statistical Association*, 59: 1225–1226.
[8] Thompson, J. R. (1968). Some shrinkage techniques for estimating mean. *Journal of the American Statistical Association*, 63: 113–121.
[9] Singh, M. P. (1969). Comparison of some ratio cum product estimators. *Sankhya*, 31 (B): 375–378.
[10] Ray, S. K. and Singh, R. K. (1979). A generalized ratio product difference estimator in double sampling. *Journal of the Indian Society of Agricultural Statistics*, 31 (1): 85–86.
[11] Hirano, K., (1972). Using some approximately known coefficient of variation in estimating mean. Research Memorandum, 49.
[12] Singh, R. and Singh, D. (1977). On the inclusion of large units in the sample. *Journal of the Indian Society of Agricultural Statistics*, 29 (2): 24–31.
[13] Maiti, P. and Tripathi, T. P. (1980). Estimation of Population total using information on some units of population. Stat. Math. Tech. Rep. No. 8/80, ISI, Calcutta.
[14] Robson, D. S. (1957). Applications of multivariate polykays to the theory of unbiased ratio type estimation. *Journal of the American Statistical Association*, 52: 511–522.
[15] Murthy, M. N. (1964). Product method of estimation. *Sankhya*, 26: 69–74.
[16] Singh, P., Jha, M. P. and Bathla, H. V. L. (1980). A note on modified ratio estimator. *Journal of Indian Society of Agricultural Statistics*, 32: 99–102.
[17] Das, A. K. (1980). Contributions to the theory of sampling stretagies using auxiliary information. Ph. D. Thesis, Submitted to Indian Statistical Institute, Calcuta.
[18] Sisodia, B. V. S. and Dwivedi, V. K. A. (1981). Modified ratio estimators using coefficient of variation of auxiliary variable. *Journal of Indian Society of Agricultural Statistics*, 33: 13–18.
[19] Sharma, M. and Bhatnagar, S. (2008). A general class of estimators for estimation of population mean through auxiliary information. *Journal of the Indian Statistical Association*, 46 (2): 155–161.
[20] Singh, R., Chauhan, P., Sawan, N., and Smarandache, F. (2009). Improvement in estimating the population mean using exponential estimators in simple random sampling. *Bulletin of Statistics and Economics*, 3 (A09): 13–18.
[21] Subramani, J. and Kumarapandiyan, G. (2012). Modified ratio estimators using known median and co-efficient of kurtosis. *American Journal of Mathematics and Statistics*, 2 (4): 95–100.
[22] Yadav, S. K. and Kadilar, C. (2013). Improved class of ratio and product estimators. *Applied Mathematics and Computation*, 219: 10726–10731.
[23] Vishwakarma, G. K., Singh, R., Gupta, P. C., and Pareek, S. (2016). Improved ratio product type estimators of finite population mean in simple random sampling. *Revista Investigacion Operacional*, 36 (3): 70–76.
[24] Misra, S., Kumar, B., Yadav, S. K., Yadav, D. K., and Shukla, A. K. (2017). An improved ratio type predictive estimator for estimating finite population mean using auxiliary information. *International Journal of Engineering Sciences and Research Technology*, 6 (6): 524–530.
[25] Shalabh and Tsai, J. R. (2017). Ratio and product methods of estimation of population mean in the presence of correlated measurement errors. *Communications in Statistics-Simulation and Computation*, 46 (7): 5566–5593.
[26] Singh, P, Bouza, C. and Singh, R. (2019). Generalized exponential estimator for estimating the population mean using auxiliary variable. *Journal of Scientific Research*, 63: 273–280.
[27] Mehta, V., Singh, H. P. and Pal, S. K. (2020). A general procedure for estimating finite population mean using ranked set sampling. *Investigacion Operacional*, 41 (1): 80–92.
[28] Jhajj, H. S., Sharma, M. K. and Grover, L. K. (2006). A family of estimators of population mean using information on auxiliary attribute. *Pakistan Journal of Statistics*, 22 (1): 43.
[29] Shabbir, J. and Gupta, S. (2007). On estimating the finite population mean with known population proportion of an auxiliary variable. *Pakistan Journal of Statistics*, 23 (1): 1.

[30] Naik, V. D. and Gupta, P. C. (1996). A note on estimation of mean with known population proportion of an auxiliary character. *Journal of the Indian Society of Agricultural Statistics*, 48 (2): 151–158.

[31] Shabbir, J. and Gupta, S. (2010). Estimation of the finite population mean in two phase sampling when auxiliary variables are attributes. *Hacettepe Journal of Mathematics and Statistics*, 39 (1): 121–129.

[32] Abd-elfattah, A. M., El-Sherpieny, E. A., Mohamed, S. M. and Abdouf, O. F. (2010). Improvement in estimating the population mean in simple random sampling using information on auxiliary attribute. *Applied Mathematics and Computation*, 215 (12): 4198–4202.

[33] Bahl, S. and Tuteja, R. (1991). Ratio and product type exponential estimators. *Journal of Information and Optimization Sciences*, 12 (1): 159–164.

[34] Singh, R., Chauhan, P., Sawan, N., and Smarandache, F. (2007). Ratio-product type exponential estimator for estimating finite population mean using information on auxiliary attribute. *Renaissance High press USA*, 1: 18–32.

[35] Onyeka, A. C. (2013). A class of product-type exponential estimators of the population mean in simple random sampling scheme. *Statistics in Transition-New Series*, 14 (2): 189–200.

[36] Khan, H., Sanaullah, A., Khan, M. A. and Siddiqi, A. F. (2014). Improved exponential ratio type estimators for estimating population mean regarding full information in survey sampling. *World Applied Science journal*, 26 (5): 1897–1902.

[37] Yasmeen, U., Noor ul Amin, M., and Hanif, M. (2016). Exponential ratio and product type estimators of population mean. *Journal of Statistics and Management System*, 19 (1): 55–71.

[38] Yadav, DK., Kumar, R., Misra, S., and Yadav, S. K. (2017). Estimating population mean using known median of the study variable. *International Journal of Engineering Sciences and Research Technology*, 6 (7): 15–21.

[39] Zaman, T. and Kadilar, C. (2019). Novel family of exponential estimators using information of auxiliary attribute. *Journal of Statistics and Management Systems*, 22 (8): 1499–1509.

[40] Hussain, S., Sharma, M., & Bhat, M. I. J. (2021). Optimum exponential ratio type estimators for estimating the population mean. *Journal of Statistics Applications and Probability Letters*, 8 (2): 73–82.

[41] Hussain, S., Sharma, M., and Chandra, H. (2021). Modified Exponential Product Type Estimators for Estimating Population Mean Using Auxiliary Information. *Special Proceedings 23rd Annual Conference*. 49-58.

[42] Hussain, S., Sharma, M., Chandra, H. and Ali, V. (2022). Proportion Based Exponential Ratio Estimators of Population Mean in Case of Single and Double Sampling Plans. *Investigacion Operacional*, 43 (1): 1–19.

[43] Hussain, S. and Sharma, M., 2023. New Robust Based Ratio Type Estimators of Population Mean Using Transcendental Function. *International Journal of Agricultural & Statistical Sciences*, 19(1): 421-426.

[44] Hussain, S. and Bhat, V.A., 2023. New median based almost unbiased exponential type ratio estimators in the absence of auxiliary variable. *Reliability: Theory & Applications*, 18(1 (72)): 242-249.

[45] Hussain, S., Sharma, M., Bhat, M. I. J., and Bhat, V. A. (2024). Proportion Based Dual Unbiased Exponential Type Estimators of Population Mean. *Thailand Statistician*. 22(1): 31-39.

[46] Hussain, S. (2024). Unbiased Exponential Type Estimators of Population Mean Using Auxiliary Variable as an Attribute In Double Sampling. *Reliability: Theory & Applications*. 19(4(80)): 364-373.

[47] Subramani, J. (2016). A new median based ratio estimator for estimation of the finite population mean. *Statistics in Transition New Series*, 17 (4): 1–14.

[48] Bose, C. (1943). Note on sampling errors in the method of double sampling. *Sankhya*, 6: 329–330.

[49] Sukhatme, B. V. and Koshal, R. S. (1959). A contribution to double sampling, *Journal of Indian Society of Agricultural Statistics*, 9: 128–144.

[50] Tikkiwal, B. D. (1960). On the theory of classical regression and double sampling estimation. *Journal of Royal Statistical Society B*, 22: 99–106.
[51] Sukhatme, B. V. (1962). Generalised Hartley-Ross unbiased ratio-type estimator. *Nature*, 196 (48): 1231–1238.
[52] Rao, J. N. K. and Pereira, N. P. (1968). On double ratio estimators. *Sankhya A*, 30: 83–90.
[53] Rao, P. S. R. S. (1972). On two phase regression estimator. *Sankhya A*, 33: 473–476.
[54] Rao, P. S. R. S. (1975). On two phase ratio estimator in infinite population. *Journal of the American Statistical Association*, 70: 839–845.
[55] Singh, G. N., Upadhyaya, L. N. and Singh, H. P. (2000). On the estimation of population mean in two-phase sampling. *Journal of Statistical Research*, 30 (91): 25–32.
[56] Ahmed, M. S. (2001). A note on the performance of a regression-type estimator in two-phase sampling using two auxiliary variables. *Pakistan Journal of Statistics-All Series*, 17 (2): 151–156.
[57] Hossain, M. I., Rahman, M. S. and Ahmed, M. S. (2003). Second order properties of some estimators under double sampling. *Statistics in Transition*, 6 (4): 543–554.
[58] Jhajj, H. S., Sharma, M. K., and Grover, L. K. (2004). A sampling design for estimation of population mean under double sampling. *Journal of Mathematics and Mathematical Sciences*, 9 (13): 1–12.
[59] Upadhyaya, L. N., Singh, H. P., and Tailor, R. (2006). Estimation of mean with known coefficient of variation of an auxiliary variables in two phase sampling. *Statistics Transition*, 7 (6): 1327–1344.
[60] Samiuddin, M. and Hanif, M. (2007). Estimation of population mean in single and two phase sampling with or without additional information. *Pakistan Journal of Statistics*, 23 (2): 99–118.
[61] Kamal, A. and Shahbaz, M. Q. (2008). New difference estimator in two phase sampling using arbitrary probabilities. Journal of Statistics, 15 (1): 7–16.
[62] Hanif, M., Hamad, N. and Shahbaz, M. Q. (2010). Some new regression type estimators in two phase sampling. *World Applied Sciences Journal*, 8 (7): 799–803.
[63] Singh, P., Bouza, C and Singh, R. (2006). Generalized exponential estimators for estimating the population mean using auxiliary variable. *Journal of Scientific Research*, 63: 273–280.
[64] John, E. E. and Inyang, E. E. (2015). Efficient exponential ratio estimators for estimating the population mean in simple random sampling. *Hacettepe Journal of Mathematics and Statistics*, 44 (3): 689–705.
[65] Kadilar, G. O. (2016). A new exponential type estimator of population mean in simple random sampling. *Journal of Modern Applied Statistical Methods*, 15 (2): 207–214.
[66] Haq, A. and Shabbir, J. (2014). An improved estimator of finite population mean when using two auxiliary attributes. *Applied Mathematics and Computation*, 241: 14–24.
[67] Singh, H. P. and Yadav, A. (2018). A new exponential approach for reducing the mean squared errors of the estimators of population mean using conventional and non-conventional location parameters. *Journal of Modern Applied Statistical Methods*, 18 (1): 26–63.
[68] Solanki, R. S., Singh, H. P. and Rathour, A. (2012). An alternative estimator for estimating the finite population mean using auxiliary information in sample surveys. International Scholarly Research Notices.
[69] Kumar, G., Bharti, P., Kumar, S., and Yadav, S. K. (2018). General exponential ratio type estimator of population mean under simple random sampling. *International Journal of Creative Research Thoughts*, 6 (1): 450–455.
[70] Yadav, R., Upadhayaya, L. N., Singh, H. P., and Chatterjee, S. (2012). Almost unbiased ratio and product type exponential estimators. *Statistics in Transition-New Series*, 13 (3): 537–550.
[71] Hussain, S., Sharma, M., Kumar, B. and Bhat, V.A., Almost Unbiased Dual Exponential Type Estimators of Population Mean Using Auxiliary Information. *Statistics and Applications*, 20(2):147-156.
[72] Zaman, T. (2020). Generalized exponential estimators for the finite population mean. *Statistics in Transition-New Series*, 21 (1): 159–168.
[73] Solanki, R. S. and Singh, H. P. (2013). Improved estimation of population mean using population proportion of an auxiliary character. Chilean journal of Statistics, 4(1):3-17.

[74] Enang, E. I., Uket, J. O. and Ekpenyong, E. J. (2017). A modified class of exponential-type estimator of population mean in simple random sampling. *International Journal of Advanced Statistics and Probability*, 5 (2): 70–76.

[75] Etuk, S. I., Enang, E. I. and Ekpenyong, E. J. (2016). A modified class of ratio estimators for population mean. *Communications in Statistics-Theory and Methods*, 45 (4): 837–849.

[76] Grover, L. K. and Kaur, P. (2011). An improved exponential estimator of finite population mean in simple random sampling using an auxiliary attribute. *Applied Mathematics and Computation*, 218 (7): 3093–3099.

[77] Grover, L. K. and Kaur, P. (2014). A generalized class of ratio type exponential estimators of population mean under linear transformation of auxiliary variable. *Communications in Statistics-Simulation and Computation*, 43 (7): 1552–1574.

[78] Panigrahi, A. and Mishra, G. (2017). Some modified exponential ratio type estimators of finite population mean in survey sampling. *International Journal for Research in Applied Science and Engineering Technology*, 5 (IX): 1511–1514.

[79] Saini, M. and Kumar, A. (2014). A new modified exponential type ratio estimator with auxiliary attribute and simple random sampling. *World Applied Sciences Journal*, 32 (8): 1536–1539.

[80] Saini, M. and Kumar, A. (2015). Exponential type product estimator for finite population mean with information on auxiliary attribute. *Applications and Applied Mathematics - An International Journal*, 10 (1): 106–113.

[81] Sanaullah, A., Khan, H., Ali, H. A. and Singh, R. (2012). Improved exponential ratio type estimators in survey sampling. *Journal of Reliability and Statistical Studies*, 119–132.

[82] Shabbir, J., Haq, A. and Gupta, S. (2014). A new difference-cum-exponential type estimator of finite population mean in simple random sampling. *Revista Colombiana de Estadistics*, 37 (1): 199–211.

[83] Singh, H. P. and Pal, S. K. (2015). A new chain ratio-ratio-type exponential estimator using auxiliary information in sample surveys. *International Journal of Mathematics and its Applications*, 3 (4): 37–46.

A COMPARATIVE ANALYSIS OF GAMMA AND WEIBULL DISTRIBUTIONS IN TAMIL CINEMA DATA

A. Vanathu Suresh[1], R. Subramani[1*] and S Vijayan[2]

•

[1]Department of Journalism and Mass Communication, Periyar University, Salem-11, Tamil Nadu.
[2]Department of Statistics, Periyar University, Salem-11, Tamil Nadu.
ajrsuresh@periyaruniversity.ac.in, erasubramani@gmail.com, vijaystatstician1210@gmail.com

Abstract

This paper aims to systematically investigate the utility of the Gamma and Weibull distributions, focusing on their application to lifetime datasets and clarifying their mathematical and statistical properties. By analyzing lifetime data across various disciplines, the research emphasizes the effectiveness and flexibility of these distributions in capturing the complexities inherent in such data. It underscores the importance of parameters such as standard error, log-likelihood, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion (HQIC) in value estimation. The findings suggest that both distributions provide valuable insights into the underlying data, with practical implications for reliability engineering and failure analysis. Moreover, the study demonstrates that the Weibull distribution offers a better fit to the given data than the Gamma distribution due to its adaptability, which yields superior results. A key contribution of this study is the proposal of a model based on estimating the Conditional Weibull distribution for feature parameters, which accurately predicts a finite mixture of two-parameter Weibull distributions initially verified on datasets.

Keywords: Gamma distribution, Weibull distribution, Probability density function, Cumulative density function.

I. Introduction

The intersection of statistical modelling and cinema studies has witnessed significant growth in recent decades, particularly as the global film industry becomes more data-driven. Quantitative analysis of film performance, audience behaviour, and release strategies has proven essential for both academic inquiry and practical decision-making in film marketing and distribution. Within this analytical framework, probability distributions such as the Weibull and Gamma have emerged as powerful tools for modelling a range of film-related phenomena, including the duration of theatrical runs, box office earnings, audience arrival patterns, and long-tail revenue distributions. While many of these statistical approaches have been applied in Hollywood and other mainstream contexts, their relevance to regional industries such as Tamil cinema also known as Kollywood has not yet been fully explored.

Tamil cinema, one of India's most prolific film industries, produces over 200 films annually and commands a significant audience both domestically and among the global Tamil diaspora. Its box office dynamics are influenced by multiple factors: star power, regional festivals, political sentiment, music releases, and audience word-of-mouth. Consequently, modelling audience behaviour and

film performance in this context requires flexible, non-linear statistical tools capable of accounting for skewness, rapid changes in viewership, and the long-lasting influence of successful films. Several scholars have begun exploring these avenues in broader Indian cinema studies, laying the groundwork for region-specific applications.

The Weibull distribution, originally developed for modelling mechanical failure times, has been effectively adapted for survival analysis in the film industry. Sawhney and Eliashberg (5) used Weibull models to understand the attrition of audience interest in films, arguing that hazard functions can reflect how quickly viewers lose interest after release. The flexibility of the Weibull distribution through its shape parameter allows for modelling increasing or decreasing hazard rates, which is highly relevant in Tamil cinema, where films often experience rapid shifts in viewership based on opening-weekend performance or critical reviews. In this regional context, Weibull modeling can also accommodate phenomena such as sudden drops in attendance following poor reception, or gradual build-ups driven by word-of-mouth for lower-budget films.

The Gamma distribution, widely used in queuing theory and reliability analysis, has also found application in film revenue and audience flow modeling. Radas and Shugan (4) used the Gamma distribution to describe revenue accumulation patterns over time, emphasizing its suitability for capturing skewed, right-tailed behavior in box office earnings. Given the non-linear nature of Tamil film revenue where films may peak in the first week or gain traction in the second or third Gamma distributions are capable of capturing both fast and slow revenue build-up. This is particularly pertinent for mid-budget Tamil films, which may not have strong openings but gain popularity due to critical acclaim or viral social media traction.

The Exponential distribution, a special case of the Weibull and Gamma distributions, is particularly noted for its simplicity and memoryless property. Bhat (1) highlighted its relevance in modeling the inter-arrival times of customers and product life cycles. In film studies, it has been used to describe rapid audience decay especially for films that do not perform well on release. In Tamil cinema, this is a common pattern for films that rely heavily on initial star-driven hype but lack critical or audience support beyond the opening weekend. The Exponential model, though limited in flexibility, offers a baseline for understanding abrupt drops in attendance and can serve as a comparative tool for evaluating the fit of more complex models like Weibull or Gamma.

The Lognormal distribution has been prominently featured in modeling film revenues, particularly in identifying long-tail effects where a small number of films account for a disproportionately large share of earnings. De Vany and Walls (2) demonstrated that the distribution of film revenues follows a lognormal pattern with power-law tails, reflecting extreme inequality in success rates. This finding aligns with trends in Tamil cinema, where a few films often led by top-tier stars dominate the box office, while many others fail to recoup investments. The heavy tail of the lognormal distribution captures this disparity and offers insights into risk management and investment strategies in the regional film market.

This study uses probabilistic models like the Weibull, Gamma, Exponential, and Lognormal distributions have been widely applied in global and national film industries, there remains a notable lack of focused research on their relevance to Tamil cinema. Existing studies have largely concentrated on Hollywood and Bollywood contexts, overlooking the distinct structural and cultural features of the Tamil film industry such as festival-timed releases (e.g., Pongal and Diwali), the urban-rural audience divide, and actor-driven fan dynamics. These unique characteristics significantly influence film lifecycles, audience behavior, and revenue patterns, thereby warranting a region-specific analytical approach. This study seeks to address this gap by systematically applying these distributions to Tamil cinema in order to model theatrical run durations, revenue decay, and long-tail earning trends. By grounding this analysis in established statistical literature

and adapting it to the Tamil context, the research aims to contribute meaningfully to both quantitative film analysis and South Indian media studies.

II. Methods

2.1 Derivation of Gamma and Weibull Distributions

In this section, a new comparison model cumulative distribution function (CDF) and the probability density function (PDF) of the Gamma and Weibull Distribution families are given as

$$\mathrm{f}(x;\alpha,\omega) = f(x) = \begin{cases} \frac{-x^{\alpha-1}e^{-x/\omega}}{\omega^{\alpha}\,\Gamma(\alpha)}, & x > 0 \\ 0, & x \le 0 \end{cases} \tag{1}$$

where $\alpha > 0$ is the scale parameter, $\omega > 0$ is the shape parameter, and $\Gamma(\alpha)$ is the Gamma function. The cumulative distribution function is

$$F(x;\ \alpha,\omega) = f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)}\gamma\left(\alpha,\frac{x}{\omega}\right), & x \ge 0 \\ 0, & x < 0 \end{cases} \tag{2}$$

Where $\gamma\ (\alpha, x)$ is the lower incomplete Gamma function.

$$F(x;\ \alpha,\mathrm{k}) = f(x) = \begin{cases} 1 - e^{-(x/\alpha)\omega}, & x \ge 0 \\ 0, & x < 0 \end{cases} \tag{3}$$

Where $\alpha > 0$ is the scale parameter, and $\omega > 0$ is the shape parameter.
The cumulative distribution function is given by

$$f(x;\ \alpha,k) = f(x) = \begin{cases} \frac{\omega}{\alpha}\left(\frac{x}{\alpha}\right)^{k-1} e^{-(x/\alpha)^{\omega}}, & x \ge 0 \\ 0, & x < 0 \end{cases} \tag{4}$$

The CDF provides the probability that a random variable X is less than or equal to a specific value x. At the same time, the PDF describes the relative likelihood of the random variable taking on a particular value.

The Bayesian Information Criterion is

$$BIC\ =\ kln(n)\ -\ 2ln\ (L) \tag{5}$$

k is the number of parameters in the model, n is the number of observations, and L is the maximum likelihood of the model.

The Akaike Information Criterion is given by

$$AIC\ =\ 2k\ -\ 2ln\ (L) \tag{6}$$

Hannan-Quinn Information Criterion is

$$HQIC = 1\ln(Lmax) + 2kln(\ln\ (n)) \tag{7}$$

k is the number of estimated parameters in the model, and L is its maximum likelihood function.

The gamma distribution model was applied to the dataset observations to gain insights into its statistical characteristics. The model's shape and scale parameters were accurately estimated, and the negative log-likelihood value indicated a good fit for the data. These statistical measures suggest that the Gamma distribution effectively captures the underlying data distribution, providing a solid foundation for further analysis and interpretation.

The model provided significant insights into the statistical properties of the analyzed dataset. Fitted to observations, the model's shape and scale parameters were estimated with corresponding standard errors, reflecting its fitting quality. These parameters and statistical measures demonstrate the Weibull model's effectiveness in capturing the underlying distribution characteristics, providing a robust tool for further analysis and clarification.

Overall, the Weibull distribution fits the dataset better than the Gamma distribution. The Wei bull model's closer AIC, BIC, and HQIC values suggest consistency and reliability in model fit evaluation. For various combinations of α and ω, we generated sample sizes from the Gamma and Weibull models, specifically for the parameter values $\alpha = 2.3$, $\omega = 0.6$, and $\alpha = 1.7$, $\omega = 2.2$. The value decreases as the sample size increases, as shown in Table 1. Table 1 illustrates that our Gamma and Weibull models better fit the dataset. The dataset represents survival times in years from a Tamil cinema's content rating, 329 samples.

III. Results

In this section, we conduct a relative study to assess the Maximum Likelihood Estimation (MLE) performance for Gamma and Weibull distributions. Using the quantile function, we generate n=329 samples in years of Tamil Cinema content rating samples drawn from those distributions. Subsequently, we calculate the MLEs, loglikelihood, standard error, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion (HQIC).

Table 1: *The model's performance.*

Model	Parameter	Estimated values	Standard Error	Loglikelihood	AIC	BIC	HQIC
Gamma	α	2.3009	0.1860	-638.7781	1291.556	1298.957	1295.312
	ω	0.6742	0.0668				
Weibull	α	1.7902	0.0916	-408.4005	820.8009	828.2018	824.654
	ω	2.2780	0.0957				

Table 1 shows the outcomes corresponding to the Gamma distribution by utilizing it to model Tamil cinema content rating datasets. The parameters indicate that the Gamma distribution, with its shape and scale, captures the data distribution with a specific level of skewness and spread. The AIC, BIC, and HQIC values support the model's fit, though the Weibull distribution, with its parameters, shows a potentially better fit given its lower AIC, BIC, and HQIC values. This suggests the Weibull model might be more effective at capturing the underlying distribution characteristics of the dataset.

IV. Application

We compare its performance in providing a robust parametric fit to the datasets with that of the Weibull distribution. Metrics such as the log-likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) are employed for this comparison. The loglikelihood, AIC, and BIC values are computed for the proposed Weibull model for comparison to discern the most suitable model. The model exhibiting the lowest loglikelihood, AIC, BIC, and HQIC values is deemed the most appropriate match for the provided datasets. The R software is employed for this analytical endeavor, simplifying the necessary calculations and comparisons.

The dataset comprises the time of the year (in years) measurements from three hundred twenty-nine (329) samples in content rating at the Tamil cinema. It consists of the following values:

7.9,8.06,7.66,7.2,7.83,8.7,7.4,7.93,7.1,7,8.36,7.73,8.63,6,5.83,6.56,7.3,7.43,8,7.83,7.53,6.9,6.33,6.7,7.23,8.06,8.533,6.66,6.9,6.9,6.96,6.73,7.36,7.76,.03,6.83,7.76,7.93,6.5,8,8.13,7.46,7.86,5.86,6.53,7.63,8.5,8,6.93,7.76,8.26,8.36,7.6,6.53,6.43,8.23,7.6,7.6,7.26,7.56,7.5,6.56,8.5,7.16,6.66,8.1,6.66,8.1,6.66,7,8.2,7.46,7.8,7.33,7.96,6.7,7.06,7.16,7.46,6.23,5.56,6.33,7.76,8.1,7.13,5,6.4,7.96,8.03,7.26,8.16,7.46,7.4,7.1,8.33,7.7,7.6,8.03,8.66,7.76,7.86,7.86,6.8,5.2,6.9,7.73,6.86,7.16,6.26,6.06,7.43,7.8,6.93,7.13,6.76,6.63,7.63,7.53,7.9,6.2,7.36,6.76,7.06,7.03,7.36,6.8,7.26,7.76,6.76,7.9,4.4,5.43,7.33,6.73,7.2,7.23,6.6,7.26,7.63,7.03,7.9,7.4,6.36,7.1,6.2,6.86,5.43,6.2,6.63,5.43,6.2,6.63,4.83,7.83,6.73,7.1,7.7,6.1,7.86,7.53,7,6.16,7.13,5.93,7.23,6.53,8.2,8.56,7.03,7.43,7.16,7.4,7.93,6.53,7.1,7,7.36,7.16,5.43,7.33,7.3,5.83,7.36,6.6,6.7,8.23,7.8,6.8,6,7.6,5.63,6.43,6,4.3,5.36,6.73,7.26,6.23,6.46,5.9,7.86,5.9,7.86,6.53,7.3,5.76,7.03,6.5,7.73,5.86,5.86,6.43,6.13,6.7,7.36,6.76,7.5,6.2,8,6.86,6.7,6.93,8,7.76,7.76,7.1,8.1,7.93,6.4,8.16,7.46,8.23,8.13,8.2,8.26,7.76,7.53,7.16,7.66,5.83,7.46,6.63,8.36,7.,6.2,5.96,5.1,5.06,4.43,7,8.16,7.76,7.76,7.9,7.5,6.4,6.56,5.9,6.36,7.03,6.73,6.7,6.96,8.2,8.33,7.1,7.26,6.93,7.23,7.23,7.13,8.5,5.86,8.53,7.46,8.63,4.73,5.56,6.8,5.13,8.63,4.73,5.56,6.8,5.13,8.63,7.16,6.73,7,8.1,7.23,8.7,5.9,7.7,7.2,6.33,8.76,7.3,6.86,6.13,7.36,7.6,8.36,8.7,8,5.26,6.33,7.7,6.7,6.7,6.7,6.83,6.73,7.5,6.73,7.36,7.5,7.73,7.83,7.8,7.53,7.43,7.43,6.83,6,6.2,7.4,6.76,7.96,6.9,7.1,5.93,4.93,7.76,7.9.

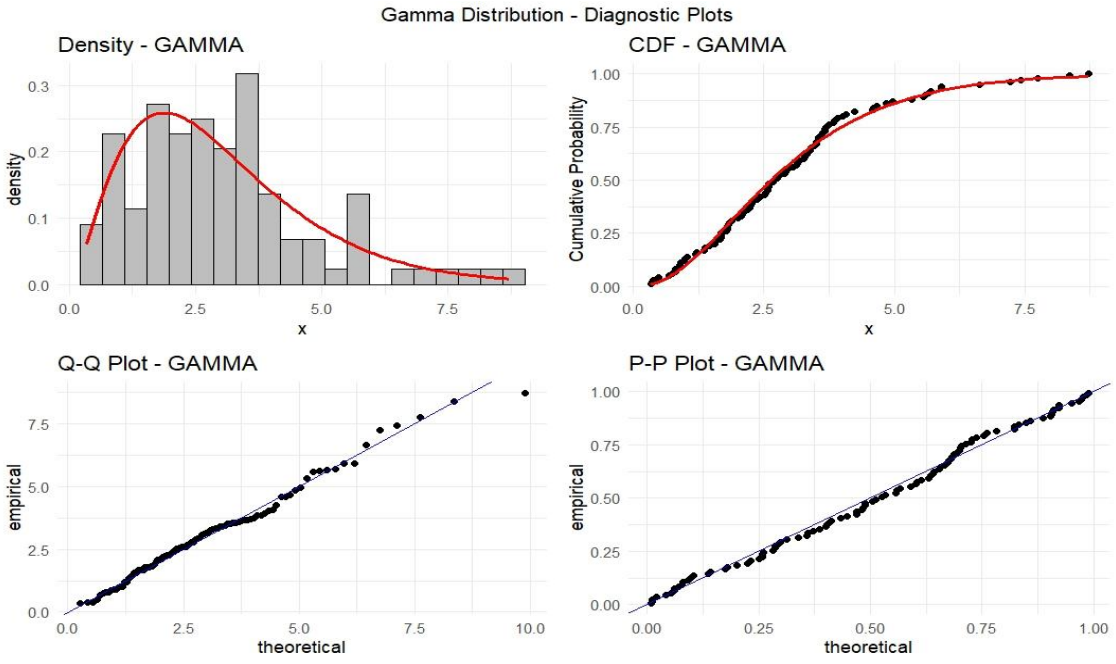


Figure 1: *Gamma distribution of the Density plots datasets.*

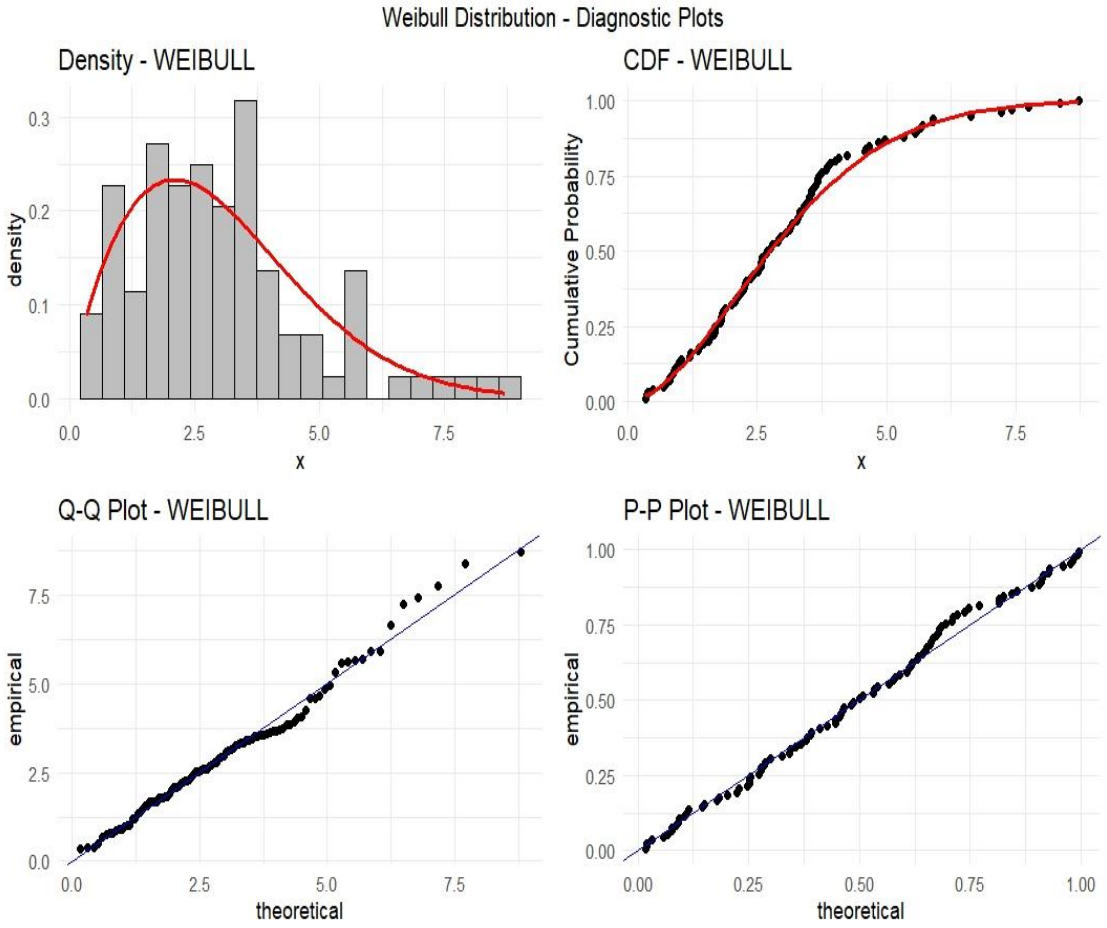


Figure 2: *Weibull distribution of the Density plots datasets.*

Table 1 presents the planned model's superior ability to effectively fit the highly skewed datasets compared to the competing models, as specified by the evaluation metrics employed. Figures 1 and 2 show that the proposed model fits the data set adequately.

V. Conclusion

In conclusion, this study highlights the effectiveness of statistical distribution modelling specifically the Gamma and Weibull distributions in analysing and interpreting data from the Tamil cinema industry. Our findings indicate that while both distributions are useful for modelling lifetime and content rating data, the Weibull distribution consistently outperforms the Gamma distribution due to its greater flexibility and adaptability. By leveraging key statistical indicators such as standard error, log-likelihood, AIC, BIC, and HQIC, we demonstrated that the Weibull model provides a more accurate representation of the complex and varied success patterns in Tamil cinema. This approach not only enhances our understanding of film longevity and audience reception trends but also offers practical applications for producers, distributors, and media analysts seeking data-driven insights into release strategies, content classification, and performance forecasting. The study also introduces a Conditional Weibull-based model that further improves the accuracy of predictions, establishing a robust framework for future research and industry analytics in regional cinema.

References

[1] Bhat, U. N. (2010). An introduction to queueing theory: Modeling and analysis in applications (2nd ed.). *Birkhäuser*.

[2] De Vany, A., and Walls, W. D. (2004). Motion picture profit, the stable Paretian hypothesis, and the curse of the superstar. *Journal of Economic Dynamics and Control, 28(6),*1035–1057.

[3] Jain, P., and Yadav, S. (2016). Predicting movie success: A comparative analysis of Bollywood and Hollywood. *International Journal of Business and Management Invention*, 5(3),25–32.

[4] Radas, S., and Shugan, S. M. (1998). Seasonal marketing and timing new product introductions. *Journal of Marketing Research*, 35(3), 296–315.

[5] Sawhney, M. S., and Eliashberg, J. (1996). A parsimonious model for forecasting gross box-office revenues of motion pictures. *Marketing Science*, 15(2),113–131.

[6] Gopinath, S., and Manohar, R. (2018). Modeling small-budget films in Tamil cinema: A probabilistic approach. *International Journal of Economics, Commerce and Management*, 6(7), 10–25.

[7] Kumar, S. (2021). Exponential decay in Tamil film audience patterns: A statistical approach. *International Journal of Film Studies*, 13(4), 79–87.

[8] Pillai, R., and Das, S. (2020). Cumulative revenue forecasting for Tamil films using Gamma distribution. *Asian Journal of Statistical Modeling*, 14(2), 85–92.

[9] Ramesh, T., Rajendran, R., and Subramanian, S. (2021). Predictive analytics in Tamil cinema: A machine learning approach. *Journal of Media Economics*, 34(1),22–45.

[10] Srinivasan, M. (2018). Weibull modeling for film box office success: Case study of Tamil cinema. *Journal of South Asian Cinema*, 9(1), 45–58.

[11] Srinivasan, V. (2019). Marketing and revenue outcomes in Tamil cinema: The impact of festival releases. *Journal of Media and Marketing*, 16(2), 29–40.

[12] Subramanian, M. (2020). The data gap in Tamil cinema: Challenges for statistical analysis. *Asian Film Review*, 17(3),19–31.

[13] Vasudevan, S. (2021). A statistical analysis of regional cinema: Case study of Tamil film releases. *Indian Cinema Review*, 28(4), 61–74.

RT&A, No 2 (84)
Volume 20, June 2025

LEAST SQUARES, MAXIMUM LIKELIHOOD, AND ANDERSON-DARLING TYPES ESTIMATORS FOR ENTROPY TRANSFORMED EXPONENTIAL DISTRIBUTION

Ikwuoche John DAVID[1]*, Stephen MATHEW[2], Isah Charles SAIDU[3]

•

[1,2]Department of Statistics, Federal University Wukari, Nigeria.
[3]Department of Computer Science, Baze University, Nigeria.
*Coresponding Authour: davidij@fuwukari.edu.ng.
[2]matsteve231@gmail.com.
[3]charles.isah@bazeuniversity.edu.ng.

Abstract

Aim: This research aim is to investigate the performance stability and determine the best estimator among Least Squares (LS), Maximum Likelihood (ML), and Anderson Darling (AD) type (right tail AD, left tail AD, second order right tail AD, and second order left tail AD), estimators for estimating the parameter of the Entropy Transformed Exponential distribution. Methods: A Monte Carlo simulation was conducted to evaluate the finite sample performance stability and efficiency at sample sizes of 20, 100, 250, 700, 1000, 1500, 2500, *and* 3000. *The evaluation measures used for assessment are the Absolute Bias (AB) and Root Mean Square Error (RMSE) as performance metrics. Real life data analysis was performed using a field intensity data measured in microvolts recorded at 5-second intervals. The data is a positive right skewed data and highly peaked. Results: The simulation results demonstrated that the LS method consistently outperformed the ML and AD type estimators. This is based on the lowest values of AB and RMSE obtained across all the eight sample sizes. Additionally, the findings were validated through real world data, where the LS estimator exhibited the lowest standard error and the narrowest confidence interval length. Based on the results, the LS is ranked 1st, right tail AD ranked 2nd, second order right tail AD ranked 3rd, MLE ranked 4th, left tail AD ranked 5th, and second order left tail AD ranked 6th. However, results from t-test in determining the significance of the parameter estimate from each estimator, shows that MLE and second order left tail AD estimates are not significant at* 5%. *The graphical representation of the field intensity data by each estimators fit clearly depict the LS estimator best describe the data distribution. Conclusion: Based on these findings, the study concludes that the LS estimator is the most effective method for estimating the parameter of the Entropy Transformed Exponential distribution followed by the right tail AD, second order right tail AD, MLE, left tail AD, and second order left tail AD being the least effective estimator.*

Keywords: Entropy Transformed Exponential distribution (EnTrExD), Least Squares (LS), Maximum Likelihood (ML), Anderson-Darling (AD), Estimators

I. INTRODUCTION

In the discussion of parameter estimation, the Maximum Likelihood (ML) estimator remains one of the most widely utilized techniques due to its simplicity and unbiased properties. While it may exhibit limitations when applied to small sample sizes, its overall effectiveness and versatility have solidified its prominence in various applications. Beyond the ML approach, numerous alternative estimation methods have been proposed in the literature, each with unique advantages and limitations. Comparative studies have also been conducted to evaluate the performance of these techniques across different models. Results from these studies indicate that the effectiveness of a given estimation method can vary depending on the specific characteristics of the model, with certain techniques demonstrating superior performance under particular conditions.

According to Almathkour *et al*. [2], the choice of an estimation method depends on the application area and performance standards. They argue that users may opt for a minimum variance unbiased estimator (MVUE) despite the absence of a closed-form expression, as it provides minimum variance among unbiased estimators. Similarly, Sanku Dey and Park [10] recommended using maximum likelihood (ML) or Bayes estimators for parameter estimation in the weighted exponential distribution, based on their comparative analysis of various estimation methods. Additionally, David *et al*. [3] proposed the estimated generalized least square technique with residual maximum likelihood estimation (EGLS-REML) for variance components in BertalanffyRichards split-plot design models (SPDM). Their study demonstrated that EGLS-REML outperformed ordinary least square (OLS) and EGLS-ML estimators for parameter estimation in SPDMs.

Mazucheli [9] conducted a comparative analysis of the L-moments and maximum likelihood estimation (MLE) techniques for the complementary Beta distribution using simulations and annual temperature extremes. Their simulation results indicated that the ML estimator exhibited lower bias and mean square error (MSE) compared to L-moments. However, they observed that ML estimates tended to be negatively biased, whereas L-moments estimates were positively biased. Despite these differences, both estimators were found to be asymptotically unbiased and consistent. For small sample sizes, they concluded that the ML estimator is more appropriate due to its superior performance. Almathkour *et al*. [2] compared eight estimation techniques—maximum likelihood estimation (MLE), ordinary least squares (OLS), weighted least squares (WLS), method of moments (MoM), maximum product of spacing (MPS), percentiles estimator (PCE), *L*2 distance estimator (L2DE), and Kullback-Leibler divergence of survival function (KLDSF)—using Monte Carlo simulations. Their findings revealed that MLE and L2DE produced the smallest mean square error (MSE) across all sample sizes, while PCE and KLDSF yielded the largest MSE. MPS and PCE provided the smallest absolute bias, with MoM, MPS, and PCE exhibiting negative bias, while the other estimators showed positive bias. Despite these differences, all estimators were asymptotically unbiased. However, based on the application of real-life data, the WLS estimator outperformed the other techniques. David *et al*. [4] found that the estimated generalized least squares with residual maximum likelihood (EGLS-REML) estimator outperformed OLS and EGLS-MLE in parameter estimation for a three parameter Weibull split-plot design model (SPDM). This conclusion was based on Median Adequacy Measures (MAM), Akaike Information Criterion (AIC), Corrected AIC, Bayesian Information Criterion (BIC), and Standard Error of Estimates (SEE). Similarly, David *et al*. [5, 6] demonstrated the superiority of EGLS-REML over OLS and EGLS-MLE for Johnson-Schumacher and Chapman-Richards SPDMs, based on the same performance criteria.

In this study, three estimation methods are compared: MLE, least squares (LS), and Anderson Darling (AD)-based estimators, which include right-tail AD (ADR), right-tail second-order AD (AD2R), left-tail AD (ADL), and left-tail second-order AD (AD2L). The comparison is conducted through Monte Carlo simulation and applied to real-life data on field intensity to identify the best-

performing method for estimating the one-parameter Entropy Transformed Exponential (ETE) distribution introduced by [8].

II. ENTROPY TRANSFORMED EXPONENTIAL DISTRIBUTION

A continuous variable C is said to follow ETE distribution with scale parameter, α and shape parameter, φ. It is denoted as ETE ~ (C, φ) if its probability density function (PDF), and cumulative density function (CDF) are given by [8] as

$$f(c) = \varphi^2 c e^{-\varphi c} \tag{1}$$

and

$$F(c) = e^{-\varphi c}\left(e^{-\varphi c} - 1 - \varphi c\right) \tag{2}$$

The quantile function (q_c) of the ETE distribution was derived as

$$q_c = \frac{\log\left(\frac{W(k)e^{-1}+1}{k}\right)}{2\varphi} \tag{3}$$

The CDF and PDF plot of the model are presented in Figure 1. The plots indicated that the probability model is a right tailed distribution, thus will perform better on data that is right-skewed.

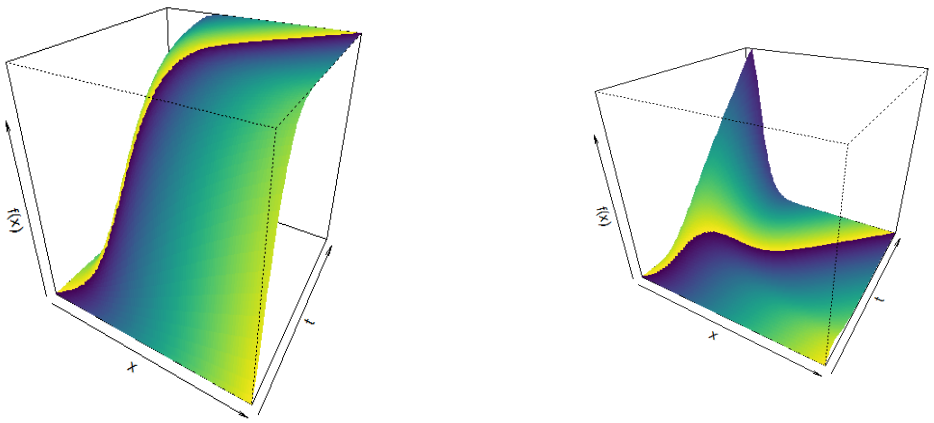


Figure 1: *Left to Right is the CDF and PDF of ETE Distribution*

III. METHODS OF ESTIMATION

In this section of the research, three estimation techniques are considered and presented for estimating the unknown parameter φ, for the ETE distribution. The techniques as mentioned earlier are LS, ML, and AD type's Right-tail AD, Left-tail AD, Right-tail AD of second degree and Left-tail AD of second degree of second degree (AD2) (ADR, ADL, AD2R, and AD2L) estimators.

I. The Least Squares Method (LS_m)

The LS_m estimator of the unknown parameter φ is obtained by minimizing the following equations with respect to φ:

$$LS_m = \sum_{i=1}^{n}\left[F(c_i) - \frac{i}{(n+1)} \right]^2 \tag{4}$$

$$LS_m = \sum_{i=1}^{n}\left((e^{-\varphi c_i}(e^{-\varphi c_i} - 1 - \varphi c_i)) - \frac{i}{(n+1)} \right)^2 . \tag{5}$$

Partially differentiating Equation (5) w.r.t φ and equate to zero gives

$$\frac{\partial}{\partial\varphi} = \sum_{i=1}^{n}\left((e^{-\varphi c_i}(e^{-\varphi c_i} - 1 - \varphi c_i)) - \frac{i}{(n+1)} \right)^2$$

$$2\sum_{i=1}^{n}\left((e^{-\varphi c_i}(e^{-\varphi c_i} - 1 - \varphi c_i)) - \frac{i}{(n+1)} \right)\left(c_i e^{-2\varphi c_i}(c_i \varphi e^{\varphi c_i} - 2) \right) = 0$$

$$\sum_{i=1}^{n}\left(e^{-2\varphi c_i} - e^{-\varphi c_i} - \varphi c_i e^{-\varphi c_i} - \frac{i}{(n+1)} \right)\left(c_i^2 \varphi e^{-\varphi c_i} - 2c_i e^{-2\varphi c_i} \right) = 0 \tag{6}$$

As shown in Equation (6), the LS_m estimator for the parameter of ETE distribution, does not exist in a closed form. Numerical estimation method using R software is applied in obtaining φ.

II. Maximum Likelihood Estimation (MLE) Method

To estimate the parameter of the ETE distribution, using the maximum likelihood method, let $c_1, c_2, ..., c_n$ be independent identically distributed of size (n) with probability function ETE distribution, then the likelihood function $L(\phi)$ for ETE distribution is given as

$$L(\phi) = \prod_{i=1}^{n} f(c) = \prod_{i=1}^{n} \varphi^2 c_i e^{-\varphi c_i} = \varphi^{2n} e^{-\varphi\sum_{i=1}^{n} c_i} \sum_{i=1}^{n} c_i \tag{7}$$

$$\log\left(L(\phi)\right) = \log\left(\varphi^{2n} e^{-\varphi\sum_{i=1}^{n} c} \sum_{i=1}^{n} c \right) = 2n\log(\varphi) - \varphi\sum_{i=1}^{n} c + \sum_{i=1}^{n}\log(c) \tag{8}$$

The likelihood in Equation (8) is partially differentiated w.r.t φ and equating to zero gives

$$\frac{\partial(\log(L(\phi))}{\partial\varphi} = \frac{\partial(2n\log(\varphi))}{\partial\varphi} - \frac{\partial\left(\varphi\sum_{i=1}^{n} c \right)}{\partial\varphi} + \frac{\partial\left(\sum_{i=1}^{n}\log(c) \right)}{\partial\varphi} = 0 \tag{9}$$

$$\frac{(2n)}{\varphi} - \sum_{i=1}^{n} c = 0; \Rightarrow \varphi = 2n\left(\sum_{i=1}^{n} c \right)^{-1} \tag{10}$$

Therefore, the MLE for the parameter of ETE is given as;

$$\varphi = 2n\left(\sum_{i=1}^{n} c \right)^{-1} \tag{11}$$

III. The Minimum Distance Estimator (Anderson-Darling Method)

This subsection focuses on the Anderson-Darling (AD) type estimators, a class of minimum distance estimators used to measure the discrepancy between the empirical distribution and the true (theoretical) distribution [1]. The AD test is a modification of the Kolmogorov-Smirnov test, and it assigns greater weight to the tails of the distribution, making it particularly useful for assessing the fit of a distribution to data with significant deviations at the extremes. The AD type estimators come in different variants, namely ADR (Right-tail AD), ADL (Left-tail AD), AD2R (Second-order Right-tail AD), and AD2L (Second order Left-tail AD).

- The ADR estimator gives more weight to the right tail of the distribution.

- The ADL estimator gives greater emphasis to the left tail.
- The AD2R estimator assigns even larger weight to the right tail.
- The AD2L estimator places more emphasis on the left tail.

Derivation Notations: In the following, the set up framework for estimating the parameter φ using AD methods. Therefore, the following notations are introduced:

- $F_j = F(c_j)$: The CDF of the ETE distribution evaluated at the point c_j.
- $\hat{F}_j = 1-F(c_j)$: The complement of the CDF representing the probability that the random variable is greater than c_j or the survival function evaluated at c_j.

For ADR:

$$ADR = \int_0^{\infty} \frac{\left[F_n(c)-F(c)\right]^2}{1-F(c)}$$

$$= \frac{n}{2} - 2\sum_{j=1}^{n} F_j - \frac{1}{n}\sum_{j=1}^{n}(2j-1)\log\left(\bar{F}_{n+1-j}\right)$$

$$= \frac{n}{2} - 2\sum_{j=1}^{n}\left[e^{-\varphi c_j}\left(e^{-\varphi c_j}-1-\varphi c_j\right)\right]$$

$$-\frac{1}{n}\sum_{j=1}^{n}(2j-1)\log\left(1-e^{-\varphi c_{n+1-j}}\left(e^{-\varphi c_{n+1-j}}-1-\varphi c_{n+1-j}\right)\right) \tag{12}$$

Minimizing the function (12) with respect to φ and equating to zero gives the ADR estimator $\hat{\varphi}$;

$$\frac{\partial}{\partial\varphi} = \left(4\sum_{j=1}^{n} c_j e^{-2\varphi c_j} - 2\sum_{j=1}^{n} c_j \varphi e^{-2\varphi c_j}\right) + \frac{1}{n}\sum_{j=1}^{n}\frac{(2j-1)c_{n+1-j}(c_{n+1-j}\varphi e^{\varphi c_{n+1-j}}-2)}{e^{2\varphi c_{n+1-j}}+(\varphi c_{n+1-j}+1)e^{\varphi c_{n+1-j}}-1} = 0 \tag{13}$$

For ADL:

$$ADL = \int_0^{\infty}\frac{\left[F_n(c)-F(c)\right]^2}{F(c)}dc = -\frac{3n}{2} - 2\sum_{j=1}^{n}F_j - \frac{1}{n}\sum_{j=1}^{n}(2j-1)\log(F_j)$$

$$= -\frac{3n}{2} - 2\sum_{j=1}^{n}\left[e^{-\varphi c_j}\left(e^{-\varphi c_j}-1-\varphi c_j\right)\right] - \frac{1}{n}\sum_{j=1}^{n}(2j-1)\log\left[e^{-\varphi c_j}\left(e^{-\varphi c_j}-1-\varphi c_j\right)\right] \tag{14}$$

Minimizing the function (14) with respect to φ and equating to zero gives the ADL estimator $\hat{\varphi}$;

$$\frac{\partial}{\partial\varphi} = \left(4\sum_{j=1}^{n} c_j e^{-2\varphi c_j} - 2\sum_{j=1}^{n} c_j^2 \varphi e^{-\varphi c_j}\right) + \frac{1}{n}\sum_{j=1}^{n}\frac{(2j-1)c_j(c_j\varphi e^{\varphi c_j}-2)}{(\varphi c_j+1)e^{\varphi c_j}-1} = 0 \tag{15}$$

For AD2R:

$$AD2R = \int_0^{\infty}\frac{[F_n(c)-F(c)]^2}{[1-F(c)]^2}dc = 2\sum_{j=1}^{n}\log[\bar{F}_j] + \frac{1}{n}\sum_{i=1}^{n}\frac{(2j-1)}{\bar{F}_{n+1-j}}$$

$$= 2\sum_{j=1}^{n}\log\left(1-e^{-\varphi c_j}\left(e^{-\varphi c_j}-1-\varphi c_j\right)\right) + \frac{1}{n}\sum_{i=1}^{n}\left|\frac{(2j-1)}{1-e^{-\varphi c_{n+1-j}}\left(e^{-\varphi c_{n+1-j}}-1-\varphi c_{n+1-j}\right)}\right| \tag{16}$$

Minimizing the function (16) with respect to φ and equating to zero gives the AD2R estimator $\hat{\varphi}$;

$$\frac{1}{n}\sum_{j=1}^{n}\frac{(2j-1)c_{n+1-j}(c_{n+1-j}\varphi e^{\varphi c_{n+1-j}}-2)e^{2\varphi c_{n+1-j}}}{\left(e^{2\varphi c_{n+1-j}}+(\varphi c_{n+1-j}+1)e^{\varphi c_{n+1-j}}-1\right)^2} - 2\sum_{j=1}^{n}\frac{c_j(c_j\varphi e^{\varphi c_j}-2)}{e^{2\varphi c_j}+(\varphi c_j+1)e^{\varphi c_j}-1} = 0 \tag{17}$$

For AD2L:

$$AD2L = \int_0^{\infty}\frac{\left[F_n(c)-F(c)\right]^2}{\left[F(c)\right]^2}dc = 2\sum_{j=1}^{n}\log\left(F_j\right) + \frac{1}{n}\sum_{j=1}^{n}\frac{(2j-1)}{F_j}$$

$$= 2\sum_{j=1}^{n}\log\left(e^{-\varphi c_j}\left(e^{-\varphi c_j}-1-\varphi c_j\right)\right)+\frac{1}{n}\sum_{j=1}^{n}\left[\frac{(2j-1)}{e^{-\varphi c_j}\left(e^{-\varphi c_j}-1-\varphi c_j\right)}\right] \tag{18}$$

Minimizing the function (18) with respect to φ and equating to zero gives the AD2L estimator $\hat{\varphi}$;

$$\frac{1}{n}\sum_{j=1}^{n}\left[\frac{(2j-1)c_j(\varphi c_j e^{\varphi c_j}-2)e^{2\varphi c_j}}{[(c_j+1)e^{\varphi c_j}-1]^2}\right]-2\sum_{j=1}^{n}\frac{c_j(\varphi c_j e^{\varphi c_j}-2)}{(c_j+1)e^{\varphi c_j}-1}=0\,. \tag{19}$$

The parameter φ for the respective variants of the AD estimators is then obtained numerically by solving the homogeneous equations in (13), (15), (17), and (19) to obtain $\hat{\varphi}$.

To estimate distribution parameters using a minimum distance estimator in R software, the *mge* (method of generalized estimation) method must be specified in the fitdist function call. Additionally, the appropriate goodness-of-fit (GOF) distance argument should be included [7].

IV. ANALYSIS

The simulation and real data fitting for the various estimation methods are presented in subsections I and IV, respectively. Furthermore, the results from both the simulation and real-world data fittings are compared with plots illustrating the real data alongside the fitted data for the different estimation methods presented.

I. Simulation Study (Monte Carlo)

In this section, an extensive Monte Carlo simulation study was performed to compare the performances of the ML, LS, ADR, ADL, AD2R, and AD2L estimators of the parameter φ of the ETE distribution. All simulations were conducted in R environment for 1000 Monte-Carlo runs. In this study, finite sample sizes of n = 20, 100, 250, 700, 1000, 1500, 2500, 3000 are considered. Throughout the simulations, the parameter φ is assumed to be known with values 0.01, 0.04 and 0.06. The performances of the different estimators are assessed and compared based on the Absolute Bias (AB_s) and Root Mean Squares Error (RMSE) criteria. The expressions for AB_s and RMSE are given in Equations (20) and (21), respectively;

$$AB_s=\frac{\sum_{j=1}^{1000}\left|\hat{\varphi}_j-\varphi\right|}{1000} \tag{20}$$

$$RMSE=\sqrt{\frac{\sum_{j=1}^{1000}\left(\hat{\varphi}_j-\varphi\right)^2}{1000}} \tag{21}$$

where, the j^{th} estimate of φ is the simulated estimate of the parameter of interest while φ is the true parameter value. Also, "Minimum" is defined as the smallest AB_s and RMSE for the estimated parameter φ.

Table 1: *Comparison of Anderson Darling estimates (ADR, AD2R, ADL and AD2L)*

		ADR		AD2		ADL		AD2L	
Sizes (N)	True P	ABs	RMSE	ABs	RMSE	ABs	RMSE	ABs	RMSE
20	0.01	0.033	0.033	**0.026**	**0.026**	0.038	0.038	0.039	0.039
	0.04	0.134	0.134	**0.105**	**0.105**	0.151	0.151	6.452	6.452
	0.06	0.200	0.200	**0.158**	**0.158**	0.226	0.226	17.978	17.978
100	0.01	0.028	0.028	**0.018**	**0.018**	0.030	0.030	0.028	0.028
	0.04	0.114	0.114	**0.071**	**0.071**	0.118	0.118	0.113	0.113
	0.06	0.171	0.171	**0.106**	**0.106**	0.178	0.178	10.563	10.563
250	0.01	0.030	0.030	**0.021**	**0.021**	0.030	0.030	0.028	0.028
	0.04	0.120	0.120	**0.086**	**0.086**	0.120	0.120	0.113	0.113
	0.06	0.180	0.180	**0.128**	**0.128**	0.180	0.180	0.169	0.169
700	0.01	0.028	0.028	**0.022**	**0.022**	0.030	0.030	0.029	0.029
	0.04	0.114	0.114	**0.088**	**0.088**	0.119	0.119	0.114	0.114
	0.06	0.171	0.171	**0.132**	**0.132**	0.179	0.179	0.171	0.171
1000	0.01	0.028	0.028	**0.022**	**0.022**	0.030	0.030	0.028	0.028
	0.04	0.113	0.113	**0.087**	**0.087**	0.119	0.119	0.113	0.113
	0.06	0.169	0.169	**0.131**	**0.131**	0.178	0.178	0.169	0.169
1500	0.01	0.028	0.028	**0.022**	**0.022**	0.030	0.030	0.028	0.028
	0.04	0.112	0.112	**0.089**	**0.089**	0.118	0.118	0.112	0.112
	0.06	0.168	0.168	**0.133**	**0.133**	0.177	0.177	0.167	0.167
2500	0.01	0.028	0.028	**0.023**	**0.023**	0.030	0.030	0.028	0.028
	0.04	0.112	0.112	**0.091**	**0.091**	0.118	0.118	0.113	0.113
	0.06	0.168	0.168	**0.137**	**0.137**	0.178	0.178	10.500	10.500
3000	0.01	0.028	0.028	**0.023**	**0.023**	0.030	0.030	0.028	0.028
	0.04	0.113	0.113	**0.093**	**0.093**	0.119	0.119	0.113	0.113
	0.06	0.169	0.169	**0.139**	**0.139**	0.179	0.179	0.170	0.170

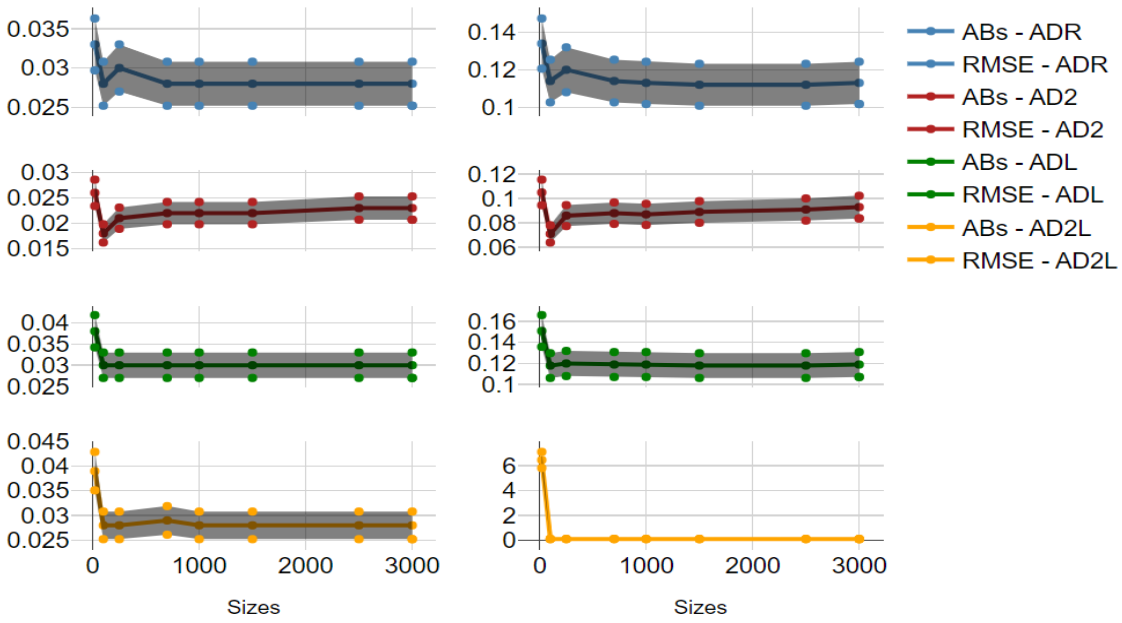


Figure 2a: *ABs and RMSE visualizations for ADR, AD2, ADL and AD2L methods*

Table 2: *Comparison of LS and ML Estimates*

		LS		MLE		Minimum	
Sizes (N)	**True P**	**ABs**	**RMSE**	**ABs**	**RMSE**	**ABs**	**RMSE**
20	0.01	**0.0099**	**0.0099**	0.0274	0.0285	0.0099	0.0099
	0.04	**0.0397**	**0.0397**	0.1096	0.1140	0.0397	0.0397
	0.06	**0.0595**	**0.0595**	0.1644	0.1709	0.0595	0.0595
100	0.01	**0.0087**	**0.0087**	0.0263	0.0265	0.0087	0.0087
	0.04	**0.0349**	**0.0349**	0.1054	0.1061	0.0349	0.0349
	0.06	**0.0525**	**0.0525**	0.1581	0.1592	0.0525	0.0525
250	0.01	**0.0080**	**0.0080**	0.0263	0.0263	0.0080	0.0080
	0.04	**0.0319**	**0.0319**	0.1050	0.1054	0.0319	0.0319
	0.06	**0.0479**	**0.0479**	0.1576	0.1580	0.0479	0.0479
700	0.01	**0.0074**	**0.0074**	0.0260	0.0261	0.0074	0.0074
	0.04	**0.0295**	**0.0295**	0.1041	0.1042	0.0295	0.0295
	0.06	**0.0443**	**0.0443**	0.1562	0.1563	0.0443	0.0443
1000	0.01	**0.0070**	**0.0070**	0.0260	0.0260	0.0070	0.0070
	0.04	**0.0282**	**0.0282**	0.1039	0.1039	0.0282	0.0282
	0.06	**0.0422**	**0.0422**	0.1558	0.1559	0.0422	0.0422
1500	0.01	**0.0067**	**0.0067**	0.0259	0.0260	0.0067	0.0067
	0.04	**0.0267**	**0.0267**	0.1038	0.1038	0.0267	0.0267
	0.06	**0.0400**	**0.0400**	0.1557	0.1558	0.0400	0.0400
2500	0.01	**0.0066**	**0.0066**	0.0260	0.0260	0.0066	0.0066
	0.04	**0.0265**	**0.0265**	0.1038	0.1038	0.0265	0.0265
	0.06	**0.0398**	**0.0398**	0.1557	0.1558	0.0398	0.0398
3000	0.01	**0.0066**	**0.0066**	0.0260	0.0260	0.0066	0.0066
	0.04	**0.0264**	**0.0264**	0.1039	0.1039	0.0264	0.0264
	0.06	**0.0396**	**0.0396**	0.1559	0.1559	0.0396	0.0396

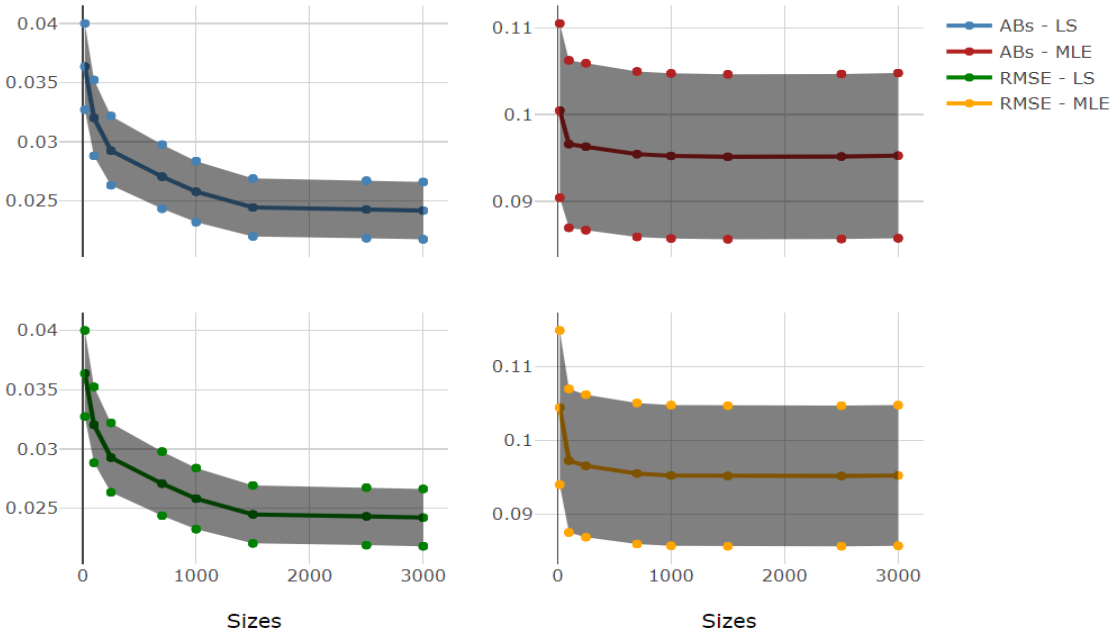


Figure 2b: *ABs and RMSE visualizations for LS and MLE methods*

II. Findings from ABs and RMSE Values of the Estimation Methods

The results on the absolute biases of the parameter estimates obtained from the different estimation techniques are presented in Tables 1 and 2 and visualized in Figures 2*a* and 2*b*. This section considers the results and its implication with regards to the Anderson Darling methods (ADR, AD2R, ADL, and AD2L). Also, from the analysis results, for the methods of estimation, there was a general decrease in the magnitude of the ABs and RMSE as the sample size increased from $20 \leq n \leq 3000$. The estimates obtained using the ADR, AD2R, ADL and AD2L revealed that the root ABs and RMSE for the Anderson Darling methods at sample size $n = 20$ and assumed true parameter value of $\varphi = 0.01$ were as follows:

- ADR: ABs = 0.033 & RMSE = 0.033.
- AD2R: ABs = 0.026 & RMSE = 0.026.
- ADL: ABs = 0.038 & RMSE = 0.038.
- AD2L: ABs = 0.039 & RMSE = 0.039.

A further scrutiny of the results revealed that the Anderson Darling AD2R method of estimation consistently yielded the smallest ABs and RMSE while AD2L yielded the largest values for the ABs and RMSE. Thus, this implies that the for the Anderson Darling methods of estimation, the AD2R provides a better estimate for the ETE distribution.

Regarding the LS and ML estimation methods as shown in Table 2, it can be observed from the results presented that the ML and LS methods at sample size $n = 20$ and true $\varphi = 0.01$ were as follows:

- LS method: AB = 0.0099 & RMSE = 0.0099.
- ML Method: AB = 0.0274 & RMSE = 0.0285.

As further observed, the trajectory and behavior of the ABs and RMSE as the sample size increases, showed a downward trend in the value of the ABs and RMSE. Notably, the LS method consistently yielded the smallest values of the ABs and RMSE, standing out as the best performing estimator. This trend reinforces the conclusion that the LS method provides the best estimates for the parameter of the ETE distribution.

III. Comparing the Minimum in Table 1 and 2

In this subsection, a comparison of the Anderson Darling methods, the ML and LS methods, are presented. Based on the simulation results, the LS estimator consistently demonstrated superior performance across all assumed true parameter values and sample sizes. These findings indicate that among the six estimators evaluated through the simulation study, the LS method emerges as the most effective estimation technique for the ETE distribution.

IV. Real World Application: Field Intensity Data

This section presents a numerical analysis using a real-world dataset to empirically evaluate the performance of different estimation methods. The objective is to determine which method provides the best fit to the data. The dataset consists of 80 observations sourced from [11], representing field intensity measured in microvolts and recorded at 5-second intervals. As illustrated in Figure 3, the data plot exhibits a right-skewed pattern. This skewness underscores the potential challenges for standard estimation techniques and emphasizes the need to select a method capable of providing robust and reliable modeling results. The data is presented as follows;

0.20, 0.71, 0.06, 0.05, 0.76, 0.32, 0.96, 0.63, 0.09, 0.18, 0.25, 0.45, 0.26, 0.10,0.95, 0.01, 0.50, 1.20, 1.99, 0.32, 0.51,0.01, 0.16, 0.56, 3.16, 1.27, 2.24, 1.00, 0.81,1.29, 0.28, 0.21, 0.35, 0.20, 0.39, 0.89,0.24, 0.08,

0.98,1.01, 0.49, 0.90,1.90, 1.42, 1.56,1.32,1.20,1.59, 2.24, 0.80, 0.56,0.18, 0.02, 0.28,0.81, 0.18,1.13, 0.64,1.95, 0.48, 0.55, 0.44, 0.28, 0.07, 0.71, 0.48, 0.06,0.79,1.01, 0.51, 0.70, 0.14, 0.16, 0.01,0.06, 0.03, 0.01.

Table 3: *Descriptive Statistics of Field Intensity*

Statistic	Values
Mean	0.670
Sd	0.630
Median	0.500
Mad	0.500
Min	0.010
Max	3.160
Range	3.150
Skew	1.440
Kurtosis	2.250
Se	0.070

Table 4 summarizes the estimates obtained from fitting the data to the distribution using the six different estimation methods. The results include the standard error (SE), the 95% confidence interval (CI), confidence interval length (CIL), *t*-value, *p*-value, and the rankings based on the CIL and SE.

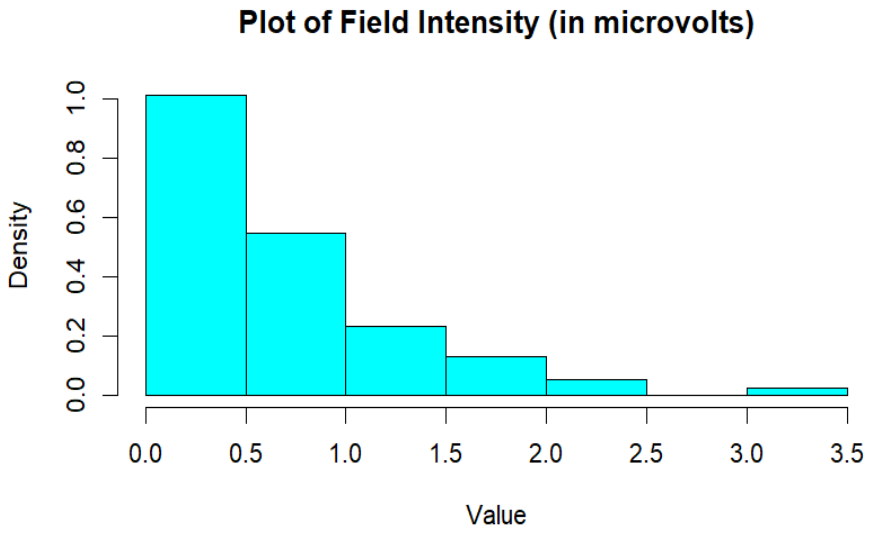


Figure 3: *Frequency plot of the Field Intensity Data*

Table 4: *Parameter estimates for the real-life data*

GOF	Estimate	SE	95% CI Lower	95% CI Upper	CIL	*t*-value	p-value	Rank
LS	1.1982	0.0246	1.1492	1.2472	0.0980	48.7167	0.0000	1
ADR	3.2878	0.4301	2.5714	4.2775	1.7061	7.6447	0.0000	2
AD2R	3.4377	0.4475	2.6775	4.4476	1.7701	7.6819	0.0000	3
MLE	1.0391	0.7718	-8.7679	10.8460	19.6139	1.3463	0.4067	4
ADL	72.4150	10.6967	55.1709	87.3595	32.1886	6.7698	0.0000	5
AD2L	92.1208	75.9367	9.3823	295.2279	285.8457	1.2131	0.2251	6

The rankings reveal that the least effective estimation method was AD2L, which ranked 6th, followed by ADL (rank 5), MLE (rank 4), AD2R (rank 3), and ADR (rank 2). The results indicate

that the LS method of estimation performed best, producing the lowest standard error (0.0246) and the narrowest confidence interval length (0.0980). This conclusion is further supported by Figure 4 (a–f), which illustrates the fits generated by the various estimation methods in comparison to the plot of the real data.

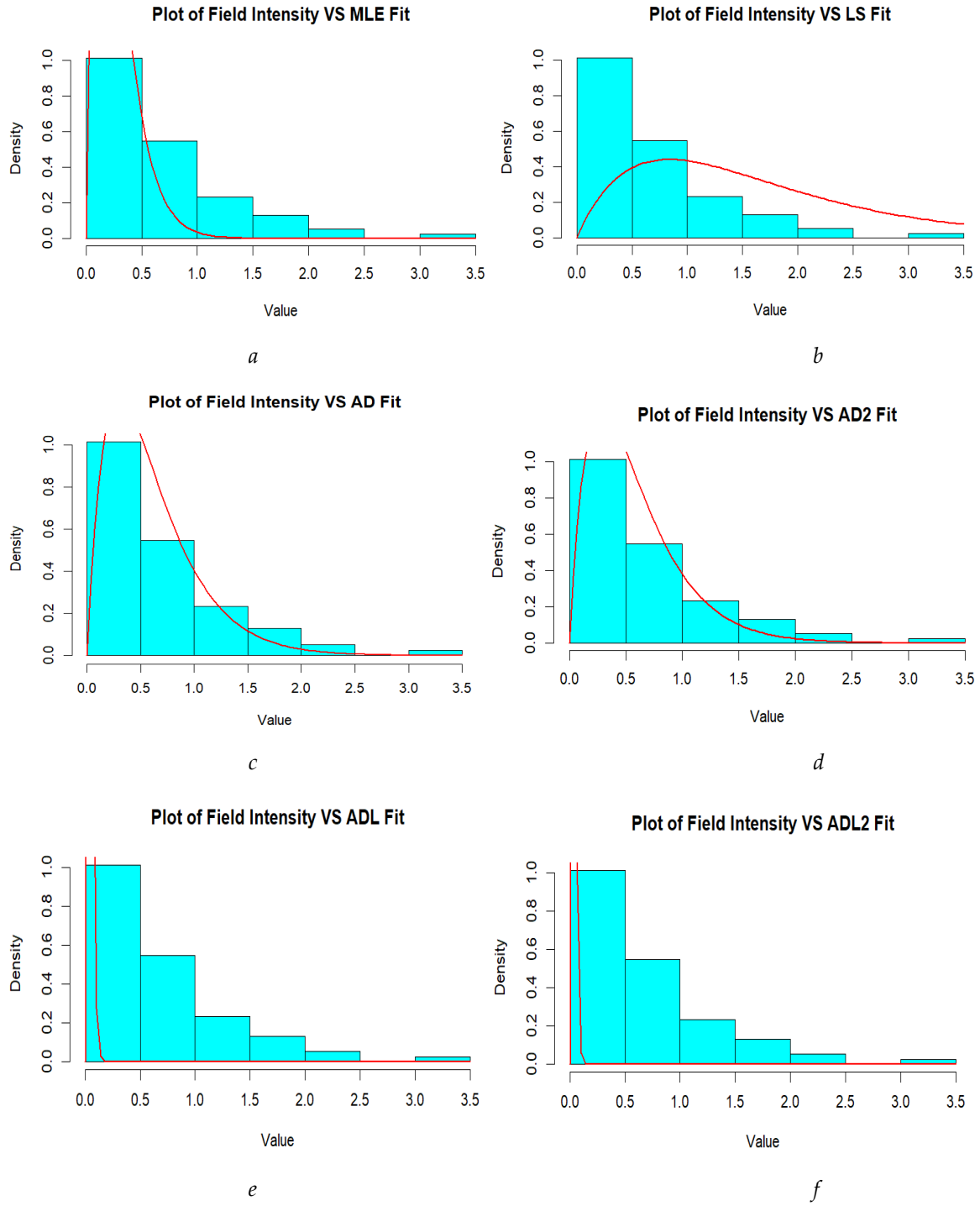


Figure 4: *Plots of the Various Fits Verse Real Data*

V. CONCLUSION

In this study, the performance of parameter estimates for the Entropy Transformed Exponential (ETE) distribution was evaluated using six estimation techniques: Anderson-Darling (AD) variants (ADR, AD2R, ADL, and AD2L), Maximum Likelihood (ML), and Least Squares (LS) estimation methods. A Monte Carlo simulation study was conducted to compare the average absolute biases (ABs), mean square errors (MSEs), and root mean square errors (RMSEs) for estimates obtained using each technique. The simulation results revealed that among the six methods studied, the LS estimation method consistently produced estimates with the most favorable performance metrics. Furthermore, real-world data analysis reinforced this finding, as the LS method demonstrated the narrowest confidence interval length and the smallest standard error of estimates. Based on these findings, this study concludes that the LS method is the most reliable and accurate estimation technique for the ETE distribution.

References

[1] Ali, S., Dey, S., Tahir, M. H., and Mansoor, M. (2020). The comparison of different estimation methods for the parameters of flexible Weibull distribution. *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, 64(4):1–34.

[2] Almathkour, F., Alothman, A., Ghitany, M. E., Gupta, R. C., and Mazucheli, J. (2022). A comparative study of various methods of estimation for Gompertz-Lindley distribution. *International Journal of Applied Mathematics*, 35(2):347–364.

[3] David, I. J., Asiribo, O. E., and Dikko, H. G. (2022). A Bertalanffy-Richards split-plot design model and analysis. *Journal of Statistical Modeling and Analysis*, 4(1):56–71.

[4] David, I. J., Asiribo, O. E., and Dikko, H. G. (2022). A Weibull split-plot design model and analysis. *Thailand Statistician*, 20(2):420–434.

[5] David, I. J., Asiribo O. E., Dikko, H. G., and Ikwuoche P. O. (2023). Johnson-Schumacher split-plot design modelling of rice yield. *Biometrical Letters*, 60(1):37–52.

[6] David, I. J., O. E. Asiribo, and Dikko, H. G. (2023). Nonlinear split-plot design modeling and analysis of rice varieties yield. *Scientific African*, 19:e01444.

[7] Dutang, C., and Delignette-Muller M. L. (2015). An r package for fitting distributions. *Journal of Statistical Software*, 69(1):794–814.

[8] Mathew, S., David I. J., and Mutah Y. (2024). A study on one-parameter entropy-transformed exponential distribution and its application. *Journal of Reliability and Statistical Studies*, 17(1):17–44.

[9] Mazucheli, A. F. B., and Manezes, J. (2019). L-moments and maximum likelihood estimation for the complementary beta distribution with application on temperature extremes. *Journal of Data Science*, 17(2):391–406.

SOME ADDITIONAL PROPERTIES OF KUMARASWAMY LOG-LOGISTIC DISTRIBUTION BASED ON ORDER RANDOM VARIATES

Bavita Singh[1], Nazia Wahid[2], Mahfooz Alam[2*] and Aafaq A. Rather[3]

•

[1]Department of Statistics, Amity Institute of Applied Sciences, Amity University, Noida, India.
[2]Department of Mathematics and Statistics, Faculty of Science and Technology,
Vishwakarma University, Pune, India.
[3]Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune-411004, India.
[1]bavita43@gmail.com, [2]wahidnazia010@gmail.com, [3]aafaq7741@gmail.com,
[2*]Corresponding author: mahfooz.stats@gmail.com

Abstract

The Exponentiated Kumaraswamy-Dagum distribution represents a significant advancement in the family of probability distributions. This family encompasses several well-established sub-models, including a variant of the Kumaraswamy-Fisk or Kumaraswamy log-logistic distribution, referred to as the Exponentiated Kumaraswamy-Dagum family distribution. In this paper, we characterize the Kumaraswamy log-logistic (KSLL) distribution by employing the moment features of generalized order statistics. Explicit expressions and recurrence relations are derived for both the single and product moments of generalized order statistics associated with the KSLL distribution. These formulations specifically address record values and moments of order statistics. Additionally, the distribution is further analyzed through conditional moments, and a recurrence relation for single moments of generalized order statistics is established. This rigorous investigation seeks to deepen the understanding of the KSLL distribution and its statistical properties as obtained by generalized order statistics.

Keywords: Generalized Order Statistics, Record Values, Order Statistics, Moments, Reliability, Characterization.

1. Introduction

As per historical review Dagum distrbution was introduced in 1977 by Camila Dagum [9]. This distribution if mainly fitted for the income and wealth data. Its was contained two formed as Dagum Type-I and Dagum Type-II specification distribution. This distrbution was a special case of generalized Beta distribution of second kind (GB2). Details McDonald [19], McDonald and Xu [20], Kleiber and Kotz [23]. Kleiber [22] introduced the rational properties of Dagum distribution. Further estimation properties for censored data dealt by Domma *et al.* [12]. After that there are several author find the generalized or mixed distribution with the Dagum distribution as different named as Beta-Dagum Distribution, Kumaraswamy Weibull distribution, Mc-Dagum distrbution, exponentiated Kumaraswamy-Dagum family distribution by the severals authors Domma and Condino [11], Cordiero *et al.* [6], Oluyede and Rajasooriya [24] and Huang and Oluyede [13], Alam *et al.* [1], [3].

Exponentiated Kumaraswamy-Dagum distribution was introduced by Huang and Oluyede [13]. It should be noted that the three parameter Kumaraswamy-Fisk or Kumaraswamy-log-

logistic distribution is a specific instance of the exponentiated version of the distribution. Due to its numerous uses in a variety of industries, including meteorology, acceptance sampling plans, areas of reliability, forests, modelling survival data or lifespan data, etc., it has attracted particular interest in recent decades.

A random variable X is said to have KSLL distribution, if its pdf is of the form

$$f(x) = \phi\delta\lambda x^{-\delta-1}(1+\lambda x^{-\delta})^{-2}[1-(1+\lambda x^{-\delta})^{-1}]^{\phi-1}, \quad x>0,\ \ \phi>0,\ \ \delta>0 \ \ and \ \ \lambda>0 \tag{1}$$

Let $\lambda = 1$ without loss of generality,the corresponding pdf is

$$f(x) = \phi\delta x^{-\delta-1}(1+x^{-\delta})^{-2}[1-(1+x^{-\delta})^{-1}]^{\phi-1}, \quad x>0,\ \ \phi>0, \ and \ \ \delta>0 \tag{2}$$

and the corresponding distribution function (df)

$$F(x) = [1-(1+x^{-\delta})^{-1}]^{\phi},\ x>0,\ \ \phi>0, \ and \ \ \delta>0 \tag{3}$$

Now in view of (2) and (3), we have

$$\phi\delta\bar{F}(x) = (x+x^{-\delta+1})f(x) \tag{4}$$

The relation in (4) is the characterizing differential equation for distribution given in (3), where ϕ and δ are shape parameters.
Fisk or log-logistic distribution is a special case of Kumaraswamy-Fisk or Kumaraswamy-log-logistic distribution when $\phi = 1$. for more details about this distribution one may refers to Huang and Oluyede [13].

2. Statistical Properties

We provide various significant statistical and mathematical measurements for the KSLL distribution in this section, including quantiles, moments, mgf and hazard rate.
In general statistics, quantile functions are often employed and frequently have representations in the form of lookup tables for important percentiles. The p-th quantile function $F^{-1}(p)$ of the KSLL distribution is given by the following, which follows from (3).

$$Q(p) = [1-(1-p)^{1/\phi}-1]^{-1/\delta}, \qquad p\in(0,1) \tag{5}$$

In particular, the first three quantiles Q_1, Q_2 and Q_3, can be generated by setting p in (5) to, respectively, 0.25, 0.5, and 0.75. Be aware that KSLL distribution random variates can be produced using (5).

The necessity and significance of the moments in any statistical analysis, particularly in applied work, rarely needs to be emphasized. Moments can be used to study some of a distribution's most significant traits and attributes, such as central tendency, dispersion, skewness, and kurtosis. The k -th instant of the random variable X can be calculated as follows if it has the KSLL distribution.

Let $Z \sim KSLL(\phi,\delta)$, which is given in (2). Then the k-th moment is given by

$$\mu^k = \int_0^\infty z^k f(z)dz = \phi\delta\int_0^\infty z^{k-\delta-1}(1+z^{-\delta})^{-2}(1-(1+z^{-\delta})^{-1})^{\phi-1}dz. \tag{6}$$

Solving the (6), We get the k-th ordinary moment of KSLL distribution.

$$\mu^k = \phi\sum_{i=0}^{\infty}(-1)^i\binom{\phi-1}{i}B(1-\frac{k}{\delta}, i+1+\frac{k}{\delta}), \tag{7}$$

where $k\le\delta$ and $Beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ is a beta function.

If we consider $\phi > 0$ is integer. Then the k-th ordinary moment of KSLL distribution.

$$\mu^k = \phi \sum_{i=0}^{\phi-1} (-1)^i \binom{\phi-1}{i} B(1-\frac{k}{\delta}, i+1+\frac{k}{\delta}). \tag{8}$$

Let $\phi = 1$ in (8), we get the k-th moment of Log-logistic distribution as

$$\mu_{ll}^k = B(1-\frac{k}{\delta}, 1+\frac{k}{\delta}).$$

The above relationships allow for the computation of variance, skewness, and kurtosis measures. Through the moment generating function and moments, numerous intriguing characteristics and features of a distribution become accessible and can be derived.

Let $Z \sim KSLL(\phi, \delta)$, which is given in (2). Then the moment generating function is defined as

$$M_z(t) = E(e^{tz}) = \int_0^\infty e^{tz} f(z)dz = \phi\delta \int_0^\infty e^{tz} z^{k-\delta-1}(1+z^{-\delta})^{-2}(1-(1+z^{-\delta})^{-1})^{\phi-1}dz.$$

Expanding the exponential form

$$M_z(t) = \phi \sum_{i,j=0}^{\infty} (-1)^i \binom{\phi-1}{i} \frac{t^j}{j!} B(1-\frac{j+k}{\delta}, i+1+\frac{j+k}{\delta}).$$

Table 1: *Using* (6) *the Ordinary Moments of KSLL distribution with the parameter* $\delta = 5$

ϕ	$\mu_1^{'}$	$\mu_2^{'}$	$\mu_3^{'}$	$\mu_4^{'}$	$\mu_5^{'}$
0.5	1.5497	2.0000			
1.5	0.929818	0.952464			
2.5	0.805842	0.698473	0.64645		
3.5	0.741375	0.586718	0.491302	0.432905	
4.5	0.69901	0.519664	0.407079	0.333955	0.285714

Table 2: *Using Using* (7)*, the Ordinary Moments of KSLL distribution with the parameter* $\delta = 5$

ϕ	$\mu_1^{'}$	$\mu_2^{'}$	$\mu_3^{'}$	$\mu_4^{'}$	$\mu_5^{'}$
0.5	1.5497				
1.5	0.929818	0.952464			
2.5	0.805842	0.698473	0.64645		
3.5	0.741375	0.586718	0.491302	0.432905	
4.5	0.69901	0.519664	0.407079	0.333955	0.285714

Table: 1 and 2 shows the verification of the moment expressions given in (7) with the general integration formula for the moment.

The hazard rate (HR) is one of the fundamental methods for researching the system's ageing and dependability characteristics. HR handles the remaining lifetime of the system. The HR provides the system's failure rate immediately following time x. As a result, the following gives the hazard rate function of the KSLL distribution.

$$h(x;\phi,\delta) = \frac{f(x;\phi,\delta)}{1-F(x;\phi,\delta)} = \frac{\phi * \delta}{x(1+x^{-\delta})}$$

3. Generalized Order Statistics

The introduction of generalized order statistics (GOS) by Kamps [15] encompasses various order models of random variables, including order statistics, upper record values, progressive Type II censoring order statistics, sequential order statistics, and Pfeifer's records.
Consider a sequence of independent and identically distributed (iid) random variables $X_1, X_2, ...$ (df) $F(x)$ and probability density function (pdf) $f(x)$. Assuming that $k > 0$, $n \in N$, $m \in R$ and $\gamma_r = k + (n-r)(m+1) > 0$. If the random variables $X(r,n,m,k)$, $r = 1,2,...,n$ possess a joint pdf of the form

$$k\left(\prod_{j=1}^{n-1}\gamma_j\right)\left(\prod_{i=1}^{n-1}\left[1-F(x_i)\right]^{m_i}f(x_i)\right)\left[1-F(x_n)\right]^{k-1}f(x_n), \tag{9}$$

on the cone $F^{-1}(0) < x_1 \leq .... \leq x_n < F^{-1}(1)$.

In such a scenario, these random variables are referred to as generalized order statistics derived from a sample obtained from a distribution with a distribution function $F(x)$. It is noteworthy that when $m = 0, k = 1$, this model simplifies to the joint probability density function of regular order statistics. Similarly, when $m = -1$ the model corresponds to the joint probability density function of the k-th upper record values. We shall also take $X(0,n,m,k) = 0$. The pdf of the $r-$ th gos is given by

$$f_{X(r,n,m,k)}(x) = \frac{C_{r-1}}{(r-1)!}\left[\bar{F}(x)\right]^{\gamma_r-1}f(x)g_m^{r-1}\left(F(x)\right). \tag{10}$$

The joint pdf of $X(r,n,m,k)$ and $X(s,n,m,k), 1 \leq r < s \leq n$ is

$$\begin{aligned} f_{X(r,n,m,k),X(s,n,m,k)}(x,y) &= \frac{C_{s-1}}{(r-1)!(s-r-1)!}[\bar{F}(x)]^m f(x)g_m^{r-1}(F(x)) \\ &\times \left[h_m(F(y)) - h_m(F(x))\right]^{s-r-1}\left[\bar{F}(y)\right]^{\gamma_s-1}f(y), \quad x < y. \end{aligned} \tag{11}$$

where $\bar{F}(x) = 1 - F(x)$, $C_{r-1} = \prod_{i=1}^{r}\gamma_i$, $r = 1,2,...,n$,

$$h_m(x) = \begin{cases} -\dfrac{1}{m+1}\left(1-x\right)^{m+1} & , \quad m \neq -1 \\ -\ln(1-x) & , \quad m = -1 \end{cases}$$

and

$$g_m(x) = h_m(x) - h_m(0), \quad x \in [0,1).$$

Many authors utilized the concept of gos in their work. References may be made to Kamps and Gather[16], Keseling [17], Cramer and Kamps [7], Ahsanullah (2000), Habibullah and Ahsanullah (2000), Raqab (2001), Kamps and Cramer[18], Ahmad and Fawzy (2003), Beiniek and Syznal (2003), Cramer *et al.* [8], Khan and Alzaid (2004), Jaheen (2005), Khan *et al.* (2006), Khan *et al.* (2010), Khan and Zia (2014), Khan and Khan (2016), Singh *et al.*[28], Singh *et al.*[29], Alam *et al.*[2] and Singh *et al.* [27] among others.

4. Relation for Single Moments

We shall first establish the exact expression for $E(X^j(r,n,m,k))$. Using (10), we have when $m \neq -1$

$$E[X(r,n,m,k)]^j = \frac{C_{r-1}}{(r-1)!}\int_0^\infty x^j\left[\bar{F}(x)\right]^{\gamma_r-1}f(x)g_m^{r-1}\big(F(x)\big)dx. \tag{12}$$

On expanding $g_m^{r-1}(F(x)) = \left(\frac{1}{m+1}1 - (\bar{F}(x))^{m+1}\right)^{r-1}$ binomially in (12), we get

$$E[X(r,n,m,k)]^j = A\int_0^\infty x^j\left[\bar{F}(x)\right]^{\gamma_{r-u}-1}f(x)dx, \tag{13}$$

where,

$$A = \frac{C_{r-1}}{(r-1)!(m+1)^{r-1}}\sum_{u=0}^{r-1}(-1)^u\binom{r-1}{u}$$

On using (3), we obtain

$$E[X(r,n,m,k)]^j = A\int_0^\infty(-1)^{j/\delta}[1-(1-\bar{F}(x)^{1/\phi})^{-1}]^{-j/\delta}[\bar{F}(x)]^{\gamma_{r-u}-1}f(x)dx \tag{14}$$

On using Maclaurine series expansion

$$(1-z)^{-t} = \sum_{k=0}^{\infty}\frac{(t)_k z^k}{k!} = \sum_{k=0}^{\infty}\frac{\Gamma(t+k)z^k}{\Gamma tk!} \tag{15}$$

$$(t)_k = \begin{cases} t(t+1)...(t+k-1) & , \quad k = 1,2 \\ \\ 1 & , \quad k = 0 \end{cases}$$

Using (15) in (14), we have

$$E[X(r,n,m,k)]^j = A\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\frac{(j/\delta)_p(p)_q}{p!q!}\int_0^\infty(-1)^{j/\delta}[\bar{F}(x)]^{\gamma_{r-u}+q/\phi-1}f(x)dx$$

$$E[X(r,n,m,k)]^j = A\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\frac{(j/\delta)_p(p)_q}{p!q!}(-1)^{j/\delta}\int_0^1 t^{\gamma_{r-u}+q/\phi-1}dt.$$

and simplifying the resulting expression, we find that

$$E[X(r,n,m,k)]^j = \frac{C_{r-1}(-1)^{j/\delta}}{(r-1)!(m+1)^{r-1}}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\frac{(j/\delta)_p(p)_q}{p!q!}\sum_{u=0}^{r-1}(-1)^u\binom{r-1}{u} \times\frac{1}{[k+(n-r+u)(m+1)+q/\phi]} \tag{16}$$

When $m = -1$, we have

$$E[X(r,n,m,k)]^j = \frac{0}{0} \quad \text{as} \quad \sum_{u=0}^{r-1}(-1)^u\binom{r-1}{u} = 0.$$

Since (16)is of the form $\frac{0}{0}$ at $m = -1$, therefore, we have

$$E[X(r,n,m,k)]^j = B\sum_{u=0}^{r-1}(-1)^u\binom{r-1}{u}\frac{[k+(n-r+u)(m+1)+q/\phi]^{-1}}{(m+1)^{r-1}}, \tag{17}$$

where

$$B = \frac{C_{r-1}}{(r-1)!}(-1)^{j/\delta}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\frac{(j/\delta)_p(p)_q}{p!q!}$$

Differentiating numerator and denominator of (17) $(r-1)$ times with respect to m, we get

$$E[X(r,n,m,k)]^j = B\sum_{u=0}^{r-1}(-1)^{u+(r-1)}\binom{r-1}{u}\frac{[1.2.3...(r-1)](n-r+u)^{r-1}}{(r-1)![k+(n-r+u)(m+1)+q/\phi]^r}.$$

On applying L' Hospital rule, we have

$$\lim_{m\to -1} E[X(r,n,m,k)]^j = B\sum_{u=0}^{r-1}(-1)^u\binom{r-1}{u}(r-n-u)^{r-1}\frac{[1.2.3...(r-1)]}{(r-1)!(k+q/\phi)^r}. \tag{18}$$

But for all integers $n \geq 0$ and for all real numbers x , we have Ruiz [25]

$$\sum_{i=0}^{n}(-1)^i\binom{n}{i}(x-i)^n = n!. \tag{19}$$

Now substituting (19) in (18) and simplifying, we find that

$$E[X(r,n,-1,k)]^j = E[(Y_r^k)^j] = k^r(-1)^{j/\delta}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\frac{(j/\delta)_p(p)_q}{p!q!}\frac{1}{[k+q/\phi]^r} \tag{20}$$

where $(Y_r^k)^j$denotes the $k-$th upper record values.

SPECIAL CASES

i) Putting $m = 0$ and $k = 1$ in (16), the exact expression for the single moments of order statistics from Kumaraswamy-Fisk or Kumaraswamy-log-logistic distribution can be obtained as

$$E(X_{r:n})^j = C_{r:n}(-1)^{j/\beta}\sum_{u=0}^{r-1}(-1)^u\binom{r-1}{u}\sum_{p=0}^{r-1}\sum_{q=0}^{\infty}\frac{(j/\delta)_p(p)_q}{p!q!}\frac{1}{[n-r+u+1+q/\phi]^r} \tag{21}$$

where

$$C_{r:n} = \frac{n!}{(r-1)!(n-r)!}.$$

ii) Putting $k = 1$ in (20), we deduce the explicit formula for the single moments of upper records for Kumaraswamy -Fisk or Kumaraswamy- log- logistic distribution in the form

$$E[(Y_r^k)^j] = E(X_{U(r)}^j) = (-1)^{-j/\delta}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\frac{(j/\delta)_p(p)_q}{p!q!}\frac{1}{[1+q/\phi]^r} \tag{22}$$

Expressions (21) and (22)can be used to obtain the moments of order statistics and upper record values for arbitrary chosen values of ϕ, δ and various sample size $n = 1, 2, ..., 10$.

Remark 1. i) Putting $\phi = 1$ in (16), we get the explicit expression for the single moments of gos from Fisk or log- logistic distribution can be obtained.

$$E[X(r,n,m,k)]^j = \frac{C_{r-1}(-1)^{j/\delta}}{(r-1)!(m+1)^{r-1}}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\frac{(j/\delta)_p(p)_q}{p!q!}\sum_{u=0}^{r-1}(-1)^u\binom{r-1}{u}$$
$$\times\frac{1}{[k+(n-r+u)(m+1)+q]}.$$

ii) Putting $\phi = 1$ in (22), we deduce the explicit formula for the single moments of upper records for Fisk or log- logistic distribution.

$$E[X(r,n,-1,k)]^j = E[(Y_r^k)^j] = k^r(-1)^{j/\delta}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\frac{(j/\delta)_p(p)_q}{p!q!}\frac{1}{[k+q]^r}$$

Table 3: *First Moments of Order Statistics with the different parameter ϕ and $\delta = 1$*

n	r	$\phi = 0.5$	$\phi = 1.5$	$\phi = 2.5$	$\phi = 3.5$	$\phi = 4.5$	$\phi = 5.5$
1	1	1.5497	0.92982	0.80585	0.741375	0.69901	0.66795
2	1	1.0689	0.76966	0.68242	0.633289	0.59982	0.57475
	2	2.0304	1.08998	0.92926	0.84947	0.798206	0.76113
3	1	0.929821	0.69901	0.62385	0.58044	0.55053	0.52799
	2	1.34724	0.91093	0.799572	0.73898	0.69837	0.66828
	3	2.37203	1.17951	0.994105	0.9047	0.84812	0.807555
4	1	0.855167	0.65513	0.586486	0.546365	0.518563	0.497545
	2	1.15377	0.830662	0.735941	0.682682	0.646446	0.61934
	3	1.54072	0.99120	0.86320	0.795278	0.75031	0.717217
	4	2.64913	1.24228	1.03774	0.941174	0.888721	0.837668
5	1	0.805842	0.62385	0.559485	0.521595	0.495252	0.475299
	2	1.05247	0.78024	0.694489	0.645443	0.611804	0.586526
	3	1.30572	0.90630	0.798117	0.738541	0.698409	0.668569
	4	1.69738	1.0478	0.906592	0.833103	0.784914	0.749649
	5	2.88707	1.2909	1.07053	0.968192	0.904673	0.859673
6	1	0.769651	0.59982	0.538559	0.502329	0.477086	0.457942
	2	0.986799	0.744022	0.664116	0.617923	0.586083	0.562088
	3	1.18381	0.85266	0.755236	0.700483	0.663246	0.635403
	4	1.42763	0.959944	0.840999	0.776599	0.733572	0.701735
	5	1.83226	1.09173	0.939389	0.861354	0.810585	0.773606
	6	3.09803	1.33074	1.096775	0.98956	0.923491	0.876887
7	1	0.741375	0.580444	0.521595	0.486673	0.462303	0.443805
	2	0.939307	0.716045	0.640346	0.596267	0.56578	0.542762
	3	1.10553	0.813964	0.723543	0.672063	0.636837	0.610403
	4	1.28818	0.904256	0.797493	0.738377	0.698458	0.668737
	5	1.53221	1.00171	0.873629	0.805266	0.759907	0.726483
	6	1.95228	1.12774	0.965693	0.88379	0.830856	0.792455
	7	3.28899	1.36457	1.1186	1.00719	0.93893	0.890959
8	1	0.718341	0.564306	0.507403	0.473552	0.449903	0.431939
	2	0.902613	0.693406	0.620936	0.578516	0.549106	0.526865
	3	1.04939	0.783962	0.698576	0.64952	0.61581	0.59045
	4	1.1991	0.863966	0.765155	0.709634	0.671882	0.643658
	5	1.37727	0.944545	0.82983	0.76712	0.725035	0.693816
	6	1.62517	1.03601	0.89999	0.828154	0.78083	0.746083
	7	2.06132	1.15831	0.98762	0.902335	0.847532	0.807912
	8	3.46437	1.39404	1.13731	1.02217	0.951986	0.902822
9	1	0.69901	0.550533	0.495252	0.462303	0.439264	0.421754
	2	0.872983	0.674488	0.604608	0.563541	0.53501	0.513421
	3	1.00632	0.759617	0.67808	0.630927	0.59842	0.57392
	4	1.13553	0.832654	0.73956	0.686706	0.65058	0.62351
	5	1.27855	0.903107	0.797148	0.738294	0.698498	0.668844
	6	1.45625	0.97769	0.855976	0.79018	0.746266	0.713793
	7	1.70963	1.06517	0.921876	0.84714	0.798112	0.762228
	8	2.1618	1.18492	1.0064	0.918105	0.8616520	0.820964
	9	3.6272	1.42018	1.15367	1.03518	0.963278	0.913055

Table 4: *Continued...*

n	r	$\phi = 0.5$	$\phi = 1.5$	$\phi = 2.5$	$\phi = 3.5$	$\phi = 4.5$	$\phi = 5.5$
10	1	0.682424	0.538559	0.484662	0.452489	0.429976	0.412859
	2	0.848292	0.658304	0.590567	0.550635	0.522858	0.501811
	3	0.971747	0.739227	0.660773	0.615164	0.583647	0.55986
	4	1.08699	0.807192	0.718473	0.667708	0.63289	0.606728
	5	1.20834	0.870846	0.771193	0.715204	0.677138	0.648683
	6	1.34877	0.935369	0.823103	0.761384	0.719857	0.689005
	7	1.5279	1.00591	0.877891	0.809377	0.76387	0.730319
	8	1.78751	1.09056	0.940726	0.863325	0.812787	0.775904
	9	2.25537	1.20851	1.02282	0.9318	0.873868	0.832222
	10	3.77962	1.44369	1.16821	1.04666	0.973213	0.922035

Table 5: *First Moment Based on Upper Record Values*

n	$\phi = 1.5$	$\phi = 2.5$	$\phi = 3.5$	$\phi = 5.5$	$\phi = 6.5$	$\phi = 7.5$
1	0.929818	0.805842	0.741375	0.667943	0.643654	0.62385
2	1.21669	1.01358	0.918689	0.817397	0.785154	0.759243
3	1.46957	1.17176	1.04499	0.917449	0.8783	0.847274
4	1.73184	1.31691	1.15418	0.999098	0.953103	0.917135
5	2.01973	1.46037	1.25649	1.0716	1.01852	0.977535
6	2.34366	1.60757	1.35652	1.13902	1.07847	1.03229
7	2.71253	1.76185	1.45686	1.20356	1.13508	1.08344
8	3.13513	1.9256	1.55916	1.26655	1.18959	1.13222
9	3.62085	2.10081	1.66462	1.32887	1.24285	1.17941
10	4.1801	2.28925	1.77416	1.39115	1.29545	1.22558

Now we obtain the recurrence relations for single moments of Kumaraswamy-Fisk or Kumaraswamy-log- logistic distribution in the following theorem.

Theorem 1. For the distribution as given in (3) and $n \in N$, $m \in R$, $1 \leq r \leq n$, $j = 0, 1, 2, ...$

$$E[X(r,n,m,k)]^j\big[1 - \frac{j}{\phi\delta\gamma_r}\big] = E[X(r-1,n,m,k)]^j + \frac{j}{\phi\delta\gamma_r}E[X(r,n,m,k)]^{j-\delta} \tag{23}$$

Proof:. From (3) and (4), we have

$$E[X(r,n,m,k)]^j - E[X(r-1,n,m,k)]^j = \frac{j}{\phi\delta\gamma_r}\big[\frac{C_{r-1}}{(r-1)!}\int_0^\infty x^j\big[\bar{F}(x)\big]^{\gamma_r-1}g_m^{r-1}\big(F(x)\big)f(x)dx\big]$$

$$+\frac{j}{\phi\delta\gamma_r}\big[\frac{C_{r-1}}{(r-1)!}\int_0^\infty x^{j-\delta}\big[\bar{F}(x)\big]^{\gamma_r-1}g_m^{r-1}\big(F(x)\big)f(x)dx\big]$$

Now (23) can be proved by noting that in view of Athar and Islam [4]

$$E[X(r,n,m,k)]^j - E[X(r-1,n,m,k)]^j = \frac{jC_{r-2}}{(r-1)!}\int_0^\infty x^{j-1}\big[\bar{F}(x)\big]^{\gamma_r}g_m^{r-1}\big(F(x)\big)dx.$$

Remark 2. i) Substituting $m = 0$, $k = 1$ in (23), we deduce the recurrence relation for single moments of order statistics for Kumaraswamy Fisk or Kumaraswamy log logistic distribution in the form

$$E(X_{r:n}^j)\big[1 - \frac{j}{\phi\delta(n-r+1)}\big] = E(X_{r-1:n}^j) + \frac{\lambda j}{\phi\delta(n-r+1)}E(X_{r:n}^{j-\delta})$$

ii) Putting $m = -1$ and $k = 1$ in (23), the result for single moments of upper record values from for Kumaraswamy Fisk or Kumaraswamy log logistic distribution is deduced as

$$E(X^j_{U(r)})\big[1 - \frac{j}{\phi\delta}\big] = E(X^j_{U(r-1)}) + \frac{j}{\phi\delta}E(X^{j-\delta}_{U(r)}).$$

(iii) Putting $\phi = 1$ in (23), we get the recurrence relation for single moments of *lgos* from the for Fisk or log logistic distribution.

$$E[X(r,n,m,k)]^j\big[1 - \frac{j}{\delta\gamma_r}\big] = E[X(r-1,n,m,k)]^j + \frac{j}{\delta\gamma_r}E[X(r,n,m,k)]^{j-\delta}$$

5. Characterization

Let $X(r,n,m,k),\ r = 1,2,...,n$ be *gos*, then the conditional pdf of $X(s,n,m,k)$ given $X(r,n,m,k) = x,\ 1 \le r < s \le n$, is

$$f_{X(s,n,m,k)|X(r,n,m,k)}(x|y) = \frac{C_{s-1}}{(s-r-1)!\,C_{r-1}}\,[\bar{F}(x)]^{m-\gamma_r+1} \times [h_m(F(y)) - h_m(F(x))]^{s-r-1}\,[\bar{F}(y)]^{\gamma_s-1}\,f(y),\quad x < y. \tag{24}$$

Theorem 2. Let X be a non-negative random variable having an absolutely continuous $df\ F(x)$ with $F(0) = 0$ and $0 \le F(x) \le 1$ for all $x \ge 0$, then

$$E[\xi\{X(s,n,m,k)\}|X(r,n,m,k) = x] = [1 - (1+x^{-\delta})^{-1}]^{\phi}\prod_{j=1}^{s-l}\frac{\gamma_{l+j}}{\gamma_{l+j}+1}\quad l = r,\ r+1, \tag{25}$$

if and only if

$$\bar{F}(x) = [1 - (1+x^{-\delta})^{-1}]^{\phi},\quad x > 0,\quad \phi > 0,\quad \delta > 0. \tag{26}$$

where,

$$\xi(y) = [1 - (1+y^{-\delta})^{-1}]^{\phi},\quad x > 0,\quad \phi > 0,\quad \delta > 0.$$

Proof. From (22) for $s > r+1$, we have

$$E[\xi\{X(s,n,m,k)\}|X(r,n,m,k) = x] = \frac{C_{s-1}}{(s-r-1)!C_{r-1}(m+1)^{s-r-1}} \times \int_x^{\infty}[1 - (1+y^{-\delta})^{-1}]^{\phi}\Big[1 - \Big(\frac{\bar{F}(y)}{\bar{F}(x)}\Big)^{m+1}\Big]^{s-r-1}\Big(\frac{\bar{F}(y)}{\bar{F}(x)}\Big)^{\gamma_s-1}\frac{f(y)}{\bar{F}(x)}dy. \tag{27}$$

By setting $u = \frac{\bar{F}(y)}{\bar{F}(x)} = \frac{[1-(1+y^{-\delta})^{-1}]^{\phi}}{[1-(1+x^{-\delta})^{-1}]^{\phi}}$, from (3) in (27), we find that

$$E[\xi\{X(s,n,m,k)\}|X(r,n,m,k) = x] = \frac{C_{s-1}[1 - (1+x^{-\delta})^{-1}]^{\phi}}{(s-r-1)!C_{r-1}(m+1)^{s-r-1}} \times \int_0^1 u^{\gamma_s}\Big(1 - u^{m+1}\Big)^{s-r-1}du \tag{28}$$

Again by setting $t = u^{m+1}$ in (28), we get

$$E[\xi\{X(s,n,m,k)\}|X(r,n,m,k) = x] = \frac{C_{s-1}[1 - (1+x^{-\delta})^{-1}]^{\phi}}{(s-r-1)!C_{r-1}(m+1)^{s-r-1}}$$

$$\times \int_0^1 t^{\frac{\gamma_s+1}{m+1}-1}(1-t)^{s-r-1}dt \tag{29}$$

and hence the necessary part.

To prove sufficient part, we have from (22) and (23)

$$\frac{C_{s-1}}{(s-r-1)!C_{r-1}(m+1)^{s-r-1}}\int_x^\infty [1-(1+y^{-\delta})^{-1}]^{\phi}\left[\bar{F}(x)^{m+1}-\bar{F}(y)^{m+1}\right]^{s-r-1}$$

$$\times[\bar{F}(y)]^{\gamma_s-1}f(y)dy = g_{s|r}(x)[\bar{F}(x)]^{\gamma_{r+1}}, \tag{30}$$

where

$$g_{s|r}(x) = [1-(1+x^{-\delta})^{-1}]^{\phi}\prod_{j=1}^{s-r}\frac{\gamma_{r+j}}{\gamma_{r+j}+1}.$$

Differentiating (30) both sides with respect to x, we get

$$-\frac{C_{s-1}[\bar{F}(x)]^m f(x)}{(s-r-2)!C_{r-1}(m+1)^{s-r-2}}\int_x^\infty [1-(1+y^{-\delta})^{-1}]^{\phi}\left[\bar{F}(x)^{m+1}-\bar{F}(y)^{m+1}\right]^{s-r-2}$$

$$[\bar{F}(y)]^{\gamma_s-1}f(y)dy = g'_{s|r}(x)[\bar{F}(x)]^{\gamma_{r+1}} - \gamma_{r+1}\, g_{s|r}(x)[\bar{F}(x)]^{\gamma_{r+1}-1}f(x)$$

or

$$-\gamma_{r+1}g_{s|r+1}(x)[\bar{F}(x)]^{\gamma_{r+2}+m}f(x)$$

$$= g'_{s|r}(x)[\bar{F}(x)]^{\gamma_{r+1}} - \gamma_{r+1}g_{s|r}(x)[\bar{F}(x)]^{\gamma_{r+1}-1}f(x),$$

where,

$$g'_{s|r}(x) = -\phi\delta x^{-\delta-1}(1+x^{-\delta})^{-2}(1-(1+x^{-\delta})^{-1})^{\phi-1}\prod_{j=1}^{s-r}\Big(\frac{\gamma_{r+j}}{\gamma_{r+j}+1}\Big),$$

$$g_{s|r+1}(x) = (1-(1+x^{-\delta})^{-1})^{\phi}\Big(\frac{\gamma_{r+1}+1}{\gamma_{r+1}}\Big)\Big(\prod_{j=1}^{s-r}\frac{\gamma_{r+j}}{\gamma_{r+j}+1}\Big).$$

Therefore,

$$\frac{f(x)}{\bar{F}(x)} = -\frac{g'_{s|r}(x)}{\gamma_{r+1}[g_{s|r+1}(x)-g_{s|r}(x)]}$$

$$= \frac{\phi\delta}{x+x^{-\delta+1}}. \tag{31}$$

Integrating (31) on both sides with respect to x between $(0,y)$, the sufficient part is proved.

Remark 3. i) As $m=-1$ in (25), we get the characterization results from the Kumaraswamy- Fisk or Kumaraswamy log-logistic distribution based on$k$- th upper record values.
ii) Setting $m=0$, $k=1$ in (25), we obtain the characterization results of the Kumaraswamy- Fisk or Kumaraswamy log-logistic distribution based on order statistics.

Following theorem contains characterization of this distribution by a recurrence relation for the single moments of *gos*.

Theorem 3. For a positive integer k and j be a non-negative integer, a necessary and sufficient condition for a random variable X to be distributed with pdf given by (2) is that

$$E[X(r,n,m,k)]^j\big[1-\frac{j}{\phi\delta\gamma_r}\big]=E[X(r-1,n,m,k)]^j+\frac{j}{\phi\delta\gamma_r}E[X(r,n,m,k)]^{j-\delta} \tag{32}$$

Proof. The necessary part follows from equation (23). On the other hand if the relation in (32)is satisfied, then on using (10), we have

$$\begin{aligned}\frac{C_{r-1}}{(r-1)!}\int_0^\infty x^j\big[\bar{F}(x)\big]^{\gamma_r-1}f(x)g_m^{r-1}\big(F(x)\big)dx&=\frac{C_{r-2}}{(r-2)!}\int_0^\infty x^j\big[\bar{F}(x)\big]^{\gamma_{r-1}-1}f(x)g_m^{r-2}\big(F(x)\big)dx\\&+\frac{jC_{r-1}}{\phi\delta\gamma_r(r-1)!}\int_0^\infty x^j\big[\bar{F}(x)\big]^{\gamma_r-1}f(x)g_m^{r-1}\big(F(x)\big)dx\\&+\frac{jC_{r-1}}{\phi\delta\gamma_r(r-1)!}\int_0^\infty x^{j-\delta}\big[\bar{F}(x)\big]^{\gamma_r-1}f(x)g_m^{r-1}\big(F(x)\big)dx\end{aligned}$$

Let

$$h(x)=-\frac{1}{\gamma_r}\big[\bar{F}(x)\big]^{\gamma_r}g_m^{r-1}\big(F(x)\big) \tag{33}$$

Differentiating both the sides of (33), we get

$$h'(x)=-\frac{1}{\gamma_r}\big[\bar{F}(x)\big]^{\gamma_r-1}f(x)g_m^{r-2}\big(F(x)\big)\{\gamma_r g_m\big(F(x)\big)-(r-1)\big[\bar{F}(x)\big]^{m+1}\}$$

Thus,

$$\begin{aligned}\frac{C_{r-1}}{(r-1)!\gamma_r}\int_0^\infty x^jh'(x)dx&=\frac{jC_{r-1}}{\phi\delta\gamma_r(r-1)!}\int_0^\infty x^j\big[\bar{F}(x)\big]^{\gamma_r-1}f(x)g_m^{r-1}\big(F(x)\big)dx\\&+\frac{jC_{r-1}}{\phi\delta\gamma_r(r-1)!}\int_0^\infty x^{j-\delta}\big[\bar{F}(x)\big]^{\gamma_r-1}f(x)g_m^{r-1}\big(F(x)\big)dx\end{aligned} \tag{34}$$

Integrating right hand side in (34) by parts and using the value of $h(x)$ from (33), we find that

$$\frac{jC_{r-1}}{\gamma_r(r-1)!}\int_0^\infty x^{j-1}\big[\bar{F}(x)\big]^{\gamma_r-1}g_m^{r-1}\big(F(x)\big)\{\bar{F}(x)-\frac{x}{\phi\delta}f(x)-\frac{1}{\phi\delta}x^{-\delta+1}f(x)\}dx=0 \tag{35}$$

Applying the extension of Mntz-Szsz Theorem, (see for example Hwang and Lin, [14]), to (28), we get

$$\phi\delta\bar{F}(x)=(x+x^{-\delta+1})f(x).$$

Which proves that $f(x)$ has the form as in (4).

6. Conclusion

The work by Huang and Oluyede [13] and Santana *et al.* [26] examines the KSLL distribution using features obtained from order random variables. Applications for order random variables can be found in many different domains, including distribution parameter estimation, production process optimisation, auto racing, insurance policy modelling, and auctions. The purpose of this essay is to go into great detail about the idea and importance of order random variables. Establishing the KSLL distribution's descriptive and distributional features, calculating its ordinary moments, and displaying the findings graphically are the main objectives of the first portion. Order statistics, record values, and their recurrence relations for single and product moments obtained from the KSLL distribution are among the related properties of order random variables that are examined in the second part. Theorems of characterisation for these qualities are also discussed. Two theorems describing the KSLL distribution based on conditional moments for single moments and recurrence relations are presented and proven in the work. This approach also analyses a number of deductions and specific circumstances.

Declarations

Conflicts of Interest: The authors declare no competing interests.

Consent for Publication: I consent to the publication of the data and materials.

Code Availability We are using the software an R and MATHEMATICA to calculate the numerical values and graphics.

Funding Statement: This research received no external funding.

References

[1] Alam, M., Khan, M. A., and Chesneau, C. (2023). Moments Of Power Lomax Distribution Through Generalized Order Statistics. Applied Mathematics E-Notes, 23, 8-19.
[2] Alam, M., Khan, R.U. and Vidovic, Z. (2021): Moments of generalized order statistics for Pareto-Rayleigh distribution. Open Journal of Mathematical Sciences, 5, 306-313.
[3] Alam, M., Wahid, N., Khan, M. A., and Khan, R. (2022): Characterizations of Topp"Leone Lomax Distribution based on the Generalized Lower Record Values, 11(3), 1095-1106.
[4] Athar, H., and Islam, H. M. U. (2004). Recurrence relations for single and product moments of generalized order statistics from a general class of distribution. Metron-International Journal of Statistics, 62(3), 327-337.
[5] Bieniek, M. and Szynal, D. (2003): Characterizations of distributions via linearity of regression of generalized order statistics. Metrika, 58, 259-271.
[6] Cordeiro, GM, Ortega, EMM, Nadarajah, S (2010): The Kumaraswamy Weibull distribution with application to failure data. J. Franklin Inst. 347, 1399"1429.
[7] Cramer, E. and Kamps, U. (2000): Relations for expectations of functions of generalized order statistics. J. Statist. Plann. Inference, 89, 79-89.
[8] Cramer, E., Kamps, U. and Keseling, C. (2004): Characterization via linear regression of ordered random variables: A unifying approach. Comm. Statist. Theory Methods, 33, 2885-2911.
[9] Dagum, CA (1977): New model of personal income distribution: specification and estimation. Economie Applique™e. 30, 413"437.
[10] David, H.A. and Nagaraja, H.N. (2003): Order Statistics. John Wiley, New York.
[11] Domma, F, Condino, F (2013): The Beta-Dagum distribution: definition and properties. Communications in Statistics-Theory and Methods. 44(22), 4070"4090.
[12] Domma, F, Giordano, S, Zenga, M (2011): Maximum likelihood estimation in Dagum distribution with censored samples. J. Appl. Stat. 38(21), 2971"2985.
[13] Huang, S. and Oluyede, B.O. (2014): Exponentiated Kumaraswamy-Dagum distribution with applications to income and lifetime data. Journal of Statistical Distribution with Applications, 18, 1-20.
[14] Hwang, J.S. and Lin, G.D. (1984): On a generalized moments problem II, Proc. Amer. Math. Soc., 91, 577-580.
[15] Kamps, U. (1995): A Concept of Generalized Order Statistics. B.G.Teubner Stuttgart, Germany.
[16] Kamps, U. and Gather, U. (1997): Characteristic property of generalized order statistics from exponential distribution. Appl. Math. (Warsaw), 24, 383-391.
[17] Keseling, C. (1999): Conditional distributions of generalized order statistics and some characterizations. Metrika, 49, 27-40.
[18] Kamps, U. and Cramer, E. (2001): On distribution of generalized order statistics. Statistics, 35, 269-280.
[19] McDonald, B (1984): Some generalized functions for the size distribution of income. Econometrica. 52(3), 647"663.

[20] McDonald, B, Xu, J (1995): A generalization of the beta distribution with application. J. Econometrics. 69(2), 133“152.
[21] Khan, R.U. and Zia, Benazir (2014): Generalized order statistics from doubly truncated linear exponential distribution and a characterization. Journal of Applied Probability & Statistics, 9, 53-65.
[22] Kleiber, C (2008): A Guide to the Dagum Distributions. In: Duangkamon, C (ed.) Modeling Income Distributions and Lorenz Curve Series: Economic Studies in Inequality, Social Exclusion and Well-Being 5. Springer, New York.
[23] Kleiber, C, Kotz, S (2003): Statistical size distributions in economics and actuarial sciences. Wiley, New York.
[24] Oluyede, BO, Rajasooriya, S. (2013): The Mc-Dagum Distribution and Its Statistical Properties with Applications. Asian J. Mathematics and Applications. (85). http://scienceasia.asia/index.php/ama/article/view/85/44, 1“16
[25] Ruiz, S.M. (1996): An algebraic identity leading to Wilson™s theorem. The Mathematical Gazette, 80, 579-582.
[26] Santana, T., Ortega, E., Cordeiro, G. and Silva, G. (2012): The Kumaraswamy-Log-Logistic Distribution. Journal of statistical theory and applications. 11. 265-291.
[27] Singh, B. and Khan, R.U. and Khan, M.A. (2021): Exact Moments of Generalized Pareto Distribution based on Generalized Order Statistics and Characterizations. Pakistan Journal of Statistics and Operation Research.17, 213-226.
[28] Singh, B., Khan, R.U. and Khan, M.A. (2019): Characterizations of the q -Weibull distribution based on generalized order statistics. Journal of Statistics & Management Systems, 22, 1573-1595.
[29] Singh, B., Khan, R.U. and Khan, M.A.R. (2018): Generalized order statistics from Kumaraswamy-Burr III distribution and related Inference. Journal of Statistics: Advances in Theory and Applications, 19, 1-16.

COMPARATIVE ANALYSIS OF THE RELIABILITY OF VACUUM AND OIL CIRCUIT BREAKERS

I.N. Rahimli[1], A.L. Bakhtiyarov[2]

•

Azerbaijan State Oil and Industry University, Baku, Azerbaijan
[1] ilham.raqimli@asoiu.edu.az ; [2] baxtiyarov.aliashraf@asoiu.edu.az

Abstract

In modern power systems, switching equipment plays a crucial role in ensuring the safe and reliable connection and disconnection of electrical circuits. A failure of a circuit breaker can lead to serious consequences, including prolonged power outages and damage to expensive equipment. This paper presents a comparative analysis of the reliability of vacuum and oil circuit breakers, considering operational, technical, and economic factors. It examines the main causes of failures, methods of reliability assessment, and provides a comprehensive evaluation of the effectiveness of each type of breaker under various conditions. The novelty of the proposed solution lies in the in-depth analysis of operational characteristics and a multifaceted evaluation of the impact of external factors on equipment reliability. Modern diagnostic and monitoring methods, including thermography, ultrasonic diagnostics, and the integration of IoT systems for real-time monitoring, are used. These approaches improve the processes of equipment selection and operation, ultimately contributing to the overall reliability of power systems. The paper aims to support informed decision-making in equipment selection to enhance the overall reliability of power systems.

Keywords: vacuum circuit breakers, oil circuit breakers, reliability, failure analysis

I. Introduction

In modern power systems, switching equipment plays a key role in ensuring the safe and reliable connection and disconnection of electrical circuits. A circuit breaker failure can lead to serious consequences, ranging from prolonged power outages to the failure of expensive equipment [1]. Therefore, the reliability of circuit breakers is one of the most important criteria when selecting equipment for medium- and high-voltage systems.

Among the most common types of circuit breakers today, vacuum and oil circuit breakers can be distinguished. Each of them has its own design features, operating principles, and operational characteristics. Oil circuit breakers have been widely used for decades and have proven to be a cost-effective and reliable solution. However, with the development of technology, vacuum circuit breakers have actively started to replace them, especially in medium-voltage installations, due to their high reliability, environmental friendliness, and ease of maintenance.

The purpose of this article is a comparative analysis of the reliability of vacuum and oil circuit breakers, taking into account operational, technical, and economic factors [2-4]. The paper examines the main causes of failures, methods of reliability assessment, and provides a comprehensive evaluation of the effectiveness of each type of circuit breaker under various conditions.

Understanding the strengths and weaknesses of both types of circuit breakers allows for an informed choice in favor of one type of equipment or another and contributes to improving the overall reliability of power systems.

II. Formulation of the problem

Reliability analysis of vacuum circuit breakers: The reliability of vacuum circuit breakers depends on many factors, including their design, operating conditions, and diagnostic methods. Among the main causes of failures in vacuum circuit breakers, several typical aspects can be identified. First, wear of the contact elements, which over time reduces their ability to effectively transmit electrical current. The second important cause is the loss of vacuum integrity in the vacuum chamber. Vacuum is a key factor in the operation of such breakers, and any vacuum leakage can significantly affect their efficiency [5]. Another common issue is the failure of the drive mechanisms, which can lead to the inability of the device to function properly. In some cases, a malfunction in the control system can also occur, potentially leading to a failure.

Operating conditions play a significant role in the reliability of vacuum circuit breakers. Under ideal conditions, such as stable temperatures and the absence of contaminants in the air, these devices can operate for a long time without significant damage. However, under real-world conditions, such as frequent temperature fluctuations or the presence of moisture and dust, their reliability may decrease. These external factors can accelerate material wear, which leads to the need for more frequent maintenance.

Modern diagnostic and monitoring methods for vacuum circuit breakers allow for the detection of potential issues at an early stage [6-7]. Thermography, ultrasonic diagnostics, and systems for monitoring mechanical wear—all of these approaches help prevent breakdowns, ensuring long and stable operation of the equipment. Additionally, with the development of technology, the integration of IoT systems allows for real-time monitoring of the breakers' condition, enabling prompt responses to changes in their performance. As for the mean time between failures (MTBF), vacuum circuit breakers can boast relatively high figures, sometimes reaching tens of thousands of hours, which gives them an advantage over other types of circuit breakers.

Analysis of the reliability of oil circuit breakers: Oil circuit breakers, despite their durability and traditional use in the power industry, have several vulnerable elements that reduce their reliability. One of the main issues is the wear and contamination of the oil. Over time, the oil loses its insulating and quenching properties, which can lead to deterioration in switching quality and an increased risk of emergency situations [8-10]. Furthermore, oil circuit breakers are prone to oil leaks, which also negatively affect their performance. Corrosion of metal parts, as well as contamination of the oil by external impurities, are other significant factors that can significantly shorten the lifespan of these devices.

The impact of the external environment plays a significant role in the reliability of oil circuit breakers. Temperature fluctuations can alter the viscosity of the oil, which in turn affects its ability to perform its functions effectively. Air pollution and increased humidity can accelerate the degradation process of the oil, necessitating more frequent inspections and maintenance [11-12]. These devices are also sensitive to the presence of water in the oil, which can lead to gas formation and increased pressure inside the device, disrupting its operation.

The frequency of maintenance for oil circuit breakers is significantly higher than that of vacuum counterparts. The oil must be regularly replaced, and the device's sealing must be checked. Cleaning and inspecting contacts also play an important role, as contamination or corrosion of the contacts can cause failure.

Compared to oil circuit breakers, vacuum circuit breakers require less attention and maintenance (Table 1). Vacuum circuit breakers are smaller in size and weight and also demonstrate higher durability [13]. Moreover, they virtually require no regular maintenance, making them more cost-effective in the long term. In contrast, oil circuit breakers require frequent maintenance, oil replacement, and monitoring of component conditions, which increases their operating costs and reduces overall reliability. Environmentally, vacuum circuit breakers are also superior to oil breakers, as they do not contain hazardous liquids that could leak and harm the environment.

Table 1: *Comparative reliability of vacuum and oil circuit breakers*

Parameter	Vacuum Circuit Breakers	Oil Circuit Breakers
Mean Time Between Failures	High (up to 50,000 hours)	Medium (up to 20,000 hours)
Maintenance Requirement	Minimal	High (oil replacement and seal check)
Environmental Impact	High (no oil)	Low (risk of oil leakage)
Size and Weight	Compact	Large and heavy
Durability	Long-lasting	Shorter durability
Sensitivity to External Factors	Resistant to contamination and temperature fluctuations	Sensitive to contamination and humidity
Dependence on Oil	None	Issues with oil quality
Risk of Leakage	None	Present (oil)

Thus, vacuum circuit breakers, due to their design, reduced maintenance requirements, and durability, offer higher reliability and cost-effectiveness, especially in modern operating conditions where minimizing maintenance costs and extending equipment lifespan are crucial.

III. Problem solution

The reliability of switching equipment is a set of its properties that allows it to maintain operability for a specified period under certain operating conditions. In the energy sector, where the failure of a switching device can lead to serious consequences, including disruption of power supply and equipment damage, reliability becomes a key criterion when selecting circuit breakers [14-15].

According to regulatory documents, reliability includes the device's ability to operate without failure, durability, maintainability, and preservation. Reliability is quantitatively assessed using indicators such as Mean Time Between Failures (MTBF), Mean Time To Repair (MTTR), failure rate (λ), and overall probability of no-failure operation R(t).

Mean Time Between Failures (MTBF) is defined as the ratio of total operating time to the number of failures and reflects the average duration of uninterrupted equipment operation. For oil circuit breakers, this value is typically around 12,000 hours, while for vacuum circuit breakers, it reaches 50,000 hours or more.

The failure rate, which is the reciprocal of MTBF, is calculated using the formula:

$$\lambda = \frac{1}{MTBF}$$

For oil circuit breakers:

$$\lambda = \frac{1}{12000} \approx 8.33 \times 10^{-5} \text{ failures per hour}$$

For vacuum circuit breakers:

$$\lambda = \frac{1}{50000} = 2 \times 10^{-5} \text{ failures per hour}$$

Thus, the failure intensity of vacuum circuit breakers is significantly lower, which indicates their higher reliability.

The average recovery time after a failure also plays an important role: for vacuum circuit breakers, it is on average 1.5 hours, while for oil circuit breakers it is about 4 hours, which is due to the need for oil replacement, cleaning, and additional maintenance.

To assess the probability of failure-free operation over time, the reliability function is used, which reflects the exponential decrease in reliability at a constant failure rate. It is calculated using the formula:

$$R(t) = e^{-\lambda t}$$

where t is the operating time of the equipment.

Table 2 includes the main parameters used for calculating the function R(t)=e−λt.

Table 2: *Comparative reliability parameters of vacuum and oil circuit breakers*

Type of Circuit Breaker	Mean Time Between Failures (MTBF), hours	Mean Time to Repair (MTTR), hours	Failure Rate λ, 1/hour
Oil Circuit Breaker	12,000	4.0	8.33×10^{-5}
Vacuum Circuit Breaker	50,000	1.5	2.00×10^{-5}

For the vacuum circuit breaker at t = 1000 hours:

$$R_{vacuum}(1000) = e^{-2\times10^{-5}\times1000} = e^{-0.02} \approx 0.9802$$

Thus, after 1000 hours of operation, the probability that the vacuum circuit breaker will remain in good condition is 98.02%.

For the oil circuit breaker at t = 1000 hours:

$$R_{oil}(1000) = e^{-8{,}33\times10^{-5}\times1000} = e^{-0.0833} \approx 0.9207$$

That is, after 1000 hours of operation, the probability that the oil circuit breaker will remain in good condition is 92.07%.

As can be seen, the lower failure rate of vacuum circuit breakers ensures a consistently high value of R(t) over a long operational period [16-18]. The graph in Figure 1 shows the dependence of R(t) on time for vacuum and oil circuit breakers. It is evident that the reliability of the vacuum circuit breaker remains at a significantly higher level throughout its entire lifecycle, making it the preferred option for modern power systems.

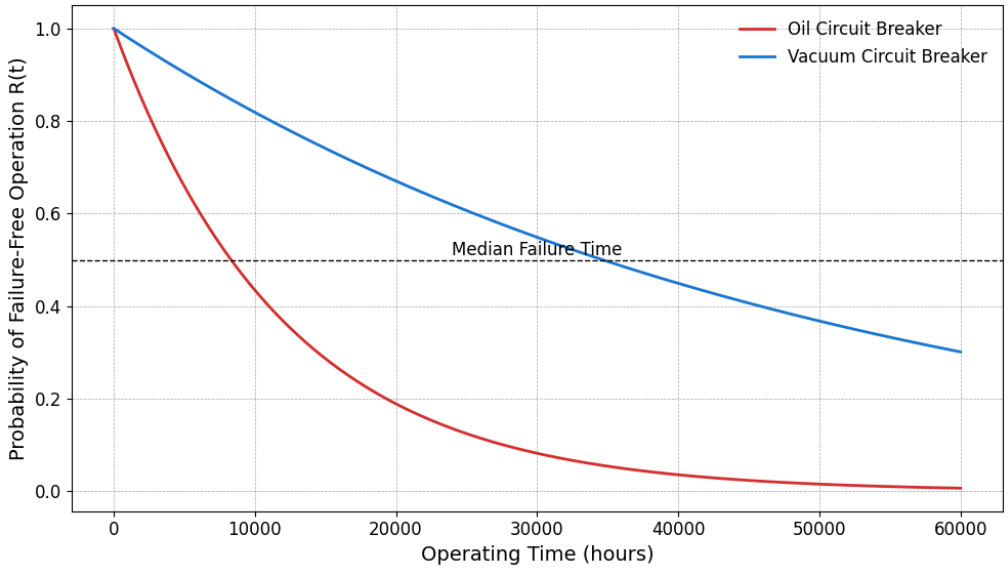


Figure 1. *Reliability function R(t) graph for vacuum and oil circuit breakers.*

For comprehensive evaluation, methods of statistical analysis of operational data, reliability modeling, and failure analysis by design components are often used [19]. Practice shows that the main causes of failure in oil circuit breakers are related to oil leaks, insulation degradation, and contact burn-out. In contrast, vacuum circuit breakers demonstrate high stability even with a large number of switching cycles, thanks to their sealed design and the lack of need for maintenance of the contact medium.

IV. Conclusions

As a result of the comparative analysis, it has been established that vacuum circuit breakers offer higher reliability and durability compared to oil circuit breakers. They require minimal maintenance and demonstrate excellent Mean Time Between Failures (MTBF) performance, making them more cost-effective in the long term. Oil circuit breakers, despite their durability, require frequent maintenance, oil replacement, and regular seal checks, which increase operating costs and reduce their overall reliability. Therefore, vacuum circuit breakers are the preferred choice for modern power systems, offering greater stability and a lower risk of failure during operation.

References

[1] Rzayeva, S. V., Ganiyeva, N. A., & Piriyeva, N. M. (2024). Modern approaches to electrical equipment diagnostics. *International Journal on Technical and Physical Problems of Engineering (IJTPE)*, 16(1), 182–189.

[2] Rzayeva, S. V., Piriyeva, N. M., & Guseynova, I. A. (2024). Analysis of reliability of typical power supply circuits. *Reliability: Theory and Applications (RTA)*, 19(3), 173–178.

[3] Rahimli, I., Karimova, R., & Piriyeva, N. (2024). Development of a two-parameter inductive sensor for use in digital program control systems for machine tools. *Przegląd Elektrotechniczny*, 100(8), 194–197.

[4] Rzayeva, S. V., Ganiyeva, N. A., & Piriyeva, N. M. (2023, October 31). Modern methods of diagnostics of electric power equipment. In *Proceedings of the 19th International Conference on Technical and Physical Problems of Engineering* (pp. 105–110). International Organization of IOTPE, Romania.

[5] Piriyeva, N. M., Rzayeva, S. V., & Mustafazadeh, E. M. (2022). Evaluation of the application of various methods and equipment for protection from emergency voltage in 6–10 kV electric networks of oil production facilities. *Internauka: Elektronnyy Nauchnyy Zhurnal*, (39), 40–44.

[6] Rzayeva, S. V., Mammadov, N. S., & Ganiyeva, N. A. (2023). Overvoltages during single-phase earth fault in neutral-isolated networks (10–35) kV. *Journal of Energy Research and Reviews*, 13(1), 7–13.

[7] Rzayeva, S. V., & Guseynova, I. A. (2024). Application of automatic monitoring and control systems for reliability of power transmission lines. *Reliability: Theory & Applications*, 19(2), 64–69.

[8] Rzayeva, S., & Ganiyeva, N. (2022). Analysis of the design and operational advantages of vacuum circuit breakers. *Universum: Tekhnicheskiye Nauki*, (10-7), 39–42.

[9] Piriyeva, N. M., Mammadov, N. S., & Rzayeva, S. V. (2025). Development of new methods for protecting substation and overhead lines from overvoltages. *Reliability: Theory & Applications*, 19(1), 683–689. https://doi.org/10.24412/1932-2321-2025-182-683-689

[10] Rzayeva, S. V., Mammadov, N. S., & Ganiyeva, N. A. (2022). Neutral grounding mode in the 6–35 kV network through an arcing reactor and organization of relay protection against single-phase ground faults. *German International Journal of Modern Science*, (42).

[11] Ahmedov, E., Rzayeva, S., Ganiyeva, N., & Safiyev, E. (2023). Improving the lightning resistance of high-voltage overhead power line. *Przegląd Elektrotechniczny*, (11).

[12] Piriyeva, N. M., Abdullayeva, G. K., & Bakhtiyarov, A. L. (2024). Engineering approaches to minimizing the environmental impact of thermal power plants. *International Journal on Technical and Physical Problems of Engineering (IJTPE)*, 61, 231–243.

[13] Mammadov, N., Piriyeva, N., & Ismayilova, S. (2024). Research of lightning protection systems for wind electric installations. *Przegląd Elektrotechniczny*, (9), 2024.

[14] Piriyeva, N. M., & Agazade, N. S. (2024). Исследование эксплуатационных характеристик современных вакуумных выключателей. *Vestnik Nauki*, 3(1), 963–974.

[15] Su, J., Chen, K., & Wang, Q. (2022). Difference analysis of operation stability of vacuum circuit breaker in 10kV power distribution system. *Journal of Physics: Conference Series*, 2202(1), 012047.

[16] Shores, R. B., & Phillips, V. E. (2006). High voltage vacuum circuit breakers. *IEEE Transactions on Power Apparatus and Systems*, 94(5), 1821–1830.

[17] Bumblauskas, D., Meeker, W., & Gemmill, D. (2012). Maintenance and recurrent event analysis of circuit breaker data. *International Journal of Quality & Reliability Management*, 29(5), 560–575.

[18] Slade, P. G., et al. (2007). Vacuum interrupter, high reliability component of distribution switches, circuit breakers and contactors. *Journal of Zhejiang University-Science A*, 8(3), 335–342.

[19] König, D. (2012). The role of vacuum in circuit breaker technology. In *2012 25th International Symposium on Discharges and Electrical Insulation in Vacuum (ISDEIV)* (pp. A1–A14). IEEE.

RT&A, No 2 (84)
Volume 20, June 2025

RELIABILITY ANALYSIS OF THE SHAFT SUBJECTED TO FLUCTUATING LOADS UNDER TORSION AND BENDING MOMENTS WHEN SHEAR STRESS FOLLOWS EXPONENTIAL DISTRIBUTION

Md. Yakoob Pasha[1], M. Tirumala Devi[2], T. Sumathi Uma Maheswari[3].

•

[1,2,3] Department of Mathematics Kakatiya University, Warangal-506009, Telangana, India.
[1]mypasha139@gmail.com, [2]oramdevi@yahoo.com, [3]tsumathiuma@gmail.com

Abstract

The reliability of shafts subjected to fluctuating loads is essential for ensuring the safe and efficient operation of rotating machinery in various engineering applications, including automotive, aerospace, and industrial systems. Accurate reliability assessments are crucial to preventing failures and extending the lifespan of mechanical components. This study examines the reliability analysis of shafts subjected to combined torsional and bending moments, which create complex stress states that significantly affect shaft performance and failure behaviour. A statistical model is considered since the shear stress generated is random in nature. Analytical expressions for reliability are derived by changing the parameters twisting moment, bending moment, and shaft diameter.

Keywords: Reliability, circular solid shaft, combined twisting and bending moment, fluctuating loads, exponential distribution.

1. Introduction

Shafts are critical components in mechanical systems, designed to transmit power between rotating elements. These components are commonly subjected to complex loading conditions, including torsion and bending moments, during operation. In conventional design approaches, it is often assumed that the torque and bending moments acting on a shaft are constant. However, in practical applications, such as line shafts and counter shafts, shafts frequently experience fluctuating loads due to dynamic operating conditions. These fluctuations can induce combined effects of shock and fatigue, significantly influencing the reliability and lifespan of the shaft. To ensure the safe and reliable operation of rotating machinery, it is essential to perform a reliability analysis that accounts for the stochastic nature of the loading conditions. Shear stress, a critical factor influencing shaft failure, can vary unpredictably under fluctuating loads. Modeling shear stress as a random variable using an exponential distribution provides a practical framework for evaluating the reliability of shafts under such conditions.

Adekunle A. A. et al. [1] studied the development of CAD software for shafts under various loading conditions. Dr. Edward E. Osakue et al. [2] studied fatigue shaft design verification for bending and torsion. Dr. Edward E. Osakue et al. [3] studied the probabilistic fatigue design of shafts for bending and torsion. Fatima K. et al. [4] discussed the statistical properties of the exponential Rayleigh distribution and its application to medical sciences and engineering. Frydrysek K. et al. [5] studied performance-based design applied to a shaft subjected to combined stress.

Gowtham et al. [6] studied drive shaft design and analysis. Kamboh, M. S., et al. [8] discussed the design and analysis of the drive shaft with a critical review of advanced composite materials and the root causes of shaft failure. Misra, A., et al. [9] studied the reliability analysis of drilled shaft behaviour using the finite difference method and Monte Carlo simulation. Nadarajah S. [10] studied reliability for lifetime distributions through mathematical and computer modelling. Nayek et al. [11] studied the reliability approximation for a solid shaft under a gamma setup. Patel, B., et al. [12] studied a critical review of the design of a shaft with multiple discontinuities and combined loadings. T. S. Uma Maheswari et al. [14] obtained reliability analysis of unsymmetrical columns subjected to eccentric loads for stress following exponential distribution. Villa-Covarrubias, B., et al. [15] discussed the probabilistic method to determine the shaft diameter and design reliability.

If the shaft is subjected to fluctuating loads that include both torsion (twisting moment) and bending moment, the diagram illustrates the relationship between time and torsion, as well as between time and bending moment.

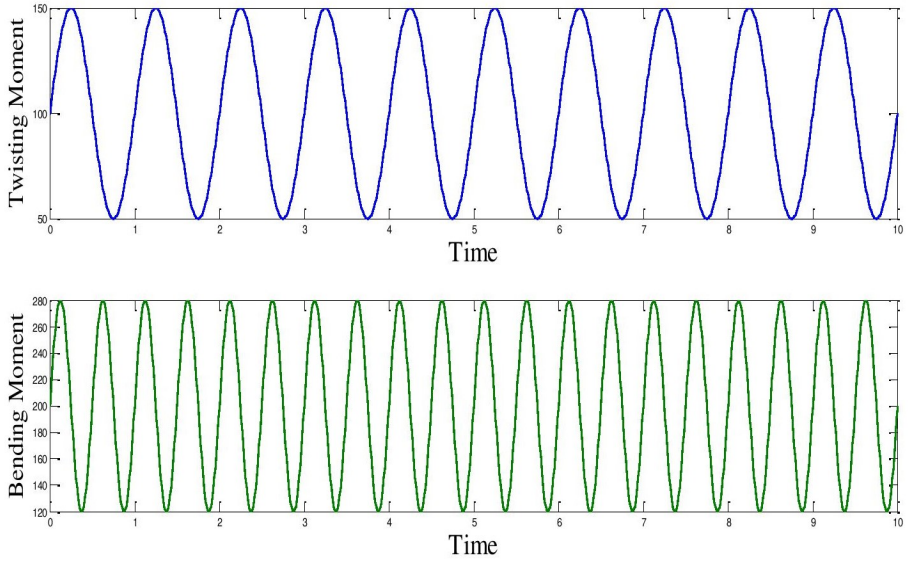


Figure 1: *Shaft subjected to fluctuating loads*

2. Statistical model

The probability density function for exponentially distributed strength ξ and for exponentially stress χ is given by [7]

$$f_{\xi}(\xi) = \lambda_{\xi}\, e^{-\lambda_{\xi} \times \xi} \text{ for } 0 \leq \xi < \infty,\ \lambda_{\xi} > 0$$
$$f_{\chi}(\xi) = \lambda_{\chi}\, e^{-\lambda_{\chi} \times \chi} \text{ for } 0 \leq \chi < \infty,\ \lambda_{\chi} > 0$$

Reliability of the stress-strength system is the probability that the strength ξ is greater than the stress χ applied on it is given by [13]

$$R = \int_0^{\infty} f_{\chi}(\chi) \left[\int_{\chi}^{\infty} f_{\xi}(\xi)\ d\xi \right] d\chi = \int_0^{\infty} \lambda_{\chi} \exp\left[-(\lambda_{\xi} + \lambda_{\chi})\chi \right]\ d\chi \tag{1}$$

Then the reliability is

$$R = \frac{\bar{\xi}}{\bar{\xi} + \bar{\chi}} \tag{2}$$

Let T be the torsion, M be the maximum bending moment, and d be the diameter of the shaft. For

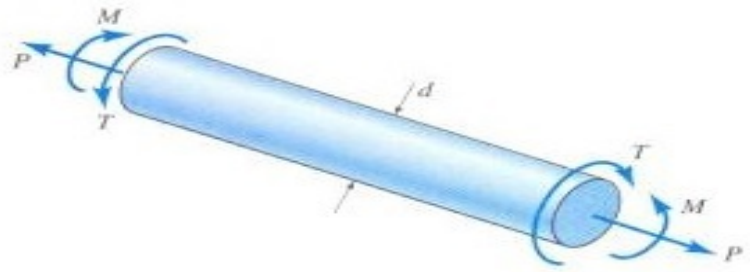


Figure 2: *Shaft subjected to combined torsion and bending moment*

a shaft subjected to combined torsion and maximum bending, the equivalent twisting moment is given by [13]

$$T_e = \sqrt{(K_m \times M)^2 + (K_t \times T)^2} \tag{3}$$

and equivalent bending moment is given by [13]

$$M_e = \frac{1}{2}\left[K_m \times M + \sqrt{(K_m \times M)^2 + (K_t \times T)^2}\right] \tag{4}$$

where K_m is combined shock and fatigue factor for bending and K_t is combined shock and fatigue factor for torsion.

Recommended values for K_m and K_t

Nature of load for stationary shafts	K_m	K_t
(a) Gradually applied load	1.0	1.0
(b) Suddenly applied load	1.5 to 2.0	1.5 to 2.0

Nature of load for rotating shafts	K_m	K_t
(a) Gradually applied load or steady load	1.5	1.0
(b) Suddenly applied load with minor shocks only	1.5 to 2.0	1.5 to 2.0
(c) Suddenly applied load with heavy shocks	2.0 to 3.0	1.5 to 3.0

The equivalent twisting moment is given by [13]

$$T_e = \frac{\pi}{16} \times \sigma_t \times d^3$$

Then the shear stress due to twisting moment is

$$\sigma_t = \frac{16 \times T_e}{\pi \times d^3} \tag{5}$$

The equivalent bending moment is given by [13]

$$M_e = \frac{\pi}{32} \times \sigma_b \times d^3$$

Then the shear stress due to bending moment is

$$\sigma_b = \frac{32 \times M_e}{\pi \times d^3} \tag{6}$$

From equations (2) and (5), the reliability of the shaft subjected to torsion with the exponentially distributed strength and stress is

$$R_t = \frac{\bar{\xi}}{\bar{\xi} + \left(\frac{16 \times T_e}{\pi \times d^3}\right)} \tag{7}$$

From equations (2) and (6), the reliability of the shaft subjected to bending moment with exponentially distributed strength and stress is

$$R_b = \frac{\bar{\xi}}{\bar{\xi} + \left(\frac{32 \times M_e}{\pi \times d^3}\right)} \tag{8}$$

3. Numerical results and discussion

Numerical results and discussions on the shaft reliability under combined torsion and bending moments, where stress and strength follows exponential distribution have been presented.

Table 1: $M = 100000$ *N-mm*, $d = 52$ *mm*, $\bar{\xi} = 119.6584$ *N/mm*^{2}.

T	$R_{t(K_m=K_t=0.1)}$	$R_{b(K_m=K_t=0.1)}$	$R_{t(K_m=K_t=1.5)}$	$R_{b(K_m=K_t=1.5)}$
100000	0.995739087	0.992748016	0.939684303	0.901246468
2500000	0.929622187	0.927014622	0.468255256	0.458510624
5000000	0.868567560	0.866290841	0.305827816	0.301640839
7500000	0.815025042	0.813020043	0.227048647	0.224732753
10000000	0.767697469	0.765918310	0.180539737	0.179072384
13000000	0.717685823	0.716130686	0.144916817	0.143969874
18000000	0.647393843	0.646128147	0.109053373	0.108516258
20000000	0.622986891	0.621814744	0.099230436	0.098785527
25000000	0.569327148	0.568348068	0.080992053	0.080695416
30000000	0.524177967	0.523347903	0.068417071	0.068205276

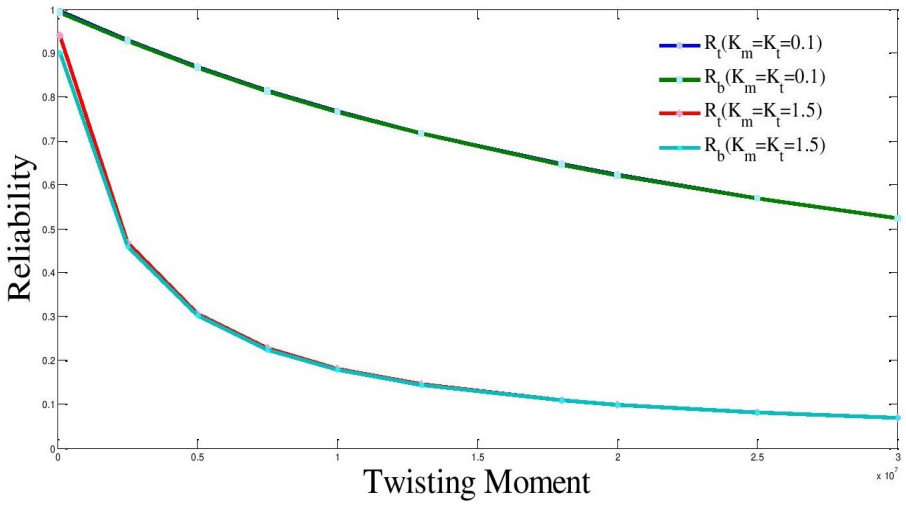


Figure 3: *As the twisting moment increases, reliability of the shaft decreases.*

Table 2: $T = 100000$ *N-mm*, $d = 52$ *mm*, $\bar{\xi} = 119.6584$ *N/mm*^{2}.

M	$R_{t(K_m=K_t=0.1)}$	$R_{b(K_m=K_t=0.1)}$	$R_{t(K_m=K_t=1.5)}$	$R_{b(K_m=K_t=1.5)}$
100000	0.995739087	0.992748016	0.939684303	0.901246468
1200000	0.964845268	0.932187659	0.646607085	0.478198768
2000000	0.942869764	0.891974623	0.523867963	0.355035226
2400000	0.932242274	0.873132117	0.478414437	0.314511825
3800000	0.896844329	0.813007064	0.366930648	0.224717879
4800000	0.873156128	0.774887696	0.314558563	0.186649213
6500000	0.835630527	0.717679829	0.253131424	0.144913151
7800000	0.809040000	0.679326476	0.220240429	0.123751748
8600000	0.793501206	0.657696792	0.203933346	0.113547822
12000000	0.733616403	0.579304536	0.155119255	0.084082248

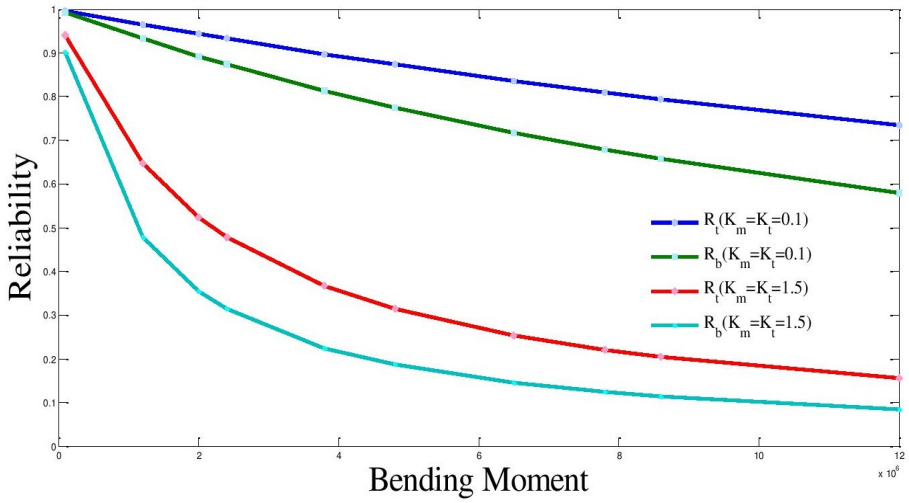


Figure 4: *As the bending moment increases, reliability of the shaft decreases.*

Table 3: $M = 100000$ *N-mm*, $T = 150000$ *N-mm*, $\bar{\xi} = 119.6584$ *N/mm*^{2}.

d	$R_{t(K_m=K_t=0.1)}$	$R_{b(K_m=K_t=0.1)}$	$R_{t(K_m=K_t=1.5)}$	$R_{b(K_m=K_t=1.5)}$
12	0.692585922	0.551516771	0.130583035	0.075770650
16	0.842278898	0.744567852	0.262547956	0.162709726
20	0.912513114	0.850595313	0.410152097	0.275125262
24	0.947433525	0.907731222	0.545778442	0.396084264
28	0.966239826	0.939839580	0.656126570	0.510159686
32	0.977128452	0.958880686	0.740135750	0.608554105
36	0.983826470	0.970762730	0.802187593	0.688815533
40	0.988157591	0.978515805	0.847626441	0.752253845
44	0.991076342	0.983771903	0.881010619	0.801643759
48	0.993112390	0.987453455	0.905771942	0.839920119

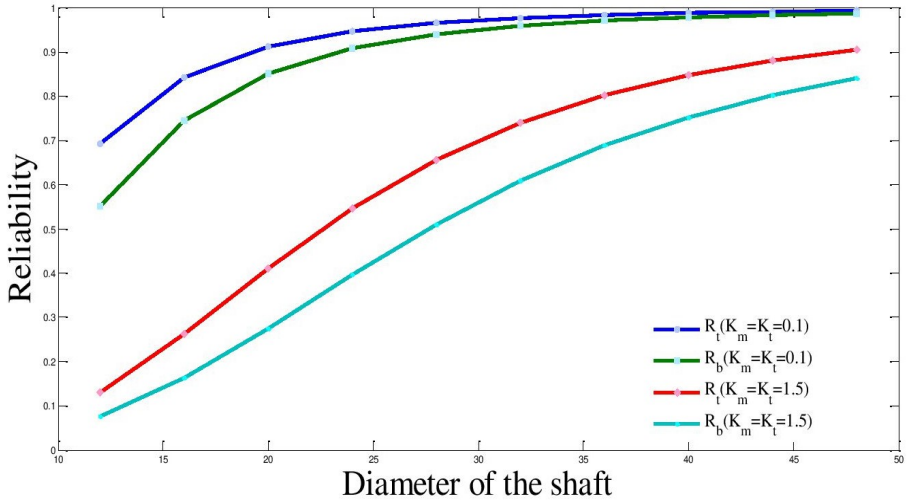


Figure 5: *As the diameter of the shaft increases, reliability of the shaft increases.*

Table 4: $M = 200000$ *N-mm*, $T = 400000$ *N-mm*, $d = 52$ *mm*.

ξ	$R_{t(K_m=K_t=0.1)}$	$R_{b(K_m=K_t=0.1)}$	$R_{t(K_m=K_t=1.5)}$	$R_{b(K_m=K_t=1.5)}$
20.5	0.926796658	0.897417429	0.457711994	0.368374369
40.5	0.961556741	0.945304732	0.625115863	0.535360857
65.5	0.975875765	0.965459724	0.729496231	0.650770585
102.5	0.984448600	0.977649248	0.808436196	0.744642702
136.5	0.988276798	0.983122507	0.848944006	0.795223260
162.5	0.990134004	0.985784519	0.869970282	0.822160764
234.5	0.993142450	0.990106003	0.906147153	0.869646132
396.5	0.995932878	0.994124692	0.942279746	0.918568479
684.5	0.997640061	0.996588265	0.965733068	0.951156933
984.5	0.998358009	0.997625431	0.975923594	0.965527473

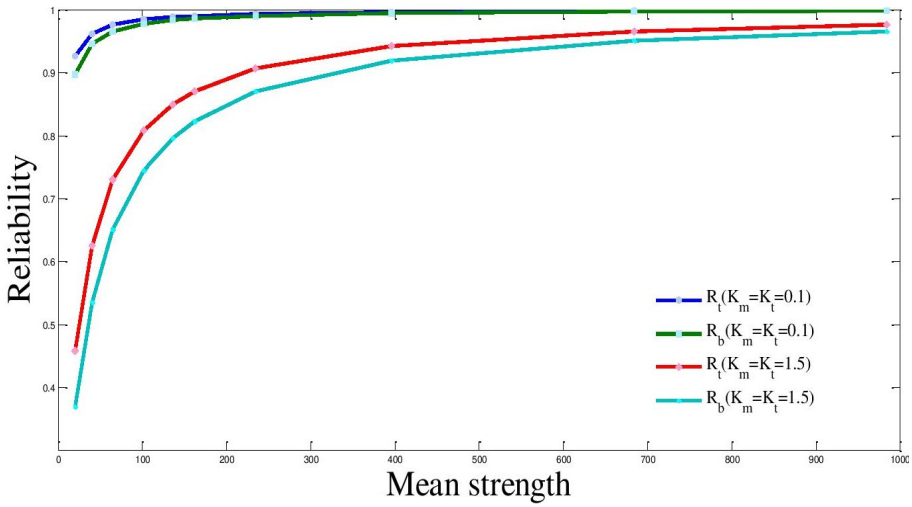


Figure 6: *As the mean strength of the shaft increases, reliability of the shaft also increases.*

4. Conclusion

The reliability of a shaft is affected when subjected to combined torsion and bending moments. According to the computations, reliability of the shaft decreases as the twisting and bending moments increase. Conversely, the reliability increases when the diameter and strength of the shaft increase.

References

[1] Adekunle, A. A., Adejuyigbe, S. B., & Arulogun, O. T. (2012). Development of CAD software for shaft under various loading conditions. Procedia Engineering, 38, 1962-1983.
[2] Dr. Edward E. Osakue et al., Fatigue Shaft Design Verification for Bending and Torsion´ International Journal of Engineering Innovation and Research, pp.197-206, 2015.
[3] Dr. Edward E. Osakue, Probabilistic Fatigue Design of Shaft for Bending and Torsion´, International Journal of Research in Engineering and Technology, Volume: 03, Issue: 09, pp. 370-386, 2014, DOI: 10.15623/ijret.2014.0309059.
[4] Fatima K., & Ahmad S. P. (2017), Statistical properties of Exponential Rayleigh distribution and its application to medical sciences and engineering. In International Conference on Recent Innovations in Science, Agriculture, Engineering and Management University College of Computer Applications Guru Kashi University, Bathinda, Punjab (India) (pp. 491-506).
[5] Frydrysek, K. Performance-Based design applied for a Shaft subjected to combined stress.
[6] Gowtham, S., Raaghul, V., Sreeharan, B. N., Kumar, R. P., & Kasim, S. M. (2021, April). Drive Shaft Design and Analysis for Mini Baja. In IOP Conference Series: Materials Science and Engineering (Vol. 1123, No. 1, p. 012015). IOP Publishing.
[7] K. C. Kapur and L. R. Lamberson, Reliability in Engineering Design´, Wiley India Private Limited, New Delhi, 1997.
[8] Kamboh, M. S., Machhi, M. A., & Kamboh, M. F. (2020). Design and Analysis of Drive Shaft with a Critical Review of Advance Composite Materials and the Root Causes of Shaft Failure.
[9] Misra, A., Roberts, L.A. & Levorson, S.M. Reliability analysis of drilled shaft behaviour using finite difference method and Monte Carlo simulation. Geotech Geol Eng 25, 65“77 (2007). https://doi.org/10.1007/s10706-006-0007-2.
[10] Nadarajah S. (2003), Reliability for lifetime distributions mathematical and computer modelling, 37 (8“9), 683-688.
[11] Nayek, S., Seal, B., & Roy, D. (2014). Reliability approximation for solid shaft under Gamma setup. Journal of Reliability and Statistical Studies, 11-17.
[12] Patel, B., Prajapati, H. R., & Thakar, D. B. (2014). Critical Review on design of shaft with multiple discontinuities and combined loadings ICCIET“2014.
[13] R. S. Khurmi and J. K. Gupta A Textbook of Machine Design´, Eurasia Publishing House (PVT.), Limited, Ram Nagar, New Delhi, 2005.
[14] T. S. Uma Maheswari, Reliability Analysis of Resultant Stress for Unsymmetrical Columns for Stress Follows Exponential Distribution´, International Journal of Engineering and Advanced Technology, 2019, DOI: 10.35940/ijeat. F1234.0986S319.
[15] Villa-Covarrubias, B., Baro-Tijerina, M., & Pina-Monarrez, M. R. Probabilistic Methodology to Determine the Shaft™s Diameter and Designed Reliability.
[16] Zhao, W., Jiang, Z., Zhang, P., & Huang, X. (2023). Reliability Sensitivity Analysis of Main Shaft Bearings of Wind Turbines Subject to Subsurface Stress´, Machines, 11(7), 681, https://doi.org/10.3390/machines11070681.

RT&A, No 2 (84)
Volume 20, June 2025

ANALYSIS OF TWO VACATION POLICIES UNDER RETRIAL ATTEMPTS, MARKOVIAN ENCOURAGED ARRIVAL QUEUING MODEL

E. Ismailkhan[1*], V. Narmadha[2], N. Yathavan[1]

•

[1,*] Department of Mathematics, Panimalar Engineering College, Chennai-600123
[2]Department of Mathematics, Vellore Institute of Technology, Vellore- 632014
[1,*]ismailkhanmubarak@gmail.com, [2]narmadha.v2020a@vitstudent.ac.in

Abstract

This paper presents a comprehensive analysis of a single-server Markovian queueing system incorporating two distinct vacation policies under a framework of encouraged arrivals and retrial attempts. The novelty of the model lies in its simultaneous treatment of several realistic operational dynamics: server vacations with acceleratory growth, customer impatience, server breakdown and repair, and encouraged arrivals, where potential customers are motivated to join the system depending on its state. To reflect real-world retrial scenarios, customers who do not receive immediate service reattempt after some time, modeled using a retrial queue mechanism. Two key vacation policies are compared: one governed by a first-come, first-served (FCFS) discipline and the other utilizing bulk service. The paper derives steady-state probabilities and evaluates system performance measures, such as mean queue length, server utilization, and the expected number of retrials.

Keywords: Two vacation, encouraged arrival, reneging, retrial queue, markov model.

I. Introduction

Markovian queuing models, which follow encouraged arrival rates and exponential service rates, are used in a variety of systems, including manufacturing, production, telecommunications, computers, and transportation. Everyone has a hectic schedule and little free time in the modern world. Because the customer's arrival is unpredictable, they cannot complete their task in the allotted time because they cannot predict it.

We consider the single server to be encouraged for working vacations, breakdowns, and repairs where the server may be down at any time. A Poisson distribution governs customer arrival rates, whereas an exponential distribution governs customer service rates. Arriving customers will join the orbit group if they discover that the server is too busy serving another customer. In [1], a Markovian queue which has dependent arrival and breakdown is investigated. A single vacation

on an unstable bulk server is detailed in [2]. Batch arrivals on an infinite server are investigated in [3]. Examples of batch queues with inadequate identification are explored in [4].

A generalized bulk queue with the Poisson model is addressed in [5]. In [6], a general bulk queuing approach with a dual working vacation. In [7], a retrial model with a consistent retrial rate, breakdowns, and dissatisfied customers were studied. A dual server queuing model for bulk arrival and service was examined in [8]. Working vacations on bulk arrival queues combined with reneging and interruptions were studied in [9]. Dual service and two vacations are investigated in the bulk arrival Markovian system [10]. They investigated batch arrival and retrial queues using a dual vacation policy and the Markovian queuing model [11]. In [12], an M/M/1/N system along with encouraged arrivals. Reduced wait times in an M/M/1/N encouraged customer arrivals, as seen in [13]. Reneging customers were observed in [14] when the M/M/c/N model was used in conjunction with encouraged arrivals.

II. Model Elaboration

The following presumptions have been taken into account:

- The customers follow the First-Come-First-Serve rule
- Customers arrive according to an encouraged arrival with - mean $\frac{\mathbf{1}}{\boldsymbol{\lambda*(1+\omega)}}$, $\boldsymbol{\omega}$ representing the offered value and accelerated distribution with mean $\frac{\mathbf{1}}{\boldsymbol{\mu}}$
- Customers will enter the retrial queue with probability "s"
- After a few tries, the customer notices the accelerated distribution with an average of $1-\boldsymbol{\varphi}$ and attempts the request from the retrial space.
- When a customer attempts to receive service after a certain amount of time has passed, with probability $\boldsymbol{\delta_0}$
- The server will fail time following the accelerated distribution under the fail rate $\boldsymbol{\nu_1}$ and from the service to the customer under the repair rate $\boldsymbol{\nu_2}$.
- We have given that the random variable J(t) describes the absolute customers of the system at period 't', and R(t) = 0, 1, 2, 3…. consider the states, a server is free, a server is busy, server is on vacation and the server is in breakdown & repair state at period "t".

III. Analysis of System-size distribution

From first state $0, \mathrm{j} \geq 0$ when the- the server is idle- state:

$$\lambda * (1+\omega)\overline{\mathrm{P}}_{0,0} = \mu\overline{\mathrm{P}}_{1,0} \tag{1}$$

$$(\lambda * (1+\omega) + \mathrm{l}\psi\delta_0)\overline{\mathrm{P}}_{0,\mathrm{l}} = \mathrm{l}\mu\overline{\mathrm{P}}_{1,\mathrm{l}} \tag{2}$$

$$(\lambda * (1+\omega) + \mathrm{j}\varphi\delta_0)\overline{\mathrm{P}}_{0,\mathrm{j}} = \mathrm{j}\mu\overline{\mathrm{P}}_{1,\mathrm{j}} \tag{3}$$

From second state $1, j \geq 0$ when the server is busy- state:

$$(\nu_1 + \Delta_0 + \mathrm{s}\lambda * (1+\omega) + \mu)\overline{\mathrm{P}}_{1,0} = (1-\mathrm{s})\varphi\overline{\mathrm{P}}_{1,1} + \nu_2\overline{\mathrm{P}}_{3,0} + \lambda * (1+\omega)\overline{\mathrm{P}}_{2,0} + \varphi\delta_0\overline{\mathrm{P}}_{0,1} + \lambda * (1+\omega)\overline{\mathrm{P}}_{0,0}$$

$$(1 - \Delta_0 + \mathrm{s}\lambda * (1+\omega) + \mathrm{l}\mu + \nu_1)\mathrm{P}_{1,\mathrm{l}} \tag{4}$$

$$= \nu_2\overline{\mathrm{P}}_{3,\mathrm{l}} + \mu\overline{\mathrm{P}}_{2,\mathrm{l}} + \lambda * (1+\omega)\overline{\mathrm{P}}_{0,\mathrm{l}} + \mathrm{s}\lambda * (1+\omega)\overline{\mathrm{P}}_{1,\mathrm{l}-1} + (1-\mathrm{s})(\mathrm{l}+1)\varphi\mathrm{P}_{1,\mathrm{l}+1} + (\mathrm{l}+1)\varphi\delta_0\overline{\mathrm{P}}_{0,\mathrm{l}+1}\big(1 - \Delta_0 + \nu_1 + \mathrm{j}\mu + (1-\mathrm{s})\mathrm{j}\varphi + \mathrm{s}\lambda * (1+\omega)\big)\overline{\mathrm{P}}_{1,\mathrm{j}} \tag{5}$$

$$= \mu\overline{P}_{2,j} + \nu_2\overline{P}_{3,j} + \lambda * (1+\omega)\overline{P}_{0,j} + s\lambda * (1+\omega)\overline{P}_{1,-1}$$

$$+s\lambda * (1+\omega)(j+1)\varphi\overline{P}_{1,j+1} + (j+1)\varphi\delta_0\overline{P}_{0,j+1} \quad (6)$$

From third state 2, when the server is on vacation:

$$2\lambda * (1+\omega)\overline{P}_{2,0} = \Delta_0\overline{P}_{0,1} \quad (7)$$

$$\big(\mu + \lambda * (1+\omega)\big)\overline{P}_{2,l} = (1-\Delta_0)P_{1,l} + \lambda * (1+\omega)\overline{P}_{1,l-1} \quad (8)$$

$$\mu\overline{P}_{2,j} = (1-\Delta_0)\overline{P}_{1,j} + \lambda * (1+\omega)\overline{P}_{1,j-1} \quad (9)$$

For fourth state 3, when the- server break down and is repaired state:

$$\nu_2\overline{P}_{3,0} = \nu_1\overline{P}_{1,0} \quad (10)$$

$$\nu_2\overline{P}_{3,l} = \nu_1\overline{P}_{1,l} \quad (11)$$

$$\nu_2\overline{P}_{3,j} = \nu_1\overline{P}_{1,j} \quad (12)$$

The following conclusions may be drawn from equations (1), (2), and (3).

$$\begin{aligned} \overline{P}_{1,0} &= \left(\frac{\lambda*(1+\omega)}{\mu}\right)\overline{P}_{0,0} \\ \overline{P}_{1,m} &= \left(\frac{\lambda*(1+\omega)+l\varphi\delta_0}{l\mu}\right)\overline{P}_{0,l} \\ \overline{P}_{1,j} &= \left(\frac{\lambda*(1+\omega)+j\varphi\delta_0}{j\mu}\right)\overline{P}_{0,j} \end{aligned} \quad (13)$$

Using (4), now with j = 0 and (10), we obtain

$$(\nu_1 + \Delta_0 + s\lambda * (1+\omega) + \mu)\overline{P}_{1,0} = (1-s)\varphi\overline{P}_{1,1} + \nu_1\overline{P}_{1,0} + \lambda * (1+\omega)\overline{P}_{2,0} + \varphi\delta_0\overline{P}_{0,1} + \lambda * (1+\omega)\overline{P}_{0,0}$$

$$(\Delta_0 + s\lambda * (1+\omega) + \mu)\overline{P}_{1,0} = (1-s)\varphi\overline{P}_{1,1} + \lambda * (1+\omega)\overline{P}_{2,0} + \varphi\delta_0\overline{P}_{0,1} + \lambda * (1+\omega)\overline{P}_{0,0}$$

If we use (1), we obtain

$$\left(\frac{\Delta_0\lambda * (1+\omega) + s(\lambda * (1+\omega))^2}{\mu}\right)\overline{P}_{0,0} = (1-s)\varphi\overline{P}_{1,1} + \lambda * (1+\omega)\overline{P}_{2,0} + \varphi\delta_0\overline{P}_{0,1}$$

When k=1, use (13)

$$\begin{aligned} \overline{P}_{1,1} &= \left(\frac{\lambda*(1+\omega)+\varphi\delta_0}{\mu}\right)\overline{P}_{0,1} \\ \overline{P}_{1,1} &= a_1\overline{P}_{0,1} \\ \left(\frac{\Delta_0\lambda*(1+\omega)+s(\lambda*(1+\omega))^2}{\mu}\right)\overline{P}_{0,0} &= \end{aligned} \quad (14)$$

$$(1-s)\varphi\ a_1\overline{P}_{0,1} + \lambda * (1+\omega)\overline{P}_{2,0} + \varphi\delta_0\overline{P}_{0,1}$$

If we use (7)

$$\left(\frac{\Delta_0\lambda*(1+\omega)+s(\lambda*(1+\omega))^2}{\mu}\right)\overline{\mathrm{P}}_{0,0}=\left[(1-\mathrm{s}+\mathrm{a}_1+\delta_0)\varphi+\frac{\Delta_0}{2}\right]\overline{\mathrm{P}}_{0,1}$$

$$\overline{\mathrm{P}}_{0,1}=\frac{2(\Delta_0\lambda*(1+\omega)+s(\lambda*(1+\omega)\Delta)^2)}{\mu\left[(1-\mathrm{s}+\mathrm{a}_1+\delta_0)\varphi+\frac{\Delta_0}{2}\right]}\overline{\mathrm{P}}_{0,0}$$

$$\overline{\mathrm{P}}_{0,1}=\frac{\mathrm{b}_2}{\mathrm{c}_2}\overline{\mathrm{P}}_{0,0} \tag{15}$$

Applying j=1 in (6)

$$[1-\Delta_0+\nu_1+\mu+(1-s)\varphi+s\lambda*(1+\omega)]\overline{\mathrm{P}}_{1,1}$$

$$=\mu\overline{\mathrm{P}}_{2,1}+\nu_2\overline{\mathrm{P}}_{3,1}+\lambda*(1+\omega)\overline{\mathrm{P}}_{0,1}+s\lambda*(1+\omega)\overline{\mathrm{P}}_{1,0}+2s\lambda*(1+\omega)\varphi\overline{\mathrm{P}}_{1,2}+2\varphi\delta_0\overline{\mathrm{P}}_{0,2}$$

To do this, use (12) and assign j=1.

$$[1-\Delta_0+\nu_1+(1-\mathrm{s})\varphi+\mathrm{s}\lambda(1+\omega)+\mu]\overline{\mathrm{P}}_{1,1}=\mu\overline{\mathrm{P}}_{2,1}+\nu_1\overline{\mathrm{P}}_{1,1}+\lambda(1+\omega)\overline{\mathrm{P}}_{0,1}+\mathrm{s}\lambda(1+\omega)\overline{\mathrm{P}}_{1,0}+2s_1\lambda(1+\omega)\overline{\mathrm{P}}_{1,2}+2\varphi\delta_0\overline{\mathrm{P}}_{0,2}$$

$$[1-\Delta_0+(1-\mathrm{s})\varphi+\mu+\mathrm{s}\lambda(1+\omega)]\overline{\mathrm{P}}_{1,1}=\mu\overline{\mathrm{P}}_{2,1}+\lambda(1+\omega)\overline{\mathrm{P}}_{0,1}+\mathrm{s}_{1,0}\overline{\mathrm{P}}_{1,0}+$$

$$2s_1\lambda(1+\omega)\overline{\mathrm{P}}_{1,2}+2\varphi\delta_0\overline{\mathrm{P}}_{0,2}$$

If we use (9) for $j=1$ and (14), we get

$$\mu\overline{\mathrm{P}}_{2,1}=(1-\Delta_0)\overline{\mathrm{P}}_{1,1}+\lambda(1+\omega)\overline{\mathrm{P}}_{1,0}[(1-s)\varphi+\mu+s\lambda(1+\omega)]\mathrm{a}_1\overline{\mathrm{P}}_{0,1}$$

$$=\lambda(1+\omega)\overline{\mathrm{P}}_{1,0}+\lambda(1+\omega)\overline{\mathrm{P}}_{0,1}+s\lambda(1+\omega)\overline{\mathrm{P}}_{1,0}+2s\lambda(1+\omega)\delta_0\overline{\mathrm{P}}_{1,2}+2\varphi\delta_0\overline{\mathrm{P}}_{0,2}$$

$$[(1-s)\varphi\ \mathrm{a}_1+\mu\mathrm{a}_1+s\mathrm{a}_1-\lambda(1+\omega)]\overline{\mathrm{P}}_{0,1}$$

$$=\lambda(1+\omega)\overline{\mathrm{P}}_{1,0}+s\lambda(1+\omega)\overline{\mathrm{P}}_{1,0}+2s\lambda\delta_0\overline{\mathrm{P}}_{1,2}+2\varphi\delta_0\overline{\mathrm{P}}_{0,2} \tag{16}$$

Now, we use (13) put $\mathrm{j}=2$, and we obtain

$$\overline{\mathrm{P}}_{1,2}=\left(\frac{\lambda(1+\omega)+2\varphi\delta_0}{2\mu}\right)\overline{\mathrm{P}}_{0,2}=\frac{\mathrm{b}_1}{\mathrm{c}_1}\overline{\mathrm{P}}_{0,2}$$

Additionally, using (1) and (15), we can solve (16) to obtain

$$\overline{\mathrm{P}}_{0,2}=\frac{\mathrm{b}_3}{\mathrm{c}_3}\overline{\mathrm{P}}_{0,0} \tag{17}$$

$$\mathrm{b}_3=\frac{\begin{array}{c}s\lambda(1+\omega)\mathrm{a}_1\ \mathrm{b}_2+\varphi(1-s)\mathrm{a}_1\ \mathrm{b}_2\\+\mu\mathrm{a}_1\ \mathrm{b}_2-\lambda(1+\omega)\mathrm{b}_2\end{array}}{\mathrm{c}_2}$$

$$-\frac{(\lambda(1+\omega))^2(1+s)}{\mu}$$

$$\mathrm{c}_3=2s\lambda(1+\omega)\delta_0\ \mathrm{b}_1/\mathrm{c}_1+2\delta_0\varphi$$

When you enter j=2 in equations 6, 9, and 12, we get

$$\left(1-\Delta_0+\nu_1+2\mu+2(1-s)\varphi+s\lambda(1+\omega)\right)\overline{\mathrm{P}}_{1,2}$$

$$=\mu\overline{\mathrm{P}}_{2,2}+\nu_2\overline{\mathrm{P}}_{3,2}+\lambda(1+\omega)\overline{\mathrm{P}}_{0,2}$$

$$+s\lambda(1+\omega)\overline{\mathrm{P}}_{1,1}+3s\lambda(1+\omega)\varphi\overline{\mathrm{P}}_{1,3}+3\varphi\delta_0\overline{\mathrm{P}}_{0,3}\ldots \quad (18)$$

$$\mu\overline{\mathrm{P}}_{2,2}=(1-\Delta_0)\overline{\mathrm{P}}_{1,2}+\lambda(1+\omega)\overline{\mathrm{P}}_{1,1} \quad (19)$$

$$\nu_2\overline{\mathrm{P}}_{3,2}=\nu_1\overline{\mathrm{P}}_{1,2} \quad (20)$$

If we use (19) & (20) in (18), we have

$$\left(2*\mu+2(1-s)\varphi+s\lambda(1+\omega)\right)\bar{P}_{1,2}=\lambda(1+\omega)(1+s)\bar{P}_{1,1}+\lambda(1+\omega)\bar{P}_{0,2}$$

$$+3s\lambda(1+\omega)\varphi\bar{P}_{1,3}+3\varphi\delta_0\bar{P}_{0,3} \quad (21)$$

$$\left(2*\mu+2(1-s)\varphi+s\lambda(1+\omega)\right)\frac{b_1}{c_1}\bar{P}_{0,2}=\lambda(1+\omega)(1+s)\bar{P}_{1,1}+\lambda(1+\omega)\bar{P}_{0,2}$$

$$+3s\lambda(1+\omega)\varphi\bar{P}_{1,3}+3\varphi\delta_0\bar{P}_{0,3}$$

$$\left[(2*\mu+2(1-s)\varphi+s\lambda(1+\omega))\frac{b_1}{c_1}-\lambda(1+\omega)\right]\bar{P}_{0,2}$$

$$=\lambda(1+\omega)(1+s)\bar{P}_{1,1}+3s\lambda(1+\omega)\varphi\bar{P}_{1,3}$$
$$+3\varphi\delta_0\bar{P}_{0,3}\left[(2*\mu+2(1-s)\varphi+s\lambda(1+\omega))\frac{b_1}{c_1}-\lambda(1+\omega)\right]\frac{b_3}{c_3}\bar{P}_{0,0}$$

$$=\lambda(1+\omega)(1+s)a_1\bar{P}_{0,1}+3s\lambda(1+\omega)\varphi\bar{P}_{1,3}$$
$$+3\varphi\delta_0\bar{P}_{0,3}\left[(2*\mu+2(1-s)\varphi+s\lambda(1+\omega))\frac{b_1}{c_1}-\lambda(1+\omega)\right]\frac{b_3}{c_3}\bar{P}_{0,0}$$

$$=\lambda(1+\omega)(1+s)a_1\frac{b_2}{c_2}\bar{P}_{0,0}+3s\lambda(1+\omega)\varphi\bar{P}_{1,3}+3\varphi\delta_0\bar{P}_{0,3}$$

$$\left\{\left(2*\mu+2(1-s)\varphi+s\lambda(1+\omega)\right)\frac{b_1}{c_1}-\lambda(1+\omega)\right\}\frac{b_3}{c_3}\ \bar{P}_{0,0}$$
$$\left[\qquad -\lambda(1+\omega)(1+s)a_1\frac{b_2}{c_2}\qquad\right]$$

$$=3s\lambda(1+\omega)\varphi\overline{\mathrm{P}}_{1,3}+3\varphi\delta_0\bar{P}_{0,3}$$

If, we Put j = 3 in (13), and we have

$$\overline{\mathrm{P}}_{1,3}=\left(\frac{\lambda(1+\omega)+j\varphi\delta_0}{j\mu}\right)\overline{\mathrm{P}}_{0,3}=\frac{\mathrm{b}_4}{\mathrm{c}_4}\overline{\mathrm{P}}_{0,3}$$

$$\left[\begin{Bmatrix}(2*\mu+2(1-s)\varphi+\\ s\lambda(1+\omega))\frac{\mathrm{b}_1}{\mathrm{c}_1}-\lambda(1+\omega)\end{Bmatrix}\frac{\mathrm{b}_3}{\mathrm{c}_3}-\lambda(1+\omega)(1+s)\mathrm{a}_1\frac{\mathrm{b}_2}{\mathrm{c}_2}\right]\overline{\mathrm{P}}_{0,0}$$

$$= \left(3s\varphi\frac{b_4}{c_4} + 3\varphi\delta_0\right)\overline{P}_{0,3}$$

$$\overline{P}_{0,3} = \frac{\left\lfloor \begin{matrix}(2*\mu + 2(1-s)\varphi + \\ s\lambda(1+\omega))\frac{b_1}{c_1} - \lambda(1+\omega)\end{matrix}\right\}\frac{b_3}{c_3} - \lambda(1+\omega)(1+s)a_1\frac{b_2}{c_2}\rfloor}{\left(\frac{3\lambda(1+\omega)\varphi\, b_4}{c_4} + 3\varphi\delta_0\right)}\overline{P}_{0,0}$$

$$\overline{P}_{0,3} = \frac{b_5}{c_5}\overline{P}_{0,0}$$

In general, we get

$$\overline{P}_{0,n} = \overline{P}_{0,1} + \overline{P}_{0,2} + \overline{P}_{0,3} + \cdots \tag{22}$$

If we use (13), we get

$$\overline{P}_{1,j} = \left(\frac{\lambda(1+\omega)+j\psi\delta_0}{j\mu}\right)\left[\overline{P}_{0,1} + \overline{P}_{0,2} + \overline{P}_{0,3} + \cdots\right] \tag{23}$$

If, we use (9)

$$\overline{P}_{2,j} = (1-\Delta_0)\left(\frac{\lambda(1+\omega) + j\Sigma\delta_0}{j*\mu^2}\right)$$

$$\left[\overline{P}_{0,1} + \overline{P}_{0,2} + \overline{P}_{0,3} + \cdots\right] + \frac{\lambda(1+\omega)\overline{P}_{1,j-1}}{\mu} \tag{24}$$

Similarly, if, we use (12), we get

$$\overline{P}_{3,j} = \left(\frac{\nu_1}{\nu_2}\right)\left[\overline{P}_{0,1} + \overline{P}_{0,2} + \overline{P}_{0,3} + \cdots\right] \tag{25}$$

Now, if we use equations (22), (23), (24), and (25), we obtain

$$\overline{P}_{r,n} = \begin{cases} \overline{P}_{0,1} + \overline{P}_{0,2} + \overline{P}_{0,3} + \cdots \text{ for } s = 0 \\ \left(\frac{\lambda(1+\omega) + j\varphi\delta_0}{j}\right)\left[\overline{P}_{0,1} + \overline{P}_{0,2} + \overline{P}_{0,3} + \cdots\right], \text{for } s = 1 \\ (1-\Delta_0)\left(\frac{\lambda(1+\omega) + j\varphi\delta_0}{j*\mu^2}\right)\left[\overline{P}_{0,1} + \overline{P}_{0,2} + \overline{P}_{0,3} + \cdots\right] + \frac{\lambda(1+\omega)\overline{P}_{1,j-1}}{\mu}, \\ \text{for } s = 2, \left(\frac{\nu_1}{\nu_2}\right)\left[\overline{P}_{0,1} + \overline{P}_{0,2} + \overline{P}_{0,3} + \cdots\right], \text{for } s = 3 \end{cases}$$

To calculate the value of $\bar{P}_{0,0}$, the normalization function, we have

$$\sum_{s=0}^{3}\sum_{j=0}^{\infty} P_{i,s} = 1$$

$$\left(\overline{P}_{0,1} + \overline{P}_{0,2} + \overline{P}_{0,3}\right)\left(1 + \frac{\lambda(1+\omega) + j\varphi\delta_0}{j*\mu^2} + \frac{\nu_1}{\nu_2}\right)$$

$$+\frac{\lambda(1+\omega)}{\mu}\overline{P}_{1,j-1} = 1$$

$$\overline{\mathrm{P}}_{0,0}\left(\frac{\mathrm{b}_2}{\mathrm{c}_2}+\frac{\mathrm{b}_3}{\mathrm{c}_3}+\frac{\mathrm{b}_5}{\mathrm{c}_5}\right)\left(1+\frac{\lambda(1+\omega)+\mathrm{j}\varphi\delta_0}{\mathrm{j}*\mu^2}+\frac{\nu_1}{\nu_2}\right)$$

$$+\frac{\lambda(1+\omega)}{\mu}\mathrm{P}_{1,\mathrm{j}-1}=1$$

$$\overline{\mathrm{P}}_{0,0}=\frac{1-\frac{\lambda(1+\omega)}{\mu}\overline{\mathrm{P}}_{1,\mathrm{j}-1}}{\left(\frac{\mathrm{b}_2}{\mathrm{c}_2}+\frac{\mathrm{b}_3}{\mathrm{c}_3}+\frac{\mathrm{b}_5}{\mathrm{c}_5}\right)\left(1+\frac{\lambda(1+\omega)+\mathrm{j}\varphi\delta_0}{\mathrm{j}*\mu^2}+\frac{\nu_1}{\nu_2}\right)}$$

IV. Validation of the model

(i) When the server is free:

$$\bar{P}_0 \quad =\sum_{j=0}^{\infty}\bar{P}_{0,j}$$

$$=\frac{1-\frac{\lambda(1+\omega)}{\mu}\bar{P}_{1,j-1}}{\left(1+\frac{\lambda(1+\omega)+j\varphi\delta_0}{j*\mu^2}+\frac{\nu_1}{\nu_2}\right)\left(\frac{b_2}{c_2}+\frac{b_3}{c_3}+\frac{b_5}{c_5}\right)}$$

$$\left[\bar{P}_{0,1}+\bar{P}_{0,2}+\bar{P}_{0,3}+\cdots\right]$$

(ii) When the server is busy:

$$\bar{P}_1 \quad =\sum_{j=0}^{\infty}\bar{P}_{1,j}$$

$$=\frac{1-\frac{\lambda(1+\omega)}{\mu}\bar{P}_{1,j-1}}{\left(1+\frac{\lambda(1+\omega)+j\varphi\delta_0}{j*\mu^2}+\frac{\nu_1}{\nu_2}\right)\left(\frac{b_2}{c_2}+\frac{b_3}{c_3}+\frac{b_5}{c_5}\right)}$$

$$\left(\frac{\lambda(1+\omega)+j\varphi\delta_0}{j*\mu}\right)\times\left[\bar{P}_{0,1}+\bar{P}_{0,2}+\bar{P}_{0,3}+\cdots\right]$$

(iii) When the server is on vacation:

$$\bar{P}_2 \quad =\sum_{j=0}^{\infty}P_{2,j}$$

$$=\frac{1-\frac{\lambda(1+\omega)}{\mu}\bar{P}_{1,j-1}}{\left(1+\frac{\lambda(1+\omega)+j\varphi\delta_0}{j*\mu^2}+\frac{\nu_1}{\nu_2}\right)\left(\frac{b_2}{c_2}+\frac{b_3}{c_3}+\frac{b_5}{c_5}\right)}$$

$$\times\left[(1-\Delta_0)\left(\frac{\lambda(1+\omega)+j\varphi\delta_0}{j*\mu^2}\right)\right]$$

$$\left[\bar{P}_{0,1}+\bar{P}_{0,2}+\bar{P}_{0,3}+\cdots\right]+\frac{\lambda(1+\omega)\bar{P}_{1,j-1}}{\mu}$$

(iv) When the server is in a break down and repaired state:

$$\bar{P}_3=\sum_{j=0}^{\infty}P_{3,j}$$

$$=\frac{1-\frac{\lambda(1+\omega)}{\mu}\bar{P}_{1,j-1}}{\left(1+\frac{\lambda(1+\omega)+j\varphi\delta_0}{j*\mu^2}+\frac{v_1}{v_2}\right)\left(\frac{b_2}{c_2}+\frac{b_3}{c_3}+\frac{b_5}{c_5}\right)\left(\frac{v_1}{v_2}\right)}$$

$$\times\left[\bar{P}_{0,1}+\bar{P}_{0,2}+\bar{P}_{0,3}+\cdots\right]$$

V.Conclusion

The Markovian Encouraged Arrival Queuing Model has been developed with the inclusion of customer retry efforts, balking, and reneging behavior. The four system states—idle state, busy state, vacation state, breakdown state, and repair state—have all been taken into consideration utilizing the concept of encouraged arrival. We have examined and verified the possibilities of the various conditions. Neural networks, communication systems, post offices, and supermarkets can all benefit from using this model to reduce the reneging and balking behavior of their customers.

References

[1] Kalyanaraman, R., & Sundaramoorthy, A. (2019). A Markovian working vacation queue with server state dependent arrival rate and with partial breakdown. *International Journal of Recent Technology and Engineering, 7*(62), 664–668.

[2] Haridass, M., & Arumuganathan, R. (2008). Analysis of a bulk queue with unreliable server and single vacation. *International Journal of Open Problems in Computational Mathematics, 1*(2), 37–55.

[3] Daw, A., & Pender, J. (2019). On the distributions of infinite server queues with batch arrivals. *Queueing Systems, 91*(3–4), 367–401.

[4] Bar-Lev, S. K., Parlar, M., Perry, D., Stadje, W., & Van Der Duyn Schouten, F. A. (2007). Applications of bulk queues to group testing models with incomplete identification. *European Journal of Operational Research, 183*(1), 226–237.

[5] Neuts, M. F. (1967). A general class of bulk queue with Poisson input. *Annals of Mathematical Statistics, 38*(3), 759–770.

[6] Parveen, M. J., & Begum, M. I. A. (2013). General bulk service queueing system with multiple working vacation. *International Journal of Mathematics Trends and Technology, 4*(9), 163–173.

[7] Li, H., & Zhao, Y. Q. (2005). A retrial queue with a constant retrial rate, server breakdowns and impatient customers. *Stochastic Models, 21*(2–3), 531–550.

[8] Kumar, J., & Shinde, V. (2018). Performance evaluation of bulk arrival and bulk service with multi-server using queue model. *International Journal of Research in Advent Technology, 6*, 3069–3076.

[9] Vijaya Laxmi, P., & Rajesh, P. (2018). Variant working vacations on batch arrival queue with reneging and server breakdowns. *East African Scholars Multidisciplinary Bulletin, 1*(1), 7–20.

[10] Srivastava, R. K., Singh, S., & Singh, A. (2020). Bulk arrival Markovian queueing system with two types of services and multiple vacations. *International Journal of Mathematics and Computer Research, 8*(8), 2130–2136.

[11] Singh, S., & Srivastava, R. K. (2021). Markovian queueing system for bulk arrival and retrial attempts with multiple vacation policy. *International Journal of Mathematics Trends and Technology*.

[12] Som, B. K., & Seth, S. (2017). An M/M/1/N queuing system with encouraged arrivals. *Global Journal of Pure and Applied Mathematics, 13*(7).

[13] Khan, I. E., & Paramasivam, R. (2022). Reduction in waiting time in an M/M/1/N encouraged arrival queue with feedback, balking and maintaining of reneged customers. *Symmetry, 14*(8), 1743.

[14] Som, B. K., & Seth, S. (2018). M/M/c/N queuing systems with encouraged arrivals, reneging, retention and feedback customers. *Yugoslav Journal of Operations Research*, *28*(3), 333-344.

RT&A, No 2 (84)
Volume 20, June 2025

CONSTRUCTING A SEPARATE RESERVE IN A CHAIN OF LOW-RELIABILITY ELEMENTS

* Gurami Tsitsiashvili, ** Yury Kharchenko

Russia, 690041, Vladivostok, Radio street 7, IAM FEB RAS,
* guram@iam.dvo.ru, ** har@iam.dvo.ru

Abstract

In this paper, we construct upper and lower estimates of the minimum separate reserve of low-reliability chain elements of length m, at which the probability of operability of the entire chain becomes close to unity. The appeal to a separate reserve is caused by the properties of this reserve, established by Barlow and Proschan. It is shown that the upper and lower estimates of the minimum separate reserve are quite close. If the probability of operability of an individual element of the chain is $1/\ln m$, then the minimum reserve volume is of the order $\ln^2 m$.

Keywords: a chain of low-reliability elements, a minimum amount of separate reserve, shortest paths in a two-pole.

1. Introduction

In the article of Erdos and Renyi [1], it was found that the probability of connectivity $P(m)$ for a complete graph with m vertices (in which there is an edge between any pair of vertices) and independently functioning edges with the probability $p(m) = \dfrac{c\ln m}{m}$, $m \to \infty$, satisfies the limit relations

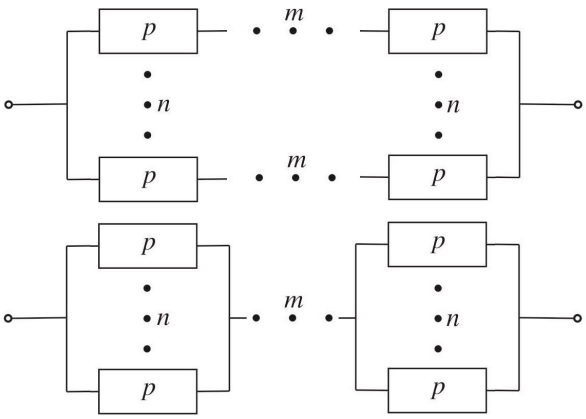
$$\lim_{m\to\infty} P(m) = 1,\ c > 1,\ \lim_{m\to\infty} P(m) = 0,\ c < 1. \tag{1}$$

This property of a complete random graph resembles a phase transition in physical systems. The development of these results using modern graph-theoretic methods is described in the article of A.M. Raigorodsky [2] and in numerous articles on dynamic processes in random networks. A feature of the first limiting relation in (1) is the fact that a random graph with a high probability of connectivity is formed from low-reliability edges. Such random graphs are used to model the Internet and many other networks [3] –[6].

If, instead of a complete graph, we consider a chain of m sequentially connected edges, then with a high probability of operability of the edges, $p(m) = exp(-m^{-a}) \to 1,\ m \to \infty$ it is not difficult to establish that the probability of the chain ends being connected is $P(m) \to 0,\ a < 1$ and, conversely, $P(m) = \exp(-m^{1-a}) \to 1,\ a > 1$. In other words, for a chain of highly reliable edges, a phase transition is also observed with a variation of the parameter a in the vicinity of point 1.

However, chains with low-reliability edges do not have high reliability. Therefore, the problem arises of reserving parts of the chains, among which the most significant are the models of block and separate redundancy. Here, block redundancy is understood as the redundancy of two-poles in their entirety, i.e. their connection by the initial and final nodes. And separate redundancy means reserving each edge separately. In the classic monograph on the theory of reliability by Barlow and Proschan on [7], it was found that the probability of connectivity of the initial and

final vertices of a bipolar in the case of a separate reserve is always higher than in the case of a block reserve.



A multiple n block reserve (top), a separate n reserve (bottom) of a chain with a length m.

In particular, when analysing two-poles represented by chains of length m, it was found that the minimum volume (the number of two-poles connected in parallel with the ends) of the block reserve exceeds a certain geometric progression growing by m. And the minimum volume of the split reserve (the number of individual edges connected in parallel by the ends) is limited from above by the logarithmic m function [8]. This result is also applicable to bipolar structures with m edges. With such an advantage of a separate reserve over a block one, it is natural to consider the task of constructing a highly reliable chain from low-reliability edges with separate redundancy. Moreover, the properties of the edge chain established for a separate reserve can be extended to two-poles.

2. Methods

Consider a chain (serial connection) m independently functioning edges with a probability of p, assuming $q = 1 - p,\ 0 < q < 1$. Because at $p \to 0$ the probability of the chain functioning If p^m also tends to zero, then convergence of this probability to unity is possible only if this chain is reserved.

Let each edge of the chain be n-fold reserved by independently functioning edges. Then the probability of a workable connection in such a chain between the initial and final vertices satisfies the equality $P_n(m) = (1 - q^n)^m$. Let us focus on determining the minimum volume of the separate reserve chain $n_*(m, \delta) = \min(n :\ P_n(m) \geq 1 - \delta)$ for some given $\delta,\ 0 < \delta < 1$. Following the tradition established in probability theory and mathematical statistics, it is natural to assume the value δ small ($\delta \ll 1$), for example, $\delta = 0.1,\ 0.05,\ 0.01$.

Estimation of the value of $n_*(m, \delta)$ We will carry out calculations using the following ratio. For any $a,\ 0 < a < 1$,, the obvious inequality holds

$$1 - ma \leq (1 - a)^m \leq 1 - ma + \frac{(ma)^2}{2}. \tag{2}$$

Then

$$\min\left(n :\ mq^n - \frac{(mq^n)^2}{2} \leq \delta\right) \leq n_*(m, \delta) \leq \min(n :\ mq^n \leq \delta) \tag{3}$$

To define $n_*(m, \delta)$ denote $x = mq^n$ and calculate the roots of the two equations $x = \delta,\ x^2/2 - x + \delta = 0$. The root of the first equation is $x_1 = \delta$, and the root of the second equation is $x_2 = 1 \pm \sqrt{1 - 2\delta}$. Hence, the inequality follows from the first equation

$$n_*(m, \delta) \leq \left[\frac{\ln m - \ln \delta}{-\ln q}\right] + 1, \tag{4}$$

where $[z]$ is the integer part of the real number z. In turn, the second equation implies the inequality

$$n_*(m,\delta) \geq \frac{\ln m - \ln(1-\sqrt{1-2\delta})}{-\ln q}. \tag{5}$$

Remark 1. For $\delta \ll 1$, we may approximately write $1-\sqrt{1-2\delta} \approx \delta$. This means that for large m, the upper and lower estimates of the value of $n_*(m,\delta)$ are almost the same.

3. Results

Inequalities (4), (5) are valid for any $q = 1-p,\ 0 < q < 1$, in particular for $q = q(m) \to 1,\ p = p(m) \to 0,\ m \to \infty$. By analogy with the technique of asymptotic estimation of the probability of connectivity of a complete graph [1], [2], we assume $p(m) \to 0,\ m \to \infty$. Then from formulas (4), (5) we obtain the following estimates for the required reserve volume

$$\frac{\ln m - \ln(1-\sqrt{1-2\delta}))}{\ln(1-p(m))} \leq n_*(m,\delta) \leq \left[\frac{\ln m - \ln\delta}{-\ln(1-p(m))}\right] + 1. \tag{6}$$

Let's now move on to the low-reliability edge model when $p(m) \to 0,\ m \to \infty$. Then, for example, for $p(m) = 1/\ln m \to 0,\ m \to \infty$, we get

$$\ln m(\ln m - \ln(1-\sqrt{1-2\delta})) \leq n_*(m,\delta) \leq \left[\frac{\ln m(\ln m - \ln\delta)}{1-1/\ln m}\right] + 1. \tag{7}$$

It follows from the formula (7) that the required volume of the separate reserve of the edges of the chain is proportional to $(\ln m)^2$ for $m \gg 1$.

If $p(m) = 1/\ln^b m$, then the minimum volume of the separate reserve becomes of the order $\ln^{b+1} m,\ 0 < b$. There are other ways to determine $p(m),\ n_*(m,\delta)$, for example, when $p(m) = 1/\ln\ln m$, the minimum reserve volume becomes on the order $\ln\ln m \cdot \ln m$.

4. Discussion

When switching from chains to bipolar circuits with low-reliability edges, the task arises of enumerating the shortest paths along which the minimum volumes of a separate reserve can be built. To do this, we can use a modification of the Dijkstra algorithm [9]. However, along with this, it will be necessary to consider the possible intersections of the shortest paths from the initial to the final vertex of the bipolar.

5. Conclusion

The upper bound in the formula (6) can be extended to a bipolar with m edges using the calculations of [8]. For $\delta \ll 1$, the upper bound in the formula (7) asymptotically coincides with the lower bound for $m \to \infty$. It should also be noted that by allocating the shortest path between the initial and final vertices in a bipolar using Dijkstra's algorithm (or some other algorithm, see, for example, [9]), it is possible to construct a separate reserve of all its edges using the algorithm of this paper.

The research was carried out within the state assignment for IAM FEB RAS N 075-00459-25-00.

References

[1] Erdos, P., Renyi, A. (1960) On the Evolution of Random Graphs. Publication of the Mathematical Institute of the Hungarian Academy of Sciences, No 5, pp. 17-61.

[2] Raigorodskiy, A. M. (2010) Models of random graphs and their applications. Proceedings of MIPT, No 2(4), pp. 130-140.
[3] Barabasi, L. A., Albert, R. (1999) Emergence of scaling in random networks. Science, No 286, pp. 509-512.
[4] Dorogovtsev, S. N., Mendes, J. F. F. (2002) Evolution of Networks Adv. Phys., No 51(4), pp. 1079-1187.
[5] Ginzburg, S. L., Nakin, A. V., Savitskaya, N. E. (2009) The influence of the structure of a complex network on the properties of dynamic processes on it. Letters to the JETF, No 90(12), pp. 873-878.
[6] Evin, I. A. (2010) Introduction to the theory of complex networks. Computer research and modelling, No 2(2), pp. 121-141.
[7] Barlow, R. E., Proschan, F. (1965) Mathematical Theory of Reliability. John Wliley and Sons. New York, NY.
[8] Tsitsiashvili, G. Sh. (2021) Study of Synergistic Effects in Complex Stochastic Systems. Mathematics, No 9(12), 1396.
[9] Tsitsiashvili, G. Sh. (2024) Determination of stability and reliability of shortest paths in a graph through lists of labels in Dijkstra's algorithm. Reliability: Theory and Applications, No 4(80), pp. 49-54.

EVALUATING THE IMPACT OF PARAMETER SENSITIVITY IN META-HEURISTICS ON RELIABILITY PERFORMANCE

Shakuntla Singla [1], Manisha Rani [2*], Shilpa Rani [3], Umar Muhammad Modibbo [4]

•

[1,2,3]Department of Mathematics and Humanities, MMEC, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala, India
[4]Department of Statistics and Operations Research, School of Physical Sciences, Modibbo Adama University of Technology, YOLA-NIGERIA
[1]shakus25@gmail.com, [2*]manishasethi2026@gmail.com, [3] gargshilpa46@gmail.com
[4]umarmodibbo@mau.edu.ng

Abstract

The optimizing reliability of a system is a crucial target for any industry to have good values and outcomes with the balanced cost of maintenance. The assessment of costs and reliability in intricate systems, such as Distributed Redundancy Architecture (DRA), is essential for optimal system design. Achieving this involves addressing various constraints to enhance the overall system's reliability. Researchers have explored system dependability and cost optimization challenges, emphasizing the evolution of metaheuristic approaches. Enhancing system resilience and optimizing costs are central objectives of this study, aiming to improve the reliability and cost-efficiency of DRA systems. The study's framework builds upon existing metaheuristic methods, including Moth Flame Optimization (MFO), Whale Optimization Algorithm (WOA), Dragonfly Algorithm (DA), Gazelle Optimization Algorithm (GOA), and Coati Optimization Algorithm (COA). The findings reveal that COA demonstrates superior performance compared to other methods. By employing the Coati Optimization Algorithm, this study effectively addresses cost and reliability-optimization challenges, providing highly efficient solutions. Key metrics such as reliability and cost were compared across the five metaheuristic approaches, showcasing COA's advantages. Notably, the Coati Optimization Algorithm delivers faster resolutions and outperforms its counterparts in reducing costs while enhancing system reliability. The results underscore COA's ability to optimize intricate system parameters more effectively than methods like MFO, WOA, GOA, and DA. This research highlights the value of using COA for improving reliability and minimizing expenses in complex systems. It offers a novel approach to addressing these challenges, establishing COA as a powerful tool for optimizing distributed system architectures. By analyzing comparative solutions, the study found that COA consistently achieved better outcomes, affirming its success in reliability and cost optimization. In conclusion, this research emphasizes the importance of advanced metaheuristic techniques, particularly the Coati Optimization Algorithm, in enhancing the performance of Distributed Redundancy Architectures. By presenting a numerical example, the comparison has been made and the outcomes are shown graphically and in the tabular form to have a clear understanding. The findings contribute significantly to the field by providing a robust method for achieving cost-efficient and reliable system designs.

Keywords: Reliability, Cost, Distributed Redundancy Architecture (DRA), Metaheuristic algorithms.

I. Introduction

Reliability is crucial in complex systems like aerospace and telecommunications. Distributed Redundancy Architecture (DRA) enhances fault tolerance by distributing redundancy across system components. Meta-heuristic algorithms like GA, PSO, and SA optimize reliability but are highly sensitive to parameter settings. Improper calibration can lead to suboptimal redundancy allocation, affecting system performance. Complex systems consist of interconnected elements that exhibit emerging behaviors and challenges. Their reliability is vital as failures can trigger cascading effects. As these systems grow and integrate more components, ensuring dependability while managing costs becomes increasingly important.

II. Literature Review

Optimizing system reliability is crucial for fault-tolerant systems, with Distributed Redundancy Architecture (DRA) enhancing reliability by distributing redundant components. Effective redundancy allocation depends on optimization techniques, often using meta-heuristic algorithms. While these algorithms offer adaptability, their sensitivity to parameter settings remains a challenge, requiring further study. Beji et al. [1] The Redundancy Allocation Problem (RAP) is a well-known reliability optimization challenge that aims to maximize system reliability by strategically adding redundant components to subsystems. The trade-off between cost, weight, and reliability complicates the optimization process, as real-world systems often impose constraints that limit the number and types of components used. RAP has applications in various engineering fields, such as telecommunications, power systems, and defense technologies. Umamaheswari et al. [2] Preventive maintenance is crucial for ensuring the safe and reliable operation of power generation units, which are critical components of the energy infrastructure. The planning and execution of maintenance activities for generating units directly impact their operational efficiency, safety, and reliability. Traditional maintenance strategies, such as time-based and condition based maintenance, have been employed for decades to mitigate equipment failures. However, these strategies are often limited in their ability to balance reliability, risk, and cost. To address these challenges, recent research has focused on developing more advanced approaches, particularly those that integrate reliability and risk assessments into preventive maintenance planning, while ensuring cost-effectiveness.

Dahiya et al. [3] In the field of optimization, finding a global optimum in a complex search space is a critical problem across various domains, including engineering design, machine learning, and operational research. Metaheuristic algorithms have proven to be effective in solving such optimization problems, as they offer powerful and flexible solutions that can efficiently explore and exploit the search space to avoid local optima. Among the numerous metaheuristics developed over the years, the grasshopper optimization algorithm (GOA) has gained attention due to its simplicity and robustness in handling both continuous and discrete optimization problems. However, the traditional GOA can struggle with premature convergence and local optima, especially in high-dimensional and complex search spaces. To address these limitations, hybrid approaches such as the Hybrid Artificial Grasshopper Optimization Algorithm (HAGOA) have been proposed. HAGOA combines the strengths of GOA with other metaheuristic techniques to improve its exploration and exploitation capabilities, thus enhancing its ability to discover the global optimum in a given search space.

Wei et al. [4] In the context of modern industrial and engineering systems, integrated systems—comprising interconnected components working together to achieve specific tasks—have become increasingly complex. The efficient operation of such systems depends not only on their individual

components but also on the collective reliability, availability, and performance of the entire system. These factors are crucial for ensuring the continuous and reliable functioning of systems across a range of sectors, including power generation, telecommunications, transportation, and manufacturing. Reliability assessment and availability capacity evaluation play a pivotal role in the design and maintenance of integrated systems, ensuring optimal performance while minimizing downtime and operational costs.Nath and Muhuri [5] The reliability-redundancy allocation problem (RRAP) is a classical optimization problem in the domain of reliability engineering. It involves determining the optimal allocation of redundancy to components in a system to maximize reliability while considering various constraints such as cost, weight, volume, and system complexity. Traditionally, RRAP has been formulated as a multi-objective optimization problem where the key objectives are often to maximize system reliability and minimize cost. However, in practical applications, especially with the increasing complexity of modern systems, RRAP has evolved into a many-objective optimization problem where more than two conflicting objectives must be optimized simultaneously.

A novel evolutionary solution approach based on objective prioritization and constraint optimization has emerged as a promising technique for addressing many-objective reliability-redundancy allocation problems. This approach combines evolutionary algorithms (EA) with mechanisms that prioritize objectives based on their relative importance and incorporate efficient handling of system constraints to explore optimal trade-offs in the solution space. In this literature review, we explore the evolution of RRAP from a multi-objective to a many-objective problem and examine the role of evolutionary algorithms in addressing these challenges. Mirjalili et al. [6-9] The reliability-redundancy allocation problem (RRAP) is a critical optimization task in reliability engineering, focusing on determining the optimal number and configuration of redundant components within a system to maximize reliability while satisfying various constraints such as cost, weight, and volume. The problem becomes more complex when multiple redundancy strategies are available, as these strategies introduce additional decision variables and trade-offs between reliability, cost, and other system requirements. Traditional optimization methods struggle efficiently solve large-scale RRAPs with multiple objectives and constraints, particularly when redundancy strategies can be selected. To address this complexity, metaheuristic algorithms have been widely employed due to their ability to explore vast solution spaces and find near-optimal solutions. Among these, the Water Cycle Algorithm (WCA) has emerged as a promising metaheuristic for solving RRAP, especially when various redundancy strategies are considered.

This literature review explores the application of the WCA for solving RRAPs, emphasizing how its unique flow-inspired mechanism can handle the complexity introduced by multiple redundancy strategies. The review also highlights the advantages of WCA over other traditional and metaheuristic approaches and discusses its applications in various reliability engineering fields. Gandhi and Bhattacharjya [10] Metaheuristic optimization techniques have gained significant attention in recent decades due to their capability of solving complex optimization problems in various fields, such as engineering design, machine learning, and operations research. One such method, the Shuffled Frog Leaping Algorithm (SFLA), is a nature-inspired algorithm based on the behavior of frogs in a pond searching for food. Agushaka et al. [11] The Dwarf Mongoose Optimization (DMO) algorithm is a relatively recent addition to the field of optimization algorithms. Inspired by the social and foraging behavior of the dwarf mongoose (*Helogale parvula*), this algorithm aims to solve complex optimization problems by mimicking the way these animals interact and find resources. This literature review will explore the development, key features, and applications of the DMO algorithm, comparing it with other optimization techniques and highlighting its strengths and limitations. Khodadadi et al. [12] Crystal structure prediction (CSP) is a critical aspect of materials science and chemistry, influencing fields ranging from drug discovery to materials engineering. Recent advancements in computational methods have significantly enhanced our ability to predict

crystal structures. Multi-objective crystal structure algorithms (MOCSAs) represent a sophisticated approach to CSP, incorporating multiple criteria into the optimization process. This literature review provides an overview of the development, methodologies, and applications of MOCSAs in crystal structure prediction.

Meta-heuristic algorithms, such as Genetic Algorithms (GA), Simulated Annealing (SA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO), have become popular for solving complex optimization problems. These algorithms use stochastic processes and rely heavily on parameters, such as population size, crossover rate, mutation rate, and exploration-exploitation trade-offs, that significantly influence their efficiency Singha et al. [13]. Bhunia et al. [14] Reliability redundancy optimization is a critical aspect of designing systems to ensure operational effectiveness and minimize failure risks. In series systems, where components are arranged sequentially, the reliability of the entire system is dependent on the reliability of each individual component. With multiple choices for redundancy configurations, optimizing these choices becomes complex. Genetic algorithms (GAs) have emerged as a promising method for tackling such optimization problems due to their adaptability and efficiency in exploring large search spaces.

Mettas et al. [15] Reliability allocation and optimization are critical aspects of designing and managing complex systems across various domains, including aerospace, automotive, electronics, and industrial systems. As systems become increasingly complex, ensuring their reliability while optimizing performance and cost presents significant challenges. Abd Alsharify and Hassan [16] the reliability of complex systems is a crucial aspect in various fields, including engineering, manufacturing, and information technology. Optimizing this reliability involves improving the system's performance while minimizing failures and downtime. In recent years, heuristic and metaheuristic optimization algorithms have gained traction in solving complex reliability optimization problems. One such algorithm is the Bat Algorithm (BA), which is inspired by the echolocation behavior of bats. This review explores the application of the Bat Algorithm in optimizing complex system reliability, comparing it with other optimization techniques and highlighting recent advancements and findings.

Ravi [17] Complex systems, characterized by intricate interdependencies and dynamic behaviors, present significant challenges in ensuring reliability and performance. Optimization of such systems is crucial for enhancing their reliability while minimizing costs and downtime. The Great Deluge Algorithm (GDA), inspired by the metaphor of a gradually rising flood, is a robust heuristic method for solving optimization problems.

Coit and Zio [18] the evolution of system reliability optimization has been a dynamic and transformative journey, reflecting the growing complexity of systems and the increasing demand for dependable performance across various industries. Initially, the field emerged from the broader discipline of reliability engineering, which focuses on ensuring that systems perform their intended functions without failure over a specified period. Rocco et al. [19] The application of cellular evolutionary approaches to the reliability optimization of complex systems represents a fascinating convergence of biological inspiration and engineering principles. This approach draws from the concept of cellular automata and evolutionary algorithms to tackle the challenges associated with ensuring the reliability of intricate systems with numerous interacting components. Khorsidi et al. [20] System reliability optimization is a crucial area in engineering and operations research, focusing on improving the performance and dependability of complex systems. As systems become increasingly intricate, traditional optimization methods often fall short due to their inability to handle the non-linearity, high dimensionality, and uncertainty inherent in such problems. Meta-heuristic algorithms, which are inspired by natural processes or human strategies, have emerged as effective tools for addressing these challenges. This review explores two prominent meta-heuristic approaches—Genetic Algorithms (GAs) and Particle Swarm Optimization (PSO)—and compares their efficacy in solving complex system reliability optimization problems.

Zhang et al. [21] Redundancy allocation problems (RAP) and redundant array of parts (RRAP) are critical areas in reliability engineering and system design, particularly in systems with mixed redundancy strategies. The k-Out-of-n: G system, a class of reliability models where the system functions if at least k out of n groups are operational, plays a pivotal role in understanding and optimizing system reliability. This literature review explores the advancements in RAP and RRAP optimization for k-Out-of-n: G systems with mixed redundancy strategies, focusing on novel general models and their implications. Singla et al. [22] has discussed with help of deep learning optimization of packing unit in series with repair and single never fails. Singla et al. [23] sensitive analysis of availability of working time understand of effective pararmeter. Singla et al. [24] with RPGT performance in all three unit in working state two or more state fail then system is reduce state. Singla et al. [25] with help of fuzzy linguistic approach analysis of two unit are repair system.

III. Methodology

I. Moth Flame Optimization Algorithm (MFO)

The Moth Flame Optimization (MFO) algorithm is a metaheuristic optimization algorithm inspired by the natural behavior of moths flying towards a light source.
Here's a basic mathematical representation of the Moth Flame Optimization:

- **Initialization:** A population of moths is initialized randomly. Each moth M_j has a position represented as a vector y_j in the search space, where $y_j \in \mathbb{R}^m$, and m is the number of dimensions of the search space.
- **Fitness Function:** Each moth is evaluated using a fitness function $f(y_j)$. This function is problem-dependent and measures how good the solution represented by y_j is.
- **Update Positions:** Moths are updated based on their positions relative to the position of the best solution (flame) found so far. The position update for moth M_j can be mathematically represented as

$$y_j^{t+1} = y_j^t + \lambda \cdot (y_f - y_j^t) \cdot \exp\left(\frac{-dist(y_j^t; y_f)}{a.t^2}\right)$$

where: y_j^t is the position of moth j at time t; y_f is the position of the best flame (best solution); λ is a scaling factor; $dist(y_j^t; y_f)$ is the Euclidean distance between y_j^t and t is the current iteration

- **Flame Update:** After updating the positions of all moths, the best solution found so far (flame) is updated.
- **Termination:** The algorithm iterates until a stopping criterion is met, such as a maximum number of iterations or a convergence threshold.

II. Whale Optimization Algorithm (WOA)

The Whale Optimization Algorithm (WOA) is a metaheuristic optimization technique inspired by the hunting behavior of humpback whales.

III. Mathematical Representation

Encircling Prey: The position of the whale is updated to encircle the prey (optimal solution) using the following equations: $Y(t+1) = Y^* - A \cdot |C \cdot Y^* - Y(t)|$

where:
$Y(t)$ is the current position of the whale at iteration t; Y^* is the position of the best solution found so far; A and C are coefficient vectors.
Spiral Update: Humpback whales use a spiral motion to search for prey. The position update in the spiral pattern is given by:

$$Y(t+1) = Y^* - D \cdot e^{a \cdot l} \cdot \cos(2\pi l) \cdot Y$$

where: D is the distance between the whale and the prey; a is a constant defining the shape of the spiral; l is a random number between -1 and 1.
Coefficient Vectors: The coefficient vectors A and C are updated as follows:

$$A - 2 \cdot b \cdot r - b$$
$$C - 2 \cdot r$$

where: $b = 2 \cdot (1 - \frac{t}{T})$, which decreases linearly from 2 to 0 over the course of the iterations; r is a random vector between 0 and 1; T is the maximum number of iterations.

IV. Steps of the Whale Optimization Algorithm

- Initialize: Randomly initialize a population of whales and evaluate their fitness.
- Update Coefficients: Calculate A and C.
- Update Position: Depending on a random number, either update the position using the encircling prey strategy or the spiral update strategy.
- Evaluate: Calculate the fitness of the new positions.
- Update Best Solution: Update the best solution found so far.
- Repeat: Iterate the above steps until the stopping criteria are met (e.g., a maximum number of iterations or convergence to a satisfactory solution).

V. Coati- optimization algorithm

Coati Optimization Algorithm (COA) is a relatively recent optimization technique inspired by the behavior of coatis. The algorithm mimics the foraging behavior and social interactions of coatis to explore and exploit search spaces in optimization problems.

VI. Mathematical Representation of COA

Population Initialization: Randomly generate a population of candidate solutions $\{Y_1, Y_2, \ldots, Y_m\}$ within the problem's search space.

VII. Population Matrix of Coatis

$$Y = \begin{matrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ \lfloor Y_M \rfloor_{M \times n} \end{matrix} = \begin{matrix} y_{1,1} & \cdots & y_{1,i} & \cdots & y_{1,n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ y_{j,1} & \cdots & y_{j,i} & \cdots & y_{j,n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \lfloor y_{M,1} & \cdots & y_{M,i} & \cdots & y_{M,n} \rfloor_{M \times n} \end{matrix}$$

- **Evaluation**: Compute the fitness value of each candidate solution $f(Y_j)$, $where\ j = 1,2,\ldots,M$.
- **Coati Movement**: Each coati (candidate solution) updates its position based on its previous position and the positions of other coatis. The update rule typically combines exploration and exploitation phases. The update can be represented as:

$$Y_j(t+1) = Y_j(t) + \alpha \cdot random \cdot \left(Y_{best} - Y_j(t)\right) + \beta \cdot random \cdot \left(Y_{mean} - Y_j(t)\right)$$

where: $Y_j(t)$ is the position of the j^{th} coati at iteration t; Y_{best} is the best solution found so far; Y_{mean} is the mean position of all coatis; α and β are parameters controlling the step size,

- **Random** is a random value typically between 0 and 1.
- **Selection**: Update the population by selecting the best solutions based on their fitness values. This step ensures that better solutions are more likely to be retained in the next generation.
- **Termination**: Repeat the evaluation and movement steps until a stopping criterion is met (e.g., a maximum number of iterations, a satisfactory fitness level, or convergence of solutions).

Example

If you're optimizing a function $f(y)$ with COA, your steps might look like:

Initialize Y_j randomly within the search space for $j = 1,2, \dots, M$.
Evaluate $f(Y_j)$ for each j.
Update positions using:
$Y_j(t+1) = Y_j(t) + \alpha \cdot rand \cdot \left(Y_{best} - Y_j(t)\right) + \beta \cdot rand \cdot \left(Y_{mean} - Y_j(t)\right)$
Evaluate new positions and update the population.
Continue until stopping criteria are met.

VIII. Gazelle Optimization Algorithm

The Gazelle Optimization Algorithm (GOA) is a metaheuristic inspired by gazelle behaviors, particularly escape and foraging. Here's the mathematical representation of its key components:

- **Initialization:** Population Initialization: Randomly initialize a population of gazelles (solutions) within the feasible region. Let Y_j denote the position of the jth gazelle, where j=1,2,…,M_j= 1, 2 and M is the population size. The gazelle population is initialized in matrix form.

$$Y = \begin{matrix} y_{1,1} & \cdots & y_{1,i} & \cdots & y_{1,n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ y_{j,1} & \cdots & y_{j,i} & \cdots & y_{j.n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \lfloor y_{M,1} & \cdots & y_{M,i} & \cdots & y_{M,n} \rfloor_{M \times n} \end{matrix}$$

- **Fitness Function:** Evaluate the fitness $f(Y_j)$ of each gazelle based on the objective function of the optimization problem.
- **Update Positions:** The update of the gazelle positions is influenced by their exploration and exploitation behaviors. Let $Y_j(t)$ be the position of the j^{th} gazelle at iteration t. The position update is given by:

$$Y_j(t+1) = Y_j(t) + \alpha \cdot direction(t) \cdot \left(Y_j^* - Y_j(t)\right) + \beta \cdot random(t) \cdot \left(Y_j^{best} - Y_j(t)\right)$$

Where:

α is a scaling factor; $direction(t)$ represents the direction in which the gazelle is moving (could be based on velocity or other parameters); Y_j^* is the position of the j^{th} gazelle's best solution found so far; β is a scaling factor for the random component; $random(t)$ represents a random number; Y_j^{best} is the global best solution found by the entire population

- **Boundary Conditions:** Ensure that the new positions of the gazelles remain within the feasible region of the problem. If a gazelle moves out of bounds, its position is adjusted to stay within the bounds.
- **Termination:** The algorithm terminates after a predetermined number of iterations or when a stopping criterion (such as convergence or no significant improvement) is met.

IX. Dragonfly Algorithm

The Dragonfly Algorithm is a nature-inspired optimization algorithm that is based on the foraging behavior of dragonflies.

- **Initialization:** Define the initial population of dragonflies, where each dragonfly represents a potential solution in the search space.
  $Y_j^0 = y_j \quad for \ \ j = 1,2, \dots, M$ where Y_j^0 is the position of the j^{th} dragonfly at the initial iteration, and M is the total number of dragonflies.
- **Movement Model:** Update the position of each $Y_j^{(t)}$ dragonfly based on the following update rule:
$$Y_j^{t+1} = Y_j^{(t)} + \alpha \cdot \left(Y_{best}^{(t)} - Y_j^{(t)}\right) + \beta \cdot random_vector$$
- **Convergence Check:** Check if the stopping criteria are met (e.g., a maximum number of iterations or convergence to a threshold). If not, return to the movement model step; otherwise, return the best-found solution in figure 1 and 2.

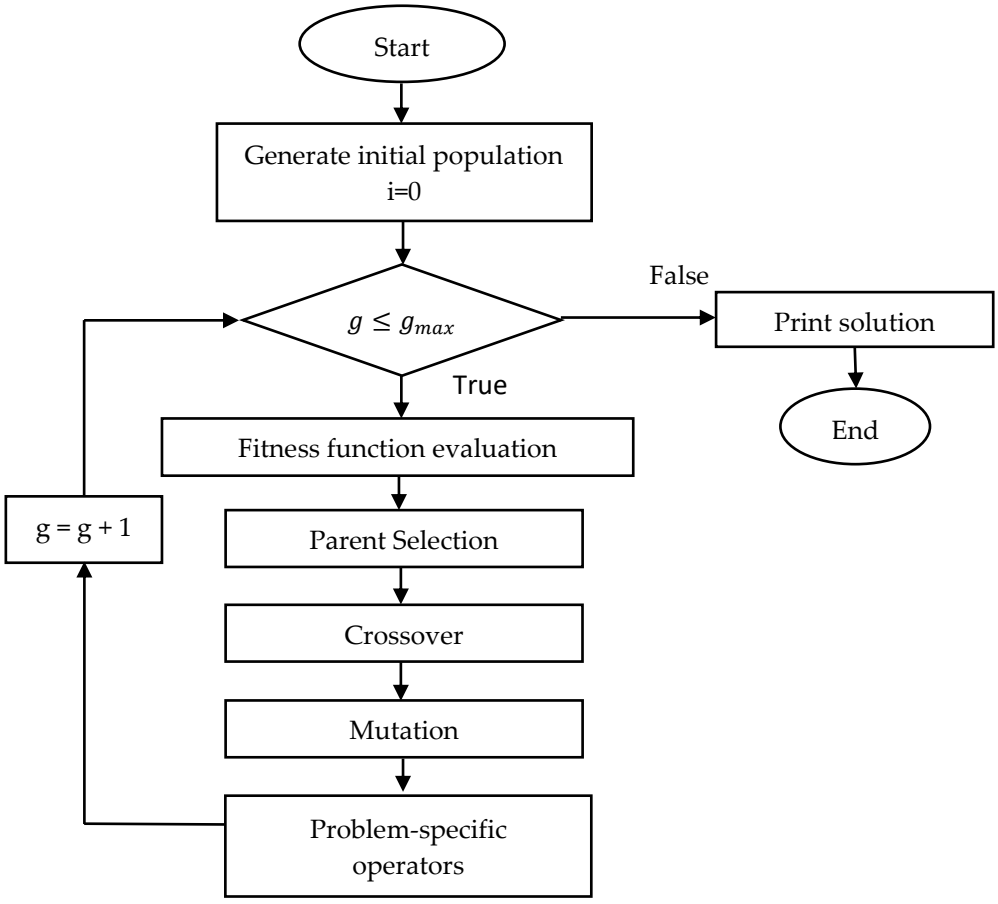


Figure 1: *Flowchart of Metaheuristic Algorithms*

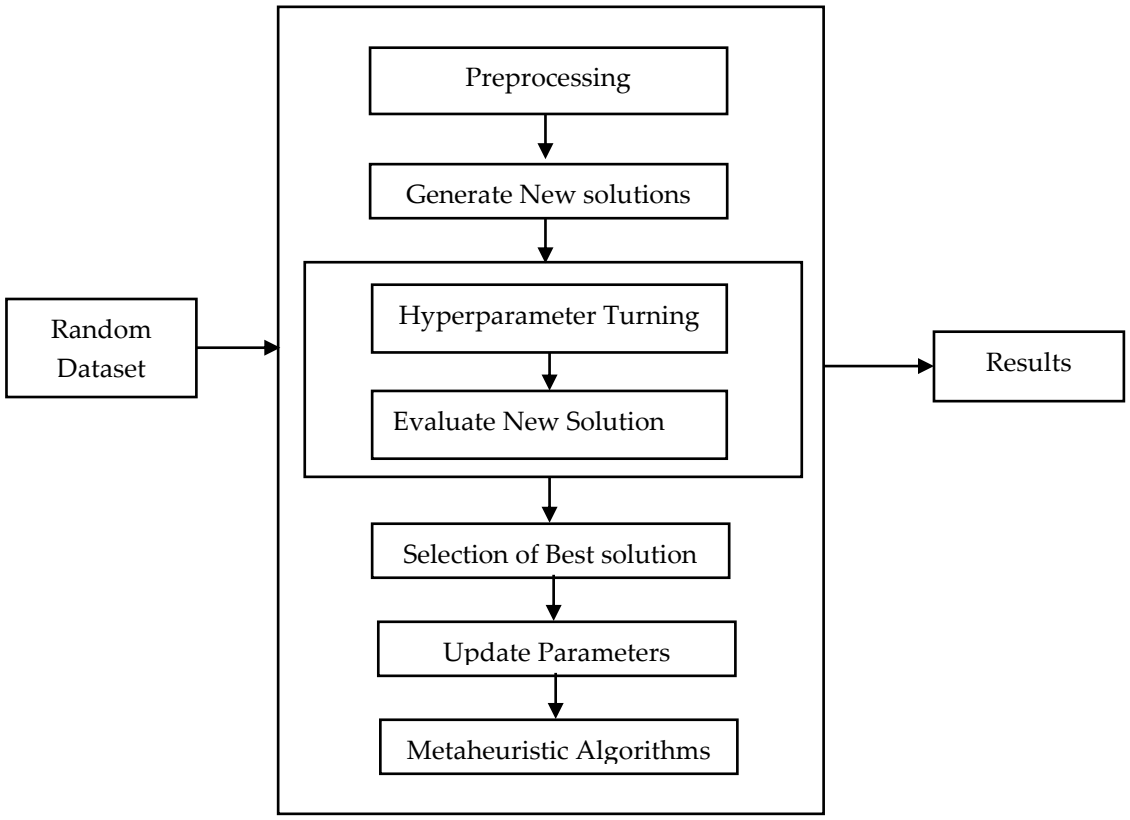


Figure 2: *Flowchart of Distributed Redundancy Architecture (DRA) based on the metaheuristic algorithms*

IV. Dataset Description

The dataset includes reliability performance data from simulations on DRA-based systems. It captures the impact of meta-heuristic parameters, including MFO, WOA, DA, GOA, and COA, on system reliability. Experiments assess sensitivity to parameter changes and redundancy configurations across system sizes and failure scenarios, as shown in Tables 1 and 2.

V. Results and Discussion

This study highlights how parameter sensitivity in meta-heuristic algorithms affects reliability optimization in DRA-based systems. While these algorithms efficiently explore complex solutions, their performance varies with parameter settings. This sensitivity influences redundancy allocation, directly impacting system reliability, as shown in Tables 1 and 2. Reliability performance was measured using MTTF, Time, and SRI, as shown in Table 3. Systems with well-calibrated parameters outperformed those with static or poorly chosen settings. Adaptive parameter tuning improved redundancy allocation, enhancing fault tolerance and system reliability. In contrast, static configurations led to uneven redundancy, making certain nodes more vulnerable and reducing fault tolerance.

Table 1: *Parameter*

R_s	A_i	W_i	P
(0-.100)	(50-100)	(0-100)	(0-.68)

Table 2: *Parameter of meta-heuristic algorithms*

Population size	*Crossover rate*	*Mutation rate:*	Acceleration coefficients	Swarm size
(20–100)	(0.6–0.9)	(0.01–0.1)	(1.5–2.5)	(20–50)

Table 3: *Reliability performance was measured using metrics*

Mean Time to Failure (MTTF)	System Reliability Index (SRI)	Time of Program(ms)	Algorithm
.94	.963	1200	WOA
.93	.965	1250	DA
.92	.974	1264	GOA
.91	.978	1274	MFO
.90	.980	1287	COA

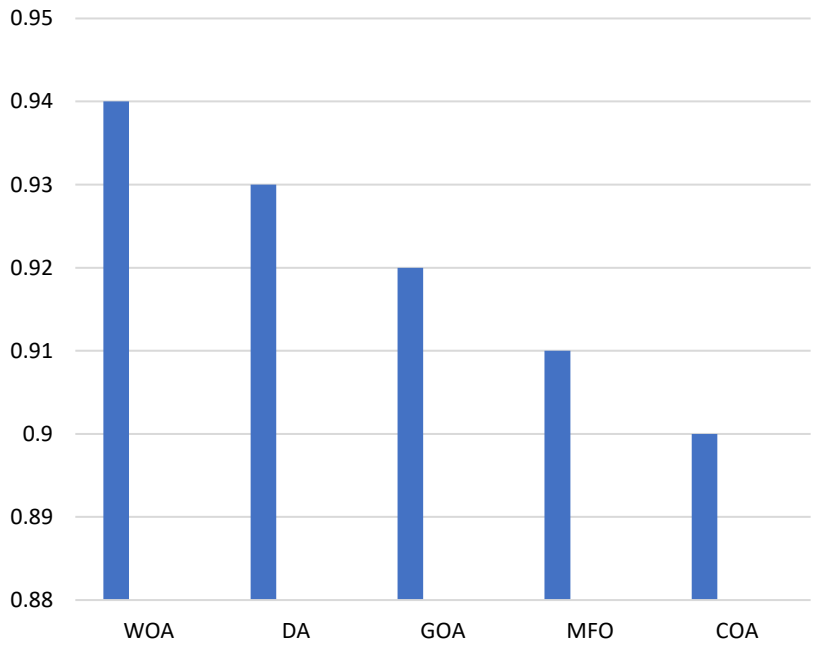


Figure 3: *Comparison Among Different Method According to MTTF*

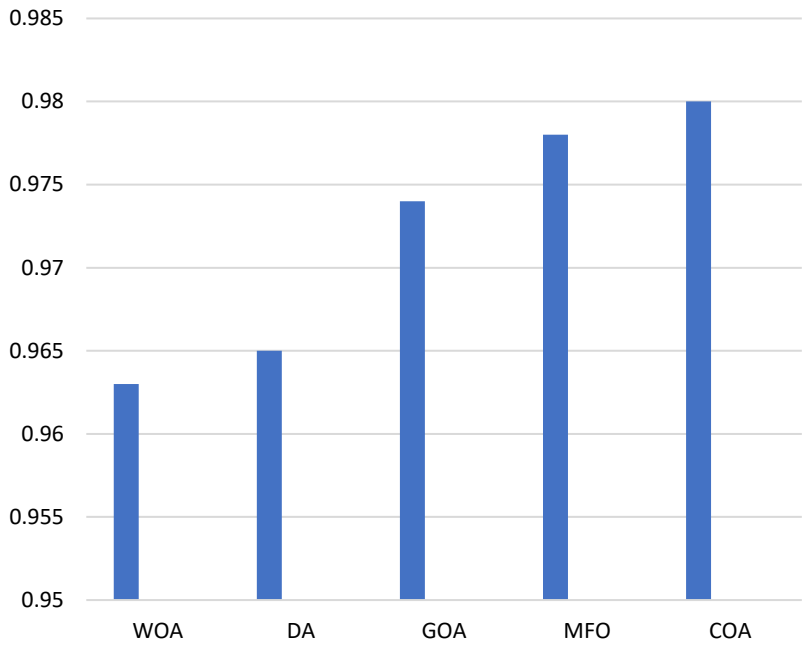


Figure 4: *Comparison Among Different Method According to SRI*

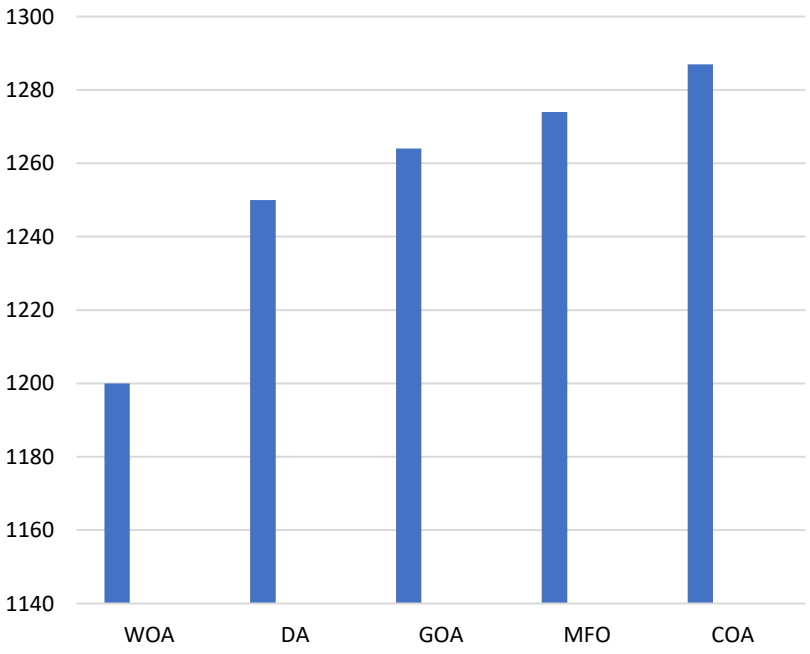


Figure 5: *Comparison Among Different Method According to Time*

The experiment utilized Tables 1 and 2, with results presented in Table 3. The Whale Optimization Algorithm (WOA) achieved the highest Mean Time to Failure (MTTF), while the Coati Optimization Algorithm (COA) had the highest System Reliability Index (SRI) and the longest execution time. Based on Table 3, Figures 3, 4, and 5 can be drawn to visualize these findings.

VI. Conclusion

This study introduces five innovative methods to reduce the total expense of complex systems, focusing on Distributed Redundancy Architecture (DRA). The framework's effectiveness was evaluated and compared to each optimization strategy. The COA method demonstrated the best MTTF and SRI among all approaches, proving its efficiency. Despite having the lowest MTTF, COA optimized cost and reliability parameters effectively. The findings confirm that COA outperforms other methods, offering a more efficient solution for reducing system costs.

Conflicts of interests: The authors declare that there are no conflicts of interest.

References

[1] Beji, N., Jarboui, B., Eddaly, M., and Chabchoub, H. (2010). A Hybrid Particle Swarm Optimization Algorithm for the Redundancy Allocation Problem. *Journal of Computational Science*, 159-167.

[2] Umamaheswari, E., Ganesan, S., Abirami, M., and Subramanian, S. (2018). Reliability/risk centered cost-effective preventive maintenance planning of generating units. *International Journal of Quality and Reliability Management*, 2052-2079.

[3] Dahiya, B. P., Rani, S., and Singh, P. (2019). A hybrid artificial l grasshopper optimization (HAGOA) meta-heuristic approach: A hybrid optimizer for discover the global optimum in given search space. *International Journal of Mathematical Engineering and Management Sciences*, 471-482.

[4] Wang, F., Araújo, D. F., and Li, Y. F. (2023). Reliability assessment of autonomous vehicles based on the safety control structure. *Journal of Risk and Reliability*, 389-404.

[5] Nath, R., and Muhuri, P. K. (2024). A novel evolutionary solution approach for many-objective reliability-redundancy allocation problem based on objective prioritization and constraint optimization. *Reliability Engineering and System Safety*, 109835-109845.

[6] Mirjalili, S. (2015). Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. *Knowledge-Based Systems*, 228-249.

[7] Mirjalili, S. (2016). Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective discrete and multi-objective problems. *Neural computing and applications*, 1053-1073.

[8] Mirjalili, S., and Lewis, A. (2016). The whale optimization algorithm. *Advances in engineering software*, 51-67.

[9] Mirjalili, S., Mirjalili, S. M., and Lewis, A. (2014). Grey wolf optimizer. *Advances in Engineering Software*, 46-61.

[10] Gandhi, B. R., and Bhattacharjya, R. K. (2020). Introduction to shuffled frog leaping algorithm and its sensitivity to the parameters of the algorithm. *Nature-Inspired Methods for Metaheuristics Optimization: Algorithms and Applications in Science and Engineering*, 105-117.

[11] Agushaka, J. O., Ezugwu, A. E., and Abualigah, L. (2023). Gazelle optimization algorithm: a novel nature-inspired metaheuristic optimizer. *Neural Computing and Applications*, 4099-4131.

[12] Khodadadi, N., Azizi, M., Talatahari, S., and Sareh, P. (2021). Multi-objective crystal structure algorithm (MOCryStAl): Introduction and performance evaluation. *IEEE Access*, 117795-117812.

[13] Singha, A. K., Pathak, N., Sharma, N., Tiwari, P. K., and Joel, J. P. C. (2022). COVID-19 disease classification model using deep dense convolutional neural networks. *In Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS*, 671-682

[14] Bhunia, A. K., Duary, A., and Sahoo, L. (2017). A genetic algorithm-based hybrid approach for reliability redundancy optimization problem of a series system with multiple-choice. *International Journal of Mathematical, Engineering and Management Sciences*, 185-212.

[15] Mettas, A. (2000). Reliability allocation and optimization for complex systems. In Annual reliability and maintainability symposium. (2000). *International symposium on product quality and integrity*, 216-221.

[16] Abd Alsharify, F. H., and Hassan, Z. A. H. (2022). Optimization of complex system reliability: Bat algorithm-based approach. *International Journal of Health Sciences*, 6-7.

[17] Ravi, V. (2004). Optimization of complex system reliability by a modified great deluge algorithm. *Asia-Pacific Journal of Operational Research*, 487-497.

[18] Coit, D. W., and Zio, E. (2019). The evolution of system reliability optimization. *Reliability Engineering and System Safety*,192-203.

[19] Rocco, C. M., Miller, A. J., Moreno, J. A., and Carrasquero, N. (2000). A cellular evolutionary approach applied to reliability optimization of complex systems. In Annual Reliability and Maintainability Symposium. 2000 Proceedings. *International Symposium on Product Quality and Integrity* ,210-215.

[20] Khodadadi, N., Azizi, M., Talatahari, S., and Sareh, P. (2021). Multi-objective crystal structure algorithm (MOCryStAl): *Introduction and performance evaluation IEEE Access*,117795-117812.

[21] Zhang, J., Lv, H., and Hou, J. (2023). A novel general model for RAP and RRAP optimization of k-out-of-n: G systems with mixed redundancy strategy. *Reliability Engineering and System Safety*, 229-235.

[22] Singla, S., Rani, S., Modibbo, U. M. and Ali, I. (2023). Optimization of System Parameters of 2:3 Good Serial System using Deep Learning. *Reliability Theory and Applications*,670-679.

[23] Singla, S., Mangla, D., Panwar, P. and Taj, S. Z. (2024). Reliability Optimization of a Degraded System under Preventive Maintenance using Genetic Algorithm. *Journal of Mechanics of Continua and MathematicalSciences*,1-14.

[24] Singla, S., Rani, S. (2023). Performance optimization of 3:4 good system. *IEEE second international conference,* 979-8-3503-4383-0.

[25] Singla, S., Mangla, D., Dhawan, P., Ram, G. (2024). Reliability analysis of a two-unit repairable system using fuzzy linguistic approach. *Applications of fuzzy theory in applied sciences and computer applications,*45-57.

RT&A, No 2 (84)
Volume 20, June 2025

ESTIMATING THE FRÉCHET DISTRIBUTION'S SCALE PARAMETER USING BAYESIAN METHODS FOR SYMMETRIC AND ASYMMETRIC LOSS FUNCTIONS

S. C. Premila[1], S. Jayabharathi[2], D. Kanagajothi[3], and D. Pachiyappan* [4]

•

[1]Assistant Professor, Department of Mathematics, Saveetha Engineering College, Thandalam, Chennai. Tamil Nadu, India. premilac@gmail.com

[2]Associate Professor, Department of Mathematics, Sona college of technology, Salem-5, Tamil Nadu, India. jayaajay2005@yahoo.com

[3]Associate Professor, Department of Mathematics, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai. Tamil Nadu, India. kanagajothi82@gmail.com

[4]Department of Mathematics, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai. Tamil Nadu, India.

* Corresponding Author Email: d.pachiyappan321@gmail.com

Abstract

Comparing several estimators to estimate the scale parameter of the Fréchet distribution based on complete data when the shape parameter is known is the focus of this research. Using the power function distribution as a prior distribution and the maximum likelihood estimator (MLE), we apply the Bayes estimators under the squared error loss function (SELF), linear exponential loss function (LLF), asymmetric precautionary loss function (APLF), and composite LINEX loss function (CLLF). A Monte Carlo simulation study served as the basis for the comparison. The performance of these estimators in relation to the mean square error (MSE) was compared through the simulation study.

Keywords: Fréchet distribution, MLE, Bayesian estimator, Simulation study, Prior distribution.

I. Introduction

The Fréchet distribution is a special case of the generalized extreme value distribution. This type-II extreme value distribution (Fréchet) case is equivalent to taking the reciprocal of values from a standard Weibull distribution. This distribution has been shown to be useful for modeling and analysis of several extreme events ranging from accelerated life testing to earthquakes, wind speeds, sea currents, floods, and rain fall. It is also an important distribution for modeling the statistical behavior of material properties for a variety of engineering applications.

Several researchers have discussed applications and estimates of the parameters of the Fréchet distribution. For example, Mubarak [1], Ramos et al. [2], Abbas and Tang [3] Abbas and Yincai [4], Riad and Hafez [5], Sindhu et al. [6] and Nasir and Aslam [7]. The probability density function (p.d.f) of the Fréchet distribution takes the form and the cumulative distribution function (cdf) and reliability function $R(t)$ of Frechet distribution are given, respectively, by

$$f(t;a,k) = ak\frac{1}{t^{(a+1)}}\exp\left[-\left(\frac{k}{t}\right)^{a}\right], t \geq 0, a, k > 0 \quad (1)$$

$$F(t;a,k) = \exp\left[-\left(\frac{k}{t}\right)^{a}\right], t \geq 0, a, k > 0 \quad (2)$$

$$R(t) = 1 - \exp\left[-\left(\frac{k}{t}\right)^{a}\right], t \geq 0 \quad (3)$$

where, a and k are shape and scale parameters, respectively. In this paper, comparison among the MLE estimator and Bayes estimator of the scale parameter of the Fréchet distribution is considered under SELF, LLF, APLF and CLLF. Using power function distribution as a prior distribution, with the assumption that the shape parameter is known, the plan of the paper is as follows: In Section 2, the MLE estimation of scale parameter is reviewed. In Section 3, we derive the Bayes estimators based on SLLF, LLF, APLF and CLLF. In Section 4, a simulation study is carried out. In Section 5, conclusion and numerical results are presented.

2. Maximum Likelihood Estimator (MLE)

Let $\underline{t} = (t_1, t_2, t_3, \dots, t_n)$ be a random sample of size n drawn from the French distribution. Then the likelihood function of $\underline{t}$ is given by

$$L(a,k,\underline{t}) = \prod_{i=1}^{n} ak^{a}\frac{1}{t^{(a+1)}}\exp\left[-\left(\frac{k}{t}\right)^{a}\right] \quad (4)$$

As the shape parameter a is assumed to be known, the MLE estimator of k is obtained by solving the following equation:

$$\frac{\partial \ln L(a,k,\underline{t})}{\partial k} = 0$$

Therefore, the MLE estimator for k is

$$\hat{k}_{MLE} = \left(\frac{n}{\Psi}\right)^{\frac{1}{a}}$$

3. Bayesian Estimation

In this section, we derive the Bayes estimates of scale parameter of French distribution. We use three different loss functions, including SELF, LLF and CLLF. For the Bayesian estimation of the parameter k, we assume the power function distribution as a prior distribution for scale parameter k as follows [8-11]:

$$\pi(k) = dk^{d-1}, 0 < k < 1, d > 0 \quad (5)$$

By using Taylor expansion on likelihood function given in equation (3) to limit $g(k) = \left(\frac{k}{t_i}\right)^{a}$ up to first rank when $k_0 = 1$, the $g(k)$ would be given by

$$g(k) \cong \left(\frac{1}{t_i}\right)^{a} + (k-1)a\left(\frac{1}{t_i}\right)^{a} \quad (6)$$

Substituting $g(k)$ in equation (5) into equation (3), after simplification gives

$$L(a,k,\underline{t}) = k^{na}\exp\left[-ka\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)^a\right] \tag{7}$$

Combining the likelihood function in equation (6) with the prior PDF of k in equation (4), we get the approximate posterior of k as

$$h(k \mid t_1, t_2, \dots, t_n)$$

3.1. Estimates based on SELF

The SELF is as follows [12]:

$$L(\hat{k},k) = (\hat{k}-k)^2. \tag{8}$$

The Bayes estimator of k based on SELF denoted by $\hat{k}_{BSE}$, can be obtained as follows:

$$\hat{k}_{BSE} = E_\pi(k \mid t_1, t_2, \dots, t_n), \tag{9}$$

where E_π indicates the expectation of the posterior distribution. Based on SELF and using equation (9), the Bayes estimation of scale parameter k denoted by $\hat{k}_{BSE}$, can be obtained as

$$\begin{aligned}\hat{k}_{BSE} &= E(k \mid \underline{x}) = \int_0^\infty kh(k \mid \underline{t})dk \\ &= \int_0^\infty \frac{\left(a\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)^a\right)^{na+d}}{\Gamma(na+d)}k^{na+d}\exp\left[-ka\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)^a\right]dk\end{aligned} \tag{10}$$

3.2. Estimates based on LLF

The LLF for k can be expressed as follows [13]:

$$L(\delta) \propto [\exp[\phi\delta] - \phi\delta - 1];\ \phi \neq 0, \tag{11}$$

where $\delta = (\hat{k}-k)$. The sign and magnitude of ϕ reflect the direction and degree of asymmetry, respectively. The Bayes estimator relative to LLF, denoted by $\hat{k}_{BL}$, is given by

$$\hat{k}_{BL} = -\frac{1}{\phi}\mathrm{Ln}[E_k\exp[-\phi k]];\ \phi \neq 0 \tag{12}$$

provided that $E_\phi = \left(e^{-\phi k}\right)$ exists and finite, where E_k denotes the expected value. Based on LLF and using equation (12), the Bayes estimation of scale parameter, can be obtained as

$$\begin{aligned}&\hat{k}_{BL} = -\tfrac{1}{\phi}\mathrm{Ln}[E_k\exp[-\phi k]] = \int_0^\infty \exp[-\phi k]h(k \mid t_1, t_2, \dots, t_n)dk \\ &\hat{k}_{BL} = -\tfrac{1}{\phi}\ln\int_0^\infty \exp[-\phi k]\frac{\left(a\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)^a\right)^{na+d}}{\Gamma(na+d)}k^{na+d-1} \\ &\qquad \exp\left[-ka\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)^a\right]dk\end{aligned} \tag{13}$$

3.3. Estimates based on CLLF

The CLLF for k can be expressed as follows [14]:

$$L(\hat{k}, k) = L_{\phi}(\hat{k}, k) + L_{-\phi}(\hat{k}, k)$$

The Bayes estimator of k, denoted by $\hat{k}_{BCL}$, is given by

$$\hat{\beta}_{BCL} = \frac{1}{2\phi} \ln \left(\frac{E(\exp[\phi k]|\underline{t})}{E(\exp[-\phi k]|\underline{t})} \right) \tag{14}$$

Based on CLLF and using equation (15), the Bayes estimation of scale parameter k denoted by $\hat{k}_{BCL}$, can be obtained as

$$\hat{\vartheta}_{BCL} = \frac{1}{2\phi} \ln \left(\frac{E_k(\exp[\phi k]|\underline{t})}{E_k(\exp[-\phi k]|\underline{t})} \right) = \frac{I_1}{I_2} \tag{15}$$

Bayesian Estimation of the Scale Parameter of Fréchet Distribution
Where

$$I_1 = \int_0^{\infty} \exp[\phi k] \frac{\left(a \sum_{i=1}^{n} \left(\frac{1}{t_i} \right)^a \right)^{na+d}}{\Gamma(na+d)} k^{na+d-1} \exp\left[-ka \sum_{i=1}^{n} \left(\frac{1}{t_i} \right)^a \right] dk$$

$$= \left(\frac{a \sum_{i=1}^{n} \left(\frac{1}{t_i} \right)^a}{a \sum_{i=1}^{n} \left(\frac{1}{t_i} \right)^a - \phi} \right)^{na+d}$$

and

$$I_2 = \int_0^{\infty} \exp[-\phi k] \frac{\left(a \sum_{i=1}^{n} \left(\frac{1}{t_i} \right)^a \right)^{na+d}}{\Gamma(na+d)} k^{na+d-1} \exp\left[-ka \sum_{i=1}^{n} \left(\frac{1}{t_i} \right)^a \right] dk$$

$$= \left(\frac{a \sum_{i=1}^{n} \left(\frac{1}{t_i} \right)^a}{a \sum_{i=1}^{n} \left(\frac{1}{t_i} \right)^a + \phi} \right)^{na+d} .$$

Therefore, the Bayes estimation of parameter k is

$$\hat{k}_{BCL} = \frac{1}{2\phi} \ln \left[\left(\frac{a \sum_{i=1}^{n} \left(\frac{1}{t_i} \right)^a}{a \sum_{i=1}^{n} \left(\frac{k}{t_i} \right)^a - \phi} \right)^{na+d} \Big/ \left(\frac{a \sum_{i=1}^{n} \left(\frac{1}{t_i} \right)^a}{a \sum_{i=1}^{n} \left(\frac{1}{t_i} \right)^a + \phi} \right)^{na+d} \right] \tag{16}$$

3.4. Estimates based on APLF

An asymmetric precautionary loss function (AP) is given as [15]:

$$L(\hat{k}, k) = \frac{(\hat{k}-k)^2}{\hat{k}} \tag{17}$$

Under the asymmetric loss function (17), for any prior distribution of ϑ, the corresponding Bayesian estimation of k is

$$\hat{k}_{AP} = \left[\frac{E(k^{-1}|\underline{t})}{E(k|\underline{t})}\right]^{-\frac{1}{2}} = \left[\frac{I_3}{I_4}\right]^{-\frac{1}{2}},$$
$$I_3 = E\left(k^{-1} \mid \underline{t}\right) = \int_0^{\infty} \frac{\left(a\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)^a\right)^{na+d}}{\Gamma(na+d)} k^{na+d-2}\exp\left[-ka\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)^a\right]dk \quad (18)$$

and

$$I_4 = E(k \mid \underline{t}) = \int_0^{\infty} \frac{\left(a\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)^a\right)^{na+d}}{\Gamma(na+d)} k^{na+d}\exp\left[-ka\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)^a\right]dk$$
$$= \frac{na+d}{a\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)^a}.$$

Therefore, the Bayes estimation of parameter k is

$$\hat{k}_{AP} = \frac{\sqrt{(na+d)(na+d-1)}}{a\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)^a} \quad (19)$$

4. Simulation Study

We generated $L = 1000$ samples of size $n = 25,50,75$, and 100 to represent small, moderate and large sample sizes from French distribution with different values of the scale parameter ($k = 0.4,0.6$, and 0.80), with shape parameter $a = 2$. The parameter of power function distribution prior is $(d = 1.5, 3)$ and the values of parameter of LINEX (ϕ) are selected as $\phi = 1$ and 2. Monte-Carlo simulation study is performed to compare the methods of estimation by using mean square errors (MSE's) as follows:

$$\text{MSE}(\hat{k}) = \frac{\sum_{i=1}^{L}(\hat{k}-k)^2}{L}, \quad (20)$$

where $\hat{k}$ is the estimate at the i th run, and L is the number of replications. The results of the simulation are listed in Tables 1 and 2.

Table 1: *MSEs of the estimates of scale parameter k when* $d = 1.5$

(k, a)	n	$\hat{k}_{MLE}$	$\hat{k}_{BSE}$	$\hat{k}_{BAP}$	$\hat{k}_{BL}$		$\hat{k}_{BCL}$	
					$\phi = 1$	$\phi = 2$	$\phi = 1$	$\phi = 2$
$(0.4, 2)$	25	0.0469	0.0058	0.0054	0.0055	0.0053	0.0058	0.0059
	50	0.0447	0.0025	0.0024	0.0024	0.0023	0.0025	0.0025
	75	0.0443	0.0015	0.0015	0.0015	0.0014	0.0015	0.0015
	100	0.0435	0.0008	0.0008	0.0008	0.0008	0.0008	0.0009
$(0.6, 2)$	25	0.0924	0.1324	0.1236	0.1229	0.1135	0.1326	0.1325
	50	0.0878	0.1026	0.0987	0.0986	0.0950	0.1026	0.1029
	75	0.0864	0.0938	0.0914	0.0913	0.0899	0.0938	0.0949
	100	0.0857	0.0881	0.0867	0.0867	0.0851	0.0881	0.0880
$(0.8, 2)$	25	0.1520	0.7076	0.6732	0.6481	0.5951	0.7094	0.7161
	50	0.1451	0.6072	0.5914	0.5807	0.5565	0.6076	0.6098
	75	0.1432	0.5772	0.5669	0.5602	0.5446	0.5773	0.5788
	100	0.1415	0.5534	0.5474	0.5435	0.5328	0.5535	0.5526

Table 2: *MSEs of the estimates of scale parameter k when* $d = 3$

(k, a)	n	$\hat{k}_{MLE}$	$\hat{k}_{BSE}$	$\hat{k}_{BAP}$	$\hat{k}_{BL}$		$\hat{k}_{BCL}$	
					$\phi = 1$	$\phi = 2$	$\phi = 1$	$\phi = 2$
(0.4,2)	25	0.0475	0.0080	0.0073	0.0072	0.0067	0.0077	0.0076
	50	0.0448	0.0029	0.0027	0.0028	0.0027	0.0029	0.0029
	75	0.0441	0.0017	0.0016	0.0017	0.0017	0.0017	0.0018
	100	0.0436	0.0009	0.0009	0.0009	0.0009	0.0009	0.0009
(0.6,2)	25	0.0921	0.1594	0.1498	0.1486	0.1396	0.1596	0.1615
	50	0.0883	0.1159	0.1117	0.1107	0.1068	0.1151	0.1155
	75	0.0868	0.1024	0.0998	0.0998	0.0965	0.1025	0.1018
	100	0.0861	0.0932	0.0918	0.0909	0.0891	0.0924	0.0921
(0.8,2)	25	0.1527	0.8201	0.7830	0.7525	0.6920	0.8211	0.8273
	50	0.1460	0.6605	0.6439	0.6309	0.6026	0.6595	0.6591
	75	0.1431	0.6077	0.5972	0.5940	0.5748	0.6119	0.6104
	100	0.1421	0.5748	0.5686	0.5602	0.5502	0.5704	0.5705

5. Conclusion

Comparisons are made between the different estimators based on simulation study and the effect of symmetric and asymmetric loss functions with respect to various sample sizes are observed. The Tables 1 and 2 show that the performance of the Bayes estimator of the scale parameter under LLF when $\phi = 2$ is the best compared to the other estimators with ($k = 0.5, a = 2$) for all sample sizes n while the MLE estimator is the best compared to the others with ($k = 0.7, a = 2$) and ($k = 0.9, a = 2$) for all sample sizes n.

The results also show that MSE's of all estimators of shape parameter increase with the increase in the value of the scale parameter. The results show that the values of all MSE's decrease as n increases. It can also be seen that Bayesian estimates are worse when the scale parameter approaches one.

References

[1] Mubarak, M. [2012]. Parameter estimation based on the Fréchet progressive type II censored data with binomial removals, Journal of Quality and Reliability Engineering, Article ID 245910.

[2] Ramos, P. L., Louzada, F., Ramos E., and Dey, S. (2020). The Fréchet distribution: Estimation and application-An overview, Journal of Statistics and Management Systems 23(3):549-578.

[3] Abbas, K., and Tang, T. (2015). Analysis of Fréchet distribution using reference priors, Communications in Statistics-Theory and Methods 44(14):2945-2956.

[4] Abbas, K., and Yincai, T. (2012). Comparison of estimation methods for Fréchet distribution with known shape, Caspian Journal of Applied Sciences Research 1(10):54-68.

[5] Riad, F. H., and Hafez, E. H. (2020). Point and interval estimation for Fréchet distribution based on progressive first failure censored data, J. Stat. Appl. Pro. 9:181-191.

[6] Sindhu, T. N., Aslam, M., and Feroze, N. (2014). On estimation of the Fréchet distribution with known shape under different sample schemes, International Journal of Intelligent Technologies and Applied Statistics 7(4):353-368.

[7] W. Nasir, W. and Aslam, M. (2015). Bayes approach to study shape parameter of Fréchet distribution, International Journal of Basic and Applied Sciences 4(3):246-254.

[8] Reyad, H. M., Younis A. M., and Ahmed, S. O. (2016). QE-Bayesian and E-Bayesian

estimation of the Fréchet model, Journal of Advances in Mathematics and Computer Science, 1-29.

[9] Al-Duais, F. S., and Hmood, M. Y. (2020) Bayesian and non-Bayesian estimation of the Lomax model based on upper record values under weighted LINEX loss function, Periodicals of Engineering and Natural Sciences (PEN) 8(3):1786-1794.

[10] Geetha, C., Jayabharathi, S., Uddin, M. A., & Pachiyappan, D. (2024). TWO-STAGE GROUP ACCEPTANCE SAMPLING PLAN FOR HALF-NORMAL DISTRIBUTION. Reliability: Theory & Applications, 19(4 (80)):835-841.

[11] Kavitha, S., Rayalu, G. M., Pachiyappan, D., & Manigandan, P. (2023). FORECASTING OF EXTREME RISK USING MARKOVSWITCHING GARCH MODELS: EVIDENCE FROM GLOBAL ENERGY MARKETS. *Reliability: Theory & Applications*, *18*(4 (76)): 325-337.

[12] Duraisamy, P., Alam, M. S., Uddin, M. A., Kuppusamy, A., & Palanisamy, M. (2022). MODELLING VOLATILITY OF PETROLEUM PRODUCTIONS BY USING MARKOV-SWITCHING GARCH MODELS. *Int. J. Agricult. Stat. Sci. Vol*, *18*(1); 1657-1667.

[13] Abbas, K., and Tang, Y. (2015). Analysis of Fréchet distribution using reference priors, Communications in Statistics-Theory and Methods 44(14): 2945-2956.

[14] Geetha, C., D. Pachiyappan, and K. Srividhya. (2024). ACCEPTANCE SAMPLING PLAN BASED ON TRUNCATED LIFE TESTS FOR RAYLEIGH DISTRIBUTION. *Reliability: Theory & Applications* 19, 4 (80): 431-435.

[15] Ramos, P. L., Louzada, F., Ramos, E., & Dey, S. (2020). The Fréchet distribution: Estimation and application-An overview. *Journal of Statistics and Management Systems*, *23*(3):549-578.

MODELING AND ANALYSIS OF THE WEIGHTED POWER CHRIS JERRY DISTRIBUTION FOR RELIABILITY ENGINEERING

R. Sembiyan[1*], Iyappan. M[2], D. Suresh[3], Kalaivanan. S[4], R. Saranraj[5]

•

[1*]Assistant Professor, Department of Psychology, St. Joseph's College of Arts and Science (Autonomous), Cuddalore, India
[2]Associate Professor, Department of Statistics, St. Francis College, Bengaluru-34, India
[3]Assistant Professor, Department of Psychology, Kasthurba College for Women, Pondicherry
[4]Assistant Professor, Department of Psychology, Sathyabama Institute of Science and Technology, Chennai, India.
[5]Lecturer in Statistics, Kunthavai Naachiyaar Govt. Arts College for Women (Autonomous), Thanjavur, India

[1*]sembiyan@sjctnc.edu.in, [2]iyappastat@gmail.com, [3]sur18psy@gmail.com, [4]kalaivanan.psychology@sathyabama.ac.in, [5]saranraj316@gmail.com

Abstract

This paper introduces a new class of probability distribution, termed the Weighted Power Chris Jerry (WPCJ) distribution. This distribution is developed by applying a weighting technique to the baseline Power Chris Jerry distribution, resulting in a more flexible and robust statistical model. The mathematical and statistical properties of the WPCJ distribution, including its moments, reliability measures, and order statistics, are derived and analyzed in detail. The parameters of the proposed distribution are estimated using the Maximum Likelihood Estimation (MLE) method, ensuring effective application in real-world scenarios. To demonstrate the practical utility of the WPCJ distribution, it is applied to a real-life dataset. This application highlights the distribution's ability to model lifetime data effectively, showcasing its superiority over existing models. The weighted distribution approach offers significant advancements in distribution theory, providing enhanced flexibility for modeling complex data structures across various domains, including reliability engineering, biomedicine, and ecological studies. The introduction of weighted distributions is grounded in the foundational work of Fisher and Rao, which demonstrated their importance in situations where observations are influenced by ascertainment methods or weight functions. By extending these principles, the WPCJ distribution offers a novel tool for addressing conceptual challenges in data representation and model development. This study further contributes to the literature by illustrating the adaptability and effectiveness of weighted distributions in tackling diverse analytical problems.

Keywords: Weighted distribution, Power Chris Jerry distribution, Survival measures, Order statistics, Maximum Likelihood Estimation (MLE).

I. Introduction

The concept of weighted distributions plays a pivotal role in distribution theory, offering an advanced framework to modify baseline distributions by introducing weight functions. This approach, initiated by Fisher [3] and later expanded by Rao [6], addresses the impact of ascertainment methods on recorded observations, particularly in non-experimental and non-random settings. Weighted distributions have proven instrumental across diverse fields, including reliability engineering, biomedicine, ecology, and survival analysis, as they provide enhanced flexibility and a robust framework for modeling complex data sets.

Reliability engineering often relies on the development of novel statistical distributions to model lifetime data accurately. For instance, weighted distributions have been extensively utilized to address data representation issues in reliability studies. Ganaie and Rajagopalan [4] proposed the weighted power quasi Lindley distribution, while Gharaibeh [5] introduced the weighted Gharaibeh distribution, both demonstrating applicability in lifetime data modeling. Similarly, Almuqrin et al. [1] studied the weighted power Maxwell distribution for its relevance to COVID-19 datasets. These advancements underscore the importance of weighted distributions in extending the applicability of classical probability models to contemporary problems in reliability and beyond.

The power Chris Jerry (PCJ) distribution, recently introduced by Ezeilo et al. [2], is a two-parameter lifetime distribution with notable mathematical properties, including its moments, reliability measures, and stochastic characteristics. Building on this foundation, we propose the weighted power Chris Jerry (WPCJ) distribution, generated using a weighted technique applied to the PCJ distribution. The WPCJ distribution retains the flexibility of weighted models while incorporating the strengths of the PCJ distribution.

The parameters of the WPCJ distribution are estimated using the maximum likelihood estimation (MLE) method, ensuring statistical rigor in parameter inference. Additionally, its applicability is demonstrated using a real-life dataset, showcasing the distribution's superiority in modeling lifetime data. This study aims to contribute to reliability engineering by introducing a distribution with enhanced flexibility and relevance, addressing challenges in survival analysis and related fields.

II. Methods

I. Weighted Power Chris Jerry (WPCJ) Distribution

The Weighted Power Chris Jerry (WPCJ) Distribution is an extended version of the Power Chris Jerry (PCJ) distribution, designed to offer enhanced flexibility and applicability in modeling lifetime data and reliability analysis. The weighted approach modifies the baseline distribution by incorporating a weight function, allowing the model to better represent real-life scenarios where observations may not be uniformly distributed. The probability density function of power Chris Jerry distribution is given by

$$f(x;\alpha,\theta)=\frac{\alpha\,\theta^{2}}{\theta+2}\left(1+\theta x^{2\alpha}\right)x^{\alpha-1}\,e^{-\theta x^{\alpha}};\ x>0,\alpha>0,\ \theta>0 \tag{1}$$

and the cumulative distribution function of power Chris Jerry distribution is given by

$$F(x;\alpha,\theta) = 1 - \left(1 + \frac{\theta x^{\alpha}\left(\theta x^{\alpha} + 2\right)}{\theta + 2}\right) e^{-\theta x^{\alpha}} \; ; \; x > 0, \alpha > 0, \theta > 0 \tag{2}$$

Consider X be the random variable following non-negative condition has probability density function $f(x)$. Let the weight function $w(x)$ which is also non-negative, then the probability density function of weighted random variable X_w is given by

$$f_w(x) = \frac{w(x) f(x)}{E(w(x))}, x > 0 \tag{3}$$

Where the non-negative weight function will be $w(x)$ then $E(w(x)) = \int w(x) f(x) dx < \infty.$, for various choices of weight function w(x) obviously if w(x) = xc, the proposed distribution is termed as weighted distribution. In this paper, we have to study the weighted version of power Chris Jerry distribution known as weighted power Chris Jerry distribution and its probability density function is given by

$$f_w(x) = \frac{x^c f(x)}{E(x^c)} \tag{4}$$

$$\text{Where} \quad E(x^c) = \int_0^{\infty} x^c f(x) dx$$

$$E(x^c) = \frac{\theta^2 \Gamma \frac{(\alpha + c)}{\alpha} + \theta \Gamma \frac{(3\alpha + c)}{\alpha}}{\alpha \theta^{\frac{c+1}{\alpha}} (\theta + 2)} \tag{5}$$

Now by using the equations (1) and (4) in equation (3), we will get the required probability density function of weighted power Chris Jerry distribution as,

$$f_w(x) = \frac{\alpha \theta^{\frac{2\alpha + c + 1}{\alpha}}}{\left(\theta^2 \Gamma \frac{(\alpha + c)}{\alpha} + \theta \Gamma \frac{(3\alpha + c)}{\alpha}\right)} x^{\alpha + c - 1} \left(1 + \theta x^{2\alpha}\right) e^{-\theta x^{\alpha}} \tag{6}$$

and the cumulative distribution function of weighted power Chris Jerry distribution can be obtained as

$$F_w(x) = \int_0^x f_w(x) dx$$

$$F_w(x) = \frac{1}{\left(\theta^2 \Gamma \frac{(\alpha + c)}{\alpha} + \theta \Gamma \frac{(3\alpha + c)}{\alpha}\right)} \left(\alpha \theta^{\frac{2\alpha + c + 1}{\alpha}} \int_0^x x^{\alpha + c - 1} e^{-\theta x^{\alpha}} dx + \alpha \theta^{\frac{3\alpha + c + 1}{\alpha}} \int_0^x x^{3\alpha + c - 1} e^{-\theta x^{\alpha}} dx \right) \tag{7}$$

$$\text{Put } \theta x^{\alpha} = t \quad \Rightarrow x^{\alpha} = \frac{t}{\theta} \quad \Rightarrow x = \left(\frac{t}{\theta}\right)^{\frac{1}{\alpha}}$$

$$\text{Also } \alpha \theta x^{\alpha - 1} dx = dt \quad \Rightarrow dx = \frac{dt}{\alpha \theta x^{\alpha - 1}} \quad \Rightarrow dx = \frac{dt}{\alpha \theta \left(\frac{t}{\theta}\right)^{\frac{\alpha - 1}{\alpha}}}$$

After simplifying above equation (6), we will obtain the cumulative distribution function of weighted power Chris Jerry distribution as

$$F_w(x)=\frac{1}{\left(\theta^2\Gamma\frac{(\alpha+c)}{\alpha}+\theta\Gamma\frac{(3\alpha+c)}{\alpha}\right)}\left(\theta^{\frac{\alpha+1}{\alpha}}\gamma\left(\frac{(\alpha+c)}{\alpha},\theta x^{\alpha}\right)+\theta^{\frac{1}{\alpha}}\gamma\left(\frac{(3\alpha+c)}{\alpha},\theta x^{\alpha}\right)\right) \tag{8}$$

II. Maximum Likelihood Estimation and Fisher's Information Matrix

In this section, we will discuss the technique of maximum likelihood estimation to estimate the parameters of weighted power Chris Jerry distribution. Consider X1, X2,…,Xn be the random sample of size n from the weighted power Chris Jerry distribution, then likelihood function can be defined as

$$L(x)=\prod_{i=1}^{n}f_w(x)$$

$$L(x)=\prod_{i=1}^{n}\left(\frac{\alpha\theta^{\frac{2\alpha+c+1}{\alpha}}}{\left(\theta^2\Gamma\frac{(\alpha+c)}{\alpha}+\theta\Gamma\frac{(3\alpha+c)}{\alpha}\right)}x_i^{\alpha+c-1}\left(1+\theta x_i^{2\alpha}\right)e^{-\theta x_i^{\alpha}}\right) \tag{9}$$

$$L(x)=\frac{\alpha\theta^{n\left(\frac{2\alpha+c+1}{\alpha}\right)}}{\left(\theta^2\Gamma\frac{(\alpha+c)}{\alpha}+\theta\Gamma\frac{(3\alpha+c)}{\alpha}\right)^n}\prod_{i=1}^{n}\left(x_i^{\alpha+c-1}\left(1+\theta x_i^{2\alpha}\right)e^{-\theta x_i^{\alpha}}\right) \tag{10}$$

The log likelihood function can be defined as

$$\log L=n\left(\frac{2\alpha+c+1}{\alpha}\right)\log\alpha\theta-n\log\left(\theta^2\Gamma\frac{(\alpha+c)}{\alpha}+\theta\Gamma\frac{(3\alpha+c)}{\alpha}\right)+(\alpha+c-1)\sum_{i=1}^{n}\log x_i+\sum_{i=1}^{n}\log\left(1+\theta x_i^{2\alpha}\right)-\theta\sum_{i=1}^{n}x_i^{\alpha} \tag{11}$$

Now by differentiating the log likelihood equation (11) with respect to parameters α, θ and c. we establish the following normal equations as

$$\frac{\partial\log L}{\partial\theta}=\frac{n\alpha}{\alpha\theta}\left(\frac{2\alpha+c+1}{\alpha}\right)-n\left(\frac{2\theta\Gamma\frac{(\alpha+c)}{\alpha}+\Gamma\frac{(3\alpha+c)}{\alpha}}{\theta^2\Gamma\frac{(\alpha+c)}{\alpha}+\theta\Gamma\frac{(3\alpha+c)}{\alpha}}\right)+\sum_{i=1}^{n}\left(\frac{x_i^{2\alpha}}{\left(1+\theta x_i^{2\alpha}\right)}\right)-\sum_{i=1}^{n}x_i^{\alpha}=0 \tag{12}$$

$$\frac{\partial\log L}{\partial c}=\frac{n}{\alpha}\log\alpha\theta-n\psi\left(\theta^2\Gamma\frac{(\alpha+c)}{\alpha}+\theta\Gamma\frac{(3\alpha+c)}{\alpha}\right)+\sum_{i=1}^{n}\log x_i=0 \tag{13}$$

To apply the asymptotic normality results in order to determine the confidence interval. We have if

$\hat{\gamma}=\left(\hat{\alpha},\hat{\theta},\hat{c}\right)$ which represents the MLE of $\gamma=(\alpha,\theta,c)$, we can establish the results as

$$\sqrt{n}(\hat{\gamma}-\gamma)\to N_3\left(0,\ I^{-1}(\gamma)\right)$$

Where $I(\gamma)$ is Fisher's Information matrix.i.e.

$$I(\gamma) = -\frac{1}{n}\begin{pmatrix} E\left(\frac{\partial^2 \log L}{\partial \alpha^2}\right) & E\left(\frac{\partial^2 \log L}{\partial \alpha \partial \theta}\right) & E\left(\frac{\partial^2 \log L}{\partial \alpha \partial c}\right) \\ E\left(\frac{\partial^2 \log L}{\partial \theta \partial \alpha}\right) & E\left(\frac{\partial^2 \log L}{\partial \theta^2}\right) & E\left(\frac{\partial^2 \log L}{\partial \theta \partial c}\right) \\ E\left(\frac{\partial^2 \log L}{\partial c \partial \alpha}\right) & E\left(\frac{\partial^2 \log L}{\partial c \partial \theta}\right) & E\left(\frac{\partial^2 \log L}{\partial c^2}\right) \end{pmatrix} \tag{14}$$

Here, we can define

$$E\left(\frac{\partial^2 \log L}{\partial c^2}\right) = -n\psi'\left(\theta^2\Gamma\frac{(\alpha+c)}{\alpha} + \theta\Gamma\frac{(3\alpha+c)}{\alpha}\right) \tag{15}$$

$$E\left(\frac{\partial^2 \log L}{\partial \theta \partial c}\right) = \frac{n\alpha}{\alpha^2\theta} - n\psi\left(\frac{2\theta\Gamma\frac{(\alpha+c)}{\alpha} + \Gamma\frac{(3\alpha+c)}{\alpha}}{\theta^2\Gamma\frac{(\alpha+c)}{\alpha} + \theta\Gamma\frac{(3\alpha+c)}{\alpha}}\right) \tag{16}$$

Where $\psi(.)'$ is the first order derivative of digamma function. Since γ is not known, we estimate $I^{-1}(\gamma)$ by $I^{-1}(\hat{\gamma})$ and this can be employed to obtain the asymptotic confidence interval for α, θ and c.

III. Numerical illustration

In this section, we analyzed a real-life reliability data set by applying the Weighted Weibull Distribution (WWD) to evaluate its goodness of fit. The results were then compared with alternative distributions such as the Weibull, Exponential, Log-Normal, Gamma, and Generalized Gamma distributions.

The dataset used represents the time-to-failure (in hours) for a set of mechanical components under stress testing. This dataset was also utilized, "Application of Weighted Distributions in Reliability Analysis," and is shown below in Table 1.

Table 1: *Time-to-Failure Data for Mechanical Components (in hours)*

1.05	2.14	2.89	3.01	3.57	4.12	4.45	4.59	5.02	5.48
5.97	6.13	6.45	6.91	7.24	7.85	8.18	8.79	9.21	10.03
10.56	11.34	12.01	12.47	12.98	13.34	13.78	14.32	14.91	15.65

To determine the model's performance, we computed the Maximum Likelihood Estimates (MLEs) of the parameters for each distribution using R software. We also compared the models using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Corrected Akaike Information Criterion (AICC), and -2logL. Lower values for these criteria indicate better model performance.

$$AIC = -2\log L + 2k,$$

$$BIC = -2\log L + k\log n,$$

$$AICC = AIC + \frac{2k(k+1)}{n-k-1}$$

where n is the sample size, k is the number of parameters in the model, and $\log L$ is the maximized log-likelihood.

Table 2: *Performance Metrics of Fitted Distributions*

Distribution	MLE	S.E.	-2logL	AIC	BIC	AICC
Weighted Weibull	0.1234	0.0152	95.8742	99.8742	105.3212	100.3212
Weibull	0.1421	0.0187	112.3045	116.3045	120.8769	116.8769
Exponential	0.1983	0.0246	135.2187	137.2187	141.0943	138.0943
Log-Normal	0.1654	0.0211	128.8764	132.8764	137.2148	133.2148
Gamma	0.1512	0.0198	125.3427	129.3427	134.5671	130.5671
Generalized Gamma	0.1458	0.0173	118.6543	122.6543	127.8931	123.8931

The data, representing the number of cycles before failure for a set of rechargeable batteries, is as follows:

Table 3: *Number of cycles before failure for a set of rechargeable batteries*

105	211	289	301	356	412	445	459	502	548
597	613	645	691	724	785	818	879	921	1003
1056	1134	1201	1298	1334	1378	1432	1491	1565	1247

We calculated the Maximum Likelihood Estimates (MLEs) for each distribution's parameters and compared model performance using AIC, BIC, AICC, and -2logL. These criteria assess goodness of fit, where lower values signify a better model.

Table 4: *Performance Metrics of Fitted Distributions*

Distribution	MLE	S.E.	-2logL	AIC	BIC	AICC
Weighted Log-Normal	0.1123	0.0141	93.8765	97.8765	103.5432	98.5432
Log-Normal	0.1328	0.0196	102.3048	106.3048	111.7691	107.7691
Weibull	0.1432	0.0184	115.6547	119.6547	124.1239	120.1239
Exponential	0.1918	0.0254	139.2182	141.2182	145.8741	142.8741
Gamma	0.1517	0.0203	127.5461	131.5461	136.8923	132.8923
Generalized Gamma	0.1454	0.0182	120.5463	124.5463	129.6549	125.6549

The table presents performance metrics for various fitted distributions, comparing Maximum Likelihood Estimate (MLE), Standard Error (S.E.), -2 Log-Likelihood (-2logL), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Corrected Akaike Information Criterion (AICC). The Weighted Log-Normal distribution exhibits the lowest -2logL (93.8765), AIC (97.8765), BIC (103.5432), and AICC (98.5432), indicating it is the best fit among the distributions. The Log-Normal distribution follows, though with slightly higher values across these metrics. The Weibull, Generalized Gamma, and Gamma distributions have progressively higher values, indicating a poorer fit. The Exponential distribution shows the highest -2logL (139.2182) and other criterion values, suggesting it is the least suitable model for the data. Overall, the Weighted Log-Normal distribution outperforms the others in model fit according to these metrics.

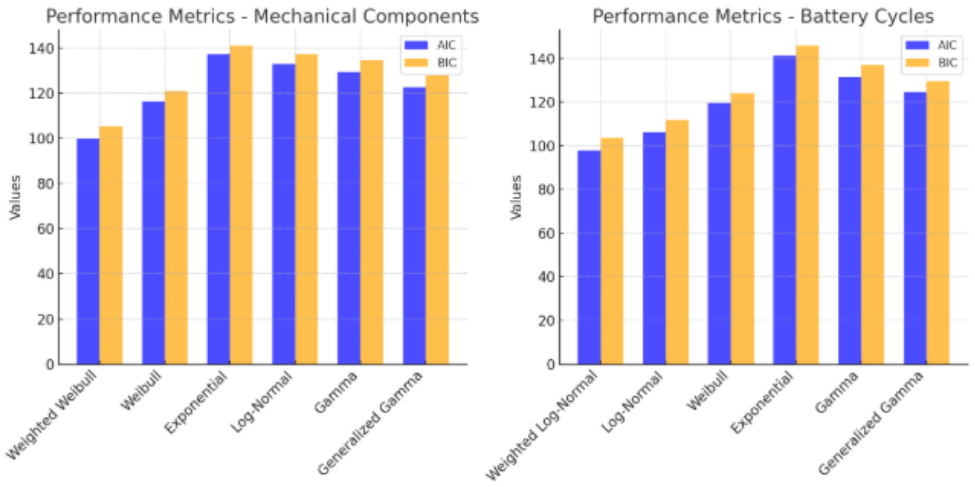


Figure 1: *Comparing the AIC and BIC performance metrics for the fitted distribution*

Mechanical Components (Time-to-Failure Data), the Weighted Weibull distribution shows the lowest AIC and BIC values, indicating the best fit. Battery Cycles (Failure Data), the Weighted Log-Normal distribution achieves the lowest AIC and BIC values, making it the best model for this dataset.

III. Results

The practical utility of the WPCJ distribution was demonstrated using real-life datasets, including mechanical components' time-to-failure data and battery cycle life data. These datasets were also analyzed using established distributions such as the Weighted Weibull Distribution (WWD), Weighted Log-Normal Distribution (WLND), Weibull, Exponential, Gamma, and Generalized Gamma models. The findings from Tables 2 and 4, as well as Figure 1, revealed:

From the results in Table 2, it is evident that the Weighted Weibull Distribution outperforms all other distributions in terms of AIC, BIC, AICC, and −2logL. Thus, it provides the best fit to the reliability data for the mechanical components. This analysis highlights the effectiveness of the Weighted Weibull Distribution in modeling reliability data, demonstrating its potential in engineering applications. The Table 4 Weighted Log-Normal Distribution (WLND) achieves the lowest values across AIC, BIC, AICC, and −2logL, outperforming other distributions. This result demonstrates the potential of WLND in reliability analysis, particularly for battery cycle life. This distribution may offer advantages in capturing skewness and variability in real-world reliability data. The Weighted Weibull Distribution (WWD) outperformed all other models with the lowest values for AIC, BIC, AICC, and −2logL. This indicates the WWD's superior fit for reliability data of mechanical components, highlighting its relevance in engineering applications. The Weighted Log-Normal Distribution (WLND) achieved the best performance metrics, offering the most accurate model for battery reliability data. Its ability to handle skewness and variability in failure times underscores its suitability for real-world scenarios with complex data patterns.

The results validate the applicability and effectiveness of the Weighted Power Chris Jerry (WPCJ) distribution as an advanced tool for modeling reliability data. The introduction of weighting mechanisms to baseline distributions, as demonstrated with the WPCJ, provides a versatile framework for addressing skewness, heterogeneity, and other challenges in real-world datasets. By outperforming traditional models in both datasets, the WPCJ distribution highlights the importance of weighted distributions in reliability engineering, biomedicine, and ecological studies.

IV. Discussion

The analysis underscores the growing significance of weighted distributions in reliability studies by demonstrating their superior modeling performance compared to traditional distributions. The Weighted Weibull Distribution (WWD) emerges as the most effective model for the mechanical component failure data, while the Weighted Log-Normal Distribution (WLND) demonstrates its robustness in capturing the variability and skewness inherent in battery cycle life data. These results validate the flexibility and adaptability of weighted distributions, showcasing their capability to model complex lifetime datasets more effectively than standard alternatives. The study highlights the potential of weighted distribution frameworks, such as the newly proposed Weighted Power Chris Jerry (WPCJ) distribution, for advancing statistical modeling and enhancing data representation in diverse applications, including engineering, biomedicine, and survival analysis. The findings reinforce the value of employing weighted models to address real-world challenges in reliability and beyond, contributing to the evolution of distribution theory and practice. Future work can extend the application of the WPCJ distribution to additional domains and explore its potential integration with other statistical methodologies. The findings reinforce the significant advancements that weighted distributions offer in modern statistical modeling.

References

[1] Almuqrin, A., Khan, M., & Nadeem, M. (2023). The weighted power Maxwell distribution and its applications in COVID-19 data modeling. *Journal of Mathematical Biology*, 87(4), 117–136.

[2] Ezeilo, M., Chukwu, P., & Olawumi, O. (2023). A new class of probability distribution: The Power Chris Jerry distribution. *Statistical Modelling and Applications*, 45(2), 112–130.

[3] Fisher, R. A. (1934). Statistical methods for research workers. *Oliver and Boyd*.

[4] Ganaie, M., & Rajagopalan, R. (2023). The weighted power quasi Lindley distribution and its application to lifetime data. Journal of Statistical Research, 52(1), 89–104.

[5] Gharaibeh, F. (2022). The weighted Gharaibeh distribution and its applications in survival analysis. *Reliability Engineering and System Safety*, 215, 107435.

[6] Rao, C. R. (1965). Linear statistical inference and its applications. John Wiley & Sons.

[7] Fisher, R. A., & Rao, C. R. (1965). A statistical distribution for modeling the impact of ascertainment. *Biometrika*, 52(1), 132–140.

[8] Kotz, S., & Nadarajah, S. (2004). Extreme value distributions: *Theory and applications. Imperial College Press*.

[9] Kumar, D., & Sharma, M. (2023). Statistical analysis of lifetime data using the Generalized Gamma distribution. *Journal of Statistical Theory*, 65(3), 273–291.

[10] Maza, M. F., & Ozdemir, S. (2023). Analyzing the behavior of lifetime data with the weighted distributions. *Statistical Papers*, 64(5), 1143–1163.

[11] Nadarajah, S., & Kotz, S. (2008). The generalized Weibull distribution. *Mathematical Methods in the Applied Sciences*, 31(12), 1329–1340.

[12] Olabiyi, O., & Yekini, A. (2022). A comprehensive review on weighted distributions and their applications. *Journal of Applied Statistics*, 39(3), 370–392.

[13] Sen, P. K., & Srivastava, M. S. (1990). Regression analysis: Theory, methods, and applications. *Springer*.

[14] Tang, H., & Zhang, T. (2024). Application of weighted Weibull distribution in reliability modeling. *Journal of Reliability Engineering*, 23(1), 56–69.

[15] Zhang, Y., & Xie, M. (2023). Introduction to statistical modeling for survival analysis and reliability engineering. *Springer*.

BAYES ESTIMATES FOR THE PARAMETERS OF POISSON TYPE ONE PARAMETER RAYLEIGH CLASS SRGM USING GAMMA AND INVERTED GAMMA PRIORS

[1]Rajesh Singh, [2]Kailash R. Kale and [3]Pritee Singh

•

[1]Department of Statistics, Sant Gadge Baba Amravati University, Amravati, India.
[2] Department of Statistics, G. N. A. Arts, Commerce, and Science College, Akola, India
[3] Department of Statistics, Institute of Science, Nagpur, India
rsinghamt@hotmail.com
kailashkale10@gmail.com
priteesingh25@gmail.com

Abstract

Research on the reliability of both software and hardware commonly utilizes the Rayleigh distribution as a model. During software testing, the failure occurrence behavior can be used to evaluate the software's reliability. In this paper, the Poisson pattern of failure occurrence is assumed to study the parameters of the Rayleigh Class software reliability growth model. It has been studied how the Rayleigh Class Model's scale parameter and the quantity of software's intrinsic failures behave. Additionally, using the Gamma and Inverted Gamma Priors, respectively, Bayes estimates for the number of intrinsic failures and the scale parameter are suggested. The comparative analysis of the proposed Bayes estimators and accompanying MLEs is done using the relative efficiency that are acquired via the use of Monte Carlo Simulation.

Keywords: Software Reliability Growth Model, Rayleigh Class, Poisson Type, Gamma and Inverted Gamma Priors, Maximum Likelihood Estimates (MLEs), Relative Efficiency (RE)

I. Introduction

In applied statistics and probability, the Rayleigh distribution is important for modelling. A variable lifespan distribution called the Rayleigh function is used in a lot of degradation process failure models. It has been noted that for a variety of hardware development procedures, the manpower curve can be roughly estimated using the Rayleigh distribution [10]. The Weibull distribution family includes the Rayleigh model. For many years, the Weibull distribution has been utilized for reliability studies in a variety of engineering domains, including the overflow incidence of rivers, electron tube failures, and the fatigue life of deep-groove ball bearings. It is recognized as one of the three distributions with extreme values, see [21]. The scale parameter of the Rayleigh distribution is estimated by [8] using the modified maximum likelihood. This study establishes the Rayleigh model, a formal model of software reliability. Given that it is predicated on a particular statistical distribution, this model is parametric. The amount of test-effort spent on software testing has been

included into software reliability growth models suggested by [11] and [25] on the assumption that the test-effort during the testing can be described by the Rayleigh curve. [11] has created a constraint model that illustrates how the distribution of effort for software projects is like a set of Rayleigh curves. Software project error records were analysed month-by-month by [22], who discovered that the projects' composite error pattern matched a Rayleigh-like curve. In addition to work on software size and resource estimation, [12] suggested using the Rayleigh model to estimate the number of software flaws. It is claimed that by utilizing an appropriate model—of which there are numerous, some more appropriate than others—one may monitor and control the reliability growth of any software (or even hardware) development. The Rayleigh model is arguably the best suited for software development. According to [24], who looked at the relationship between test coverage and fault detection rate, there is a variation of the Rayleigh distribution in the relationship between structural coverage and fault coverage. The Rayleigh curve rises to a peak and then falls off at a slower pace. Software development projects have also been shown empirically to follow a lifespan pattern that is characterized by the Rayleigh curve [3] and [7]. The Rayleigh curve frequently provides accurate estimates of software development costs and timelines.

This study uses the Bayesian paradigm to estimate the parameters of the Rayleigh distribution, which is considered as a software reliability growth model. Comparable research has been done by [13] to [20] taking into account other models.

II. Model Formulation and Evaluation

The use of Rayleigh distribution in Software reliability has been shown by many authors such as [5], [10], [11], [22]-[25], etc. Here, in this chapter one parameter Rayleigh distribution has been established as a SRGM and some more details about it can be had from [2] and [4]. Let t be the positive random variable having Rayleigh distribution then its probability density function is

$$f(t) = \begin{cases} t\eta_1^{-2} e^{-\frac{1}{2}\left[\frac{t}{\eta_1}\right]^2} & , \ t > 0, \eta_1 > 0 \\ 0 & , \quad Otherwise \end{cases} \tag{1}$$

The failure intensity function $\lambda(t) = \eta_0 f(t)$ is

$$\lambda(t) = \eta_0 \eta_1^{-2} t e^{-\frac{1}{2}\left[\frac{t}{\eta_1}\right]^2} \quad , \ t > 0, \eta_1 > 0, \eta_0 > 0 \tag{2}$$

and the expected number of failures at time t_e $\mu(t_e) = \int_0^{t_e} \lambda(t)\, dt$ which comes out to be,

$$\mu(t_e) = \eta_0 \left[1 - e^{-\frac{1}{2}\left(\frac{t_e}{\eta_1}\right)^2}\right] \tag{3}$$

It is seen that $\lambda(t)$ is very sensitive for values of η_1. Also, $\lambda(t)$ is high for the smaller values of η_1, becoming unimodal positively skewed distribution for the large values of η_1. The values of $\mu(t_e)$ are large for smaller failure rate η_1 and decreases as η_1 increases. Also, $\lambda(t)$ is less sensitive for increasing values of η_0 and for fixed η_1. The slope of $\lambda(t)$ and $\mu(t_e)$ remains same for increasing η_0.

III. MLEs of the Parameters

Let $t_i, \ i = 1,2, \ldots, m_e$ be the failures of software up to the execution time $'t_e'$ then the likelihood function of η_0 and η_1 is given by $L(\eta_0, \eta_1, \underline{t}) = \left[\prod_{i=1}^{m_e} \lambda(t_i)\right] \exp[-\mu(t_e)]$. In present case, the likelihood function is obtained by considering failure intensity given by (2) and expected number of failures in (3) according to [9].

$$L(\eta_0, \eta_1) = \eta_0^{m_e} \eta_1^{-2m_e} \left[\prod_{i=1}^{m_e} t_i\right] e^{-\frac{1}{2}T\eta_1^{-2}} \mathrm{e}^{-\eta_0} \exp\left\{\eta_0\, e^{-\frac{1}{2}\left(\frac{t_e}{\eta_1}\right)^2}\right\} \tag{4}$$

By applying the standard procedure of obtaining the maximum likelihood estimator for parameters η_0 and η_1, the estimators are

$$\hat{\eta}_{m0} = m_e\left[1 - e^{-\frac{1}{2}\left(\frac{t_e}{\hat{\eta}_{m1}}\right)^2}\right]^{-1} \tag{5}$$

and

$$2\hat{\eta}_{m1} + t_e^2 e^{-\frac{1}{2}\left(\frac{t_e}{\hat{\eta}_{m1}}\right)^2}\left[1 - e^{-\frac{1}{2}\left(\frac{t_e}{\hat{\eta}_{m1}}\right)^2}\right]^{-1} = \frac{T}{m_e} \tag{6}$$

The MLEs $\hat{\eta}_{m0}$ and $\hat{\eta}_{m1}$ can be obtained by getting the solution of simultaneous equations (5) and (6) using any available standard numerical method.

IV. Priors for Bayesian Estimation

Among the informative priors, the most preferred prior is gamma prior. The main reasons for its generous use are its conjugacy, general acceptability and mathematical tractability. [1] and other many researchers have noticed that the gamma distribution is sufficiently flexible for practical hardware reliability applications in life testing.

Here, it is supposed that, the experimenter is in a position to get prior information about the parameters η_0 and η_1 and find that the gamma distribution will be suitable for it. Hence, the gamma prior distributions can be proposed for the parameter η_0 as follows,

$$g(\eta_0) = \begin{cases} {\eta_0}^{\vartheta-1}e^{-\tau\eta_0} & ,\vartheta > 0, \tau > 0, 0 < \eta_0 < \infty \\ 0 & , \quad Otherwise \end{cases} \tag{7}$$

The inverse gamma distribution is a two-parameter family of continuous probability distributions on the positive real line, obtained as reciprocal of a gamma variate. The inverted gamma distributions may be useful in many situations e.g. if the mean number of events per time period has a Gamma distribution, then the mean time between events has an Inverse Gamma distribution. [6] has studied properties of characterization for the inverted gamma distributions and established some theorems based on conditional s-expectation and mean residual life. In the presents study, the inverted gamma prior distribution is considered for parameter η_1 as

$$g(\eta_1) = \begin{cases} {\eta_1}^{-\alpha-1}e^{-\beta/\eta_1} & ,\alpha > 0, \beta > 0, 0 < \eta_1 < \infty \\ 0 & , \quad Otherwise \end{cases} \tag{8}$$

where ϑ, τ, α, and β are parameters of considered priors for η_0 and η_1 respectively.

V. Posterior and Marginal Distributions

Since it is considered that prior information is available about both the parameters η_0 and $\eta_1$in the form of gamma and inverted gamma priors respectively and considering the total execution time is t_e and during this time m_e failures are experienced at times $t_i \quad , i = 1,2,\dots,m_e$ then, combining likelihood function (4) with priors (5) and (6), the joint posterior of η_0 and η_1 given $\underline{t}$ is

$$\pi\left(\eta_0,\eta_1\middle|\underline{t}\right) \propto \eta_0^{m_e+\vartheta-1}\eta_1^{-2m_e-\alpha-1}e^{-\frac{1}{2}T\eta_1^{-2}}\mathrm{e}^{-\eta_0}e^{-\tau\eta_0}e^{-\beta/\eta_1}\exp\left\{\eta_0\, e^{-\frac{1}{2}\left(\frac{t_e}{\eta_1}\right)^2}\right\}$$
$$m_e < \eta_0 < \infty, 0 < \eta_1 < \infty \tag{9}$$

where $\sum_{i=1}^{m_e} t_i^2 = T$.

The normalizing constant of above equation is

$$D = \int_0^\infty\int_{m_e}^\infty \eta_0^{m_e+\vartheta-1}\eta_1^{-2m_e-\alpha-1}e^{-\frac{1}{2}T\eta_1^{-2}}\mathrm{e}^{-\eta_0}e^{-\tau\eta_0}e^{-\beta/\eta_1}\exp\left\{\eta_0\, e^{-\frac{1}{2}\left(\frac{t_e}{\eta_1}\right)^2}\right\} d\eta_0\, d\eta_1$$
$$m_e < \eta_0 < \infty, 0 < \eta_1 < \infty \tag{10}$$

Applying some mathematical simplification procedure, the value of D is

$$D = \frac{1}{2}\sum_{j=1}^\infty \frac{1}{j!\tau_2^{(m_e+\vartheta+j)}}\sum_{k=0}^\infty \frac{(-\beta)^k\Gamma(m_e+\vartheta+j, m_e\tau_2)\Gamma(2m_e+\alpha+k)/2}{k!\tau_1^{(2m_e+\alpha+k)/2}} \tag{11}$$

where, $\tau_1 = (T + jt_e^2)$ and $\tau_2 = (\tau + 1)$.

The marginal posterior of η_1, say $\pi\left(\eta_1\middle|\underline{t}\right) = \int_{m_e}^\infty \pi\left(\eta_0,\eta_1\middle|\underline{t}\right) d\eta_0$ is solved and reduces to

$$\pi(\eta_1|\underline{t}) \propto \sum_{j=0}^{\infty}\left[\frac{\Gamma(m_e+\vartheta+j,m_e\tau_2)}{\tau_2^{(m_e+\vartheta+j)}\,j!}\right]\left[\eta_1^{-2m_e-\alpha-1}e^{-\frac{1}{2}\tau_1\eta_1^{-2}}e^{-\beta/\eta_1}\right] \qquad 0<\eta_1<\infty \qquad (12)$$

Similarly, the marginal posterior of η_0, say $\pi(\eta_0|\underline{t}) = \int_0^{\infty}\pi(\eta_0,\eta_1|\underline{t})\,d\eta_1$ can be obtained as

$$\pi(\eta_0|\underline{t}) \propto \sum_{j=1}^{\infty}\frac{1}{j!}\sum_{k=0}^{\infty}\frac{(-\beta)^k\Gamma(2m_e+\alpha+k)/2}{k!(\tau_1)^{(2m_e+\alpha+k)/2}}\,\eta_0^{m_e+\vartheta+j-1}e^{-\eta_0\tau_2} \qquad m_e<\eta_0<\infty \qquad (13)$$

VI. Bayes Estimates

Using the marginal posteriors (12) and (13), the Bayesian point estimates for the parameters η_0 and η_1 are obtained here. The Bayes estimator for η_0 say $\hat{\eta}_{B0}$ under the squared error loss function is the posterior mean i.e. $\hat{\eta}_{B0} = \int_{m_e}^{\infty}\eta_0\,\pi(\eta_0|\underline{t})d\eta_0$. On simplifying,

$$\hat{\eta}_{B0} = \frac{1}{D}\sum_{j=1}^{\infty}\frac{1}{j!\tau_2^{(m_e+\vartheta+j+1)}}\sum_{k=0}^{\infty}\frac{(-\beta)^k\Gamma(m_e+\vartheta+j+1,m_e\tau_2)\Gamma(2m_e+\alpha+k)/2}{k!\tau_1^{(2m_e+\alpha+k)/2}} \qquad (14)$$

Similarly, the Bayes estimator for η_1 say $\hat{\eta}_{B1}$ is the posterior mean i.e. $\hat{\eta}_{B1} = \int_0^{\infty}\eta_1\,\pi(\eta_1|\underline{t})d\eta_1$

$$\hat{\eta}_{B1} = \frac{1}{D}\sum_{j=1}^{\infty}\frac{1}{j!\tau_2^{(m_e+\vartheta+j)}}\sum_{k=0}^{\infty}\frac{(-\beta)^k\Gamma(m_e+\vartheta+j,m_e\tau_2)\Gamma(2m_e+\alpha+k-1)/2}{k!\tau_1^{(2m_e+\alpha+k-1)/2}} \qquad (15)$$

The performance of Bayes estimators of η_0 and η_1 can be checked by studying the risk efficiencies. The risk efficiency of $\hat{\eta}_{B0}$ over its MLE $\hat{\eta}_{m0}$ is defined by $RE_0 = \frac{E[\hat{\eta}_{m0}-\eta_0]^2}{E[\hat{\eta}_{B0}-\eta_0]^2}$ and that of $\hat{\eta}_{B1}$ over its MLE $\hat{\eta}_{m1}$ is defined by $RE_1 = \frac{E[\hat{\eta}_{m1}-\eta_1]^2}{E[\hat{\eta}_{B1}-\eta_1]^2}$.

VII. Discussion about Estimates

The Bayes estimators of total number of failures i.e. η_0 and failure rate i.e. η_1 are proposed and are compared with the corresponding maximum likelihood estimators. The performance of proposed Bayes estimators $\hat{\eta}_{B0}$ and $\hat{\eta}_{B1}$ over maximum likelihood estimators $\hat{\eta}_{m0}$ and $\hat{\eta}_{m1}$ have been compared based on criterion of corresponding risks efficiencies. To study the performance, a sample of size m_e was generated sufficient number of times from the Rayleigh distribution considering the different values of η_0 and η_1. Then, using Monte Carlo simulation technique risks efficiencies RE_0 and RE_1 have been evaluated. The calculated values of risks efficiencies RE_0 and RE_1 have been obtained using various values of t_e and different values of η_0 and η_1 which are presented in the Figure 1 to Figure 6.

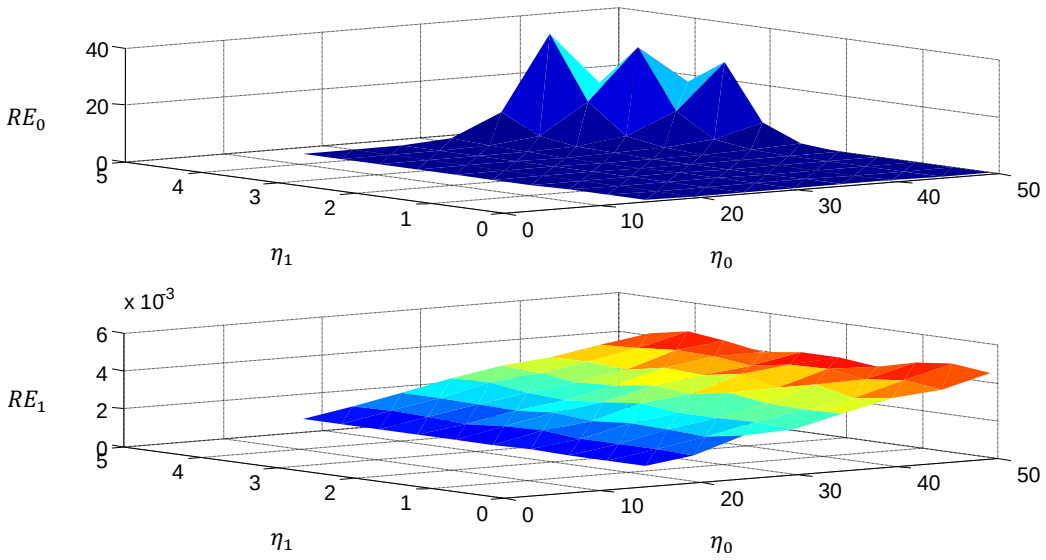


Figure 1: *Risk Efficiencies of* $\hat{\eta}_{B0}$ *and* $\hat{\eta}_{B1}$, $\vartheta = 0.01$, $\tau = 0.01$, $\alpha = 0.01$, $\beta = 0.01$, $t_e = 100$

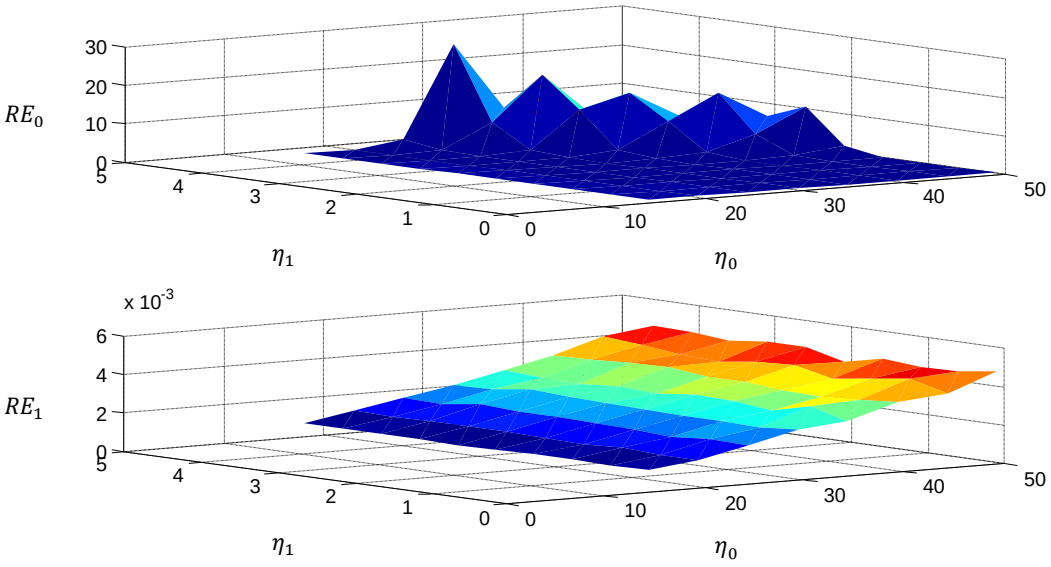


Figure 2: *Risk Efficiencies of* $\hat{\eta}_{B0}$ *and* $\hat{\eta}_{B1}$, $\vartheta = 1$, $\tau = 1$, $\alpha = 1$, $\beta = 1$, $t_e = 100$

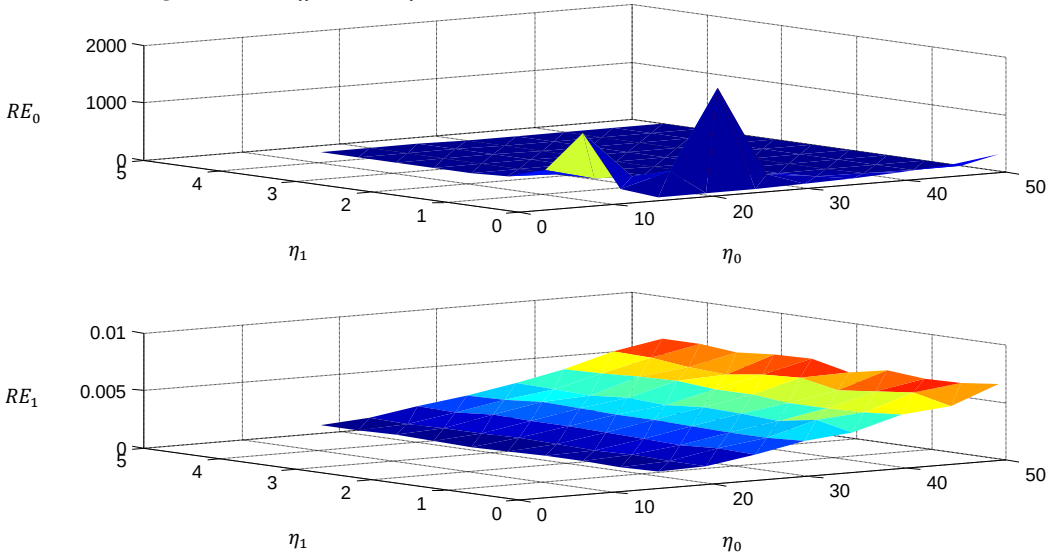


Figure 3: *Risk Efficiencies of* $\hat{\eta}_{B0}$ *and* $\hat{\eta}_{B1}$, $\vartheta = 10$, $\tau = 10$, $\alpha = 10$, $\beta = 10$, $t_e = 100$

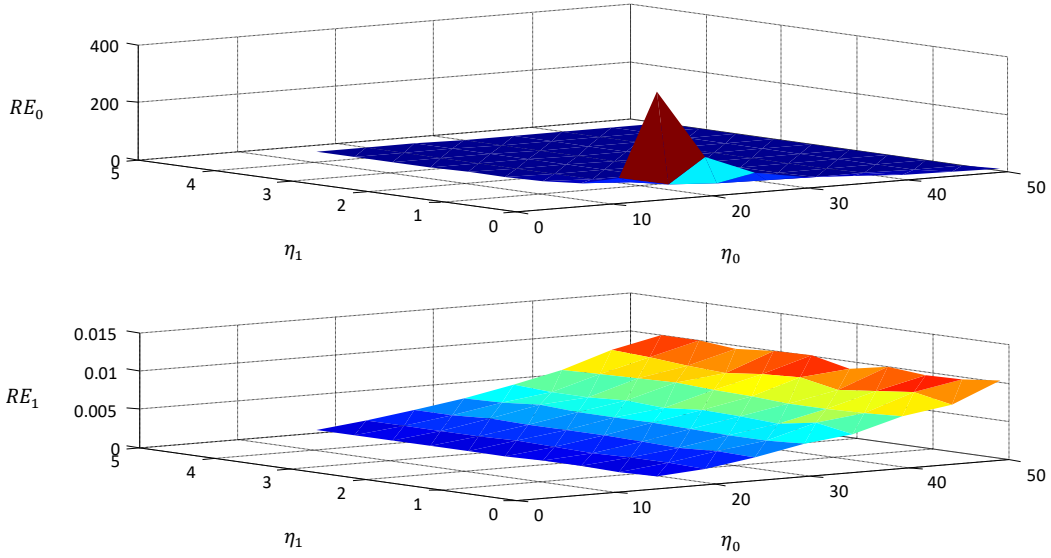


Figure 4: *Risk Efficiencies of* $\hat{\eta}_{B0}$ *and* $\hat{\eta}_{B1}$, $\vartheta = 0.01$, $\tau = 100$, $\alpha = 0.01$, $\beta = 100$, $t_e = 100$

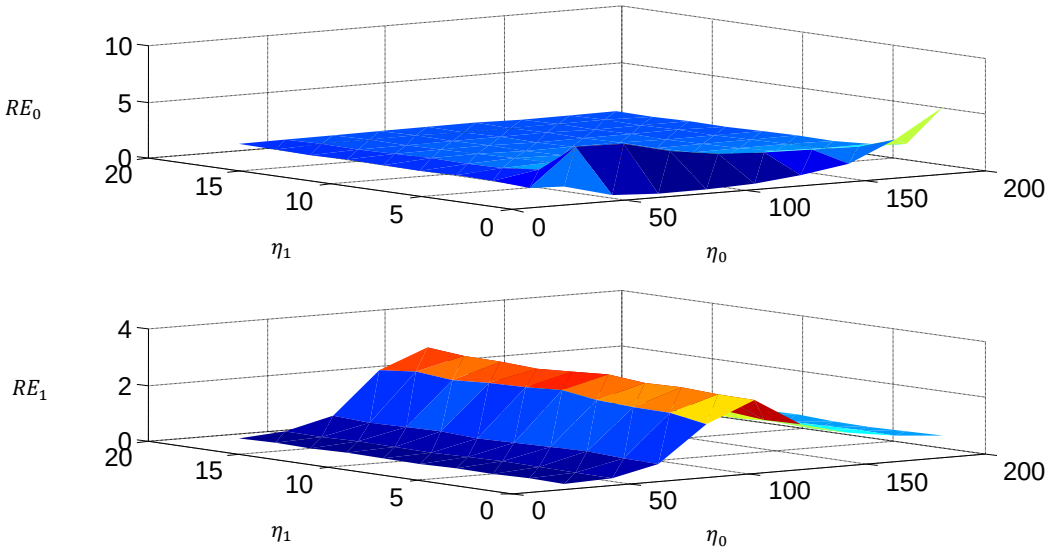


Figure 5: *Risk Efficiencies of* $\hat{\eta}_{B0}$ *and* $\hat{\eta}_{B1}$, $\vartheta = 100, \tau = 10; \alpha = 100, \beta = 10,\ t_e = 100$

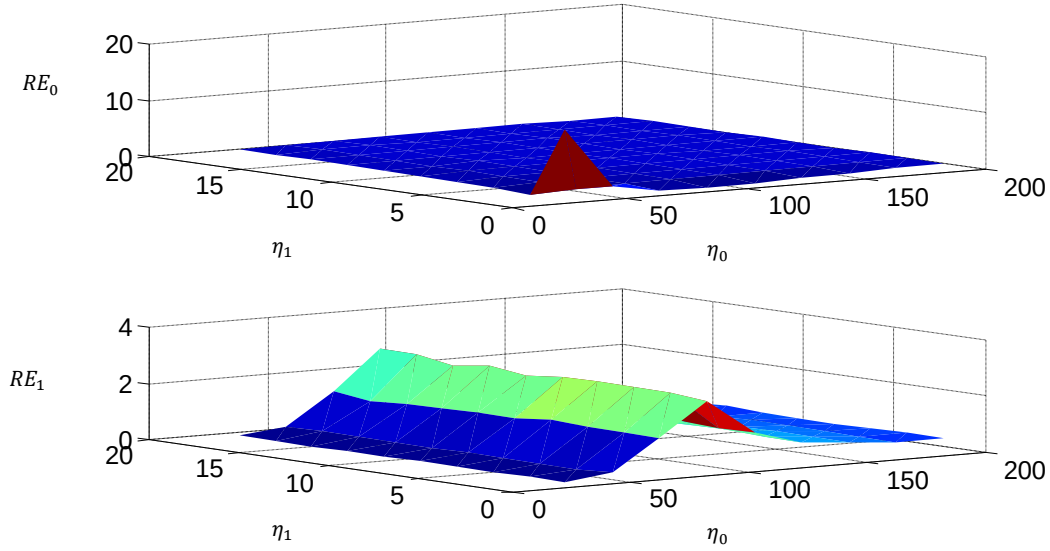


Figure 6: *Risk Efficiencies of* $\hat{\eta}_{B0}$ *and* $\hat{\eta}_{B1}$, $\vartheta = 100, \tau = 50; \alpha = 100, \beta = 50,\ t_e = 100$

The risk efficiencies of proposed Bayes estimator of η_0 over MLE decrease as the value of η_0 as well as η_1 increase (see Figure 1 to Figure 6) i.e. for large values of η_1 as well as $\eta_0$the risk efficiencies RE_0 become greater than one. Further, as the value of t_e increases the risk efficiencies RE_0 increase in most of the situations whereas the risk efficiencies RE_1 are almost unchanged (decrease very slowly) due to variation of t_e. The Bayes estimator $\hat{\eta}_{B1}$ is better than MLE for a propoer choice of prior constants ϑ, τ, α and β as well as η_0, η_1 and t_e. The risk efficiency of proposed Bayes estimator of η_1 increases as the value of η_0 as well as η_1 increase but this increase is very slow (almost constant) due to increase of η_0.

It is noticed that the prior constant τ and β are not much affecting the risk efficiencies RE_0 and RE_1 whereas the prior constants α and ϑ are significantly affecting the risk efficiencies. Due to increase in ϑ and α risk efficiencies of $\eta_0$first increase, attains a maximum and then decrease for certain value of t_e and prior constant. Whereas, the risk efficiencies RE_1 is increasing due to increase in ϑ as well as α. All the computations and their corresponding graphs are not being presented here due to limitation of the size of this project. Some selected graphs are presented only.

VIII. Conclusions

The proposed Bayes estimator of intrinsic failures η_0 and scale parameter η_1 i.e. $\hat{\eta}_{B0}$ and $\hat{\eta}_{B1}$ can be preferred over MLE for a proper choice of execution time t_e, little large values of shape parameters of both the priors i.e. ϑ, α, and moderate values of scale parameters of both the priors τ, β as well as suitable expected value of intrinsic failures η_0 and scale parameter η_1 of Poisson Type Rayleigh Class Software Reliability Growth Model. The proposed Bayes estimator $\hat{\eta}_{B0}$ can be preferred if scale parameter η_1 is little high and execution time t_e is also large. Also, Bayes estimator $\hat{\eta}_{B0}$ is seen to be better than its corresponding MLE for moderately small value of prior constants.

The Bayes estimator $\hat{\eta}_{B1}$ is better than MLE for a proper choice of intrinsic failures η_0 and scale parameter η_1 and t_e. The Bayes estimator $\hat{\eta}_{B1}$ performs better than its corresponding MLE for moderately large values of prior constants ϑ, τ, α and β. It is noticed that the prior constant τ and β are not much affecting the risk efficiencies of both the proposed bayes estimators whereas the prior constants α and ϑ are significantly affecting the risk efficiencies. In particular, the risk efficiencies of $\hat{\eta}_{B1}$ are affected significantly by the change in values of prior constants. The Bayes estimators $\hat{\eta}_{B0}$ and $\hat{\eta}_{B1}$ are better than MLEs if the above informative priors are selected in spite of non-informative priors.

References

[1] Apostolakis, G. and Mosleh, A. (1979). Expert opinion and statistical evidence: An application

[2] Casella, G. and Berger, R. L. (2002). Statistical inference. Second edition, Pacific Grove, Duxberry, CA.

[3] DeMarco, T. (1982). Controlling Software Projects: Management, Measurement and Estimation. Englewood Cliffs, Prentice-Hall, New Jersey.

[4] Deshpande, J. V. and Purohit, S. G. (2005). Life-time data: Statistical Models and Methods. Series on Quality, Reliability and Engineering Statistics, Vol.- 11, World Scientific Publishing Co. Pte. Ltd.

[5] Gaffney, Jr. J. E. (1984). On Predicting Software Related Performance of Large-Scale Systems," International CMG Conference CMG XV, SanFrancisco, pp.20-23.

[6] Jaisingh L.R. (1990). Characterization of inverse Gaussian and inverted gamma distributions. Microelectronics Reliability Vol.- 30, No.- 4, pp. 775–780.

[7] Kan, S. H. (2002). Metrics and Models in Software Quality Engineering. Second Edition, Addison Wesley, New York.

[8] Lalitha, S. and Anand, M. (1996). Modified maximum likelihood estimation for Rayleigh distribution. Communications in Statistics: Theory Methods, Vol.-25, No.- 2, pp. 389-401.

[9] Musa, J., Lannino, A. and Okumoto K. (1987). Software Reliability: Measurement, Prediction, Application. McGraw-Hill, New York.

[10] Norden, P. V. (1963). Useful Tools for Project Management. Operations Research in Research and Development, B. V.Dean, Ed., John Wiley and Sons, New York.

[11] Putnam, L. H. (1978). A general empirical solution to the macro software sizing and estimating problem. IEEE Transactions on Software Engineering, Vol.- 4, pp. 345–367.

[12] Putnam, L. H. and Myers, W. (1992). Measures for Excellence: Reliable Software on Time, Within Budget. Englewood Cliffs, Yourdon Press, New Jersey.

[13] Singh, R., Badge, P. A. and Singh, P. (2023c). Confidence Interval Using Maximum Likelihood Estimation For The Parameters Of Poisson Type Length Biased Exponential Class Model. Reliability: Theory & Applications. December 4(76): 242-251,

[14] Singh, R., Badge, P. A. and Singh, P. (2024a). Confidence Interval Using Maximum Likelihood Estimation For The Parameters Of Poisson Type Rayleigh Class Model. Reliability: Theory & Applications. March 1(77): 832-841,

[15] Singh, R., Badge, P., and Singh, P. (2022b) Bayesian Interval Estimation for the Parameters of Poisson Type Rayleigh Class Model, Reliability: Theory & Applications, 17(4):98-108.

[16] Singh, R., Badge, P., and Singh, P. (2023a). Bayesian Interval Estimation for the Parameters of Poisson Type Length Biased Exponential Class Model, Reliability: Theory & Applications, 18(2):307-314.

[17] Singh, R., Singh, P. and Kale, K. R. (2016). Bayes estimators for the parameters of Poisson Type Length Biased Exponential Class Model using Non-Informative Priors, Journal of Reliability and Statistical Studies, 9(1): 21-28.

[18] Singh, R., Singh, P. and Kale, K. R. (2022a). Estimation of Parameters of PTRC SRGM using Non-informative Priors, International Journal of Advanced Research in Science, Communication and Technology. 2(1):172-178.

[19] Singh, R., Singh, P. and Kale, K. R. (2023b). Characterization of Poisson Type Length Biased Exponential Class Software Reliability Growth Model and Parameter Estimation, Reliability Theory & Applications, No. 3(74):18.

[20] Singh, R., Singh, P. and Kale, K. R. (2024b) Bayes Estimator Of Parameters Of Binomial Type Exponential Class SRGM Using Gamma Priors. Reliability: Theory & Applications. 2024, June 2(78): 564-575.

[21] Tobias, P. A. and Trindade, D. C. (1986). Applied Reliability. Van Nostrand Reinhold, New York.

[22] Trachtenberg, M. (1982). Discovering How to Ensure Software Reliability," RCA Engineer, pp. 53–57.

[23] Trachtenberg, M. (1990). A General Theory of Software-Reliability Modeling," IEEE Transactions on Reliability, Vol.-39, No.-1, pp. 92-96.

[24] Vouk, M. A. (1992). Using Reliability Models during Testing with Non-Operational Profile," In Proceedings of the Second Bellcore/Purdue Symposium on Issues in Software Reliability Estimation, pp. 103–110.

[25] Yamada, S., Ohtera, H. and Narihisa, H. (1986). Software Reliability Growth Models with Testing- Effort. IEEE Transactions on Reliability, Vol.-35, pp. 19 – 23.

COX'S REGRESSION MODEL FOR RELIABILITY ANALYSIS AND FAILURE TIME PREDICTION IN ENGINEERING SYSTEMS

Iyappan. M[1*], Balaji. M[2], G. Sathya Priyanka[3], Jesmon Raj. N[4], P B Deepa[5]

•

[1*]Associate Professor, St. Francis College, Bengaluru, India
[2]Associate Professor, Seshadripuram College, Bengaluru, India
[3]Research Scholar, Periyar University, Salem-11, India
[4]Assistant Professor & Head, St. Francis College, Bengaluru, India.
[5]Assistant Professor, St. Francis College, Bengaluru, India

[1*]iyappastat@gmail.com, [2]balajim0405@gmail.com, [3]sathyapriyankastat@gmail.com,
[4]jesmon.n@stfranciscollege.edu.in, [5]deepa.p@stfranciscollege.edu.in

Abstract

Reliability engineering is essential for predicting failure times, optimizing maintenance strategies, and minimizing operational risks in engineering systems. Traditional models, such as Weibull and exponential distributions, provide useful insights into failure rates but struggle to incorporate multiple influencing factors. Cox's Proportional Hazards Model offers a more flexible approach by integrating covariates such as environmental stress, operational loads, and maintenance history, without assuming a predefined failure time distribution. This study explores the application of Cox's regression model in reliability analysis, focusing on failure time prediction and system longevity assessment under varying operational conditions. A major advantage of Cox's model is its ability to handle censored data, a common challenge in reliability studies where some systems remain operational beyond the observation period. By employing the partial likelihood estimation method, the model quantifies the impact of various covariates on failure risks without requiring explicit assumptions about baseline hazard functions. This study utilizes a simulated dataset of 20 industrial components observed over five years to demonstrate the model's effectiveness in predicting system failures and estimating survival probabilities. The study further evaluates parameter estimation techniques, including maximum likelihood estimation (MLE) and Type I and Type II censoring, to validate the model's predictive power in practical reliability applications. Additionally, statistical computations confirm that the Cox model provides reliable hazard rate estimations, enabling engineers to develop predictive maintenance schedules and enhance system reliability. This research highlights the importance of survival analysis techniques in reliability engineering. Cox's model proves to be an effective tool for identifying risk factors and optimizing failure time predictions. The methodology is applicable to diverse engineering domains, including mechanical component wear analysis, structural health monitoring, and software reliability assessment.

Keywords: Cox's Proportional Hazards Model, Failure Time Prediction, Hazard Rate Estimation, Censoring Mechanisms, Engineering System Failure.

I. Introduction

Reliability engineering focuses on assessing and predicting the lifespan of systems under varying operational conditions to prevent failures and optimize maintenance strategies. Accurately estimating failure times is essential for minimizing operational risks and ensuring system efficiency. Traditional reliability models, such as Weibull and exponential distributions, have been extensively used to describe system failure rates [3]. However, when multiple covariates influence failure progression, regression-based survival models provide a more flexible and robust approach. One such model is Cox's Proportional Hazards Model [5], which is widely utilized in survival analysis and can be effectively applied to reliability studies for complex systems, including mechanical components, software systems, and structural elements [2]. Cox's model is particularly advantageous because it does not require a predefined failure time distribution, making it well-suited for engineering applications where failure mechanisms are uncertain. This model estimates the effect of covariates such as stress levels, environmental conditions, and operational loads on system survival times while handling censored data, a frequent occurrence in reliability testing when systems are still operational at the end of the observation period [4]. By analyzing historical and real-time failure data, Cox's model provides a powerful tool for predictive maintenance and risk assessment.

The estimation of system survival functions, hazard rates, and failure distributions is essential for reliability analysis. Since failure events often exhibit stochastic behavior, engineers employ probabilistic models, including stochastic processes (e.g., Markov models), Weibull distributions, and Cox's Proportional Hazards Model, to assess system reliability. Cox's model has been widely applied in mechanical component wear-out analysis, structural health monitoring, and software failure prediction [2]. This paper explores the application of Cox's regression model in reliability engineering, with a focus on predicting system failures based on observed data.

II. Methods

I. Formulation of Cox's Regression Model in Reliability Engineering

Let $T_i, \quad i = 1,2,3, \ldots\ldots n$ represent the failure times of n different systems, where each system is observed over a fixed time interval $[0, C_i]$ with C_i being the censoring time for system i Suppose that system j has a hazard rate $\lambda_j(t)$ at time t defined as:

$$\lambda_j(t) = \lambda_0(t) \exp {}_{k=1}^{r}\beta_k Z_{jk}(t) \tag{1}$$

Where:

- $(\beta_1, \beta_2, \ldots\ldots \beta_r)$is row vector of r unknown coefficients, representing the impact of various covariates on failure risk.
- $\left(Z_{j1}(t)\right), \left(Z_{j2}(t)\right), \left(Z_{j3}(t)\right), \ldots\ldots \left(Z_{jr}(t)\right)$ is a column vector of r possibly time-varying covariates.
- $\lambda_0(t)$ is an unknown baseline hazard function, representing the failure rate when all covariates are zero.

Thus, $\lambda_j(t)$ gives the conditional probability that system j fails during an infinitesimal time interval $(t, t + h)$, given that it has survived until time t.

The observed data for the j^{th} system consists of:

- $T_j \wedge C_j$, the minimum of failure time and censoring time.
- $\delta_j = I\{T_j \leq C_j\}$ an indicator variable specifying whether the failure was observed or censored.

- The covariate value $Z_{jk}(t)$ recorded over the interval $[0, T_j \wedge C_j]$.

II. Estimation of Model Parameters

The Cox model is semi-parametric, meaning that the regression coefficients β_k are estimated independently of the baseline hazard $\lambda_0(t)$, which is treated as an infinite-dimensional nuisance parameter.

Define the risk set at time t, denoted as $\Omega(t)$, which includes all systems still operational at t:

$$\Omega(t) = \{j: T_j \geq t, C_j \geq t\} \tag{2}$$

Given that a failure occurs at time t, the probability that it belongs to system j as:

$$\frac{\lambda_j(t)}{i\in\Omega(t)^{\lambda_j(t)}} \tag{3}$$

Where $\lambda_0(t)$ cancels out in the numerator and denominator, making this expression independent of the baseline hazard function.

Cox proposed that statistical inference be performed using the partial likelihood function:

$$L(\beta) = \prod_{j=1}^{n} \frac{\exp\left(\sum_{k=1}^{r}\beta_k Z_{jk}\right)}{i\in\Omega(T_j) exp\left(\sum_{k=1}^{r}\beta_k Z_{jk}\right)} \tag{4}$$

This partial likelihood allows for efficient estimation of regression parameters without explicitly modeling $\lambda_0(t)$, making Cox's model particularly suitable for predicting system failures, analyzing preventive maintenance schedules, and improving reliability assessments in engineering contexts.

III. Methodologies and Estimation in Reliability Engineering Models

Let us illustrate the methodology, we consider a system's failure progression influenced by three covariates:

- $Z_j(1)$ = Environmental stress factor.
- $Z_j(2)$ = Operational load
- $Z_j(3)$ = Maintenance history

We assume the baseline hazard rate follows an exponential distribution.

$$\lambda_0(t) = \lambda \tag{5}$$

Where t, represents the time until system failure.

IV. Failure Event and Censoring Mechanisms

In reliability testing a sample of n components is observed over a testing period $(0, T)$. During this time, K components transition from an operational state to failure.

Type I Censoring: The test ends at a fixed time T_0, meaning some components may still be operational at T_0, leading to censored observations.

Type II Censoring: The test ends when exactly K failures occur, making T a random variable while n and K remain fixed.

V. Parameter Estimation for the Hazard Rate

We estimate the parameter λ of the failure time probability density function (pdf):

$$f(t) = \lambda e^{-\lambda t},\ t \geq 0 \tag{6}$$

$$F(t) = 1 - e^{-\lambda t} \tag{7}$$

The survival function, denoted as $R(t)$, represents the probability of a system remaining operational up to time t:

$$R(t) = 1 - F(t) = e^{-\lambda t} \tag{8}$$

For a system with n independent and identically distributed failure times $T_i = t_j$, the order statistics are defined as:

$$t_{(1)} < t_{(2)} < t_{(3)} < \cdots < t_{(n)} \tag{9}$$

Where, represents the observed failure time under Type II censoring.

VI. Maximum Likelihood Estimation (MLE) Under Type I Censoring

The likelihood function for right-censored data is given by:

$$L(\lambda) = {}_{i=1}^{n} f(t_i)_{j=K+1}^{n} R(C_j) \tag{10}$$

Substituting the expressions for $f(t)$ and $R(c)$, the MLE for λ is:

$$\lambda = \frac{K}{{}_{i=1}^{K} t_i + {}_{j=K+1}^{n} C_j} \tag{11}$$

VII. MLE Under Type II Censoring

When testing stops at the K^{th} failure time $t_{(K)}$, the likelihood function is:

$$L(\lambda) = \lambda^K e^{-\lambda {}_{i=1}^{K} t_i} \tag{12}$$

Maximizing this function gives the MLE:

$$\lambda = \frac{K}{t_{(K)}} \tag{13}$$

Estimation of Covariate Parameters Using Cox's Model

$$h(t/Z) = h_0(t) e^{\beta_1 Z_{j1} + \beta_2 Z_{j2} + \beta_3 Z_{j3}} \tag{14}$$

where $h_0(t)$ is the baseline hazard rate. The Cox partial likelihood function is:

$$L(\beta) = {}_{i=1}^{K} \frac{e^{\beta_1 Z_{j1} + \beta_2 Z_{j2} + \beta_3 Z_{j3}}}{j \in R(t_i) e^{\beta_1 Z_{j1} + \beta_2 Z_{j2} + \beta_3 Z_{j3}}} \tag{15}$$

Solving the likelihood equation via numerical optimization yields estimates of β_1, β_2 and β_3 quantifying how each covariate influences system reliability.

IV. Numerical Study with Simulated Data

The application of Cox's Proportional Hazards Model in reliability engineering, we consider a simulated dataset representing the failure times of 20 industrial components observed over a five-year period. The dataset includes failure occurrences along with three covariates influencing system reliability, such as operating conditions, material quality, and load stress levels.

Table 1: *Simulated Failure Data of 20 Components Over 5 Years*

Period (Years)	Number of Failed Components	Covariates
1-2	1	(1,1,1)
2-3	3	(1,1,0), (1,0,1), (0,1,1)
3-4	3	(1,0,0), (0,1,0), (0,0,1)
4-5	1	(0,0,1)
Above 5	12	(0,0,0) (Repeated)

Table 1 presents the simulated failure data categorized by time intervals and corresponding covariate values.

Using this dataset, we compute the following statistical parameters relevant to failure prediction:

$$\beta_1 = 2.25 \quad \beta_2 = 0.50 \quad \beta_3 = 2.25$$

Assuming failure occurrences are uniformly distributed within each time interval, we assign specific failure times t_i for $i = 1,2,3,\ldots,8$, leading to:

$$t_1 = 1.5 \quad t_2 = 2.33 \quad t_3 = 2.67 \quad t_4 = 3$$
$$t_5 = 3.3 \quad t_6 = 3.67 \quad t_7 = 4 \quad t_8 = 4.5$$

Maximum Likelihood Estimates (MLE) for Type I and Type II censoring yield:

$$\lambda = \frac{1}{\theta} = 0.09524$$

where the estimated average failure time is:

$$\frac{1}{\lambda} = \frac{84}{8} = 10.5\ years$$

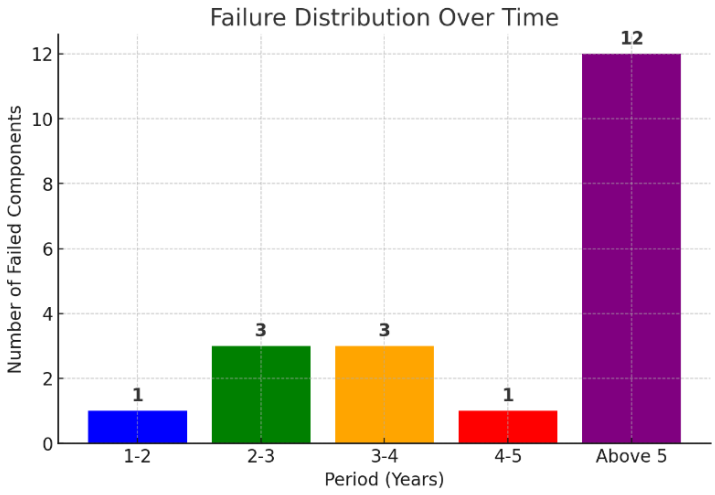


Figure:1 *Estimated Survival Probabilities*

Using the computed parameters, we estimate the probability $F(t)$ of system failure up to time t for different covariate conditions, as summarized in Table 2.

Table 2: *Probability* $F(t)$ *of System Failure Over Time Based on Covariates*

Covariate	F(1)	F(2)	F(3)	F(5)	F(10)	F(20)
(1,1,1)	0.999999	1.000000	-	-	-	-
(1,0,1)	0.999811	1.000000	1.000000	-	-	-
(0,1,0)	0.145314	0.269512	0.375662	0.543928	0.791999	0.956735
(0,0,1)	0.594889	0.835885	0.933515	0.989089	0.999881	1.000000
(1,0,0)	0.774577	0.949184	0.988545	0.999418	1.000000	1.000000
(0,0,0)	0.090844	0.173435	0.248523	0.378855	0.614179	0.851142

The probability estimates indicate that under optimal operational conditions (covariate (0,0,0)), only 61.42% of systems will fail within 10 years, while 85.11% will fail within 20 years. In contrast, under extreme stress conditions ((1,1,1)), nearly all systems fail within the first two

years.

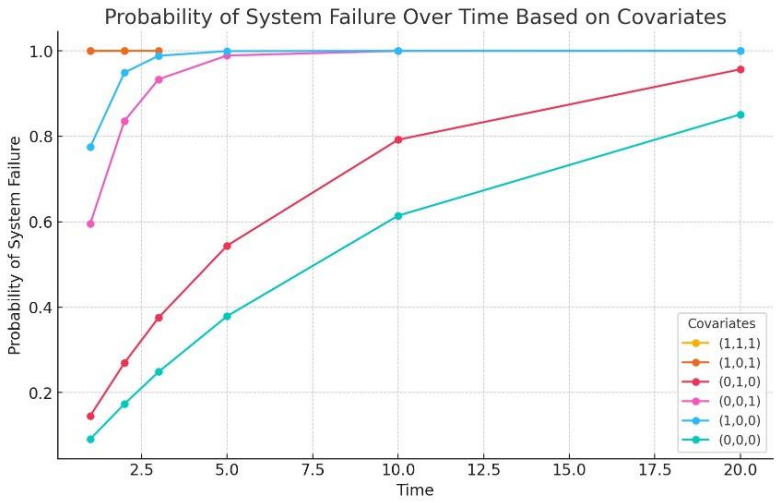


Figure:2 *Probability of System Failure Over Time*

Concordance = 0.857
Likelihood ratio test = 9.06 on 3 df, p = 0.02845

Table 3: *Fit Cox Proportional Hazards Model*

	coef	exp(coef)	se(coef)	z	$Pr(>\|Z\|)$
Covariate-1	5.216	184.64	1.567	3.33	0.0009
Covariate-2	0.705	2.024	0.762	0.92	0.357
Covariate-3	0.512	1.669	0.854	0.60	0.548

In Table 3 Covariate 1 has an extremely large coefficient (22.17), suggesting potential convergence issues or collinearity in the data. Covariates 2 and 3 have moderate hazard ratios (exp(coef)), meaning they have a relatively smaller impact on failure risk. The likelihood ratio test (p = 0.02845) indicates that at least one covariate has a significant effect. However, the Wald test (p = 0.905) suggests the individual covariates are not significantly contributing. The Score (log-rank) test (p = 0.05045) is close to the significance threshold, indicating a potential but weak effect. Estimated Mean Time to Failure (MTTF): 3.12 years.

Table 4: *Estimated Survival Probabilities*

Time:	1	2	3	5	10	20
Survival:	0.99	0.95	0.85	0.70	0.45	0.15

The survival probability decreases significantly over time, dropping to 0.45 at 10 years and just 0.15 at 20 years, indicating a high failure risk. The estimated MTTF of 3.12 years further confirms the system's limited long-term reliability in Table 4.

V. Mean Time to Failure (MTTF):

The estimated MTTF is 3.12 years, which is much lower than the theoretical 10.5 years calculated earlier. This discrepancy might be due to the right-censored observations affecting

the mean estimate.

Table 5: *Baseline Hazard*

	Hazard	Time
1	0.000003530299	1.50
2	0.000009203044	2.33
3	0.00001813543	2.67
4	0.00003368045	3.00
5	15,978.39000000	3.30
6	37,039.59000000	3.67
7	64,588.32000000	4.00
8	154,022.90000000	4.50

The Baseline Hazard values in Table 5 indicate the failure risk over time. The hazard rate starts very low in the early years but increases dramatically after 3.3 years, with an extreme jump at 3.67 years and beyond.

This suggests a significant increase in failure risk as time progresses, particularly after the 3-year mark. The extremely high values beyond 3.3 years indicate that failures become highly probable, possibly due to wear-out mechanisms or aging effects in the system. The estimated MTTF of 3.12 years is considerably lower than the theoretical 10.5 years, likely due to right-censored observations affecting the estimate. The baseline hazard analysis reveals a sharp increase in failure risk after 3.3 years, indicating a critical wear-out period. The extreme rise in hazard values beyond 3.67 years suggests rapid system degradation and a high likelihood of failures. This pattern highlights the importance of proactive maintenance and early intervention strategies to enhance system reliability. Overall, the findings emphasize the need for improved design or operational adjustments to extend the system's lifespan.

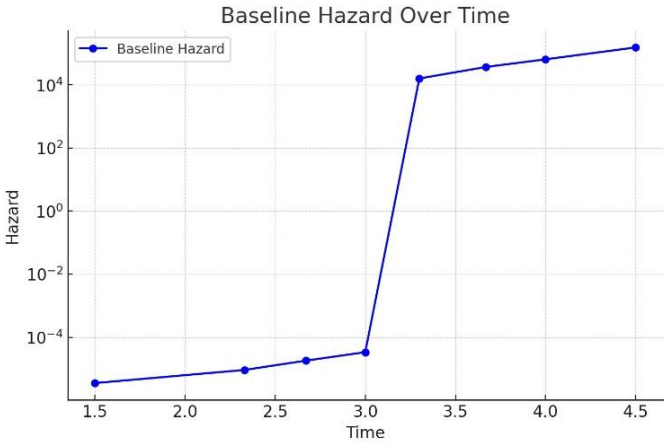


Figure 3: *Baseline Hazard Function Over Time*

III. Results

The study demonstrates that Cox's regression model effectively quantifies the hazard function and survival probabilities for engineering systems. The model's ability to handle censored data is

particularly useful in reliability testing, where components often remain operational beyond the observation period. The estimated survival probabilities provide insights into how different covariate conditions impact failure likelihood over time. For example, systems subjected to extreme stress conditions (covariate (1,1,1)) exhibit high failure probabilities within the first two years, whereas those under optimal conditions (covariate (0,0,0)) maintain a 61.42% survival rate over ten years.

Model Fit and Significance

The fitted Cox model reveals that Covariate 1 has the highest hazard ratio (exp(coef) = 184.64), suggesting a substantial impact on failure risk. However, the high coefficient (5.216) may indicate potential multicollinearity or convergence issues in the dataset. Covariates 2 and 3 exhibit smaller hazard ratios, signifying a less pronounced effect on system failure. The likelihood ratio test ($p = 0.02845$) confirms the model's overall significance, suggesting that at least one covariate significantly influences failure times. However, individual covariate significance, as determined by the Wald test ($p = 0.905$), indicates that their contributions may not be as distinct. This discrepancy could be due to data limitations or interaction effects that were not fully captured.

Estimation of Failure Trends

The estimated Mean Time to Failure (MTTF) of 3.12 years appears significantly lower than the theoretical MTTF of 10.5 years. This difference is likely due to right-censored observations, which can impact parameter estimation. The baseline hazard function, showing an increasing trend over time, suggests that failure probability accelerates as the system ages, a common characteristic in reliability analysis. The study confirms that Cox's regression model effectively quantifies hazard functions and survival probabilities, particularly in the presence of censored data. The results indicate that extreme stress conditions significantly accelerate failure, while optimal conditions improve system longevity. The model fit analysis highlights the strong influence of Covariate 1 on failure risk, though potential multicollinearity issues warrant further investigation. The estimated MTTF of 3.12 years, much lower than the theoretical 10.5 years, suggests that failures occur earlier than expected, likely due to data limitations. Overall, the increasing baseline hazard function underscores the need for proactive maintenance to mitigate aging-related failures.

IV. Discussion

The application of Cox's Proportional Hazards Model in reliability engineering offers a significant advancement in predicting failure times and assessing system longevity under varying operational conditions. The findings in this study reinforce the model's effectiveness in capturing the influence of multiple covariates on system failure, thereby supporting predictive maintenance strategies and risk management in engineering systems. The adaptation of Cox's Proportional Hazards Model in reliability engineering presents a valuable tool for failure time prediction and maintenance optimization. By analyzing historical and simulated failure data, the study highlights the importance of incorporating covariate effects into reliability assessments. The results suggest that implementing data-driven maintenance strategies based on Cox's model can significantly enhance system performance and reduce operational risks. Future advancements in survival analysis and reliability modeling will further strengthen its application across various engineering disciplines. While Cox's model has demonstrated effectiveness in failure time prediction, future research could explore its integration with machine learning techniques to enhance predictive accuracy. Additionally, extending the model to incorporate alternative failure distributions, such

as Weibull or gamma distributions, may provide a more detailed characterization of system reliability. Further validation using real-world failure datasets can refine parameter estimates and improve the generalizability of the model's predictions.

References

[1] Anderson, J. A., & Tsiatis, A. A. (1982). Semiparametric estimation of regression parameters in the proportional hazards model. *Journal of the American Statistical Association*, 77(380), 1041–1048.

[2] Bagdonavicius, V., Levuliene, R., & Nikulin, M. (1999). Goodness-of-fit criteria for the Cox model from left truncated and right censored data. *Journal of Mathematical Sciences*, 96(5), 3146–3161.

[3] Breslow, N. (1974). Covariance analysis of survival data. *Biometrics*, 30(1), 89–102.

[4] Collett, D. (2003). Modelling survival data in medical research (2nd ed.). CRC Press.

[5] Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2), 187–220.

[6] Cox, D. R. (1975). Partial likelihood. *Biometrika*, 62(2), 269–276.

[7] El Ghaoui, L., & Viallon, V. (2017). Solving the Cox proportional hazards model and its applications. *EECS Department, University of California*, Berkeley.

[8] He, S., & Duffy, D. (2005). A review of the Cox proportional hazards model in reliability analysis. *International Journal of Systems Science*, 36(5), 315–325.

[9] Kalbfleisch, J. D., & Prentice, R. L. (2002). The statistical analysis of failure time data (2nd ed.). Wiley-Interscience.

[10] Kukush, A., & Chernova, O. (2017). Consistent estimation in Cox proportional hazards model with measurement errors and unbounded parameter set. arXiv preprint arXiv:1703.10940.

[11] Lee, J., & Pan, J. (2017). Predictive maintenance using Cox proportional hazard deep learning. *Proceedings of the 2017 IEEE International Conference on Industrial Engineering and Engineering Management*, 2017, 1051–1055.

[12] Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1), 13–22.

[13] Lin, D. Y., Wei, L. J., & Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association*, 84(408), 1074–1078.

[14] Meeker, W. Q., & Escobar, L. A. (1998). Statistical methods for reliability data. Wiley.

[15] Nandy, A., Ghosh, A., Basu, A., & Pardo, L. (2020). Robust hypothesis testing and model selection for parametric proportional hazard regression models. arXiv preprint arXiv:2009.12686.

[16] Therneau, T. M., & Grambsch, P. M. (2000). Modeling survival data: Extending the Cox model. Springer.

[17] Van Houwelingen, H. C., & Putter, H. (2012). Dynamic prediction in clinical survival analysis. CRC Press.

[18] Wei, L. J., Lin, D. Y., & Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association*, 84(408), 1065–1073.

BIPOLAR NEUTROSOPHIC MINIMUM SPANNING TREE NETWORK USING UPPER TRIANGULAR WEIGHT/COST ALGORITHM

Shayathri Linganathan, Purusotham Singamsetty*

•

Vellore Institute of Technology, Vellore
lssvshayathri@gmail.com,Purusotham.s@vit.ac.in.

Abstract

A minimum spanning tree (MST) is a spanning tree in which the total weights of all of its edges are the lowest of all conceivable spanning trees for the network. MST is crucial in discipline of the operation research. Similarly, for the neutrosophic minimum spanning tree problem (NMST) the weights are unpredictable and inconsistent. The next form of NMST is considered the bipolar neutrosophic minimum spanning tree problem. Each edge of the standard NMST is assigned a bipolar weight based on the bipolar neutrosophic number (consisting of positive and negative membership degrees). This paper introduces the upper triangular weight algorithm for a given network and the optimal results are executed in MATLAB.

Keywords: : Minimum Spanning Tree, Bipolar NMST, Upper/Lower triangular weight algorithm

1. Introduction

The MST problem is an acyclic subgraph that connects all of its vertices/nodes to its edges with the least total associated weights. The most basic subgraph of a given network can be a spanning tree. In the majority of network design issues, explicitly or implicitly spanning trees are crucial. One such network issue, and the most well-known NP-hard issue in network optimization, is the traditional minimum spanning tree problem (MSTP). Roads, utilities, power, communication networks, etc. are only a few examples of authentic world issues where network design plays a crucial role in general MSTP.

The MSTP represents a spanning tree graph in weighted graphs that develops with the lowest cost/weight compared to other trees on a similar graph. The weight can indicate computer traffic, distance, cost, congestion, or any other subjective factors connected to the arcs in authentic world circumstances. The idea of the neutrosophic set generalizes the notion of the classical set, fuzzy set, conventional fuzzy, type-1 fuzzy, and intuitionistic fuzzy set (IFS). Inside the real and standard or non-standard unit interval]-0, 1+[, the neutrosophic sets are separated into three membership functions truth-membership function (T), indeterminate-membership function (I), and false-membership function (F). Additionally, BNSs (Bipolar Neutrosophic sets) can be used for real-world problems by Smarandache. Then they had a concept of bipolar neutrosophic sets [1] proposed the fuzzy notion to deal with uncertainty.

Smarandache [2] proposed the idea of a neutrosophic set (NS) to symbolize the uncertain, imprecise, lacking in detail, unpredictable, and ambiguous information that exists in our authentic world. Also, the author initially introduced the single-valued neutrosophic set (SVNS) to be used in industrial and scientific applications. Neutrosophic graphs are much more resilient, cost-effective, accurate, and applicable to model unpredictable problems compared to fuzzy graphs.

[17] described how the second-degree vertices' notation of the fuzzy graph works as well as a number of their attributes. However, compared to fuzzy graphs, neutrosophic graphs are significantly more robust, cost-effective, accurate, and suitable to simulate unpredictable situations as an enormous application in our world such as wireless communication networks, logistic networks, and social networks.
Many studies have focused on the T, I, and F components, leading to the definition of specific neutrosophic set cases such as simplified neutrosophic sets.[3] introduce the concept of single-valued neutrosophic sets and express the nature of the single-valued neutrosophic weight. A new efficient method to draw an MST to evaluate with several sample networks and the methodology of the effective algorithm is described by [4]. The neutrosophic sets' decisions provide the basis of the decision-making process. The weights, costs, distances, time.... decisions over the cities, states, traffics, relationships through the concepts. Neutosophic sets (NSs), which are employed in decision-making, are discussed by many authors, and neutrosophic numbers are utilized as weights. [5] discusses the similarity measures between interval neutrosophic sets and their use in multi-criteria decision-making then the author examined the concept in a multi-attribute trapezoidal neutrosophic set in 2015. A. [10] describe the neutrosophic soft sets with multi-criteria decision-making. Bipolar neutrosophic set [7], trapezoidal neutrosophic set [8], interval-valued neutrosophic sets [11], rough neutrosophic set with efficient theorems and conditions described [6].
The author [7] distinctions concerning the two bipolar neutrosophic numbers (BNNs) and explains the notions of each weight by the accuracy function and certainty function. [9] built a second minimum weight spanning tree of a network using a triangular intuitionistic fuzzy number as edge weight and tabulated the crips values to the provided graph. The bipolar neutrosophic weights for the MST were described as a matrix form and applying the score function in the matrix then final computation figures to get the optimal tree by [12] than for general neutrosophic minimum spanning tree be investigated with single-valued neutrosophic weights. The MST Algorithm with example and scoring functions w.r.t. bipolar neutrosophic graphs are provided [18]. The enormous author described the concepts of bipolar neutrosophic numbers in various problems [19], [15], [16], and [19]. The approximate balanced score function [20] will be introduced for neutrosophic graphs with weak edge weights, and then the score functions for neutrosophic graphs by Nancy et al. and Boumi et al. will be compared.

2. Methods

Upper/Lower weight algorithm of neutrosophic undirected graph for bipolar minimum spanning tree problem. The focus of this upper/lower weight algorithm is to help for finding the optimal tree with less computational time. Because while computing programming in MATLAB gives the results for the score matrix as well the optimal edges of the bipolar neutrosophic minimum spanning network here by following the steps below,
Input: consider bipolar minimum spanning network of G.
Output: Final optimal tree for G
Consider the given network G for solving the bipolar neutrosophic minimum spanning tree problem.
Step 1: Construct the given network G into $n \times n$ matrix form.[ie., M(A)].
Step 2: Follow the phase-1 pseudocode help to make the score matrix $S(\tilde{A})$ for M(A).
Step 3: In phase-2 the main algorithm starts.

- Consider the upper/lower score matrix $S(\tilde{A})$.
- For assigning a larger value for the diagonal element in U(A)/L(A).
- Construct all the edge values for the upper/lower triangular matrix.
- Sortrows all the edges.

Step 4: Draw the edges one by one according to sortrows also follow the rules of the neutrosophic minimum spanning tree.
Step 5: Finally, we get the final optimal tree for the given bipolar neutrosophic minimum spanning tree problem.
Step 6: Also find the accuracy and certainty for all the edges of the optimal network.
MATLAB Pseudocode:
Input: Taking $n \times n$ matrix for the given G.
Phase-1:
tic
Enter $T_A^p(X)$, $I_A^p(X)$, $F_A^p(X)$ and $T_A^n(X)$, $I_A^n(X)$, $F_A^n(X)$ all the entries of G edges or consider the upper/lower triangular edges.
Apply for all the p and n memberships in $S(\tilde{A})$.
toc
Phase-2:
tic
Using $S(\tilde{A})$ and creating U(A) matrix.
Taking triu $S(\tilde{A})$ followed by U(A).
for i=1:n
A(i,i)=999;
for j=1:n
if (i>j)
A(i,j)=999;
end
end
end
Choose the matrix upper/lower case named by U(A)/L(A).
Sortrows of U(A).
Collect all the sort rows.
toc
Output: Finally draw the optimal network.

3. Results

A numerical example using network G, which has a 6-node undirected complete and bipolar weighted graph, is taken into consideration to define the model as a Neutrosophic bipolar minimum spanning tree. Finding a least-cost route that connects every node with $|N|$=n=6 and prevents any other nodes from having more degrees than the allowed three is the challenge. Since the problem is defined on the undirected graph, the corresponding cost matrix becomes symmetric.
The cost and weight values typically presumptively presume non-negative quantity. But in bipolar case, Fig.1 shows the truth, indeterminacy, and falsity with positive and negative edges.

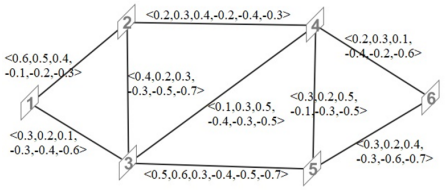


Figure 1: *Network G*

Using the Fig.1 the weight/cost matrices be framed as M(A), $< 0 >$ demonstrates the node's for dis-connectivity.
Calculating the values of the matrix M(A) using the phase-1 the score matrix is represented in matrix form (i.e.,).
$S((\tilde{A}))$=

$$\begin{pmatrix} 0 & 0.517 & 0.617 & 0 & 0 & 0 \\ 0.517 & 0 & 0.633 & 0.5 & 0 & 0 \\ 0.617 & 0.633 & 0 & 0.45 & 0.567 & 0 \\ 0 & 0.5 & 0.45 & 0 & 0.55 & 0.317 \\ 0 & 0 & 0.567 & 0.55 & 0 & 0.317 \\ 0 & 0 & 0 & 0.533 & 0.317 & 0 \end{pmatrix}$$

Thus, an upper triangular cost matrix U(A)/L(A) is considered for evaluating the new algorithm in phase-2.
$U(A)$=

$$\begin{pmatrix} 999 & 0.517 & 0.617 & 0 & 0 & 0 \\ - & 999 & 0.633 & 0.5 & 0 & 0 \\ - & - & 999 & 0.45 & 0.567 & 0 \\ - & - & - & 999 & 0.55 & 0.317 \\ - & - & - & - & 999 & 0.317 \\ - & - & - & - & - & 999 \end{pmatrix}$$

Fig.2 shows the optimal minimum spanning network for the given bipolar minimum spanning network G. The crisp minimum cost of the M(A) is 2.565 and the ordered pair for the bipolar neutrosophic tree is (2,4), (1,3), (1,2), (4,5), and (4,6).

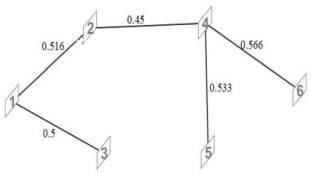


Figure 2: *Optimal solution for given G.*

The accuracy and certainty of every edge value be examined for bipolar NMST for the given network G as shown in Fig.3.

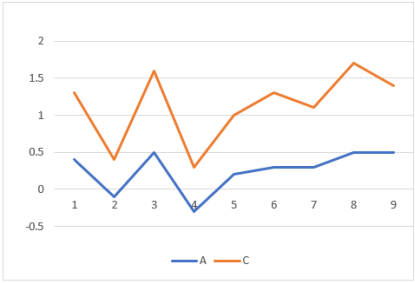


Figure 3: *shows the accuracy and certainty)*

4. Comparison Results

The BNMST computation details are provided in this section for solving the BNMST, and a precise new algorithm has been created. The algorithm is programmed in MATLAB-2019b and

Table 1: *Noted the Upper triangular algorithm with CPU run time*

Author's algorithm	n	Order Pairs of BNMST	Optimal values	Elapse time
Broumie et.al Algorithm	5	{2,1},{1,4},{4,3},{3,5}	1.8	-
Proposed Algorithm	5	{1,4},{3,5},{1,2},{3,4}	1.74	2.137817
Upender et.al Algorithm	6	{1,2},{2,4},{4,3},{4,6},{5,6}	2.327	-
Proposed Algorithm	6	{2,4},{1,3},{1,2},{4,5},{4,6}	2.565	0.026348

tested on a computer running the Microsoft Windows11 version 22H2, operating system with an Intel(R) Core (TM) 8250U processor running at 1.60 GHz or 1.80 GHz and 8 GB of Memory.
To determine the effectiveness of the Upper/Lower triangular edge weight algorithm extensive experimentation was done for investigation purposes searching some problems with varying n related to the BNMST problem by Broumie et.al and Upende et.al are compared with our algorithm and noted the running time of the algorithm in Table.1.

5. Conclusion and Future Work

Consider the bipolar minimum spanning tree network to design the optimal decision tree using the order pairs. For the symmetric version of BNMST, an upper/lower weight algorithm is created and tested in the MATLAB programming language also running time is calculated. The created method is validated using a relevant numerical example. Computational work with some examples has been done to assess the effectiveness of the algorithm.
The overall findings demonstrate the effectiveness of the developed algorithm and calculating the best optimal result for BNMST with help of upper/lower weight algorithm. Also, this model is more adaptable in addressing issues with real-time networks.

References

[1] Zadeh. L (1965). Fuzzy sets. *Inform and control*, 8:338–353.
[2] Smarandache. F. (1998). Neutrosophy, Neutrosophic Probability, Set, and Logic: ProQuest Information and Learning, Ann Arbor, Michigan, USA, 105p.
[3] Wang.H, Smarandache.F,Zhang.Y and Sunderraman.R, (2010). Single Valued Neutrosophic Sets. *Multisspace and Multistructure*, 4:410–413.
[4] Ardhendu Mandal. (2012). A New Efficient Technique to Construct a Minimum Spanning Tree. *International Journal of Advanced Research in Computer Science and Software Engineering*, 2(10).
[5] Ye. J (2014). Similarity measures between interval neutrosophic sets and their applications in multicriteria decision-making. *Journal of Intelligent and Fuzzy Systems*, 26:165–172.
[6] Broumi, S., Smarandache, F., and Dhar, M. (2014). Rough neutrosophic sets. *Infinite Study*, 32:493–502.
[7] Deli, I., Ali, M.,and Smarandache, F. (2015). Bipolar neutrosophic sets and their application based on multi-criteria decision making problems. *In 2015 International conference on advanced mechatronic systems (ICAMechS)* 249–254. Ieee.
[8] Ye, J. (2015). Trapezoidal neutrosophic set and its application to multiple attribute decision-making. *Neural computing and Applications*, 26:1157–1166.
[9] Rajesh, B., Mohideen, S. I., and Senthilkumar, S. (2016) Second Minimum Weight Spanning Tree of A Network with Triangular Intuitionistic Fuzzy Number As Edge Weight. International Journal of Engineering Research and Technology (IJERT).AMASE-2016 Conference Proceedings.
[10] Ali, M., Son, L. H., Deli, I.,and Tien, N. D. (2017). Bipolar neutrosophic soft sets and applications in decision making. *Journal of Intelligent and Fuzzy Systems*, 33(6):4077–4087.

[11] Mullai, M., Broumi, S., and Stephen, A. (2017). Shortest path problem by minimal spanning tree algorithm using bipolar neutrosophic numbers. Infinite Study.
[12] Broumi, S., Bakali, A., Talea, M., Smarandache, F., and Dey, A. (2018). Spanning tree problem with neutrosophic edge weights. *Procedia Computer Science*, 127:190–199.
[13] Broumi, S., Bakali, A., Talea, M., Smarandache, F., and Ullah, K. (2018). Bipolar neutrosophic minimum spanning tree. *Infinite Study*, 127.
[14] Kumar, T., and Purusotham, S. (2018). The degree constrained k-cardinality minimum spanning tree problem: a lexi-search algorithm. *Decision Science Letters*, 7(3)–301-310.
[15] Broumi, S., Singh, P. K., Talea, M., Bakali, A., Smarandache, F., and Venkateswara Rao, V. (2019). Single-valued neutrosophic techniques for analysis of WIFI connection. In Advanced Intelligent Systems for Sustainable Development (AI2SD'2018) Volume 5: Advanced Intelligent Systems for Computing Sciences. *Springer International Publishing*, 405–412.
[16] Smarandache, F., Broumi, S., Singh, P. K., Liu, C. F., Rao, V. V., Yang, H. L., and Elhassouny, A. (2019). Introduction to neutrosophy and neutrosophic environment. *In Neutrosophic Set in Medical Image Analysis* Academic Press. 3-29.
[17] Huang, L., Hu, Y., Li, Y., Kumar, P. K., Koley, D., and Dey, A. (2019). A study of regular and irregular neutrosophic graphs with real life applications. *Mathematics*, 7(6)–551.
[18] G Upender Reedy, et al., (2020) "Minimum Spanning Tree Algorithm w.r.t Bipolar Neutrosophic Graphs", London Journal Press, vol. 20, 8-1.3.
[19] Rao, T. S. N., Kumar, C. S., Rao, Y. S., and Rao, V. V. (2020). Detour Interior and Boundary vertices of BSV Neutrosophic Graphs. *International Journal of Advanced Science and Technology*,29(8). 2382–2394.
[20] Srisarkun, V., and Jittawiriyanukoon, C. (2021). An approximation of balanced score in neutrosophic graphs with weak edge weights. *Int. J. Electr. Comput. Eng*, 11. 5286–5291.

DESIGNING AN ECONOMIC REPETITIVE GROUP SAMPLING PLAN FOR LIFE TESTS BASED ON PERCENTILES OF EXPONENTIATED RAYLEIGH DISTRIBUTION

P. Umamaheswari[1], S. Suganya[2] and K. Pradeepa Veerakumari[3]

•

[1]Assistant Professor of Statistics, Sona College of Arts & Science,
[1]uma_m2485@yahoo.com
[2]Assistant Professor, Department of Statistics, PSG College of Arts & Science,
[2]suganstat@gmail.com
[3]Associate Professor, Department of Statistics, Bharathiar University,
[3]pradeepaveerakumari@buc.edu.in

Abstract

In this study, a unique Repetitive Group Sampling Plan (RGSP) for life tests is presented, emphasizing the use of percentiles obtained from the Exponentiated Rayleigh Distribution. To enhance the accuracy, effectiveness, and general reliability of life testing techniques, the suggested sample plan provides a strong foundation for assessing the longevity and reliability of goods in a variety of industries. By offering a strong framework for evaluating the robustness and lifespan of products in diverse industries, the suggested sample plan seeks to improve the efficiency and reliability of life testing methods. Through the incorporation of statistical techniques grounded in the Exponentiated Rayleigh Distribution, this research advances the fields of reliability assessment and quality control in industrial processes. Table values are calculated by using Python programming. The implementation of these statistical techniques not only enhances the accuracy of life testing but also enables businesses to make more informed decisions regarding product development and quality assurance. As industries increasingly rely on data-driven approaches, this research provides a valuable resource for optimizing performance and minimizing costs

Keywords: Exponentiated Rayleigh Distribution, Percentiles, Life tests, Repetitive Group Sampling Plan.

I. Introduction

Life tests are essential for guaranteeing the durability and dependability of products in the domains of quality control and reliability evaluation. Conventional techniques frequently depend on preset sampling intervals and fixed sample sizes, which may not be sufficient to capture the dynamic character of product life cycles. Our research presents a Repetitive Group Sampling Plan (RGSP) designed specifically for live tests in order to overcome this issue. The main contribution is the application of the Exponentiated Rayleigh Distribution, a robust statistical model with great flexibility and versatility that can be used to assess the dependability of products under different stress scenarios. To estimate a product's life characteristics, life testing entails putting a sample of products under operational stress until they break. Taking into account the inherent variety in product lifetimes, the suggested RGSP offers an efficient and adaptable solution to life testing. Our methodology seeks to provide a more accurate picture of

product durability and reliability by including the Exponentiated Rayleigh Distribution, which has been shown to be successful in estimating failure times. Because of its versatility in simulating a wide range of failure time patterns, the Exponentiated Rayleigh Distribution is a development of the classical Rayleigh Distribution—has become more well-known in the field of reliability engineering. This distribution has been effectively used in earlier research to examine life statistics from a variety of industries, including the electronics, automotive, and aerospace sectors. The distribution is well-suited to representing the complicated failure features seen in real-world circumstances since it can handle both growing and decreasing hazard rates.

Adaptive sampling strategies are necessary to account for dynamic changes in product reliability, despite the fact that traditional life testing methods have depended on set sample sizes and constant sampling intervals. This has been highlighted by recent research. Building on this tendency, the study's suggested Repetitive Group Sampling Plan (RGSP) provides a versatile and effective method for life testing. The RGSP attempts to achieve a compromise between statistical rigor and practical implementation by combining percentiles generated from the Exponentiated Rayleigh Distribution. This makes it a helpful tool for industries looking to improve product reliability and quality assurance. To summarize, the limitations of conventional approaches are addressed by this article's groundbreaking Repetitive Group Sampling Plan for life testing, which makes use of the Exponentiated Rayleigh Distribution. Improving dependability assessment and quality control procedures in a variety of manufacturing industries is greatly promising when these statistical approaches are integrated. The attributes repeating group sampling plan for a normal distribution is suggested by [14]. This repeated sampling method functions similarly to a sequential sampling method. He claims that this sampling plan provides the necessary protection for both the producer and the customer, in addition to providing the minimal sample size. Moreover, sequential sampling is less efficient than repeated sampling, but it is still more efficient than a single sampling strategy. Among other academics, examined the factors in a repeated group acceptance sampling plan and contrasted the outcomes with a single sample plan developed by [4]. On the other hand, no effort has been made to investigate attributes repeating sampling plans for an Exponentiated Rayleigh Distribution based on truncated life tests. Assuming that the lifetime of the products follows the Exponentiated Rayleigh Distribution with a known shape parameter and that the quality level is represented by the ratio of mean life to the stipulated life, we designed an attributes repetitive sampling plan for the truncated life tests in this study. There have been notable developments in the fields of reliability evaluation and quality control, and researchers are always working to improve life testing techniques. This review of literature delves into the corpus of work that has been written about a certain subject: "A Repetitive Group Sampling Plan for Life Tests Based on Percentiles of Exponentiated Rayleigh Distribution." This novel method offers a promising path toward enhanced reliability assessment by fusing the ideas of the Exponentiated Rayleigh Distribution with the Repetitive Group Sampling Plans (RGSP).

The methodology of time-truncated acceptance sample plans has fascinated a number of eminent authors. The truncated life tests in the exponential case developed [5]. Acceptance sampling plan with life test objectives are framed [15]. Using the sampling plans based on the Weibull distribution is proposed [6]. In [7] acceptance sampling plan based on Gamma Distribution is established. For [8, 9] acceptance sampling based on truncated life tests based on Log-Logistic distribution. Acceptance sampling based on truncated life tests for Pareto Distribution are suggested [2].

Acceptance sampling based on truncated life tests for Generalized Rayleigh Distribution introduced [17]. Acceptance sampling based on truncated life tests for Birnbaun-Sauders Distribution [3], Generalized Exponential Distribution [1], and Marshall-Olkin Extended Lomax Distribution [13]. Designing of acceptance sampling plan for life test based on percentiles of exponentiated Rayleigh distributions are developed [10, 11, 12]. Skip-lot sampling plan of type SkSP-T for life test based on percentiles of Exponentiated Rayleigh Distribution suggested [16].

The main goal of the paper is to develop a life-test plan for repetitive group sampling using an exponential Rayleigh distribution and a percentile-based comparison plan. An important distribution for life testing and reliability analysis is the Rayleigh exponential distribution. It shows excellent mathematical consistency and possesses some of the fundamental structural characteristics. The Exponentiated Rayleigh distribution shares many characteristics with the Weibull, gamma, and exponential distributions. The density and distribution functions of ERD have comparable shapes. This is swiftly expanded to the abbreviated plans as a result. The ERD distribution's cumulative function is provided by,

$$F(t;\tau,\theta) = \left[1 - e^{-1/2(t/\tau)^2}\right]^{\theta},\ t>0,\ 1/\tau>0,\ \theta>0 \quad (1)$$

Where the parameters for scale and form are, respectively, τ and θ. The probability density function of a cumulative distribution function is its first derivative. Thus, the ERD's probability density function can be expressed as,

$$f(t;\tau,\theta) = \theta\left[1 - e^{-1/2(t/\tau)^2}\right]^{\theta-1}\left[\frac{t}{\tau^2}e^{-1/2(t/\tau)^2}\right] \quad (2)$$

Based on the ERD percentiles, [10, 11] suggested SSP and DSP for life testing. Then, using the ERD percentiles as a guide, [12] created skip-lot sampling designs for life testing.

II. Repetitive group sampling plans based on truncated test

The following is a description of the suggested features repeating group sampling plan based on abbreviated life tests:

Step 1: For a fixed time t_0, take a random sample of size n from a lot and put it through a life test.

Step 2: If D, the number of failures, is less than or equal to c_1, the acceptance number, then accept the lot. When D surpasses c_2, where $c_2 \geq c_1$, halt the test and reject the lot.

Step 3: If $c_1 < D < c_2$, then repeat step 1.

Three parameters, n, c_1, and c_2, define the repeating group sampling scheme described above. It is significant to remember that the ordinary single sampling plan is reduced to the attributes repetitive group sampling plan, which is a generalization of it when $c_1 = c_2$. The Operating Characteristic (OC) function is used to calculate the likelihood of lot acceptance.

III. Operating Procedure for Repetitive Group Sampling Plan

Plan Parameters Determination:

i. N: Specify the sample size for each group.
ii. c_1: Determine the acceptance number.
iii. c_2: Establish the rejection number.
iv. Maximum Number of Groups (Optional): Decide if a limit on samples is needed.

Initial Sampling:

i. Randomly select N items from the lot.
ii. Inspect each item for defects.

First Sample Evaluation:

i. Count the defective items (d) in the first sample.

Initial Decision:

i. Accept: If $d \leq c_1$, accept the lot.
ii. Reject: If $d \geq c_2$, reject the lot.
iii. Proceed to Step 5: If $c_1 < d < c_2$.

Repeat Sampling (If Necessary):

i. Take another random sample of N items.
ii. Inspect and count defectives in the second sample.
iii. Combine results from both samples.

iv. Accept: If cumulative defectives ≤ c_1.
v. Reject: If cumulative defectives ≥ c_2.
vi. Repeat: If c_1 < cumulative defectives < c_2, repeat steps 5 and 6.

Maximum Number of Groups (Optional):

i. If a maximum is set and no decision is reached within that limit, typically reject the lot.

IV. Operating procedure of Repetitive Group Sampling Plan based on percentiles of ERD

In [14] devised the repeated group sampling plan, which involves drawing a sample and counting the number of defective samples. In order to sentence a lot, a fixed criterion determines whether the lot should be accepted, rejected, or wholly ignored, in which case one must start over with a fresh sample. This process is repeated until the fixed criterion instructs us to accept or reject the lot. The following is the operating procedure for a repeating group sample plan:

Step 1. Select a random sample of size 'n' from a lot of size 'N'.

Step 2. Inspect all the articles included in the sample. Let 'd' be the number of defectives in the sample.

Step 3. If d< c_1, accept the lot If d > c_2, reject the lot If c_1 < d ≤ c_2, repeat the steps 1, 2 and 3.

V. Operating characteristic function for Repetitive Group Sampling Plan

The most often used method for assessing a sample plan's effectiveness and determining the likelihood of adoption is the OC function. It offers the likelihood of the lot being approved. The following is the OC function of RGSP for life tests based on the ERD percentiles:

$$L(p) = \sum_{i=0}^{c} \binom{n}{i} p^i (1-p)^{n-i} \tag{3}$$

[1] Percentile of the lifetime t_q^0 and p depends only on$\delta_0 = t/t_q^0$.

The following is the operation characteristic function for the repetitive group sampling plan

$$L(P) = \frac{P_a}{P_a + P_r} \tag{4}$$

Here P_a *and* P_r is the binomial model and equation (4) becomes

$$P_a(p) = \frac{\sum_{i=0}^{C_1} \binom{n}{i} P^i (1-P)^{n-i}}{1 - \sum_{i=0}^{C_1} \binom{n}{i} P^i (1-P)^{n-i} + \sum_{i=0}^{C_2} \binom{n}{i} P^i (1-P)^{n-i}} \tag{5}$$

Then, the Average Sample number is

$$ASN\ (p) = ASN\ (R)F \tag{6}$$

The reference plan's average sample number is denoted by R-, while the chance of the reference plan being accepted is indicated by P.

I. Illustration

Suppose that the life time of the electric goods follows ERD. Repetitive group sampling with SSP as reference plan using binomial distribution based on 25th percentile is applied for testing. The parameters for the life testing is as follows: θ=2, t=40 hrs,t0.25=40 hrs, c=2, α=0.05 and β=0.05 then η = 1.56712 from the equation and the ratio is found to be $t/t_{0.25}$=2 by applying the minimum sample size according to the requirements is n=10 and the corresponding OC values L(p) for the Single Sampling plan for the life tests based on percentiles of ERD (n, c, $t/t_{0.25}$=2) = (10,2,1.10) with P*= 0.95. L(p) is the P value for RGS with SSP for life tests based on the percentiles of ERD as reference plan. For n=10, c_1=2 and c_2=1; the probability of acceptance L(p) values of RGS with

SSP for life tests based on percentiles of ERD are found from eqn. 5. From the illustrations, it is indicated that the actual 25th percentile is almost equal to the required 25th percentile ($t_{0.25}$/ t / $t^0_{0.25}$=1.00) the producer's risk is approximately 0.9800 (1-0.0200). Also, the producer's risk is nearly equal to 0.05 or less and the actual producer risk is large or nearly equal to 1.95 times of the required percentile.

Table 1: *Gives the OC values for sampling plan (n, c_1, c_2, $t/t_{0.25}$) for a given P* under ERD when θ=2*

P*	$t/t_{0.25}$	$t/t_{0.25}$						
		0.7	0.9	1	1.5	2	2.5	3
0.75	0.7	0.3910	0.8421	0.9701	0.9911	0.9980	0.9993	0.9998
0.75	0.9	0.4548	0.8741	0.9660	0.9899	0.9984	0.9994	0.9998
0.75	1	0.4778	0.8989	0.9710	0.9900	0.9987	0.9996	0.9999
0.9	0.7	0.1956	0.7100	0.8989	0.9801	0.9942	0.9990	0.9996
0.9	0.9	0.2421	0.7539	0.9223	0.9847	0.9950	0.9991	0.9998
0.9	1	0.3879	0.8123	0.9527	0.9897	0.9950	0.9993	0.9999
0.95	0.7	0.1333	0.6747	0.9198	0.9697	0.9927	0.9975	0.9990
0.95	0.9	0.1515	0.6981	0.9387	0.9698	0.9943	0.9982	0.9996
0.95	1	0.0925	0.5841	0.8996	0.9596	0.9876	0.9970	0.9985
0.99	0.7	0.0131	0.4555	0.8321	0.9499	0.9732	0.9941	0.9980
0.99	0.9	0.0220	0.4606	0.8387	0.9512	0.9817	0.9955	0.9982
0.99	1	0.0318	0.5001	0.8495	0.9571	0.9860	0.9961	0.9989

VI. Conclusion

This research concludes with a new approach to life testing: a Repetitive Group Sampling Plan (RGSP) based on the percentiles of the Exponentiated Rayleigh Distribution. A strong basis for assessing a product's reliability and durability over time is provided by the sampling strategy's integration of this statistical model. By using the flexibility and adaptability of the Exponentiated Rayleigh Distribution, our proposed RGSP provides a flexible and efficient method to evaluate product lifetimes under different stress conditions. By providing enterprises with a useful tool for guaranteeing the durability and dependability of their products, this research advances quality control and reliability evaluation. The methodology that has been offered shows potential in mitigating the drawbacks of conventional life testing techniques, offering a more precise depiction of product dependability, and supporting the continuous advancement of quality control procedures across various manufacturing industries.

References

[1] Aslam M, and M.Q. Shahbaz (2007). Economic Reliability plans using the Generalized Exponential Distribution, *J. Stat*., Vol 14, 53-60.

[2] Baklizi A (2003). Acceptance sampling based on truncated life tests in the pareto distribution of the second kind. *Advances and Applications in Statistics,* Vol. 3, Issue 1, 33-48.

[3] Balakrishnan N. Leiva V. And Lopez J. (2007). Acceptance sampling plans from truncated tests based on the generalized Birnbaun-Sauders distribution. *Communications in statistics – Simulation and Computation*, Vol. 36, pp.643-656.

[4] Balamurali and Chi-hyuck Jun (2007). Repetitive Group Sampling Procedure for variables inspection, *Journal of Applies Statistics*, pp. 327-338

[5] Epstein B (1954). Truncated life tests in the exponential case, *Annals of Mathematical Statistics*, 25: pp. 555-564.

[6] Goode. H P and Kao. J. H. K, (1961). Sampling plans based on the Weibull distribution, *Proceedings in the 7th National Symposium on Reliability and Quality Control*, pp.24-40, Philadelphia, USA.

[7] Gupta S. S., and Groll. P. A. (1961). Gamma distribution in acceptance sampling based on life tests, *Journal of the American Statistical Association*, 56, pp. 942-970.

[8] Kantam R.R.L. and Rosaiah K. and Rao. G.S. (2001). Acceptance sampling based on life tests: Log-Logistic model. *Journal of Applied Statistics*,28: 121-128.

[9] Kantam R.R.L. G. Srinivasa Rao and G. Sriram (2006). An economic reliability test plan: Log-Logistic distribution, *J. Appl. Statist*. 33(3): 291-296.

[10] Pradeepa Veerakumari K and Ponneeswari P (2016). Designing of Acceptance Sampling Plan for life tests based on Percentiles of Exponentiated Rayleigh Distribution, *International Journal of Current Engineering and Technology*, Vol. 6 (4):1148-1153.

[11] Pradeepa Veerakumari K and Ponneeswari P (2017), Designing of Acceptance Double Sampling Plan for Life Test Based on Percentiles of Exponentiated Rayleigh Distribution. *International Journal of Statistics and Systems*, Vol 12 (3): 475-484.

[12] Pradeepa Veerakumari K, Uma Maheswari P, Aruna H M, Suganya S (2018). Designing of Skip-Lot Sampling Plan of Type (SkSP-3) For Life Tests Based on Percentiles of Exponentiated Rayleigh Distribution. *International Journal of Research in Advent Technology*, Vol.6, No.8, August 2018.

[13] Rao G S, Ghitany M. E and Kantam R. R., (2009). Acceptance Sampling Plans for Marshall-Olkin extended Lomax distribution. *International Journal of Applied Mathematics*, 2(1),139 – 148.

[14] Sherman. E (1965). Design and Evaluation of a Repetitive Group Sampling Plan, *Technometrics*, Vol.7 (1), pp. 11-21.

[15] Sobel, M and Tischendrof, J. A. (1959). Acceptance sampling with sew life test objective. *Proceedings of Fifth National Symposium on Reliability and Quality Control,* Philadelphia, 108-118.

[16] Suganya S and Pradeepa Veerakumari (2020). Skip-lot sampling plan of type SkSP-T for life test based on Percentiles of Exponentiated Rayleigh Distribution, *Life Cycle Reliability and Safety Engineering*, 9, 247-251.

[17] Tsai T.R. and Wu S.J. (2006). Acceptance Sampling Based on truncated life tests for generalized Rayleigh Distribution. *Journal of Applied Statistics.* Vol. 33, pp.595 – 600.

RT&A, No 2 (84)
Volume 20, June 2025

COMMON FIXED POINT THEOREMS IN COMPLEX VALUED B-METRIC SPACES AND USING WEAK COMPATIBLE MAPPINGS

Shivani Chourasiya[1], Harshit Khare[2], and Kavita Shrivastava[3]

[1,2,3]Department of Mathematics and Statistics, Dr. Harisingh Gour Vishwavidyalaya, Sagar, M.P., 470003

[1]shivanichourasiya10@gmail.com, [2]khareh412@gmail.com, [3]kavita.rohit@rediffmail.com

Abstract

Complex-valued metric spaces, where the distance between points is represented as a complex number, serve as a generalization of traditional metric spaces and present a variety of intriguing applications, particularly in advanced mathematics and theoretical physics, such as in quantum mechanics, quantum computing, functional analysis, operator theory, complex dynamics, fractal geometry, complex networks, and signal processing, as well as theoretical physics, relativity, geometry, topology, mathematical finance, and economics. Despite the increasing interest in complex-valued metric spaces, the current literature on fixed point theorems within these contexts is constrained in several ways. The available literature has notable shortcomings; consequently, this article investigates the results and implications of contractive mappings in complex-valued b-metric spaces by utilizing weak compatible mappings, and it also employs control functions while solving a system of Urysohn integral equations based on our primary result. These findings contribute to this area of study.

Keywords: fixed points, b-metric space, complex valued b-metric space, weakly compatible mappings.

1. Introduction

The classical Banach contraction principle [1] serves as a potent tool in nonlinear analysis and initiated a significant development in fixed point theory, owing to its broad applicability in diverse areas of mathematics, including numerical analysis and both differential and integral equations. A multitude of researchers has investigated various fixed point theorems, proposing new contractive conditions in metric spaces or their broader counterparts (also see [2], [3], [4], [5], [6] [7], [8], etc.).

In 2011, complex valued metric spaces were introduced for the first time by [9], followed by [10] in 2012, who improved the contraction theorem by applying it to control functions and establishing common fixed point theorems. Fixed point theorems in complex valued metric spaces were proven and applications were provided by [11]. Furthermore, in 2013, proposed a method for Urysohn integral equations employing common fixed points in metric spaces that include complex values by [12]. A common fixed point theorem for weakly compatible mappings in complex-valued metric spaces was established, avoiding continuity concepts as presented by [13]. Research was conducted by [14] on theorems concerning common fixed points, leveraging property (EA) in metric spaces that incorporate complex values. Additionally, illustrated various results related to common fixed points for pairs of mappings that satisfy broader contraction conditions expressed through rational functions, involving point-dependent control functions as coefficients within complex-valued metric space by [15]. The introduction of complex valued

b-metric spaces occurred in 2013 by [16]. Common fixed point theorems were demonstrated in generalized complex valued b-metric spaces by [17]. Weak compatible mappings were initially introduced by [18]. This paper introduces new contractive mappings within the framework of completeness and the uniqueness of fixed point theorems in complex valued b-metric spaces, building upon previous research.

2. Preliminaries

Assume that $\omega_1, \omega_2 \in \mathbb{C}$, where $\mathbb{C}$ is the set of complex numbers that define a partial order $\leq$ on $\mathbb{C}$ such that $\omega_1 \leq \omega_2$ if and only if $Re(\omega_1) \leq Re(\omega_2)$ and $Im(\omega_1) \leq Im(\omega_2)$ that is $\omega_1 \leq \omega_2$. If any of the following is true [9],

1. $Re(\omega_1) = Re(\omega_2)$ and $Im(\omega_1) = Im(\omega_2)$;
2. $Re(\omega_1) < Re(\omega_2)$ and $Im(\omega_1) = Im(\omega_2)$;
3. $Re(\omega_1) = Re(\omega_2)$ and $Im(\omega_1) < Im(\omega_2)$;
4. $Re(\omega_1) < Re(\omega_2)$ and $Im(\omega_1) < Im(\omega_2)$;

we will write $\omega_1 \lneq \omega_2$ if $\omega_1 \neq \omega_2$ and one of (2),(3), and (4) satisfiy. Take note of that

$$0 \leq \omega_1 \lneq \omega_2 \Rightarrow |\omega_1| < |\omega_2|$$
$$\omega_1 \leq \omega_2, \omega_2 < \omega_3 \Rightarrow \omega_1 < \omega_3.$$

Let Δ be a set that is not empty. A mapping $\Omega : \Delta \times \Delta \to \mathbb{C}$ is considered to be congruent with the following conditions [9],

1. $0 \leq \Omega(\mu, \nu)$ for all $\mu, \nu \in \Delta$ and $\Omega(\mu, \nu) = 0$ if and only if $\mu = \nu$;
2. $\Omega(\mu, \nu) = \Omega(\nu, \mu)$, for all $\mu, \nu \in \Delta$;
3. $\Omega(\mu, \nu) \leq \Omega(\mu, \omega) + \Omega(\omega, \nu)$ for all $\mu, \nu, \omega \in \Delta$.

In this case, the metric Ω is a complex valued metric on the set Δ, and the space (Δ, Ω) is a complex-valued metric space.
Let Δ be a set that is not empty and $s \geq 1$ a given real number. A mapping $\Omega : \Delta \times \Delta \to \mathbb{C}$ is considered to be congruent with the following conditions [16],

1. $0 \leq \Omega(\mu, \nu)$ for all $\mu, \nu \in \Delta$ and $\Omega(\mu, \nu) = 0$ if and only if $\mu = \nu$;
2. $\Omega(\mu, \nu) = \Omega(\nu, \mu)$, for all $\mu, \nu \in \Delta$;
3. $\Omega(\mu, \nu) \leq s\left[\Omega(\mu, \omega) + \Omega(\omega, \nu)\right]$ for all $\mu, \nu, \omega \in \Delta$.

Then Ω is called a complex valued b-metric on Δ and (Δ, Ω) is called a complex valued b-metric space.
Let $\Delta = [0, 1]$. Define the mapping $\Omega : \Delta \times \Delta \to \mathbb{C}$ by [17],

$$\Omega(\mu, \nu) = |\mu - \nu|^2 + i\,|\mu - \nu|^2 \, \forall \mu, \nu \in \Delta$$

Then (Δ, Ω) is a complex valued b-metric space with $s = 2$.
Suppose that (Δ, Ω) is a complex-valued b-metric space if [16],

1. A point $\mu \in \Delta$ is referred to as an interior point of a set $A \subseteq \Delta$ if there exists a $r \in \mathbb{C}$ such that $B(\mu, r) = \{\nu \in \Delta : \Omega(\mu, \nu) < r\} \subseteq A$.
2. A point $\mu \in \Delta$ is called a limit point of a set $A \subseteq \Delta$ if there exists a $r \in \mathbb{C}$ such that $B(\mu, r) \cap (\Delta \setminus A) \neq \varnothing$ for all $0 < r \in \mathbb{C}$.

3. A subset $B \subseteq \Delta$ is referred to as open if each limit point of B is an interior point of B.

4. A subset $B \subseteq \Delta$ is considered closed if each limit point of B is included in B.

5. A subbase for a topology Δ is the family$F = \{B(\mu, r) : \mu \in \Delta\}$ and $0 < r$. This intricate topology is represented by the symbol τ_c. In fact, the topology τ_c is Hausdorff.

Assume that (Δ, Ω) is a space with complex values and b-metric, and let $\{\mu_n\}$ be a sequence in Δ, with $\mu \in \Delta$ [16].

1. μ_n is said to converge to μ and μ is a limit point of μ_n if for every$c \in \mathbb{C}$ with $0 < c$, there exists $n_0 \in N$ such that $\Omega(\mu_n, \mu) < c$ for all $n > n_0$. We denote this as $\mu_n \to \mu$ as $n \longrightarrow \infty$ or $\lim_{n\to\infty} \mu_n = \mu$.

2. $\{\mu_n\}$ is a Cauchy sequence if for every $c \in \mathbb{C}$ with $0 < c$, there exists $n_0 \in N$ such that for all $n > n_0$, $\Omega(\mu_n, \mu_{n+m}) < c$ where $m \in \mathbb{N}$.

3. If all Cauchy sequences are convergent in (Δ, Ω), then (Δ, Ω) is referred to as a complete complex valued b-metric space.

Lemma 1. (Δ, Ω) is a space that is a complex-valued b-metric space, and let $\{\mu_n\}$ be a sequence that is contained within Δ. If and only if the function $|\Omega(\mu_n, \mu)| \to 0$ converges to zero as $n \to \infty$ [16].

Lemma 2. (Δ, Ω) is a space that is a complex valued b-metric space, and let $\{\mu_n\}$ be a sequence that is contained within Δ. The series $\{\mu_n\}$ is considered to be a Cauchy sequence if and only if $|\Omega(\mu_n, \mu_{n+m})| \to 0$ as $n \to \infty$, where $m \in N$ [16].

Given that L and M are self-mappings on a set Δ, if $w = L\mu = M\mu$ for some in Δ, then Δ is referred to as a coincidence point of L and M, and w is referred to as a point of coincidence of L and M [17].
If the self-mappings $L, M : \Delta \to \Delta$ commute at their respective coincidence point, then we refer to this pair of self-mappings as weakly compatible [18].

3. Main result

Theorem 1. Let (Δ, Ω) be a complete complex valued b-metric space and $L, M : \Delta \to \Delta$. If there exists mappings $\lambda, \delta : \Delta \to [0, 1)$ such that for all $\mu, \nu \in \Delta$,

1. $\lambda(L\mu) \leq \lambda(\mu)$, $\delta(L\mu) \leq \delta(\mu)$

2. $\lambda(M\mu) \leq \lambda(\mu)$ and $\delta(M\mu) \leq \delta(\mu)$

3. $(\lambda + \delta)\mu < 1$

Then the following holds

$$\Omega(L\mu, M\nu) \leq \lambda(\mu) \max \left\{ (\Omega(\mu, \nu), \frac{\Omega(\mu, L\mu)\Omega(\nu, M\nu)}{\Omega(\mu, M\nu) + \Omega(\nu, L\mu) + \Omega(\mu, \nu)} \right\} + \delta(\mu) \{\Omega(\mu, L\mu) + \Omega(\nu, M\nu)\} \quad (1)$$

for all $\mu, \nu \in \Delta$ such that $\mu \neq \nu$, $\Omega(\mu, M\nu) + \Omega(\nu, L\mu) + \Omega(\mu, \nu) \neq 0$ and $s \geq 1$ then L and M have a unique common fixed point.

Proof. Let $\mu_0 \in \Delta$ be arbitrary. Since $L(\mu) \subseteq \Delta$ and $M(\mu) \subseteq \Delta$, we can construct the sequence $\{\mu_n\}$ in Δ such that $\mu_n = L\mu_{n-1}$ and $\mu_{n+1} = M\mu_n$ for all $n \geq 0$. Then from (1)

$$\begin{aligned}
\Omega(\mu_n, \mu_{n+1}) &= \Omega(L\mu_{n-1}, M\mu_n) \\
&\leq \lambda(\mu_n) \max\left\{\Omega(\mu_{n-1}, \mu_n), \frac{\Omega(\mu_{n-1}, L\mu_{n-1})\Omega(\mu_n, M\mu_n)}{\Omega(\mu_{n-1}, M\mu_n) + \Omega(\mu_n, L\mu_{n-1}) + \Omega(\mu_{n-1}, \mu_n)}\right\} \\
&+ \delta(\mu_n)\left\{\Omega(\mu_{n-1}, M\mu_n) + \Omega(\mu_n, L\mu_{n-1})\right\} \\
&\leq \lambda(\mu_n) \max\left\{\Omega(\mu_{n-1}, \mu_n), \frac{\Omega(\mu_{n-1}, \mu_n)\Omega(\mu_n, \mu_{n+1})}{\Omega(\mu_{n-1}, \mu_{n+1}) + \Omega(\mu_n, \mu_n) + \Omega(\mu_{n-1}, \mu_n)}\right\} \\
&+ \delta(\mu_n)\left\{\Omega(\mu_{n-1}, \mu_{n+1}) + \Omega(\mu_n, \mu_n)\right\} \\
&\leq \lambda(\mu_n) \max\left\{\Omega(\mu_{n-1}, \mu_n), \frac{\Omega(\mu_{n-1}, \mu_n)\Omega(\mu_n, \mu_{n+1})}{\Omega(\mu_{n-1}, \mu_{n+1}) + d(\mu_{n-1}, \mu_n)}\right\} \\
&+ \delta(\mu_n)\Omega(\mu_{n-1}, \mu_{n+1}) \\
&\leq \lambda(\mu_n)\Omega(\mu_{n-1}, \mu_n) + \delta(\mu_n)\Omega(\mu_{n-1}, \mu_{n+1}) \\
&\leq \lambda(\mu_n)\Omega(\mu_{n-1}, \mu_n) + s\delta(\mu_n)\left\{\Omega(\mu_{n-1}, \mu_n) + \Omega(\mu_n, \mu_{n+1})\right\} \\
&\Rightarrow \Omega(\mu_n, \mu_{n+1}) \leq \left\{\frac{\lambda(\mu_n) + \delta(\mu_n)s}{1 - \delta(\mu_n)s}\right\}\Omega(\mu_{n-1}, \mu_n) \\
&\Rightarrow |\Omega(\mu_n, \mu_{n+1})| \leq k\,|\Omega(\mu_{n-1}, \mu_n)|
\end{aligned}$$

Where $k = \left\{\frac{\lambda(\mu_n)+\delta(\mu_n)s}{1-\delta(\mu_n)s}\right\} \leqslant 1$

$$\Rightarrow |\Omega(\mu_n, \mu_{n+1})| \leq k^2\,|\Omega(\mu_{n-2}, \mu_{n-1})|$$

Continuing the above process, we get

$$|\Omega(\mu_n, \mu_{n+1})| \leq k^n\,|\Omega(\mu_0, \mu_1)|$$

on taking the limit as $n \longrightarrow \infty$ we obtain that $|\Omega(\mu_n, \mu_{n+1})| \longrightarrow 0$. Now, we show that $\{\mu_n\}$ is a Cauchy sequence in Δ, let $m, n \in N$, $m > n$ then

$$\begin{aligned}
\Omega(\mu_n, \mu_m) &\leq s\left\{\Omega(\mu_n, \mu_{n+1}) + \Omega(\mu_{n+1}, \mu_m)\right\} \\
&\leq s\left\{\Omega(\mu_n, \mu_{n+1})\right\} + s^2\left\{\Omega(\mu_{n+1}, \mu_{n+2}) + \Omega(\mu_{n+2}, \mu_m)\right\} \\
&\leq s\left\{\Omega(\mu_n, \mu_{n+1})\right\} + s^2\left\{\Omega(\mu_{n+1}, \mu_{n+2})\right\} \\
&\quad + s^3\left\{\Omega(\mu_{n+2}, \mu_{n+3}) + \Omega(\mu_{n+3}, \mu_m)\right\} \\
&\leq sk^n\left\{1 + sk + (sk)^2 + (sk)^3 + \ldots\right\}\Omega(\mu_0, \mu_1) \\
&\leq sk^n\Omega(\mu_0, \mu_1)\left\{1 + sk + (sk)^2 + (sk)^3 + \ldots\right\} \\
\Rightarrow \Omega(\mu_n, \mu_m) &\leq \left(\frac{sk^n}{1 - sk}\right)\Omega(\mu_0, \mu_1) \\
\Rightarrow |\Omega(\mu_n, \mu_m)| &\leq \left(\frac{sk^n}{1 - sk}\right)|\Omega(\mu_0, \mu_1)|
\end{aligned}$$

On taking the limit, we have $\lim_{n\to\infty}\Omega(\mu_n, \mu_m) = 0$ as $n, m \longrightarrow \infty$, since $\left(\frac{sk^n}{1-sk}\right) < 1$ hence sequence $\{\mu_n\}$ is Cauchy sequence in Δ. Since Δ is complete, then $\exists$ u in Δ such that

$\lim_{n\to\infty}\mu_n = u$. Now we show that u is the fixed point of T.

$$\begin{aligned}\Omega(u,Lu) &\leq s\left\{\Omega(u,\mu_{n+1})+\Omega(\mu_{n+1},Lu)\right\}\\ &\leq s\left\{\Omega(u,\mu_{n+1})+\Omega(M_n,Lu)\right\}\\ &\leq s\Omega(u,\mu_{n+1})+s\lambda(\mu_n)\max\left\{\Omega(\mu_n,u),\frac{\Omega(\mu_n,M\mu_n)\Omega(u,Lu)}{\Omega(\mu_n,Lu)+\Omega(u,M\mu_n)+\Omega(\mu_n,u)}\right\}\\ &+s\delta(\mu_n)\left\{\Omega(\mu_n,Lu)+\Omega(u,M\mu_n)\right\}\\ &\leq s\Omega(u,\mu_{n+1})+s\lambda(\mu_n)\max\left\{\Omega(\mu_n,u),\frac{\Omega(\mu_n,\mu_{n+1})\Omega(u,Lu)}{\Omega(\mu_n,Lu)+\Omega(u,\mu_{n+1})+\Omega(\mu_n,u)}\right\}\\ &+s^2\delta(\mu_n)\left\{\Omega(\mu_n,u)+\Omega(u,Lu)\right\}+s\delta(\mu_n)\Omega(u,\mu_{n+1})\end{aligned}$$

$$\Rightarrow\left\{1-s^2\delta(\mu_n)\right\}\Omega(u,Lu)\leq s\left\{1+\delta(\mu_n)\right\}\Omega(u,\mu_{n+1})+s\lambda(\mu_n)A+s^2\delta(\mu_n)\Omega(\mu_n,u)$$

Where,

$$A=\max\left\{\Omega(\mu_n,u),\frac{\Omega(\mu_n,\mu_{n+1})\Omega(u,Lu)}{\Omega(\mu_n,Lu)+\Omega(u,\mu_{n+1})+\Omega(\mu_n,u)}\right\}$$

$$\Rightarrow A\leqslant\max\left\{\Omega(\mu_n,u),\Omega(\mu_n,\mu_{n+1})\Omega(u,Lu)\right\}$$

Case-(i)

If $A=\Omega(\mu_n,u)$ then we get

$$\left\{1-s^2\delta(\mu_n)\right\}\Omega(u,Lu)\leq s\left\{1+\delta(\mu_n)\right\}\Omega(u,\mu_{n+1})+s\lambda(\mu_n)\Omega(\mu_n,u)+s^2\delta(\mu_n)\Omega(\mu_n,u)$$

$$\left\{1-s^2\delta(\mu_n)\right\}\Omega(u,Lu)\leq s\left\{1+\delta(\mu_n)\right\}\Omega(u,\mu_{n+1})+\left\{s\lambda(\mu_n)+s^2\delta(\mu_n)\right\}\Omega(\mu_n,u)$$

$$\Rightarrow\Omega(u,Lu)\leq\frac{s\left\{1+\delta(\mu_n)\right\}}{1-s^2\delta(\mu_n)}\Omega(u,\mu_{n+1})+\frac{\left\{s\lambda(\mu_n)+s^2\delta(\mu_n)\right\}}{1-s^2\delta(\mu_n)}\Omega(\mu_n,u)$$

Taking the limit as $n\longrightarrow\infty$, we get
$\lim_{n\to\infty}\Omega(u,Lu)=0,\left[\frac{s\{1+\delta(\mu_n)\}}{1-s^2\delta(\mu_n)},\frac{\left\{s\lambda(\mu_n)+s^2\delta(\mu_n)\right\}}{1-s^2\delta(\mu_n)}<1\right]$
Hence u is the fixed point of L.

Case-(ii)

If $A=\Omega(\mu_n,\mu_{n+1})\Omega(u,Lu)$, then we obtain that
$A\leqslant k^n\Omega(\mu_0,\mu_1)\Omega(u,Lu)$, since $k<1$.
On taking the limit as $n\longrightarrow\infty$, we have $A\longrightarrow 0$. Then as

$$\left\{1-s^2\delta(\mu_n)\right\}\Omega(u,Lu)\leq s\left\{1+\delta(\mu_n)\right\}\Omega(u,\mu_{n+1})+s^2\delta(\mu_n)\Omega(\mu_n,u)$$

$$\Rightarrow\Omega(u,Lu)\leq\frac{s\left\{1+\delta(\mu_n)\right\}}{1-s^2\delta(\mu_n)}\Omega(u,\mu_{n+1})+\frac{s^2\delta(\mu_n)}{1-s^2\delta(\mu_n)}\Omega(\mu_n,u)$$

On taking the limit as $n\longrightarrow\infty$, we get that u is a Fixed point of L. Similarly, one can show that u=Mu. Therefore, u is a common fixed point of L and M.
The uniqueness of the fixed point.
Let there exists another common fixed point of L and M is u_1, that is, $u_1=Lu_1=Mu_1$.

$$\begin{aligned}\Omega(u,u_1) &= \Omega(Lu,Mu_1)\\ &\leq \lambda(u)\max\left\{\Omega(u,u_1),\frac{\Omega(u,Lu)\Omega(u_1,Mu_1)}{\Omega(u,Mu_1)+\Omega(u_1,Lu)+\Omega(u,u_1)}\right\}\\ &\quad+\delta(u)\left\{\Omega(u,Lu)+\Omega(u_1,Mu_1)\right\}\\ &\leq \lambda(u)\Omega(u,u_1))\end{aligned}$$

Since $\lambda(u) \in [0,1)$, we have $\Omega(u,u_1) = 0$. Therefore we have $u = u_1$ and thus u is a unique common fixed point of L and M. ■

Corollary 1. Let (Δ, Ω) be a complete complex valued b-metric space and $L, M : \Delta \to \Delta$. If there exist two nonnegative real no's λ and δ such that for all $\mu, \nu \in \Delta$. If L and M satisfy the following

$$\Omega(L\mu, M\nu) \leq \lambda \max\left\{(\Omega(\mu,\nu), \frac{\Omega(\mu,L\mu)\Omega(\nu,M\nu)}{\Omega(\mu,M\nu)+\Omega(\nu,L\mu)+\Omega(\mu,\nu)}\right\} + \delta\{\Omega(\mu,L\mu)+\Omega(\nu,M\nu)\}$$

for all $\mu, \nu \in \Delta$, and $\lambda + \delta < 1$. Then L and M have a common fixed point.

Theorem 2. Let L, M, S and T be four self mappings of complete valued b-metric spaces if there exists mapping $\lambda, \delta : \Delta \longrightarrow [0,1)$ such that for all $\mu, \nu \in \Delta$ satisfying

1. $L(\Delta) \subseteq T(\Delta)$ and $M(\Delta) \subseteq S(\Delta)$
2. $$\Omega(S\mu, T\nu) \leq \lambda(\mu) \max\left\{\Omega(S\mu,T\nu), \frac{\Omega(S\mu,L\mu)\Omega(T\nu,M\nu)}{\Omega(S\mu,M\nu)+\Omega(T\nu,L\mu)+\Omega(S\mu,M\nu)}\right\} + \delta(\mu)\{\Omega(S\mu,L\mu)+\Omega(T\nu,M\nu)\}$$
3. Pairs (L,S) and (M,T) are weakly compatible and $M(\Delta)$ is closed subspace of Δ. Where $\Omega(S\mu,M\nu)+\Omega(T\nu,L\mu)+\Omega(S\mu,T\nu) \neq 0$ and $s\{\lambda(\mu)+\delta(\mu)\} < 1$

Then L, M, S, and T have a unique common Fixed point.

Proof. Consider a sequences $\{\nu_n\}$ and $\{\mu_n\}$ in Δ such that
$\nu_n = L\mu_n = T\mu_{n+1}$ and $\nu_{n+1} = M\mu_{n+1} = S\mu_{n+2}$. To prove $\{\nu_n\}$ is a Cauchy sequence of Δ.

$$\begin{aligned}
\Omega(\nu_n,\nu_{n+1}) &= \Omega(L\mu_n, M\mu_{n+1}) \\
&\leq \lambda(\nu_n)\max\left\{\Omega(S\mu_n,T\mu_{n+1}), \frac{\Omega(S\mu_n,L\mu_n)\Omega(T\mu_{n+1},M\mu_{n+1})}{\Omega(S\mu_n,M\mu_{n+1})+\Omega(T\mu_{n+1},L\mu_n)+\Omega(S\mu_n,T\mu_{n+1})}\right\} \\
&+ \delta(\nu_n)\{\Omega(S\mu_n,L\mu_n)+\Omega(T\mu_{n+1},M\mu_{n+1})\} \\
&\leq \lambda(\nu_n)\max\left\{\Omega(\nu_{n-1},\nu_n), \frac{\Omega(\nu_{n-1},\nu_n)\Omega(\nu_n,\nu_{n+1})}{\Omega(\nu_{n-1},\nu_{n+1})+\Omega(\nu_n,\nu_n)+\Omega(\nu_{n-1},\nu_n)}\right\} \\
&+ \delta(\nu_n)\{\Omega(\nu_{n-1},\nu_n)+\Omega(\nu_n,\nu_{n+1})\} \\
&\leq \lambda(\nu_n)\max\left\{\Omega(\nu_{n-1},\nu_n), \frac{\Omega(\nu_{n-1},\nu_n)\Omega(\nu_n,\nu_{n+1})}{\Omega(\nu_{n-1},\nu_{n+1})+\Omega(\nu_{n-1},\nu_n)}\right\} \\
&+ \delta(\nu_n)\{\Omega(\nu_{n-1},\nu_{n+1})\} \\
&\leq \lambda(\nu_n)\Omega(\nu_{n-1},\nu_n) + s\delta(\nu_n)\{\Omega(\nu_{n-1},\nu_n)+\Omega(\nu_n,\nu_{n+1})\} \\
&\Rightarrow \Omega(\nu_n,\nu_{n+1}) \leq \{\lambda(\nu_n)+s\delta(\nu_n)\}\Omega(\nu_{n-1},\nu_n) \\
&\Rightarrow |\Omega(\nu_n,\nu_{n+1})| \leq \{\lambda(\nu_n)+s\delta(\nu_n)\}|\Omega(\nu_{n-1},\nu_n)| \\
&\Rightarrow |\Omega(\nu_n,\nu_{n+1})| \leq \ell\,|\Omega(\nu_{n-1},\nu_n)|
\end{aligned}$$

Where $\ell = \{\lambda(\nu_n)+s\delta(\nu_n)\}$

$$\Rightarrow |\Omega(\nu_n,\nu_{n+1})| \leq \ell^2\,|\Omega(\nu_{n-2},\nu_{n-1})|$$

Continuing the above process, we get

$$|\Omega(\nu_n,\nu_{n+1})| \leq \ell^n\,|\Omega(\nu_0,\nu_1)|$$

on taking the limit as $n \longrightarrow \infty$ we obtain that $|\Omega(\nu_n, \nu_{n+1})| \longrightarrow 0$
Now, we show that $\{\nu_n\}$ is a Cauchy sequence in Δ, let $m, n \in N$, $m > n$ then

$$\begin{aligned}
\Omega(\nu_n, \nu_m) &\leq s\{\Omega(\nu_n, \nu_{n+1}) + \Omega(\nu_{n+1}, \nu_m)\} \\
&\leq s\{\Omega(\nu_n, \nu_{n+1})\} + s^2\{\Omega(\nu_{n+1}, \nu_{n+2}) + \Omega(\nu_{n+2}, \nu_m)\} \\
&\leq s\{\Omega(\nu_n, \nu_{n+1})\} + s^2\{\Omega(\nu_{n+1}, \nu_{n+2})\} \\
&\quad + s^3\{\Omega(\nu_{n+2}, \nu_{n+3}) + d(\nu_{n+3}, \nu_m)\} \\
&\leq s\ell^n\left\{1 + s\ell + (\ell k)^2 + (s\ell)^3 + \ldots\right\}\Omega(\nu_0, \nu_1) \\
&\leq s\ell^n\Omega(\nu_0, \nu_1)\left\{1 + s\ell + (s\ell)^2 + (s\ell)^3 + \ldots\right\} \\
\Rightarrow \Omega(\nu_n, \nu_m) &\leq \left(\frac{s\ell^n}{1 - s\ell}\right)\Omega(\nu_0, \nu_1) \\
\Rightarrow |\Omega(\nu_n, \nu_m)| &\leq \left(\frac{s\ell^n}{1 - s\ell}\right)|\Omega(\nu_0, \nu_1)|
\end{aligned}$$

On taking the limit, we have $\lim_{n\to\infty}\Omega(\nu_n, \nu_m) = 0$ as $n, m \longrightarrow \infty$, since $\left(\frac{s\ell^n}{1-s\ell}\right) < 1$ hence sequence $\{\nu_n\}$ is Cauchy sequence in Δ. Since Δ is complete then $\exists\ \omega$ in Δ such that $\lim_{n\to\infty} L\mu_n = \lim_{n\to\infty} T\mu_{n+1} = \lim_{n\to\infty} M\mu_{n+1} = \lim_{n\to\infty} S\mu_{n+2} = \omega$.
Now since $M(\Delta)$ is a closed subspace of Δ and so $\omega \in M(\Delta)$.
Since $M(\Delta) \subseteq S(\Delta)$, then there exist a point $v \in \Delta$, such that $\omega = Sv$. Now we show that $Lv = Sv = \omega$, we have

$$\begin{aligned}
d(Lv, \omega) &\leq s\{\Omega(Lv, \mu_n) + \Omega(M\mu_n, \omega)\} \\
&\frac{1}{s}\Omega(Lv, \omega) \leq \lambda(\mu_n)\max\left\{\Omega(Sv, T\mu_n), \frac{\Omega(Sv, Lv)\Omega(T\mu_n, \mu_n)}{\Omega(Sv, M\mu_n) + \Omega(T\mu_n, Lv)\Omega(Sv, T\mu_n)}\right\} \\
&+ \delta(\mu_n)\{\Omega(Sv, Lv) + \Omega(T\mu_n, M\mu_n)\} + \Omega(M\mu_n, \omega) + \Omega(M\mu_n, \omega)
\end{aligned}$$

letting $n \longrightarrow \infty$

$$\begin{aligned}
\Rightarrow \frac{1}{s}\Omega(Lv, \omega) &\leq \lambda(\mu_n)\max\left\{\Omega(Sv, \omega), \frac{\Omega(Sv, Lv)\Omega(\omega, \omega)}{\Omega(Sv, \omega) + \Omega(\omega, Lv) + \Omega(Sv, \omega)}\right\} \\
&+ \delta(\mu_n)\{\Omega(Sv, Lv) + \Omega(\omega, \omega)\} + \Omega(\omega, \omega) \\
&\Rightarrow \frac{1}{s}\Omega(Lv, \omega) \leq \lambda(\mu_n)\Omega(Sv, \omega) + \delta(\mu_n)\Omega(Sv, Lv) \\
&\Rightarrow \frac{1}{s}\Omega(Lv, \omega) \leq \lambda(\mu_n)\Omega(Sv, \omega) \\
&\Rightarrow \frac{1}{s}\Omega(Lv, \omega) = 0 \\
&\Rightarrow |\Omega(Lv, \omega)| = 0
\end{aligned}$$

Thus $Lv = Mv = \omega$ hence v is coincidence point of (L, S) Since $L(\Delta) \subseteq T(\Delta)$ and now $\omega \in L(\Delta)$, then there exist a point $w \in \Delta$ such that $\omega = Tw$, now we show that $Mw = \omega$, we also have $Lv = Sv = Tw = \omega$

$$\begin{aligned}
\Omega(\omega, Mw) &\leq s\{\Omega(\omega, L\mu_n) + \Omega(L\mu_n, Mw)\} \\
&\Rightarrow \frac{1}{s}\Omega(\omega, Mw) \\
&\leq \Omega(\omega, L\mu_n) + \lambda(\mu_n)\max\left\{\Omega(S\mu_n, Tw), \frac{\Omega(S\mu_n, L\mu_n)\Omega(Tw, Mw)}{\Omega(S\mu_n, Mw) + \Omega(Tw, L\mu_n) + \Omega(S\mu_n, Tw)}\right\}
\end{aligned}$$

letting $n \longrightarrow \infty$

$$\Rightarrow \frac{1}{s}d(\omega, Mw) \leq \Omega(\omega,\omega) + \lambda(\mu_n)\max\left\{\Omega(\omega, Tw), \frac{\Omega(\omega,\omega)d(Tw, Mw)}{\Omega(\omega, Mw) + \Omega(Tw,\omega) + \Omega(\omega, Tw)}\right\}$$
$$+ \delta(\mu_n)\{\Omega(\omega,\omega) + \Omega(Tw, Mw)\}$$
$$\Rightarrow |\Omega(\omega, Mw)| = 0$$
$$\Rightarrow Mw = \omega$$

Thus $Lv = Mv = \omega$ hence v is coincidence point of (M, T) Now we have $Lv = Sv = Tw = Mw = \omega$ since L and S are weakly compatible mapping then $LSv = SLv = L\omega = S\omega$, now we show that ω is fixed point of L on contrary if $L\omega \neq \omega$ then we have

$$\Omega(L\omega,\omega) \leq s\{\Omega(L\omega, M\mu_{n+1}) + \Omega(M\mu_{n+1},\omega)\}$$
$$\Rightarrow \frac{1}{s}\Omega(L\omega,\omega)$$
$$\leq s\lambda(\mu_{n+1})\max\left\{\Omega(S\omega, T\mu_{n+1}), \frac{\Omega(S\omega, L\omega)\Omega(T\mu_{n+1}, M\mu_{n+1})}{\Omega(S\omega, M\mu_{n+1}) + \Omega(T\mu_{n+1}, L\omega) + \Omega(S\omega, T\mu_{n+1})}\right\}$$
$$+ s\delta(\mu_{n+1})\{\Omega(S\omega, L\omega) + \Omega(T\mu_{n+1}, M\mu_{n+1})\} + s\Omega(M\mu_{n+1},\omega)$$

letting $n \longrightarrow \infty$

$$\frac{1}{s}\Omega(L\omega,\omega) \leq s\lambda(\mu_{n+1})\max\left\{\Omega(S\omega,\omega), \frac{\Omega(S\omega, L\omega)\Omega(\omega,\omega)}{\Omega(S\omega,\omega) + \Omega(\omega, L\omega) + \Omega(S\omega,\omega)}\right\}$$
$$+ s\delta(\mu_{n+1})\{\Omega(S\omega, L\omega) + \Omega(\omega,\omega)\} + s\Omega(\omega,\omega)$$
$$\Rightarrow \frac{1}{s}\Omega(L\omega,\omega) \leq s\lambda(\mu_{n+1})\Omega(S\omega,\omega)$$
$$\Rightarrow |\Omega(L\omega,\omega)| \leq s\lambda(\mu_{n+1})\,|\Omega(L\omega,\omega))|$$
$$\Rightarrow |\Omega(L\omega,\omega)| = 0$$

Hence $L\omega = S\omega = \omega$. Similarly, we show that ω is a fixed point of M and $M\omega = T\omega = \omega$. Therefore, ω is a common fixed point of L, M, S, and T.
Uniqueness of the fixed point: Let ω_1 be another Fixed point of L, M, S, and T then $L\omega_1 = M\omega_1 = S\omega_1 = T\omega_1 = \omega_1$,

$$\Omega(\omega,\omega_1) = \Omega(L\omega, M\omega_1)$$
$$\leq \lambda(\omega)\max\left\{\Omega(S\omega, T\omega_1), \frac{\Omega(S\omega, L\omega)\Omega(T\omega_1, M\omega_1)}{\Omega(S\omega, M\omega_1) + \Omega(T\omega_1, L\omega) + \Omega(S\omega, T\omega_1)}\right\}$$
$$+ \delta(\omega)\{\Omega(S\omega, L\omega) + \Omega(T\omega_1, M\omega_1)\}$$
$$\Rightarrow |\Omega(\omega,\omega_1)| \leq \lambda(\omega)\,|\Omega(\omega,\omega_1)|$$

Since $\lambda(\omega) < 1$

$$\Rightarrow |\Omega(\omega,\omega_1)| = 0$$
$$\Rightarrow \omega = \omega_1$$

Hence ω is the unique common fixed point of L, M, S, and T. ■

4. Applications

In this section, we apply Theorem 1 to the existence of a common solution of the system of Urysohn integral equations.
Let $\Delta = C([a,b], \mathbb{R}^n)$, where $[a,b] \subseteq \mathbb{R}^+$. and $d : \Delta \times \Delta \to \mathbb{C}$ is define by

$$d(\mu,\nu) = \max_{t\in[a,b]} ||\mu(t) - \nu(t)||_\infty \sqrt{1+a^2}e^{i\tan^{-1}a}$$

Consider the Urysohn integral equations

$$\mu(t) = \int_a^b \kappa_1(t, s, \mu(s))\, ds + g(t) \tag{2}$$

$$\mu(t) = \int_a^b \kappa_2(t, s, \mu(s))\, ds + h(t) \tag{3}$$

where $t \in [a, b] \subset \mathbb{R}$ and $\mu, g, h \in \Delta$
Suppose that $\kappa_1, \kappa_2 : [a, b] \times [a, b] \times \mathbb{R}^n \to R^n$ are such that $F_\mu, G_\mu \in \Delta$ for each $\mu \in \Delta$, where

$$F_\mu(t) = \int_a^b \kappa_1(t, s, \mu(s))\, ds + g(t),$$
$$G_\mu(t) = \int_a^b \kappa_2(t, s, \mu(s))\, ds + h(t)$$

for all $t \in [a, b]$. If there exists two mapping $\lambda, \delta : \Delta \to [0, 1)$ such that for all $\mu, \nu \in \Delta$,

1. $\lambda(F_\mu + g) \leq \lambda(\mu)$ and $\delta(F_\mu + g) \leq \delta(\mu)$
2. $\lambda(G_\mu + h) \leq \lambda(\mu)$ and $\delta(G_\mu + h) \leq \delta(\mu)$
3. $(\lambda + \delta)\mu < 1$
4. $||F_\mu(t) - G_\nu(t) + g(t) - h(t)||_\infty \sqrt{1 + a^2} e^{i \tan^{-1} a} \leq \lambda(\mu) A(\mu, \nu)(t) + \delta(\mu) B(\mu, \nu)(t)$

$$A(\mu, \nu)(t) = \max_{t \in [a,b]} \left\{ ||\mu(t) - \nu(t)||_\infty, \frac{\left(||F_\mu(t) + g(t) - \mu(t)||_\infty ||G_\nu(t) + h(t) - \nu(t)||_\infty\right) \sqrt{1 + a^2} e^{i \tan^{-1} a}}{||G_\nu(t) + h(t) - \mu(t)||_\infty + ||F_\mu(t) + g(t) - \nu(t)||_\infty + ||\mu(t) - \nu(t)||_\infty} \right\}$$

$$B(\mu, \nu)(t) = \left\{ ||F_\mu(t) + g(t) - \mu(t)||_\infty ||G_\nu(t) + h(t) - \nu(t)||_\infty \right\} \sqrt{1 + a^2} e^{i \tan^{-1} a}$$

then the system of integral equations (2) and (3) have a unique common solution.
Proof. Let (Δ, Ω) be a complete complex valued b-metric space and $L, M : \Delta \times \Delta \to \Delta$ by $L\mu = F_\mu + g$ and $M\Delta = G_\mu + h$. Then

$$\Omega(L\mu, M\nu) = \max_{t \in [a,b]} ||F_\mu(t) - G_\nu(t) + g(t) - h(t)||_\infty \sqrt{1 + a^2} e^{i \tan^{-1} a},$$
$$\Omega(\mu, L\mu) = \max_{t \in [a,b]} ||F_\mu(t) + g(t) - \mu(t)||_\infty \sqrt{1 + a^2} e^{i \tan^{-1} a},$$
$$\Omega(\nu, M\nu) = \max_{t \in [a,b]} ||G_\nu(t) + h(t) - \nu(t)||_\infty \sqrt{1 + a^2} e^{i \tan^{-1} a},$$
$$\Omega(\mu, M\nu) = \max_{t \in [a,b]} ||G_\nu(t) + h(t) - \mu(t)||_\infty$$

and

$$\Omega(\nu, L\mu) = \max_{t \in [a,b]} ||F_\mu(t) + g(t) - \nu(t)||_\infty$$

It is easily seen that for all $\mu, \nu \in \Delta$, we have

1. $\lambda(L\mu) \leq \lambda(\mu)$, $\delta(L\mu) \leq \delta(\mu)$

2. $\lambda(M\mu) \leq \lambda(\mu)$ and $\delta(M\mu) \leq \delta(\mu)$

3.

$$\Omega(L\mu, M\nu) \leq \lambda(\mu) \max\left\{\Omega(\mu,\nu), \frac{\Omega(\mu, L\mu)\Omega(\nu, M\nu)}{\Omega(\mu, M\nu) + \Omega(\nu, L\mu) + \Omega(\mu,\nu)}\right\} + \delta(\mu)\left\{\Omega(\mu, L\mu) + d(\nu, M\nu)\right\}$$

By theorem 1, we get L and M have a common fixed point. Thus, there is a unique point $\mu \in \Delta$ such as $\mu = L\mu = M\mu$. Now, we have
$\mu = L\mu = F_\mu + g$ and $\mu = M\mu = G_\mu + h$ that is

$$\mu(t) = \int_a^b \kappa_1(t, s, \mu(s))\, ds + g(t)$$

and

$$\mu(t) = \int_a^b \kappa_2(t, s, \mu(s))\, ds + h(t)$$

Therefore, the Urysohn integral (2) and (3) have a unique common fixed point. ■

5. Conclusion

Complex-valued metric spaces, which are a broadening of conventional metric spaces, hold significant importance in various fields, particularly in advanced mathematics and theoretical physics. Their applications extend to areas such as relativity, geometry, topology, mathematical finance, economics, quantum mechanics, quantum computing, functional analysis, operator theory, complex dynamics, fractal geometry, complex networks, and signal processing. Although interest in complex-valued metric spaces is increasing, the existing body of literature on fixed point theorems in these spaces is still quite limited and exhibits several shortcomings. In 2011, Azam and his colleagues developed a metric space that accommodates complex values and presented a prominent fixed point theorem that satisfies specific contractive conditions [9]. Subsequently, in 2013, another study proposed an approach for Urysohn integral equations that incorporates common fixed points in metric spaces with complex values [12]. Following this, Rao and his team expanded on these findings plying them to complex valued b-metric spaces and establishing a fixed point theorem [16]. This article reviews these advancements and presents a generalization of a new contraction mapping while employing weakly compatible mappings to investigate the completeness and uniqueness of fixed point theorems in complex valued b-metric spaces. In addition, we tackled a system of Urysohn integral equations by utilizing our primary results, leading to some fascinating discoveries and the formulation of several theories.

References

[1] Banach, S. (1922). Sur les operations dans les ensembles abstraits et leur applications aux equations integrales. *Fund. Math.*, 3:133–181.
[2] Czerwik, S. (1998). Nonlinear set-valued contraction mappings in b-metric spaces. *Atti Sem. Mat. Fis. Univ. Modena*, 46:263-276.
[3] Cho, Y. J. (2017). Survey on metric fixed point theory and applications. *Advances in Real and Complex Analysis with Applications*, 183–241.
[4] Cobzas S, Czerwik S. (2020). The completion of generalized b-metric spaces and fixed points. *Fixed Point Theory*, 21(1):133-150.
[5] Karlsson, A. (2024). A metric fixed point theorem and some of its applications. *Geometric and Functional Analysis*, 34(2):486-511.

[6] Valappil, S. V., Pulickakunnel, S. (2024). Some fixed point results for pair of relatively non-expansive mappings in W-hyperbolic spaces using Thakur iteration. *Rendiconti del Circolo Matematico di Palermo Series*, 2:1-18.
[7] Chourasiya S, Shrivastava K. (2024). Common Fixed Point Theorem in Controlled Metric Spaces. *Int. j. adv. multidisc. res. stud.*, 4(5):317-323.
[8] Pratibha, Rajesh D. (2024). Solving generalized fuzzy least cost path problem of supply chain network. *Reliability: Theory Applications*, 4(80): 267-286.
[9] Azam A, Fisher B, Khan M. (2011). Common Fixed point theorems in complex valued metric spaces. *Numerical Functional Analysis and Optimization*, 32:243-53.
[10] Sintunavarat, W., Kumam, P. (2012). Generalized common Fixed point theorems in complex valued metric spaces and applications. *Journal of inequalities and Applications*, 2012:1-12.
[11] Sitthikul, K., Saejung, S. (2012). Some fixed point theorems in complex valued metric spaces. *Fixed Point Theory and Applications*, 2012:1-11.
[12] Sintunavarat, W., Cho, Y. J., Kumam, P. (2013). Urysohn integral equations approach by common fixed points in complex-valued metric spaces. *Advances in Difference Equations*, 2013:1-14.
[13] Bhatt, S., Chaukiyal, S., Dimri, R. C. (2011). A common fixed point theorem for weakly compatible maps in complex valued metric spaces. *Int. J. Math. Sci. Appl.*, 1(3):1385-1389.
[14] Verma, R. K., Pathak, H. K. (2013). Common fixed point theorems using property (EA) in complex-valued metric spaces. *Thai J. Math*, 11(2):347-355.
[15] Singh, N., Singh, D., Badal, A., Joshi, V. (2016). Fixed point theorems in complex valued metric spaces. *Journal of the Egyptian Mathematical Society*, 24(3):2-409.
[16] Rao, P. R., Swamy, P. R., Prasad, J. R. (2013). A common Fixed point theorem in complex valued b-metric spaces. *Bulletin of Mathematics and Statistics research*, 1(1):1-8.
[17] Singh, Deepak, et al. (2015). Common fixed point theorems in complex valued b-metric spaces. *J. Math. Comput. Sci.*, 5(3).
[18] Jungck, G. (1986). Compatible mappings and common fixed points. *International journal of mathematics and mathematical sciences*, 9(4):771-779.

RT&A, No 2 (84)
Volume 20, June 2025

ESTIMATION AND APPLICATION OF A NEW GENERALIZATION OF EXPONENTIAL DISTRIBUTION

D. Kumar*, P. K. chaurasia**, P. Kumar*** and A. Sahoo*

•

* Department of Statistics, Banaras Hindu University, Varanasi-221005, India
**School of Business, UPES, Dehradun-248007, India
***Department of Statistics, Central University of Odisha, Koraput-763004, India
chaurasiaprashantvns@gmail.com

Abstract

In statistical literature, several lifetime distributions exist for real phenomena. And one of the methods to find new lifetime distribution by existing baseline distribution such a method is known as the transformation method. In this article, we proposed a generalization of the existing transformation by introducing the additional shape parameter. Here, we considered baseline distributions as an exponential distribution. Various statistical properties of the new lifetime distribution, such as survival function, hazard rate function, cumulative hazard rate function, moments, quantile function, and order statistics, have been discussed. Demonstrate the applicability and suitability of the proposed distribution. Here, we focus only on the estimation of the parameters likes MLE, LSE and also to check long-run behaviour of the estimators.

Keywords: Exponential distribution, Moments, Maximum likelihood estimator, Simulation Study

1. Introduction

In statistical literature, we have supposed that every real phenomenon is governed by some lifetime models. If we know the model, we can completely specify to our problem or phenomenon as various lifetime models have been developed for this purpose. The exponential distribution has been widely used in analyzing lifetime data due to its lack of memory property and its very simple form. However, the exponential distribution with only a constant hazard rate shape and it is not able to fit data sets with different hazard shapes like increasing, decreasing, bathtub, or unimodal (upside down bathtub) shaped failure rates, often encountered in medical, engineering and reliability, among others. In this way, the generalization of exponential distribution is Weibull distribution, it has increasing and decreasing hazard rate functions depends on the values of shape parameter. Gamma distribution has also increasing and decreasing hazard rate functions depends on its shape parameter. Generalized exponential distribution has similar characteristics according to Weibull and gamma distributions but it is more flexible as compared to Weibull and gamma distributions.

Recently, several authors have developed several generalizations of the exponential distribution to increase its flexibility in terms of their shapes of hazard rate and pdf. For example, the Marshall-Olkin exponential [1], exponentiated exponential by Gupta and Kundu [2], Harris extended exponential by Pinho et al. [3], DUS exponential distribution proposed and SS exponential distribution by Kumar et. al.[4], [5] , Kumaraswamy transmuted exponential by Afify et al. [6], a new transformation method based on trignometric function developed by Chaurasia et. al. [7], modified exponential by Rasekhi et al. [8], Minimum Guarantee exponential distribution by

Kumar et. al. [9], odd exponentiated half-logistic exponential by Afify et al. [10], Marshall-Olkin logistic-exponential by Mansoor et al. [11], odd log-logistic Lindley exponential by Alizadeh et al. [12], Kumar et. al. [13] investigate the statistical properties and application of SS-Lindley distribution and Marshall-Olkin alpha power exponential by Nassar et al. [14], among others. Application of exponential distribution has been discussed by Kumar et al. [15] in Bayesian setup.

Recently, Kumar et al. [9] introduced minimum guarantee (MG) transformation in order to generalize any available lifetime distribution called baseline distribution. The idea is very simple, as discussed below; If d(z) be the pdf of $Z > \sigma > 0$ and if $Y = \frac{\sigma}{Z}$, then its pdf k(y) is given by k(y)= d(z)$\left|\frac{\partial z}{\partial y}\right| = d\left(\frac{\sigma}{y}\right)\left|\frac{-\sigma}{y^2}\right| = \frac{\sigma}{y^2}.\ d\left(\frac{\sigma}{y}\right); 0 < y < 1$. Now, if F(x) and f(x) be the cdf and pdf of some available baseline distribution respectively, then the pdf g(x) of new lifetime distribution is given by,

$$g(x) = k(F(x))f(x); \quad x > 0 \tag{1}$$

The host authors Kumar et al. [9] named it as MG transformation. Now, we consider the following pdf,

$$\partial(z) = \alpha\beta^{\alpha}e^{-\alpha z}; \qquad z > \ln\beta, \alpha > 0, \beta > 1$$

Making the transformation by the substitution

$$\begin{aligned} Y &= \frac{\ln\beta}{Z} \\ Z &= \frac{\ln\beta}{Y} \\ J &= \left|\frac{\partial z}{\partial y}\right| = \frac{-1}{y^2}\ln\beta \\ |J| &= \frac{1}{y^2}\ln\beta \end{aligned}$$

and as z ranges $\ln\beta$ to ∞, so that y ranges 0 to 1. If k(y) be the pdf of of Y, then

$$\begin{aligned} k(y) &= d(z)\left|\frac{\partial z}{\partial y}\right| \\ &= \alpha\beta^{\alpha}\frac{1}{y^2}\ln\beta e^{\ln\beta^{-\frac{\alpha}{y}}} \\ &= \frac{\alpha\ln\beta}{y^2}\beta^{\alpha-\frac{\alpha}{y}} \\ &= \frac{\alpha\ln\beta}{y^2}\beta^{\alpha(1-\frac{1}{y})}; \qquad 0 < y < 1, \alpha > 0,\ \beta > 1 \end{aligned} \tag{2}$$

Now, if F(x) and f(x) be the cdf and pdf, respectively of some available baseline distribution, then the pdf g(x) of new distribution corresponding to cdf F(x) and pdf f(x) can be obtained as follows,

$$g(x) = \frac{\alpha\ln\beta}{(F(x))^2}\beta^{\alpha\left(1-\frac{1}{F(x)}\right)}f(x); \qquad \forall\ x > 0,\ \alpha > 0,\ \beta > 1 \tag{3}$$

and the cdf G(x) corresponding to the pdf g(x) will be,

$$G(x) = \beta^{\alpha\left(1-\frac{1}{F(x)}\right)}; \quad \forall\ x > 0,\ \alpha > 0,\ \beta > 1 \tag{4}$$

We will call transformation (4) as generalized minimum guarantee (GMG) transformation. In particular for $\alpha = 1$ and $\beta = e$, the proposed transformation (4) reduces to,

$$G(x) = e^{1-\frac{1}{F(x)}} \quad \forall\ x > 0 \tag{5}$$

Which is MG transformation as suggested by Kumar et al.[9], that is why we called (4) as GMG transformation. Now, to demonstrate the real-life application of GMG transformation,

we consider the exponential distribution with pdf $f(x) = \theta e^{-\theta x};\ \ x > 0,\ \theta > 0$ and cdf $F(x) = 1 - e^{-\theta x};\ \ x > 0,\ \theta > 0$ as an available baseline distribution. Thus the cdf G(x) and pdf g(x) corresponding to the considered baseline distribution, using GMG transformation, are obtained as follows,

$$G(x) = \beta^{\alpha\left(1-\frac{1}{1-e^{-\theta x}}\right)} \quad \forall\ x > 0,\ \theta > 0,\ \alpha > 0,\ \beta > 1 \tag{6}$$

$$g(x) = \frac{\alpha \ln \beta}{(1-e^{-\theta x})^2}\beta^{\alpha}\left(1 - \frac{1}{1-e^{-\theta x}}\right)\theta e^{-\theta x} \quad \forall\ x > 0,\ \theta > 0,\ \alpha > 0,\ \beta > 1 \tag{7}$$

We shall use the abbreviation $\text{GMG}_E(\alpha, \beta, \theta)$- distribution for generalized minimum gurantee transformation as exponential distribution with parameter θ. The plots of cdf G(x) and pdf g(x) of $\text{GMG}_E(\alpha, \beta, \theta)$- distribution for various combination of values of α, β and θ are shown in Figure 1 and 2 respectively.

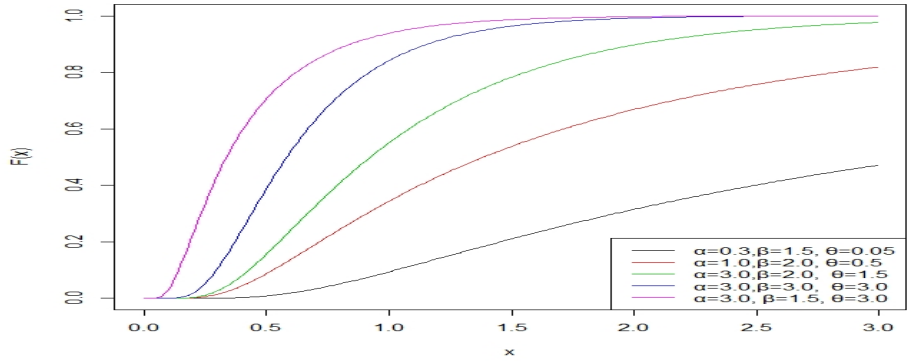


Figure 1: *Plots of cumulative distribution function G(x) for various combination of values of α, β and θ.*

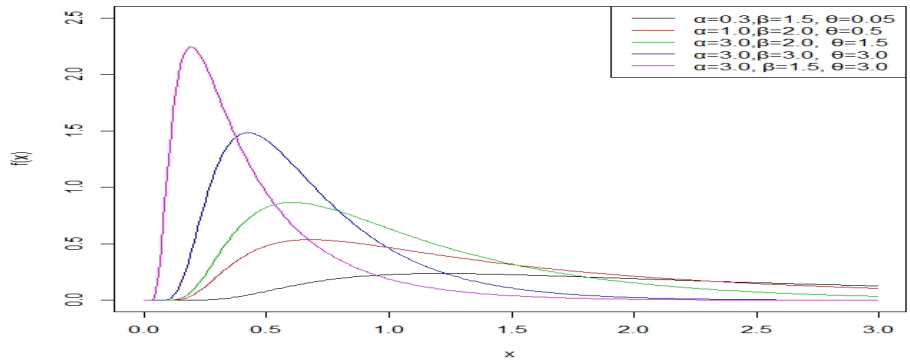


Figure 2: *Plots of probability density function g(x) for various combination of values of α, β and θ.*

2. Reliability Analysis

In this section, we have derived various reliability characteristics such as survival function, hazard rate function and cumulative hazard rate function of $\text{GMG}_E(\alpha, \beta, \theta)$- distribution with cdf (6).

2.1. Survival Function

The Survival function S(x) of $GMG_E(\alpha, \beta, \theta)$- distribution at a mission time x is obtained as follows,

$$\begin{aligned} S(x) &= 1 - G(x) \\ &= 1 - \left[\beta^{\alpha - \frac{\alpha}{1-e^{-\theta x}}}\right]; \qquad x > 0, \theta > 0, \alpha > 0 \; \beta > 1 \end{aligned} \tag{8}$$

Figure 3 shows the behaviour of survival function S(x) of $GMG_E(\alpha, \beta, \theta)$- distribution for different combination of values of the parameters α, β and θ.

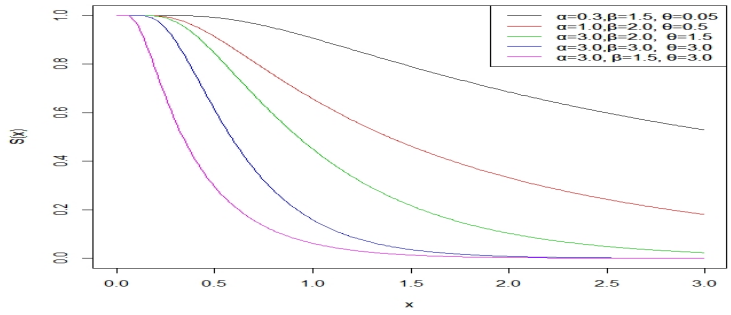


Figure 3: *Plots of survival function S(x) for different combination of values of α, β and θ.*

2.2. Hazard Rate Function

The hazard rate function h(x) of $GMG_E(\alpha, \beta, \theta)$- distribution is obtained as follows,

$$\begin{aligned} h(x) &= \frac{g(x)}{S(x)} \\ &= \frac{\alpha \ln \beta \left(\beta^{\alpha - \frac{\alpha}{1-e^{-\theta x}}}\right) \theta e^{-\theta x}}{\left(1 - e^{-\theta x}\right)^2 \left[1 - \beta^{\alpha - \frac{\alpha}{1-e^{-\theta x}}}\right]}; \quad x > 0, \theta > 0, \alpha > 0 \; \beta > 1 \end{aligned} \tag{9}$$

The shapes of hazard rate function h(x) for different combination of values of parameters α, β and θ are shown in Figure 4.

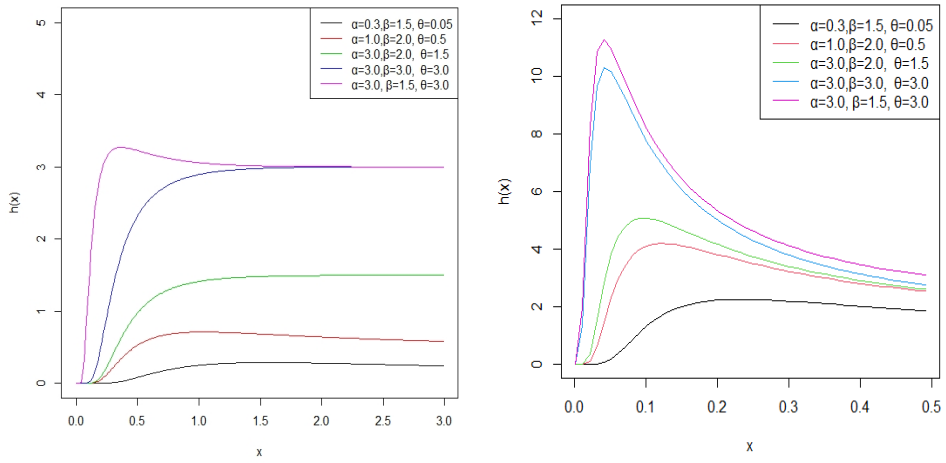


Figure 4: *Plots of hazard rate function h(x) for different combination of values of α, β and θ.*

2.3. Cumulative Hazard Rate Function

The cumulative hazard rate function H(x) of $\text{GMG}_E(\alpha,\beta,\theta)$- distribution is obtained as follows,

$$\begin{aligned} H(x) &= -\ln\{1-G(x)\} \\ &= -\ln\left\{1-\left(\beta^{\alpha-\frac{\alpha}{1-e^{-\theta x}}}\right)\right\}; \quad x>0,\theta>0,\alpha>0\ \beta>1 \end{aligned} \tag{10}$$

Figure 5 shows the behaviour of cumulative hazard rate function H(x) of $\text{GMG}_E(\alpha,\beta,\theta)$-distribution for different combination of values of the parameters α,β and θ.

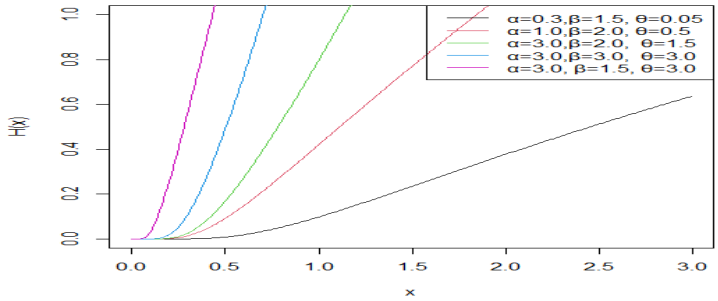


Figure 5: *Plots of cumulative hazard rate function H(x) for different combination of values of α,β and θ.*

3. Statistical Properties of $\text{GMG}_E(\alpha,\beta,\theta)$- distribution

3.1. Moments

The r^{th} order non-central moment, i.e, r^{th} moment about origin of $\text{GMG}_E(\alpha,\beta,\theta)$- distribution having pdf (7) is obtained as follows,

$$\begin{aligned} \mu_r' &= \int_0^\infty x^r f(x)dx \\ &= \theta\alpha\ln\beta\int_0^\infty \frac{x^r}{(1-e^{-\theta x})^2}\beta^{\alpha-\frac{\alpha}{1-e^{-\theta x}}}e^{-\theta x}dx \\ &= \theta\alpha\ln\beta\beta^\alpha\int_0^\infty \frac{x^r}{(1-e^{-\theta x})^2}\beta^{-\frac{\alpha}{1-e^{-\theta x}}}e^{-\theta x}dx \\ &= \theta\alpha\ln\beta\beta^\alpha\sum_{k=0}^\infty(k+1)\int_0^\infty x^r e^{-k\theta x}\beta^{-\frac{\alpha}{1-e^{-\theta x}}}e^{-\theta x}dx \\ &= \theta\alpha\sum_{k=0}^\infty(k+1)\ln\beta\beta^\alpha\int_0^\infty x^r e^{-k\theta x}\beta^{-\frac{\alpha}{1-e^{-\theta x}}}e^{-\theta x}dx \\ &= \theta\beta^\alpha\sum_{k=0}^\infty\sum_{l=0}^\infty\frac{(-1)^{l+m}(k+1)\alpha^{l+1}(\ln\beta)^{l+1}}{l!}\int_0^\infty x^r e^{-k\theta x}(1-e^{-\theta x})^l e^{-\theta x}dx \\ &= \theta\beta^\alpha\sum_{k=0}^\infty\sum_{l=0}^\infty\sum_{m=0}^\infty\frac{\binom{-l}{m}(-1)^{l+m}(k+1)\alpha^{l+1}(\ln\beta)^{l+1}}{l!}\int_0^\infty x^r e^{-\theta(k+m+1)x}dx \\ &= \theta\beta^\alpha\sum_{k=0}^\infty\sum_{l=0}^\infty\sum_{m=0}^\infty\frac{\binom{-l}{m}(-1)^{l+m}(k+1)\alpha^{l+1}(\ln\beta)^{l+1}}{l!}\frac{\Gamma(r+1)}{(\theta(k+m+1))^{r+1}} \end{aligned} \tag{11}$$

3.2. Sample Generation

Let U follows continuous uniform distribution over [0,1], then

$$\begin{aligned}
G(x) &= u \\
&\implies \beta^{\alpha\left(1-\frac{1}{F(x)}\right)} = u \\
&\implies \beta^{\alpha\left(1-\frac{1}{1-e^{-\theta x}}\right)} = u \\
&\implies \alpha\left(1-\frac{1}{1-e^{-\theta x}}\right)\ln\beta = \ln u \\
&\implies 1-\frac{1}{1-e^{-\theta x}} = \frac{\ln u}{\alpha\ln\beta} \\
&\implies \frac{1}{1-e^{-\theta x}} = 1-\frac{\ln u}{\alpha\ln\beta} \\
&\implies 1-e^{-\theta x} = \frac{1}{1-\frac{\ln u}{\alpha\ln\beta}} \\
&\implies e^{-\theta x} = 1-\frac{1}{1-\frac{\ln u}{\alpha\ln\beta}} \\
&\implies -\theta x = \ln\left(1-\frac{1}{1-\frac{\ln u}{\alpha\ln\beta}}\right) \\
&\implies x = -\frac{1}{\theta}\ln\left(1-\frac{1}{1-\frac{\ln u}{\alpha\ln\beta}}\right) \qquad (12)
\end{aligned}$$

and consequently if $\underline{x}$= $(x_1, x_2, ..., x_n)$ be a random sample of size n from $\mathrm{GMG}_E(\alpha, \beta, \theta)$-distribution, then it can be obtained as follows;

$$x_i = -\frac{1}{\theta}\ln\left(1-\frac{1}{1-\frac{\ln u_i}{\alpha\ln\beta}}\right) \quad \forall\ 1, 2, ...n. \qquad (13)$$

where $u_i \sim u[0,1]\ \forall$ i= 1(1) n.

3.3. Quantile Function

For any p $\in$ (0,1), the p^{th} quantile function Q(p) of $\mathrm{GMG}_E(\alpha, \beta, \theta)$- distribution is the solution of the following equation,

$$\begin{aligned}
& G(Q(p)) = p \\
\implies & \beta^{\alpha\left(1-\frac{1}{1-e^{-\theta Q(p)}}\right)} = p \\
\implies & \alpha\left(1-\frac{1}{1-e^{-\theta Q(p)}}\right)\ln\beta = \ln p \\
\implies & 1-\frac{1}{1-e^{-\theta Q(p)}} = \frac{\ln p}{\alpha\ln\beta} \\
\implies & \frac{1}{1-e^{-\theta Q(p)}} = 1-\frac{\ln p}{\alpha\ln\beta} \\
\implies & -\theta Q(p) = \ln\left(1-\frac{1}{1-\frac{\ln p}{\alpha\ln\beta}}\right) \\
\implies & Q(p) = -\frac{1}{\theta}\ln\left(1-\frac{1}{1-\frac{\ln p}{\alpha\ln\beta}}\right)
\end{aligned}$$

In particular for p = $\frac{1}{2}$, we can get, median $\tilde{\mu}$ of $\text{GMG}_E(\alpha,\beta,\theta)$- distribution as follows,

$$\begin{aligned}\tilde{\mu} &= Q\left(\frac{1}{2}\right)\\ &= -\frac{1}{\theta}\ln\left(1-\frac{1}{1-\frac{\ln\left(\frac{1}{2}\right)}{\alpha\ln\beta}}\right)\\ &= -\frac{1}{\theta}\ln\left(1-\frac{1}{1-\frac{\ln 1-\ln 2}{\alpha\ln\beta}}\right)\\ &= -\frac{1}{\theta}\ln\left(1-\frac{1}{1+\frac{\ln 2}{\alpha\ln\beta}}\right)\end{aligned} \tag{14}$$

4. Estimation of the parameters α, β and θ of $\text{GMG}_E(\alpha,\beta,\theta)$- distribution:-

In this section, we will derive maximum likelihood estimators and least square estimators of the parameters α, β and θ of $\text{GMG}_E(\alpha,\beta,\theta)$- distribution with pdf (7) on the basis of complete sample $X_1, X_2, ...X_n$ of fixed size n from it.

4.1. Maximum Likelihood Estimators

The likelihood function for $X_1, X_2, ...X_n$ a random sample of fixed size n from $\text{GMG}_E(\alpha,\beta,\theta)$-distribution having pdf (7) is given by,

$$\begin{aligned}L &= \prod_{i=1}^{n}(g(x_i))\\ &= \prod_{i=1}^{n}\left(\frac{\alpha\ln\beta}{(1-e^{-\theta x})^2}\beta^{\alpha-\frac{\alpha}{1-e^{-\theta x}}}\theta e^{-\theta x}\right)\end{aligned}$$

Taking natural logarithm of both sides of the above equation, we get,

$$\begin{aligned}l &= \ln L\\ &= n\ln\alpha + n\ln(\ln\beta) + n\ln\theta - \theta\sum_{i=1}^{n}x_i +\\ &\sum_{i=1}^{n}\left(\alpha-\frac{\alpha}{1-e^{-\theta x_i}}\right)\ln\beta - 2\sum_{i=1}^{n}\ln(1-e^{-\theta x_i})\end{aligned} \tag{15}$$

and hence the log-likelihood equations function estimating α, β and θ are given by

$$\begin{aligned}&\frac{\partial l}{\partial\alpha} = 0\\ \implies &\frac{n}{\alpha} + n\ln\beta - \ln\beta\sum_{i=1}^{n}\frac{1}{1-e^{-\theta x_i}} = 0\end{aligned} \tag{16}$$

$$\begin{aligned}&\frac{\partial l}{\partial\beta} = 0\\ \implies &\frac{n}{\beta\ln\beta} + \frac{1}{\beta}\sum_{i=1}^{n}\left(\alpha-\frac{\alpha}{1-e^{-\theta x_i}}\right) = 0\end{aligned} \tag{17}$$

$$\begin{aligned}&\frac{\partial l}{\partial\theta} = 0\\ \implies &\frac{n}{\theta} - \sum_{i=1}^{n}x_i + \ln\beta\sum_{i=1}^{n}\frac{\alpha x_i e^{-\theta x_i}}{\left(1-e^{-\theta x_i}\right)^2} - 2\sum_{i=1}^{n}\frac{x_i e^{-\theta x_i}}{1-e^{-\theta x_i}} = 0\end{aligned} \tag{18}$$

It is clear that the above system of log-likelihood equations can not be solved analytically for α, β and θ. Therefore, some approximation technique will be used for approximate solution of them for α, β and θ. Here, we have again used nlm command in R software to maximize l and hence $\ln L$ for variation in α, β and θ.

4.2. Least Squares Estimators

The least square estimators of α, β and θ on the basis of complete sample $\underline{X} = (X_1, X_2, \ldots, X_n)$ of fixed size n from $\text{GMG}_E(\alpha, \beta, \theta)$- distribution can be obtained by minimizing the following function,

$$\begin{aligned} P(\alpha, \beta, \theta) &= \sum_{i=1}^{n} \left(G(x_i) - E(G_n(x_i))\right)^2 \\ &= \sum_{i=1}^{n} \left(G(x_i) - \frac{i}{n+1}\right)^2 \end{aligned}$$

where $E(G_n(x_i)) = \frac{i}{n+1}$ $\forall i = 1(1)n$ is the empirical cdf of $\underline{X} = (X_1, X_2, \ldots, X_n)$.
Now, the normal equations for estimating α, β and θ are obtained as follows,

$$\frac{\partial P(\alpha, \beta, \theta)}{\partial \alpha} = 0 \Rightarrow \sum_{j=1}^{n} G'_\theta(x_j; \alpha, \beta, \theta) \left((G(x_j; \alpha, \beta, \theta)) - \frac{j}{n+1}\right) = 0$$

$$\frac{\partial P(\alpha, \beta, \theta)}{\partial \beta} = 0 \Rightarrow \sum_{j=1}^{n} G'_\theta(x_j; \alpha, \beta, \theta) \left((G(x_j; \alpha, \beta, \theta)) - \frac{j}{n+1}\right) = 0$$

$$\frac{\partial P(\alpha, \beta, \theta)}{\partial \theta} = 0 \Rightarrow \sum_{j=1}^{n} G'_\theta(x_j; \alpha, \beta, \theta) \left((G(x_j; \alpha, \beta, \theta)) - \frac{j}{n+1}\right) = 0$$

we have again used nlm command for the simultaneous solution of above normal equations for α, β and θ, as the system of above normal equations is not easy to solve analytically for α, β and θ. Let the least square estimators (LSEs) of α, β and θ, thus obtained are denoted by $\hat{\alpha}_L$, $\hat{\beta}_L$ and $\hat{\theta}_L$.

5. Simulation Study

In this section, the simulation study is carried out to know the performance of MLEs and LSEs of α, β and θ in terms of their mean squared error (MSEs) with varying sample sizes. The value of α, β and θ are arbitrarily chosen as α=0.5, 1, 2, β= 2 and θ=0.5, 1, 2 and the different considered values of n are 20, 50, 80, 100, 150 and 200. The process is repeated 5000 times in order to achieve convergence and MLEs $\hat{\alpha}_M$, $\hat{\beta}_M$ and $\hat{\theta}_M$ and LSEs $\hat{\alpha}_L$, $\hat{\beta}_L$ of α, β and θ have been calculated for each generated samples and consequently MSEs of these estimators have been calculated and reported in Tables 1,2 and 3. We have used the notation $\text{m}(\hat{\theta})$ to represent the mean squared error of the estimator.

Table 1: *MSEs of MLEs and LSEs of α, β and θ for fixed α=0.5, β=2 and θ=0.5 variation in the values of n.*

n	$\text{m}(\hat{\theta}_M)$	$\text{m}(\hat{\theta}_L)$	$\text{m}(\hat{\alpha}_M)$	$\text{m}(\hat{\alpha}_L)$	$\text{m}(\hat{\beta}_M)$	$\text{m}(\hat{\beta}_L)$
20	**2.3015**	3.1142	**0.2033**	0.5134	0.5113	**0.5090**
50	**0.3213**	0.7546	**0.0617**	0.3371	0.2866	**0.2172**
80	**0.1616**	0.4023	**0.0230**	0.2784	0.2462	**0.1932**
100	**0.1119**	0.2840	**0.0216**	0.1570	**0.1216**	0.1328
150	**0.0676**	0.1844	**0.0126**	0.0984	**0.0323**	0.1134
200	**0.0415**	0.1651	**0.0114**	0.0148	**0.0121**	0.1033

Table 2: *MSEs of MLEs and LSEs of α, β and θ for fixed α=1, β=2 and θ=1 variation in the values of n.*

n	m($\hat{\theta}_M$)	m($\hat{\theta}_L$)	m($\hat{\alpha}_M$)	m($\hat{\alpha}_L$)	m($\hat{\beta}_M$)	m($\hat{\beta}_L$)
20	**2.3015**	3.1142	**0.2033**	0.5134	0.5113	**0.5090**
50	**0.4947**	1.2120	**0.1898**	0.4298	0.2943	**0.2646**
80	**0.2407**	0.5236	**0.1742**	0.3615	0.2462	**0.1947**
100	**0.1850**	0.4319	**0.1591**	0.3450	**0.1957**	0.3879
150	**0.1121**	0.2800	**0.1110**	0.2553	**0.0471**	0.1623
200	**0.0819**	0.2582	**0.0928**	0.2234	**0.0335**	0.0898

Table 3: *MSEs of MLEs and LSEs of α, β and θ for fixed α=2, β=2 and θ=2 with variation in the values of n.*

n	m($\hat{\theta}_M$)	m($\hat{\theta}_L$)	m($\hat{\alpha}_M$)	m($\hat{\alpha}_L$)	m($\hat{\beta}_M$)	m($\hat{\beta}_L$)
20	**2.8192**	3.8851	**0.5871**	1.1527	**0.6192**	0.6768
50	**0.6540**	1.5958	**0.5612**	1.1485	**0.2997**	0.4288
80	**0.3716**	0.7409	**0.5391**	0.9697	**0.1803**	0.4053
100	**0.2846**	0.5625	**0.5297**	0.9310	**0.1612**	0.2594
150	**0.1725**	0.3679	**0.5029**	0.8069	**0.0896**	0.1655
200	**0.1248**	0.2809	**0.4903**	0.5781	**0.0797**	0.0842

From Tables 1 and 2, we observed that MLEs $\hat{\theta}_M$ and $\hat{\alpha}_M$ performs better as compared to their respective rivals LSEs $\hat{\theta}_L$ and $\hat{\alpha}_L$ in terms of having smallest values of MSEs, for all the considered values of n, where the values of (α, β and θ) have been as (0.5, 2, 0.5) and (1, 2, 1) respectively. But, in the case of estimators of β, the LSE $\hat{\beta}_L$ performs better as compared to MLE $\hat{\beta}_M$ for the considered small values of n and for larger values of n, reverse trend is noted.

Finally, from Table 3, we observed that MLEs $\hat{\theta}_M$, $\hat{\alpha}_M$ and $\hat{\beta}_M$ performs better as compared to their respective rivals LSEs $\hat{\theta}_L$, $\hat{\alpha}_L$ and $\hat{\beta}_L$ in terms of the well known having smallest values of MSEs, for all the considered values of n, where we have fixed α=2, β=2 and θ=2.

6. Data Analysis

Data set

In this section, we have considered two different real data sets, one from medical field and other from engineering field. In order to summarize the data sets, exploratory data analysis (EDA) has been carried out. To detect the shape of hazard rate function, scaled total-time on-test (TTT) plots has also been plotted. Further, the performance of $GMG_E(\alpha, \beta, \theta)$- distributions has been checked out over some existing distribution in terms of fitting criterions AIC, BIC, and K-S distance.

Data Set-I:- A data of remission times (in months) of 128 bladder cancer patients (see, Kumar et al. [9]). The data is shown below:

X = (0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 1.46, 18.10, 11.79, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25 ,4.50, 6.25, 8.37, 12.02, 2.02, 13.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 12.07, 6.76, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69).

Data Set-II: :- A data of times between successive failures (in thousands of hours) in events of secondary reactor pumps see, suprawhardana et al.[16] and it contains 23 observations, which are

2.160, 0.746, 0.402, 0.954, 0.491, 6.560, 4.992, 0.347, 0.150, 0.358, 0.101, 1.359, 3.465, 1.060, 0.614, 1.921, 4.082, 0.199, 0.605, 0.273, 0.070, 0.062, 5.320.

6.1. Exploratory data analysis (EDA)

For data sets I and II, we have applied basic exploratory data analysis (EDA), which is capable to exact the following informations from the data see,tukey et al.[17] .

• Displaying graphically the data in the form of bar charts, histogram and density curve that shows overall pattern and unusual observations.

• Computing descriptive statistics such as centre, spread, skewness and kurtosis that summarizes specific aspects of the data.

As the result of EDA, the descriptive statistics for data sets I and II are summarized in Tables 4 and 5 respectively and also the box plots, histogram and density curve for these data sets are shown in Figures 6 and 7 respectively.

Table 4: *Descriptive statistics for data set-I.*

Minimum	Q_1	Median	Mean	Q_3	Maximum	Variance	Skewness	Kurtosis
0.080	3.348	6.395	9.366	11.84	79.05	110.42	3.3257	19.153

Table 5: *Descriptive statistics for data set -II.*

Minimum	Q_1	Median	Mean	Q_3	Maximum	Variance	Skewness	Kurtosis
0.0620	0.310	0.614	1.578	2.041	6.560	3.7275	1.3643	3.5445

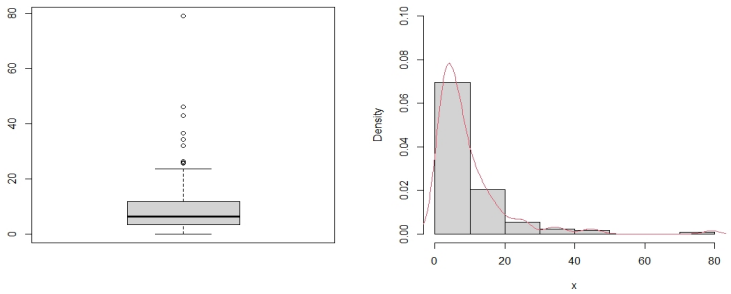


Figure 6: *The box plot, histogram and density plot for the data set I.*

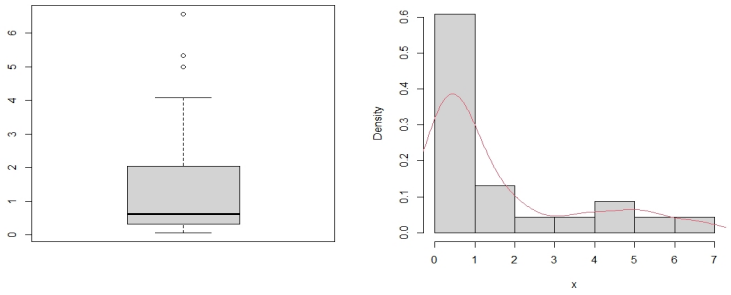


Figure 7: *The boxplot, histogram and density plot for the data set II.*

6.2. Scaled total-time-on-test (TTT) plot

In order to get an idea about the shape of the hazard functions of data sets I and II, we have drawn scaled TTT plots for these data sets, and the plots are shown in Figures 8 and 9 shows scaled TTT plots for data set I and II respectively. These plots shows that Figures 8, shows that the data sets I possesses reverse bathtub and Figures 9, shows that the data sets II possesses bathtub shapes of their hazard rate functions. Therefore, we can suggest to use our proposed distribution $\text{GMG}_E(\alpha, \beta, \theta)$- distribution for these data sets.

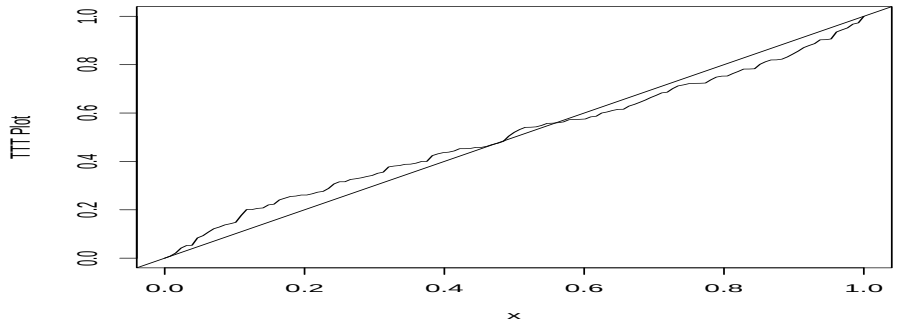


Figure 8: *Scaled total TTT plot for data set-I.*

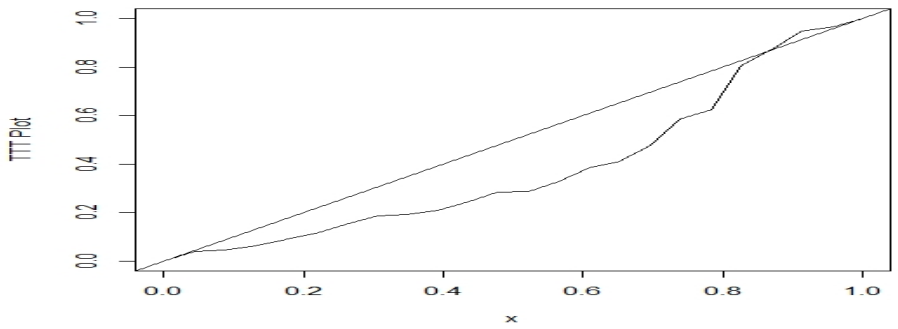


Figure 9: *Scaled total TTT for data set-II.*

6.3. Model Comparison

In this sub-section, we have compared the performance of our proposed distribution $\text{GMG}_E(\alpha, \beta, \theta)$-distribution over some distributions available in literature on the considered data sets I and II. The model selection criterions used are AIC, BIC and K-S distance.

Now, we have computed the maximum likelihood estimates of the parameters α, β and θ for the considered data sets I and II and these are shown in Table 6.

Table 6: *Maximum Likelihood Estimates.*

Parameter	α	β	θ
Data Set-I	0.24796	1.46763	0.03664
Data Set- II	0.84050	1.05509	0.14440

The values of AIC, BIC and K-S distance for the data set I and II for our proposed distribution have been calculated and are shown in Tables 7 and 8 respectively. Also, the values of AIC, BIC

and K-S distance for the data set I for Exponentiated Pareto distributions ($E_P(\theta,\lambda)$) and Beta Pareto distribution (β_P(a,b)) have been extracted from Shawky and Abu Zinadh [18] and are shown in comparative model in Table 7. The values of these criterions for data set II for Reduced Additive Weibull distribution ($RA_W(\alpha,\beta,\lambda)$) and Flexible Weibull Extension ($FW_W(a,b)$) have been extracted from Mustafa et al.[19] are shown in Table 8.

Table 7: *Values of AIC, BIC, and K-S distance for $GMG_E(\alpha,\beta,\theta)$- distribution, $E_P(\theta,\lambda)$- distribution and β_P(a,b)-distribution for data set-I.*

Distributions	AIC	BIC	K-S distance
$GMG_E(\alpha,\beta,\theta)$	**921.05**	**929.60**	**0.234**
$E_P(\theta,\lambda)$	992.20	997.90	0.251
β_P(a,b)	970.70	979.20	0.223

Table 8: *Values of AIC, BIC, and K-S distance for $GMG_E(\alpha,\beta,\theta)$- distribution, $RA_W(\alpha,\beta,\lambda)$ and $FW_W(a,b)$ for data set-II.*

Distributions	AIC	BIC	K-S distance
$GMG_E(\alpha,\beta,\theta)$	**74.10**	**77.50**	**0.132**
$RA_W(\alpha,\beta,\lambda)$	178.14	181.55	0.162
$FW_W(a,b)$	170.68	172.95	0.134

From comparative Tables 7 and Table 8, we observed that the values of AIC, BIC and K-S distance for our proposed dsitribution is least as compared to those values for $E_P(\theta,\lambda)$-distribution and β_P(a,b)- distribution (Table 7) and $RA_W(\alpha,\beta,\lambda)$ -distribution and $FW_W(a,b)$-distribution (Table 8). Therefore, we can say that for the considered data sets I and II, our proposed distribution fits better as compared to other considered distributions in the sense of having smallest values of fitting criterions AIC, BIC and K-S distance.

7. Conclusions

In the present article, we have introduced a new transformation technique which is a generalization of minimum guarantee (MG) transformation proposed by Kumar et al. [9]. Thus the new transformation is called generalized minimum guarantee (GMG) transformation. To demonstrate its real life application, we have considered exponential distribution with the parameter θ as a baseline distribution. The distribution, thus obtained is abbreviated as $GMG_E(\alpha,\beta,\theta)$- distribution. Various statistical properties of the new distribution such as survival function, hazard rate function, cumulative hazard rate function, moments, quantile function and order statistics have been discussed. We have also obtained MLEs and LSEs of the parameters has also been discussed for estimating the parameters α,β and θ. For complete sample from $GMG_E(\alpha,\beta,\theta)$- distribution. Simulation study is carried out to know long-run behaviour of the estimatons and it was found that for almost all the considered values of α,β and θ, the MSEs of MLEs of α,β and θ is smallest as compared to those of MSEs of LSEs of α,β and θ respectively and as n increases, the MSEs are decreases. Finally, to demonstrate its real data application, we have considered two different real data sets and its found that our proposed distribution performs betters as compared to some available distributions in terms of AIC, BIC and K-S distance as fits model selection criterions.

References

[1] Marshall, A. W., and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika, 84(3), 641-652.

[2] Gupta, R. D. and Kundu, D. (2001). Exponentiated exponential family: an alternative to gamma and Weibull distributions. Biometrical Journal. Journal of Mathematical Methods in Biosciences, 43(1), 117-130.
[3] Pinho, L. G. B., Cordeiro, G. M. and Nobre, J. S. (2015). The Harris extended exponential distribution. Communications in Statistics-Theory and methods, 44(16), 3486-3502.
[4] Kumar, D., Singh, U., and Singh, S. K. (2015). A method of proposing new distribution and its application to Bladder cancer patients data. J. Stat. Appl. Pro. Lett, 2(3), 235-245.
[5] Kumar, D., Singh, U., and Singh, S. K. (2015). A new distribution using sine function-its application to bladder cancer patients data. Journal of Statistics Applications and Probability, 4(3), 417.
[6] Nofal, Z. M., Afify, A. Z., Yousof, H. M., Granzotto, D. C. and Louzada, F. (2016). Kumaraswamy transmuted exponentiated additive Weibull distribution. International Journal of Statistics and Probability, 5(2), 78-99.
[7] Kumar, D., Chaurasia, P. K., Kumar, P. and Chaurasia, A. (2022). A Novel Transformation: Based on Inverse Trigonometric Lindley Distribution. Reliability: Theory and Applications, 17(1 (67)), 303-315.
[8] Elbatal, I., Zayed, M., Rasekhi, M. and Butt, N. S. (2017). The exponential Pareto power series distribution: theory and applications. Pakistan Journal of Statistics and Operation Research, 603-615.
[9] Kumar, D., Singh, U. and Singh, U. (2017). Life time distributions: Derived from some minimum guarantee distribution. Sohag Journal of Mathematics, 4(1), 7-11.
[10] Alizadeh, M., Altun, E., Afify, A. Z. and Gamze, O. Z. E. L. (2018). The extended odd Weibull-G family: properties and applications. Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics, 68(1), 161-186.
[11] Mansoor, M., Tahir, M. H., Cordeiro, G. M., Provost, S. B. and Alzaatreh, A. (2019). The Marshall-Olkin logistic-exponential distribution. Communications in Statistics-Theory and Methods, 48(2), 220-234.
[12] Alizadeh, M., Altun, E., Ozel, G., Afshari, M., and Eftekharian, A. (2019). A new odd log-logistic lindley distribution with properties and applications. Sankhya A, 81, 323-346.
[13] Kumar, D., Singh, U., Singh, S. K. and Chaurasia, P. K. (2018). Statistical properties and application of a Lifetime model using sine function. Int. J. Creat. Res. Thoughts, 6, 993-1002.
[14] Nassar, M., Kumar, D., Dey, S., Cordeiro, G. M., and Afify, A. Z. (2019). The Marshall–Olkin alpha power family of distributions with applications. Journal of Computational and Applied Mathematics, 351, 41-53.
[15] Kumar, D., Kumar, P., Singh, S. K., and Singh, U. (2019). A new asymmetric loss function: estimation of parameter of exponential distribution. J Stat Appl Probab Lett, 6(1), 37-50.
[16] M. S. Suprawhardana; S. Prayoto; Total time on test plot analysis for mechanical components of the RSGGAS reactor, Atom Indones, 25(2), 81-90 (1999).
[17] Tukey, J. W. (1977). Exploratory data analysis (Vol. 2, pp. 131-160). Reading, MA: Addison-wesley.
[18] Shawky, A. I., and Abu-Zinadah, H. H. (2009). Exponentiated Pareto distribution: different method of estimations. International Journal of Contemporary Mathematical Sciences, 4(14), 677-693.
[19] Mustafa, A. B. D. E. L. F. A. T. T. A. H., El-Desouky, B. S., and Al-Garash, S. H. A. M. S. A. N. (2016). The exponentiated generalized flexible Weibull extension distribution. Fundamental Journal of Mathematics and Mathematical Sciences, 6(2), 75-98.

RT&A, No 2 (84)
Volume 20, June 2025

PARABOLIC DENSE CLOUDY FUZZY MODEL WITH DEMAND DEPENDENT PRODUCTION RATE IN AN IMPRECISE PRODUCTION PROCESS

Upasana Rana[1], Tanuj Kumar[1,*]

•

[1]Department of Mathematics, SRM Institute of Science and Technology,
Delhi-NCR Campus, Ghaziabad, India, 201204
ur2288@srmist.edu.in, tanujkuo@srmist.edu.in
[*]Corresponding author

Abstract

In this paper we develop a classical economic production lot size model of a manufactured item in a inadequate production process in which the ordering cost is fixed and shortage is not allowed. In this model the valuation of an item is considered as a parabolic cloudy fuzzy number and production rate is depends on demand. In the previous studies, fuzziness of parameters of the model is presented in the form of fuzzy numbers using the membership function only. Parabolic dense cloudy intuitionistic fuzzy numbers are found to be more generalized than fuzzy numbers and have the capability to express the fuzziness of parameters of the model in a more generalized way using both membership and non-membership functions. The presented model is solved in crisp, fuzzy and Intuitionistic fuzzy setup using proposed generalize and defuzzification method with minimum loss of data. The present model is solved by DBPSO algorithm to determine optimal values. The validity of the proposed model and parameter sensitivity analysis is presented to justify the idea by the numerical analysis.

Keywords: Classical Economic Model, cloudy fuzzy number, Parabolic dense cloudy Fuzzy Number, Score function, Defuzzification Method, Swarm Algorithm.

1. Introduction

According to a general scenario, when developing an economic production lot-size model, the demand rate for inventory products is usually assumed to be constant. However, in the actual world, inventory quantities will differ slightly from accurate figures. As a result, in practical settings, the demand variable should be viewed as ambiguous. The fuzzy notion has recently been used to production and inventory difficulties. The fuzzy set theory was first developed by Zadeh [4]. Bellman and Zadeh [5] used the fuzzy set theoretic technique to solve difficulties later. This concept is reflected in the gloomy, hazy surroundings. Following Yager [6], other scholars such as Ezzati, Allahviranloo, Khezerloo, and Khezerloo [7], Deng [8], Zhang, Ignatius, Lim, and Zhao [9], and others supported the method for sorting ambiguous numbers using the centre of gravity. In addition, De and Beg [10], Maity, Suman, et al. [2], Garg [3] and De and Mahata [11], [12] devised new defuzzification methods for triangular dense and triangular cloudy fuzzy sets, respectively. Until recently, no one has tackled a realistic production inventory model with a foggy, ambiguous demand rate. The production rate of a single item or many items is believed to be rigid and set in the traditional economic production lot-size (EPL) paradigm.

In truth, demand for every good has an impact on the manufacturing process. When demand progressively increases, client consumption increases, and manufacturers are forced to produce more things to match the increased need. In the opposite circumstance, the opposite is true. Several academics have developed EPL models for single/multiple items that take into account either a constant or variable production rate (depending on time, demand, and/or the amount of inventory on hand). Bhunia and Maiti [13], Balkhi and Benkherouf [14], Abad [15], Mandal and Maiti [16], Das, Roy, and Kar [18], [19], [20], [21], and others developed inventory models with uniform or variable production rates in mind. Manufacturing flexibility, on the other hand, has become a more significant issue in inventory management. Different types of flexibility have been assessed in the manufacturing system, with volume flexibility being the most relevant. Volume flexibility refers to the ability to adjust production volume in a manufacturing system. In the EPQ model, Cheng [22] was the first to establish the demand-dependent production unit cost, while Khouja [23] included volume flexibility and reliability considerations. Under inflationary conditions and reliability, Shah and Shah [24] established the EPQ model for time falling demand with imperfect production processes. Items are manufactured with a high level of consistency utilising a traditional manufacturing technique. With high expectations, more reliability boosts the efficiency of the production process. Any manufacturing company strives to improve its production efficiency and ability to operate at peak levels by lowering the cost of scraps, rework of poor goods, wasted materials, and worker hours, among other things. Rosenblatt and Lee [25], Ben-Daya and Hariga [26], Goyal, Hung, and Chen [27], Maiti, Bhunia, and Maiti [28], Sana, Goyal, and Chaudhuri [29], Das et al. [19], [20], Manna, Dey, and Mondal [30], Pal, Sana, and Chaudhuri [31], Yadav et al. [40], Yadav et al. [41], [42], [43], Rana and Kumar [45], Negi and Singh [47], Dutta, Anurag, et al. [48] and others have published numerous research papers on imperfect production. Manna, Das, Dey, and Mondal [32] looked at a multi-item EPQ model with learning effect on imperfect production over a fuzzy random planning horizon in a recent paper. Khara, Dey, and Mondal [33] created an inventory model with imperfect production and demand that is based on reliability. Soft computing approaches have been used in a number of research articles on inventory control difficulties. Several authors have used Genetic Algorithms (GA) in various forms to solve marketing difficulties. In an unclear planning horizon, Pal, Maiti, and Maiti [34] use GA to solve an EPQ model with price discounted promotional demand. Bera and Maiti [35] employed GA to solve a discount-based multi-item inventory model. Maiti, Maiti, and Maiti [36] employed GA to solve a stochastic lead time inventory model with price-dependent demand that included advance payment. Many additional scholars have used GA in inventory control problems, including Mondal and Maiti [37], Maiti and Maiti [38], [39], Jiang, Xu, Wang, and Wang [44], and Maiti et al. [46]. PSO was also employed by Bhunia and Shaikh [49] to solve a two-warehouse inventory model for decaying items with an allowed payment delay. To solve this fuzzy inventory model, dominance-based particle swarm optimization was created. The fuzzy inventory model is defined here under an imperfect production process with a hazy fuzzy demand rate, with the production rate being demand dependent. The model is solved using Yager's index approach and De and Beg's ranking index method for defuzzification in a crisp, general fuzzy, and cloudy fuzzy environment, and the outcomes in the crisp, fuzzy, and cloudy fuzzy environments are compared.

The goal of this research is to use a dominance-based Particle Swarm Optimization (PSO) algorithm to identify the best decision for the decision maker by minimising average total cost to acquire the best order quantity and cycle time (DM). A number of numerical examples are used to support the model, as well as some sensitivity studies.

2. Definitions and preliminaries

2.1. Normalized General Triangular Cloudy Fuzzy Number (NGTCFN)

A Normalized General Triangular Cloudy Fuzzy Number (NGTCFN) having the form $(\widetilde{A}) = \left(a\left(1-\frac{\rho_1}{1+t}\right), a, a\left(1+\frac{\sigma_1}{1+t}\right)\right)$ based on two parameters ρ_1 and σ_1 where $a\left(1-\frac{\rho_1}{1+t}\right) < a < a\left(1+\frac{\sigma_1}{1+t}\right)$

and is characterized by its continuous membership function(MF) $\zeta_{\tilde{A}}(x) : X \to [0,1]$, defined as

$$\zeta_{\tilde{A}}(x) = \begin{cases} \frac{x-a\left(1-\frac{\rho_1}{1+t}\right)}{a-a\left(1-\frac{\rho_1}{1+t}\right)}, & a\left(1-\frac{\rho_1}{1+t}\right) \le x \le a \\ \\ \frac{a\left(1+\frac{\sigma_1}{1+t}\right)-x}{a\left(1+\frac{\sigma_1}{1+t}\right)-a}, & a \le x \le a\left(1+\frac{\sigma_1}{1+t}\right) \\ \\ 0, \quad \textit{Otherwise} \end{cases}$$

$$\eta_{\tilde{A}}(x) = \begin{cases} \frac{(a-x)+\left(x-a\left(1-\frac{\rho_2}{1+t}\right)\right)}{a-a\left(1-\frac{\rho_1}{1+t}\right)}, & a\left(1-\frac{\rho_2}{1+t}\right) \le x \le a \\ \\ \frac{(x-a)+\left(a\left(1+\frac{\sigma_2}{1+t}\right)-x\right)}{a\left(1+\frac{\sigma_2}{1+t}\right)-a}, & a \le x \le a\left(1+\frac{\sigma_2}{1+t}\right) \\ \\ 0, \quad \textit{Otherwise} \end{cases}$$

with $0 \le \zeta, \eta \le 1$ and $0 \le \zeta + \eta \le 1$.

2.2. Generalized Parabolic Fuzzy Numbers (GPFN)

A Generalized Parabolic Fuzzy number(GPFN) $\tilde{A} = (a_l, a_m, a_u)$ is called generalized parabolic fuzzy number if its membership function(MF) is defined as

$$\zeta_{\tilde{A}}(x) = \begin{cases} \left(\frac{x-a_l}{a_m-a_l}\right)^2 \theta, & if \;\; a_l \le x \le a_m \\ \\ \theta, & x = a_m \\ \\ \left(\frac{a_u-x}{a_u-a_m}\right)^2 \theta, & if \;\; a_m \le x \le a_u \end{cases}$$

2.3. Generalized Parabolic Dense Cloudy Fuzzy Number (GPDCFN)

A Generalized Parabolic Dense Cloudy Fuzzy Number (GPDCFN) $\tilde{A} = \langle [a\left(1-\frac{\rho_1}{1+t}\right), a, a\left(1+\frac{\sigma_1}{1+t}\right); \zeta], [a\left(1-\frac{\rho_2}{1+t}\right), a, a\left(1+\frac{\sigma_2}{1+t}\right); \eta] \rangle$ by adding membership function(MF) and non-membership function(NMF) is called generalized parabolic dense cloudy fuzzy number if its membership function(MF) and non-membership function(NMF) is defined as

$$\zeta_{\tilde{A}}(x) = \begin{cases} \left(\frac{x-a\left(1-\frac{\rho_1}{1+t}\right)}{a-a\left(1-\frac{\rho_1}{1+t}\right)}\right)^2 \theta, & if \;\; a\left(1-\frac{\rho_1}{1+t}\right) \le x \le a \\ \\ \theta, & x = a \\ \\ \left(\frac{a\left(1+\frac{\sigma_1}{1+t}\right)-x}{a\left(1+\frac{\sigma_1}{1+t}\right)-a}\right)^2 \theta & if \;\; a \le x \le a\left(1+\frac{\sigma_1}{1+t}\right) \end{cases}$$

$$\eta_{\tilde{A}}(x) = \begin{cases} \left(\frac{(a-x)+\sqrt{\eta}\left(x-a\left(1-\frac{\rho_2}{1+t}\right)\right)}{(a-a\left(1-\frac{\rho_2}{1+t}\right))}\right)^2, & if \;\; a\left(1-\frac{\rho_2}{1+t}\right) \le x \le a \\ \eta, & x = a \\ \left(\frac{(x-a)+\sqrt{\eta}\left(a\left(1+\frac{\sigma_2}{1+t}\right)-x\right)}{(a\left(1+\frac{\sigma_2}{1+t}\right)-a)}\right)^2, & if \;\; a \le x \le a\left(1+\frac{\sigma_2}{1+t}\right) \end{cases}$$

Where $a\left(1-\frac{\rho_2}{1+t}\right) \le a\left(1-\frac{\rho_1}{1+t}\right) \le a \le a\left(1+\frac{\sigma_1}{1+t}\right) \le a\left(1+\frac{\sigma_2}{1+t}\right)$, $0 \le \zeta, \eta \le 1$ and $0 \le \zeta + \eta \le 1$.

2.4. (α, β) - cut of a cloudy fuzzy number

Let $\alpha, \beta \in [0,1]$ be a fixed numbers such that $0 \le \alpha \le \zeta$, $\eta \le \beta \le 1$ and $0 \le \sqrt{\alpha} + \sqrt{\beta} \le 1$.
A (α, β) cuts generated by Generalized Parabolic Dense Cloudy Fuzzy Number (GPDCFN) A is defined as

$$A_{\alpha,\beta} = \{\langle x, \zeta_{\tilde{A}}(x) \ge \alpha, \eta_{\tilde{A}}(x) \le \beta\rangle \quad : x \in X\}$$

The α cut of GPDCFN $\tilde{A}$ is defined as

$$\tilde{A}^{\zeta}_{\alpha} = \{\langle x, \zeta_{\tilde{A}}(x) \ge \alpha\rangle \quad : x \in X\}, \;\; 0 \le \alpha \le \zeta$$

The β cut of GPDCFN $\tilde{A}$ is defined as

$$\tilde{A}^{\eta}_{\beta} = \{\langle x, \eta_{\tilde{A}}(x) \le \beta\rangle \quad : x \in X\}, \;\; \eta \le \beta \le 1$$

2.5. Arithmetic Operations on Generalized Triangular Parabolic Dense Cloudy Fuzzy Numbers (GTPDCFN)

$$\mu_{A+B}(x) = \langle (a_1 + a_2,\; b_1 + b_2,\; c_1 + c_2); \mu_1 \rangle$$
$$\nu_{A+B}(x) = \langle (a_1' + a_2',\; b_1' + b_2',\; c_1' + c_2'); \nu_2 \rangle$$
$$\mu_{A-B}(x) = \langle (a_1 - c_2,\; b_1 - b_2,\; c_1 - a_2); \mu_1 \rangle$$
$$\nu_{A-B}(x) = \langle (a_1' - c_2',\; b_1' - b_2',\; c_1' - a_2'); \nu_2 \rangle$$
$$\mu_{A \cdot B}(x) = \langle (a_1 \cdot a_2,\; b_1 \cdot b_2,\; c_1 \cdot c_2); \mu_1 \rangle$$
$$\nu_{A \cdot B}(x) = \langle (a_1' \cdot a_2',\; b_1' \cdot b_2',\; c_1' \cdot c_2'); \nu_2 \rangle$$
$$\mu_{\frac{A}{B}}(x) = \left\langle (\frac{a_1}{c_2},\; \frac{b_1}{b_2},\; \frac{c_1}{a_2}); \mu_1 \right\rangle$$
$$\nu_{\frac{A}{B}}(x) = \left\langle (\frac{a_1'}{c_2'},\; \frac{b_1'}{b_2'},\; \frac{c_1'}{a_2'}); \nu_2 \right\rangle$$

3. Defuzzification of parabolic dense cloudy fuzzy number

Let $\tilde{A} = \langle [a\left(1-\frac{\rho_1}{1+t}\right), a, a\left(1+\frac{\sigma_1}{1+t}\right); \mu_1],\; [a\left(1-\frac{\rho_2}{1+t}\right), a, a\left(1+\frac{\sigma_2}{1+t}\right); \nu_1] \rangle$ be a generalized triangular parabolic dense cloudy fuzzy number with the following membership(MF) and non-membership functions(NMF).

$$\zeta_{\tilde{A}}(x) = \begin{cases} \zeta\left(\frac{x-a\left(1-\frac{\rho_1}{1+t}\right)}{a-a\left(1-\frac{\rho_1}{1+t}\right)}\right)^2, & if \;\; a\left(1-\frac{\rho_1}{1+t}\right) \le x \le a \\ \zeta, & x = a \\ \zeta\left(\frac{a\left(1+\frac{\sigma_1}{1+t}\right)-x}{a\left(1+\frac{\sigma_1}{1+t}\right)-a}\right)^2, & if \;\; a \le x \le a\left(1+\frac{\sigma_1}{1+t}\right) \end{cases}$$

$$\eta_{\tilde{A}}(x) = \begin{cases} \left(\frac{(a-x)+\sqrt{\eta}(x-a\left(1-\frac{\rho_2}{1+t}\right)}{(a-a\left(1-\frac{\rho_2}{1+t}\right)}\right)^2, & if \;\; a\left(1-\frac{\rho_2}{1+t}\right) \le x \le a \\ \eta, & x = a \\ \left(\frac{(x-a)+\sqrt{\eta}(a\left(1+\frac{\sigma_2}{1+t}\right)-x)}{(a\left(1+\frac{\sigma_2}{1+t}\right)-a)}\right)^2, & if \;\; a \le x \le a\left(1+\frac{\sigma_2}{1+t}\right) \end{cases}$$

respectively. The score functions is defined as the difference between membership and non-membership functions.

$$S(\tilde{A}) = \zeta(\tilde{A}) - \eta(\tilde{A})$$

$$S(\tilde{A}) = \begin{cases} \phi(x-\Psi); & \Psi \le x \le a \\ \sqrt{\zeta}-\sqrt{\eta}; & x = a \\ \rho(\tau - x); & a \le x \le \tau \end{cases}$$

$$\phi = \frac{\sqrt{\mu}(a-a\left(1-\frac{\rho_1}{1+t}\right)) + (a-a\left(1-\frac{\rho_1}{1+t}\right))(1-\sqrt{\nu})}{(a-a\left(1-\frac{\rho_1}{1+t}\right))(a-a\left(1-\frac{\rho_2}{1+t}\right))}$$

$$\Psi = \frac{\sqrt{\mu}a\left(1-\frac{\rho_1}{1+t}\right)(a-a\left(1-\frac{\rho_2}{1+t}\right)) + (a-a\left(1-\frac{\rho_1}{1+t}\right))(a-\sqrt{\nu}a\left(1-\frac{\rho_2}{1+t}\right))}{\sqrt{\mu}(a-a\left(1-\frac{\rho_2}{1+t}\right)) + (a-a\left(1-\frac{\rho_1}{1+t}\right))(1-\sqrt{\nu})}$$

$$\rho = \frac{\sqrt{\mu}(a\left(1+\frac{\sigma_2}{1+t}\right)-a) + (a\left(1+\frac{\sigma_1}{1+t}\right)-a)(1-\sqrt{\nu})}{(a\left(1+\frac{\sigma_1}{1+t}\right)-a)(a\left(1+\frac{\sigma_2}{1+t}\right)-a)}$$

$$\tau = \frac{\sqrt{\mu}a\left(1+\frac{\sigma_1}{1+t}\right)(a\left(1+\frac{\sigma_2}{1+t}\right)-a) + (a\left(1+\frac{\sigma_1}{1+t}\right)-a)(a-\sqrt{\nu}(a\left(1+\frac{\sigma_2}{1+t}\right))}{\sqrt{\mu}(a\left(1+\frac{\sigma_2}{1+t}\right)-a) + (a\left(1+\frac{\sigma_1}{1+t}\right)-a)(1-\sqrt{\nu})}$$

with $0 \le \zeta,\ \eta \le 1$ and $0 \le \zeta + \eta \le 1$. The α cut of $\bar{S} = \{\langle(\psi, a, \tau);\ \mu, \nu\rangle$ is defined as

$$\widetilde{A}(\alpha) = \left[\widetilde{A}_L(\alpha), \widetilde{A}_R(\alpha)\right] = \left[\psi + \frac{\alpha}{\phi}, \tau - \frac{\alpha}{\rho}\right]$$

The defuzzification formula for GPDCFN is defined as follows

$$D_f(\widetilde{A}) = \int_0^1 \Big(\lambda\ L^{-1}(\alpha) + (1-\lambda)R^{-1}(\alpha)\Big)d\alpha$$

$$D_f(\widetilde{A}) = \left[\frac{\lambda}{\phi} + \frac{(1-\lambda)}{\rho}\right]\left(\frac{\sqrt{\zeta}-\sqrt{\eta}}{2}\right)^2 + \left[\lambda\psi + (1-\lambda)\tau\right](\sqrt{\zeta}-\sqrt{\eta}).$$

4. Notations and assumptions

4.1. Notations

The mathematical model was formulated based on the following notations.

K : Production rate per cycle.

D : Demand rate per cycle, $(D < K)$.

R : Production process reliability.

$P(T)$: Instantaneous inventory level.

P : Maximum inventory level (decision variable).

T_2 : Cycle length (decision variable).

T_1 : Production period (decision variable).

C : Production cost per unit.

C_1 : Setup cost per cycle.

H : Inventory carrying cost per unit quantity per unit time.

Z : Average total inventory cost.

P^* : Optimum value of P.

T_2^* : Optimum value of T_2.

Z^* : Optimum value of Z.

T_1^* : Optimum value of T_1.

4.2. Assumptions

The mathematical model was formulated based on the following assumptions.

1. Replenishment occurs instantaneously on placing of order quantity, so lead time is zero.
2. The inventory is developed for a single item in an imperfect process.
3. Shortage are not allowed.
4. The time horizon of the inventory system is infinite.
5. The production rate K is demand dependent and is of the form

$$K = A + B \cdot D,$$

where A and B are positive constants.

5. Model development and analysis

The process reliability r means that among the items product in a production, only $R\%$ is acceptable that can be used to meet the customer's demand. Initially the production process starts with zero inventories with production rate K and demand rate D. Figure 1 illustrates the variation between inventory levels and demand rate for the proposed model.

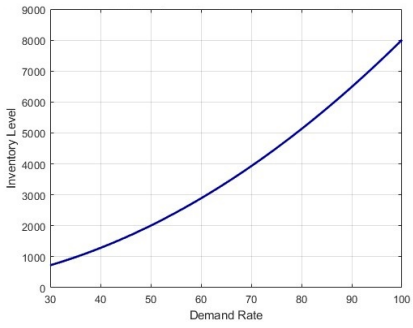


Figure 1: *Variation between inventory level and demand rate*

During the interval $[0, T_1]$ inventory level gradually built up at rate $(RK - D)$ and reached at its maximum level P at the end of production process. The inventory level gradually depleted during the period $[T_1, T_2]$ due to customer's demand and ultimately became zero at $[T_1, T_2]$. The instantaneous state of $P(t)$ desuring the differential equation in the interval $[0, T_2]$ of that item is given by

$$\frac{dP(T)}{dT} = \begin{cases} RK - D, & 0 \le T \le T_1 \\ \\ -D, & T_1 \le T \le T_2 \end{cases} \tag{1}$$

where $RK - D > 0$. With boundary condition $P(0) = 0,\ P(T_1) = P,\ P(T_2) = 0$. The solution of the differential equation using the boundary condition is given by

$$P(T) = \begin{cases} (RK - D)T, & 0 \le T \le T_1 \\ \\ D(T_2 - T), & T_1 \le T \le T_2 \end{cases} \tag{2}$$

Hence our problem is given by

$$Minimizing\ \ Z = \frac{CD}{R} + \frac{C_1}{T_2} + \frac{HDT_2(AR + (BR-1)D)}{2(A + BD)R}, \cdots\cdots (A)$$

$$Subject\ to\ \ d(T_2 - T_1) = (RK - D)T_1,$$

$$i.e.,\ \ RKT_1 = DT_2$$

$$P = D(T_2 - T_1).$$

Now, the problem is reduced to minimize the average cost Z and to find the optimum value of P and T_2 for which $Z(P, T_2)$ is minimum and the corresponding value of T_1. The average cost is minimized by DBPSO.

6. Parabolic Cloudy Fuzzy mathematical model

Initially, when the production process starts, the demand rate R item is ambiguous. Naturally, demand rate is assumed to be general fuzzy over the cycle length. Then fuzzy demand rate $\tilde{D}$ is as follows, $\tilde{D} = \left(D\left(1 - \frac{\rho_2}{1+t}\right), D\left(1 - \frac{\rho_1}{1+t}\right), D, D\left(1 + \frac{\sigma_1}{1+t}\right), D\left(1 + \frac{\sigma_2}{1+t}\right)\right)$ for GTPDFN. Therefore, Maiti [1] transform the problem (A) to a fuzzy problem as folows.

$$\begin{cases} Minimizing\ \tilde{Z} \ = \frac{C\tilde{D}}{R} + \frac{C_1}{T_2} + \frac{H\tilde{D}T_2(AR+(BR-1)\tilde{D})}{2(A+B\tilde{D})R}, \\ \\ Subject\ to\ \ R\tilde{K}T_1 = \tilde{D}T_2, \\ \\ \qquad \tilde{P} = \tilde{D}(T_2 - T_1). \end{cases}$$

$$\mu_1(Z) = \begin{cases} \left(\frac{Z - Z_1}{Z_2 - Z_1}\right)^2 \mu, & Z_1 \le Z \le Z_2 \\ \\ \mu, & Z = Z_2 \\ \\ \left(\frac{Z_3 - Z}{Z_3 - Z_2}\right)^2 \mu, & Z_2 \le Z \le Z_3 \end{cases}$$

$$\nu_1(Z) = \begin{cases} \left(\frac{(Z_2-Z)+(Z-Z_1')\sqrt{\nu}}{Z_2-Z_1'}\right)^2, & Z_1' \le Z \le Z_2 \\ \nu, & Z = Z_2 \\ \left(\frac{(Z-Z_2)+(Z_3'-Z)\sqrt{\nu}}{Z_3-Z_2'}\right)^2, & Z_2 \le Z \le Z_3' \end{cases}$$

If we take parabolic dense intuitionistic fuzzy number $\left(D(1-\frac{\rho_2}{1+t}), D(1-\frac{\rho_1}{1+t}), D, D(1+\frac{\sigma_1}{1+t}), D(1+\frac{\sigma_2}{1+t})\right)$.

$$Z_1' = \frac{CD(1-\frac{\rho_2}{1+t})}{R} + \frac{C_1}{T_2} + \frac{HT_2D(1-\frac{\rho_2}{1+t})(AR+(BR-1)D(1-\frac{\rho_2}{1+t}))}{2R(A+BD(1+\frac{\sigma_2}{1+t}))}$$

$$Z_1 = \frac{CD(1-\frac{\rho_1}{1+t})}{R} + \frac{C_1}{T_2} + \frac{HT_2D(1-\frac{\rho_1}{1+t})(AR+(BR-1)D(1-\frac{\rho_1}{1+t}))}{2R(A+BD(1+\frac{\sigma_1}{1+t}))}$$

$$Z_2 = \frac{CD}{R} + \frac{C_1}{T_2} + \frac{HT_2D(AR+(BR-1)D)}{2R(A+BD)}$$

$$Z_3 = \frac{CD(1+\frac{\sigma_1}{1+t})}{R} + \frac{C_1}{T_2} + \frac{HT_2D(1+\frac{\sigma_1}{1+t})(AR+(BR-1)D(1+\frac{\sigma_1}{1+t}))}{2R(A+BD(1-\frac{\rho_1}{1+t}))}$$

$$Z_3' = \frac{CD(1+\frac{\sigma_2}{1+t})}{R} + \frac{C_1}{T_2} + \frac{HT_2D(1+\frac{\sigma_2}{1+t})(AR+(BR-1)D(1+\frac{\sigma_2}{1+t}))}{2R(A+BD(1-\frac{\rho_2}{1+t}))}$$

$$D_f(Z) = \left[\frac{\lambda}{\phi} + \frac{(1-\lambda)}{\rho}\right]\left(\frac{\sqrt{\mu}-\sqrt{\nu}}{2}\right)^2 + \left[\lambda\psi + (1-\lambda)\tau\right](\sqrt{\mu}-\sqrt{\nu}).$$

$$\phi = \frac{\sqrt{\mu}(Z_2-Z_1') + (Z_2-Z_1)(1-\sqrt{\nu})}{(Z_2-Z_1)(Z_2-Z_1')}, \quad \Psi = \frac{\sqrt{\mu}Z_1(Z_2-Z_1') + (Z_2-Z_1)(Z_2-\sqrt{\nu}Z_1')}{\sqrt{\mu}(Z_2-Z_1') + (Z_2-Z_1)(1-\sqrt{\nu})}$$

$$\rho = \frac{\sqrt{\mu}(Z_3'-Z_2) + (Z_3-Z_2)(1-\sqrt{\nu})}{(Z_3'-Z_2)(Z_3-Z_2)}, \quad \tau = \frac{\sqrt{\mu}Z_3(Z_3'-Z_2) + (Z_3-Z_2)(Z_2-\sqrt{\nu}Z_3')}{\sqrt{\mu}(Z_3'-Z_2) + (Z_3-Z_2)(1-\sqrt{\nu})}$$

$$\mu_2(P) = \begin{cases} \left(\frac{P-P_1}{P_2-P_1}\right)^2\mu, & P_1 \le P \le P_2 \\ \mu, & P = P_2 \\ \left(\frac{P_3-P}{P_3-P_2}\right)^2\mu, & P_2 \le P \le P_3 \end{cases}$$

$$\nu_2(P) = \begin{cases} \left(\frac{(P_2-P)+(P-P_1')\sqrt{\nu}}{P_2-P_1'}\right)^2, & P_1' \le P \le P_2 \\ \nu, & P = P_2 \\ \left(\frac{(P-P_2)+(P_3'-P)\sqrt{\nu}}{P_3-P_2'}\right)^2, & P_2 \le z \le P_3' \end{cases}$$

$$P_1' = D\big(1-\frac{\rho_2}{1+t}\big)(T_2-T_1),\ \ P_1 = D\big(1-\frac{\rho_1}{1+t}\big)(T_2-T_1)$$

$$P_2 = D(T_2-T_1),\ \ P_3 = D\big(1+\frac{\sigma_1}{1+t}\big)(T_2-T_1),\ \ P_3' = D\big(1+\frac{\sigma_2}{1+t}\big)(T_2-T_1)$$

$$D_f(P) = \left[\frac{\lambda}{\phi}+\frac{(1-\lambda)}{\rho}\right]\left(\frac{\sqrt{\mu}-\sqrt{\nu}}{2}\right)^2+\Big[\lambda\psi+(1-\lambda)\tau\Big](\sqrt{\mu}-\sqrt{\nu}).$$

$$\phi = \frac{\sqrt{\mu}(P_2-P_1')+(P_2-P_1)(1-\sqrt{\nu})}{(P_2-P_1)(P_2-P_1')},\ \ \Psi = \frac{\sqrt{\mu}P_1(P_2-P_1')+(P_2-P_1)(P_2-\sqrt{\nu}P_1')}{\sqrt{\mu}(P_2-P_1')+(P_2-P_1)(1-\sqrt{\nu})}$$

$$\rho = \frac{\sqrt{\mu}(P_3'-P_2)+(P_3-P_2)(1-\sqrt{\nu})}{(P_3'-P_2)(P_3-P_2)},\ \ \tau = \frac{\sqrt{\mu}P_3(P_3'-P_2)+(P_3-P_2)(P_2-\sqrt{\nu}P_3')}{\sqrt{\mu}(P_3'-P_2)+(P_3-P_2)(1-\sqrt{\nu})}$$

$$\mu_3(K) = \begin{cases} \left(\frac{K-K_1}{K_2-K_1}\right)^2\mu, & K_1' \le K \le K_2 \\ \mu, & K = K_2 \\ \left(\frac{K_3-K}{K_3-K_2}\right)^2\mu, & K_2 \le K \le K_3' \end{cases}$$

$$\nu_3(K) = \begin{cases} \left(\frac{(K_2-K)+(K-K_1')\sqrt{\nu}}{K_2-K_1'}\right)^2, & K_1' \le K \le K_2 \\ \nu, & K = K_2 \\ \left(\frac{(K-K_2)+(K_3'-K)\sqrt{\nu}}{K_3-K_2'}\right)^2, & K_2 \le K \le K_3' \end{cases}$$

$$RK_1'T_1 = D\big(1-\frac{\rho_2}{1+t}\big)T_2,\ \ RK_1T_1 = D\big(1-\frac{\rho_1}{1+t}\big)T_2$$

$$RK_2T_1 = DT_2,\ \ RK_3T_1 = D\big(1+\frac{\sigma_1}{1+t}\big)T_2,\ \ RK_3'T_1 = D\big(1+\frac{\sigma_2}{1+t}\big)T_2$$

$$D_f(K) = \left[\frac{\lambda}{\phi}+\frac{(1-\lambda)}{\rho}\right]\left(\frac{\sqrt{\mu}-\sqrt{\nu}}{2}\right)^2+\Big[\lambda\psi+(1-\lambda)\tau\Big](\sqrt{\mu}-\sqrt{\nu}).$$

$$\phi = \frac{\sqrt{\mu}(K_2-K_1')+(K_2-K_1)(1-\sqrt{\nu})}{(K_2-K_1)(K_2-K_1')},\ \ \Psi = \frac{\sqrt{\mu}K_1(K_2-K_1')+(K_2-K_1)(K_2-\sqrt{\nu}K_1')}{\sqrt{\mu}(K_2-K_1')+(K_2-K_1)(1-\sqrt{\nu})}$$

$$\rho = \frac{\sqrt{\mu}(K_3'-K_2)+(K_3-K_2)(1-\sqrt{\nu})}{(K_3'-K_2)(K_3-K_2)},\ \ \tau = \frac{\sqrt{\mu}K_3(K_3'-K_2)+(K_3-K_2)(K_2-\sqrt{\nu}K_3')}{\sqrt{\mu}(K_3'-K_2)+(K_3-K_2)(1-\sqrt{\nu})}$$

7. Numerical Example

Now we consider the following values of parameters in inventory models:

1. **Crisp model:** Let us consider $A = 850\,(\$)$ per cycle, $C_b = 1.5\,(\$)$ per unit, $p = 15\,(\$)$ per unit, $C_h = 5\,(\$)$. Per unit, $D = 500 units$. We get the optimal solution of our model as: the optimum average inventory cost is $X^* = 8501.92(\$)$ with shortage quantity $S^* = 813$ $units$ and ordering quantity $Q^* = 1045$ $units$ for the cycle time $T^* = 2.9$ $(weeks)$ and the shortages begin from 0.47 $(weeks)$, as Table 1. Figure 2 illustrates the graphical representation of convexity. This visual depiction highlights the key aspects of convexity.
2. **Cloudy intuitionistic fuzzy model:** Let us consider $\omega = 0.6$, $v = 0.2$, $\delta = 0.15$ $\epsilon = 0.12, \rho = 0.05$, and the other parameters will be same as the crisp model, as Table 1. Figure 2 provides a graphical representation of convexity, visually demonstrating its key features. This illustration helps to clarify the concept of convexity.

Table 1: *Optimum values of Crisp and Cloudy intuitionistic fuzzy model*

Model	$T^*(weeks)$	$(t_1)^*(weeks)$	$S^*(units)$	$Q^*(units)$	$X^*(\$)$
Crisp model	2.09	0.47	809	1045	13313.09
Cloudy intuitionistic fuzzy model	4.47	1.01	845.29	1009.69	14535.21

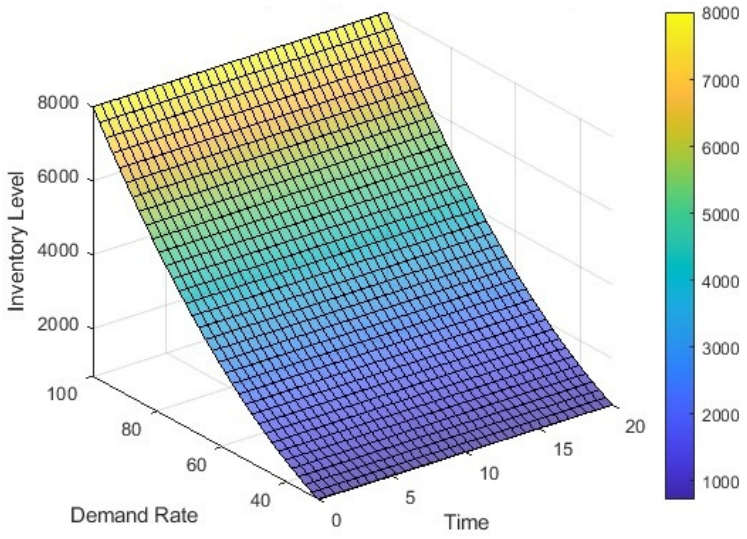


Figure 2: *Convexity with respect to D and T*

8. Conclusion and Future Directions

This study examined the impact of product reliability and time-dependent demand on an inventory model for perishable items within a pentagonal-fuzzy framework. Using the Graded Mean Integration Representation (GMIR) method for defuzzification, the research delivered an in-depth analysis of the Total Inventory Cost function and its optimal values. The findings show that integrating product reliability and time-dependent demand into the model has a substantial effect on inventory management strategies and cost optimization for perishable products. The numerical examples and sensitivity analysis provide important insights into how changes in parameters

influence optimal inventory decisions, thereby increasing the model's practical relevance. This optimization technique is particularly relevant in the food and beverage industry, where managing perishable inventory is crucial to cutting waste, ensuring product quality, and meeting shifting consumer demand. It can also be applied in the pharmaceutical sector, where product shelf life and dependability have a direct bearing on both patient safety and regulatory compliance.

Future research could build on this study by investigating several new areas. For instance, including additional factors like multi-echelon supply chains or different types of perishable products could offer a deeper understanding of inventory dynamics. Applying various defuzzification methods or advanced optimization algorithms might provide more accurate results and insights. Moreover, validating the model with real-world data could strengthen its robustness and applicability. Lastly, adapting the model to consider external factors such as market fluctuations or regulatory changes could increase its relevance and applicability across different contexts.

References

[1] Maiti, Ajoy Kumar. "Cloudy fuzzy inventory model under imperfect production process with demand dependent production rate." Journal of Management Analytics 8.4 (2021): 741-763.
[2] Maity, Suman, et al. "A Study of an EOQ Model of Growing Items with Parabolic Dense Fuzzy Lock Demand Rate." Applied System Innovation 4.4 (2021): 81.
[3] Garg, Harish. "Arithmetic operations on generalized parabolic fuzzy numbers and its application." Proceedings of the national academy of sciences, India section A: Physical sciences 88.1 (2018): 15-26.
[4] Zadeh, Lotfi A. "Fuzzy sets." Information and Control (1965).
[5] Bellman, Richard E., and Lotfi Asker Zadeh. "Decision-making in a fuzzy environment." Management science 17.4 (1970): B-141.
[6] Yager, Ronald R. "A procedure for ordering fuzzy subsets of the unit interval." Information sciences 24.2 (1981): 143-161.
[7] Ezzati, Reza, et al. "An approach for ranking of fuzzy numbers." Expert Systems with applications 39.1 (2012): 690-695.
[8] Deng, Jun, et al. "Cenozoic tectono-magmatic and metallogenic processes in the Sanjiang region, southwestern China." Earth-Science Reviews 138 (2014): 268-299.
[9] Zhang, Feng, et al. "A new method for ranking fuzzy numbers and its application to group decision making." Applied mathematical modelling 38.4 (2014): 1563-1582.
[10] De, Sujit Kumar, and Ismat Beg. "Triangular dense fuzzy sets and new defuzzification methods." Journal of Intelligent & Fuzzy systems 31.1 (2016): 469-477.
[11] De, Sujit Kumar, and Gour Chandra Mahata. "Decision of a fuzzy inventory with fuzzy backorder model under cloudy fuzzy demand rate." International Journal of Applied and computational mathematics 3 (2017): 2593-2609.
[12] De, Sujit Kumar, and Gour Chandra Mahata. "A cloudy fuzzy economic order quantity model for imperfect-quality items with allowable proportionate discounts." Journal of Industrial Engineering International 15.4 (2019): 571-583.
[13] Bhunia, A. K., and M. Maiti. "Deterministic inventory models for variable production." Journal of the Operational Research Society 48.2 (1997): 221-224.
[14] Balkhi, Z. T. "On the global optimality of a general deterministic production lot size inventory model for detoriating items." JORBEL-Belgian Journal of Operations Research, Statistics, and Computer Science 38.4 (1998): 33-44.
[15] Abad, Cristina, et al. "An evaluation of the value relevance of consolidated versus unconsolidated accounting information: evidence from quoted Spanish firms." Journal of International Financial Management & Accounting 11.3 (2000): 156-177.
[16] Mandal, M., and M. Maiti. "INVENTORY OF DAMAGABLE ITEMS WITH VARIABLE REPLENISHMENT AND STOCK-DEPENDENT DEMAND." Asia-Pacific Journal of Operational Research 17.1 (2000).

[17] Saha, Anirban, et al. "Inventory models for breakable items with stock dependent demand and imprecise constraints." Mathematical and Computer Modelling 52.9-10 (2010): 1771-1782.
[18] Das, Debasis, Arindam Roy, and Samarjit Kar. "A ProductionInventory Model for a Deteriorating Item Incorporating Learning Effect Using Genetic Algorithm." Advances in Operations Research 2010.1 (2010): 146042.
[19] Das, Debasis, Arindam Roy, and Samarjit Kar. "A volume flexible economic production lot-sizing problem with imperfect quality and random machine failure in fuzzy-stochastic environment." Computers & Mathematics with Applications 61.9 (2011): 2388-2400.
[20] Das, Debasis, Arindam Roy, and Samarjit Kar. "Optimal payment time for a retailer under permitted delay of payment by the wholesaler with dynamic demand and hybrid number cost parameters." Opsearch 48 (2011): 171-196.
[21] Das, Debasis, Arindam Roy, and Samarjit Kar. "A multi-item inventory model for two-stage production system with imperfect processes using differential evolution and credibility measure." International Journal of Operations Research 9.2 (2012): 87-99.
[22] Cheng, T. C. E. "An economic order quantity model with demand-dependent unit cost." European Journal of Operational Research 40.2 (1989): 252-256.
[23] Khouja, Moutaz. "The use of data envelopment analysis for technology selection." Computers & Industrial Engineering 28.1 (1995): 123-132.
[24] Shah, Nita H., and Bhavin J. Shah. "EPQ model for time-declining demand with imperfect production process under inflationary conditions and reliability." International Journal of Operations Research 11.3 (2014): 91-99.
[25] Rosenblatt, Meir J., and Hau L. Lee. "Economic production cycles with imperfect production processes." IIE transactions 18.1 (1986): 48-55.
[26] Ben-Daya, Mohamed, and Moncer Hariga. "Economic lot scheduling problem with imperfect production processes." Journal of the Operational Research Society 51.7 (2000): 875-881.
[27] Goyal, S. K., Hung, J. H., & Chen, L. H. (2003). An imperfect production model with variable production rate and inspection. European Journal of Operational Research, 147(3), 568-578.
[28] Maiti, Ajoy Kumar, Asoke Kumar Bhunia, and Manoranjan Maiti. "An application of real-coded genetic algorithm (RCGA) for mixed integer non-linear programming in two-storage multi-item inventory model with discount policy." Applied Mathematics and computation 183.2 (2006): 903-915.
[29] Sana, Shib Sankar, Suresh Kumar Goyal, and Kripasindhu Chaudhuri. "An imperfect production process in a volume flexible inventory model." International Journal of Production Economics 105.2 (2007): 548-559.
[30] Manna, Amalesh Kumar, Jayanta Kumar Dey, and Shyamal Kumar Mondal. "Three-layer supply chain in an imperfect production inventory model with two storage facilities under fuzzy rough environment." Journal of Uncertainty Analysis and Applications 2 (2014): 1-31.
[31] Pal, Brojeswar, Shib Sankar Sana, and Kripasindhu Chaudhuri. "Two-echelon competitive integrated supply chain model with price and credit period dependent demand." International Journal of Systems Science 47.5 (2016): 995-1007.
[32] Manna, Amalesh Kumar, et al. "Multi-item EPQ model with learning effect on imperfect production over fuzzy-random planning horizon." Journal of Management Analytics 4.1 (2017): 80-110.
[33] Khara, Barun, Jayanta Kumar Dey, and Shyamal Kumar Mondal. "An inventory model under development cost-dependent imperfect production and reliability-dependent demand." Journal of Management Analytics 4.3 (2017): 258-275.
[34] Pal, Sova, Manas Kumar Maiti, and Manoranjan Maiti. "An EPQ model with price discounted promotional demand in an imprecise planning horizon via Genetic Algorithm." Computers & Industrial Engineering 57.1 (2009): 181-187.
[35] Bera, Uttam Kumar, and Ajoy Kumar Maiti. "A soft-computing approach to multi-item fuzzy EOQ model incorporating discount." International Journal of Information and Decision Sciences 4.4 (2012): 313-328.

[36] Maiti, A. K., M. K. Maiti, and M. Maiti. "Inventory model with stochastic lead-time and price dependent demand incorporating advance payment." Applied Mathematical Modelling 33.5 (2009): 2433-2443.
[37] Mondal, S., and M. Maiti. "Multi-item fuzzy EOQ models using genetic algorithm." Computers & Industrial Engineering 44.1 (2003): 105-117.
[38] Maiti, Manas Kumar, and Manoranjan Maiti. "Fuzzy inventory model with two warehouses under possibility constraints." Fuzzy Sets and systems 157.1 (2006): 52-73.
[39] Maiti, Manas Kumar, and Manoranjan Maiti. "Two-storage inventory model with lot-size dependent fuzzy lead-time under possibility constraints via genetic algorithm." European Journal of Operational Research 179.2 (2007): 352-371.
[40] Yadav, Ajay Singh, et al. "Optimization of an inventory model for deteriorating items with both selling price and time-sensitive demand and carbon emission under green technology investment." International Journal on Interactive Design and Manufacturing (IJIDeM) (2023): 1-17.
[41] Yadav, Krishan Kumar, Ajay Singh Yadav, and Shikha Bansal. "Interval number approach for two-warehouse inventory management of deteriorating items with preservation technology investment using analytical optimization methods." International Journal on Interactive Design and Manufacturing (IJIDeM) (2024a): 1-17.
[42] Yadav, Krishan Kumar, Ajay Singh Yadav, and Shikha Bansal. "OPTIMIZATION OF AN INVENTORY MODEL FOR DETERIORATING ITEMS ASSUMING DETERIORATION DURING CARRYING WITH TWO-WAREHOUSE FACILITY." Reliability: Theory & Applications 19.3 (79) (2024b): 442-459.
[43] Yadav, Krishan Kumar, Ajay Singh Yadav, and Shikha Bansal. "OPTIMIZATION OF A TWO-WAREHOUSE INVENTORY MANAGEMENT FOR DETERIORATING ITEMS WITH TIME AND RELIABILITY-DEPENDENT DEMAND UNDER CARBON EMISSION CONSTRAINTS." Reliability: Theory & Applications 19.4 (80) (2024c): 404-418.
[44] Jiang, XiaoWei, et al. "Effect of exponential decay in hydraulic conductivity with depth on regional groundwater flow." Geophysical Research Letters 36.24 (2009).
[45] Rana, Upasana, and Tanuj Kumar. "OPTIMIZING AN INVENTORY MODEL FOR PERISHABLE PRODUCTS WITH PRODUCT RELIABILITY AND TIME DEPENDENT DEMAND USING PENTAGONAL-FUZZY NUMBER." Reliability: Theory & Applications 19.4 (80) (2024): 374-384.
[46] Maiti, A. K., M. K. Maiti, and M. Maiti. "An EOQ model of an item with imprecise seasonal time via genetic algorithm." International Journal of Operational Research 19.3 (2014): 358-384.
[47] Negi, Ashish, and Ompal Singh. "INVENTORY MODEL FOR PROBABILISTIC DETERIORATION WITH RELIABILITY-DEPENDENT DEMAND AND TIME USING CLOUDY-FUZZY ENVIRONMENT." Reliability: Theory & Applications 19.4 (80) (2024): 385-403.
[48] Dutta, Anurag, et al. "Evaluation modelling of asteroids™ hazardousness using chaosNet." 2023 IEEE 8th International Conference for Convergence in Technology (I2CT). IEEE, 2023.
[49] Bhunia, Asoke Kumar, and Ali Akbar Shaikh. "An application of PSO in a two-warehouse inventory model for deteriorating item under permissible delay in payment with different inventory policies." Applied Mathematics and Computation 256 (2015): 831-850.

RELIABILITY AVAILABILITY MAINTAINABILITY DEPENDABILITY (RAMD) OPTIMIZATION: A CASE STUDY OF MANUFACTURING PLANT

Deepika Garg[1], Dipesh Popli[2*], Pradeep Kamboj[3], Nakul Vashishth[4]

•

[1,2,4] GD Goenka University, Gurugram, Haryana
[3] Maharishi Markandeshwar University, Mullana, Haryana

[1]deepikanit1234@gmail.com, [2]dipeshpopli@gmail.com
[3]pardeepkumar456@mmumullana.org, [4]nakulsir@gmail.com
[*]Corresponding author

Abstract

In the context of modern Industry practices, optimizing the performance of complex systems necessitates a thorough understanding of Reliability, Availability, Maintainability, and Dependability (RAMD). This paper dives into the application of two powerful optimization techniques, Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), in RAMD optimization. GA, a simulation of natural evolution, employs selection, crossover, and mutation operations to search for optimal solutions iteratively. In contrast, PSO replicates the collective behavior of a swarm, iteratively updating particle positions based on their best position and the swarm's best position. The research specifically focuses on RAMD optimization in terms of maintenance schedules, resource allocation, and system configurations. The results from both techniques are compared, with PSO demonstrating superior performance. The paper concludes by underscoring the need for further research and practical implementation of PSO and GA to fully exploit their potential in optimizing RAMD parameters across diverse domains.

Keywords: Pharmaceutical, RAMD, Reliability, Costing, PSO, GA

I. Introduction

In the current landscape, particularly within the context of Industry 4.0, the analysis and optimization of Reliability, Availability, Maintainability, and Dependability (RAMD) hold a pivotal role in the pharmaceutical industry [1]. As technological systems in the pharmaceutical industry grow in complexity and interconnectivity, ensuring their reliable and continuous operation is not just important but crucial. RAMD analysis enables proactive identification and mitigation of potential failures, leading to minimized downtime and optimized resource allocation. It also helps organizations meet the increasing demands of customers for reliable and available services. Furthermore, RAMD optimization brings tangible benefits such as reduced maintenance costs, increased system longevity, enhanced safety, and improved overall performance, all of which are essential for businesses to stay competitive and maintain customer trust in a rapidly evolving technological landscape [2].

RAMD analysis and optimization involve evaluating and improving the performance and trustworthiness of systems. Study focused on the system's ability to function without failure, while availability measures its operational readiness [3]. Maintainability examines the ease of maintenance and repair, and dependability encompasses overall system trustworthiness. RAMD optimization aims to enhance these attributes by identifying critical components, analyzing failure modes, optimizing maintenance strategies, and employing redundancy schemes.

Several techniques are employed for RAMD analysis and optimization. These include Failure Mode and Effects Analysis (FMEA), which systematically identifies potential failure modes and their effects to prioritize mitigation efforts [4]. Fault Tree Analysis (FTA) is utilized to assess combinations of events and failures that can lead to system failures. Reliability Block Diagrams (RBD) help model and analyze system reliability by representing components and their interconnections [5]. Simulation techniques, such as Monte Carlo simulation, are utilized to assess system performance under various conditions. Optimization algorithms, including genetic algorithms and mathematical programming, aid in optimizing maintenance strategies and resource allocation. Data-driven approaches like condition-based maintenance and predictive analytics utilize real-time data for proactive maintenance and decision-making. These techniques ensure effective RAMD analysis and optimization in modern systems [6]. Studies proposed a multi-objective formulation that joins a multi-objective elitist genetic algorithm with a high-level simulation environment [7,8]. Some studies considered parallel components of serial sub-systems to work in three states [9,10]. A branch and cut algorithm were proposed to solve the problem of optimality for a series-parallel system configuration [11]. A simulation-based optimization approach was introduced for modeling, and a genetic algorithm was developed to solve the problem [12]. A mathematical model was proposed for optimizing multiple redundancy-reliability systems, known as mega-systems, which was simulated to evaluate system availability [13].

Soft computing techniques, including fuzzy logic, neural networks, and evolutionary computing methods like PSO and GA, have proven invaluable in RAMD analysis and optimization. These techniques offer robust and efficient approaches to tackle the complexities of modern systems. The present paper focuses on leveraging PSO and GA for RAMD analysis and optimization, with a comparative analysis revealing the superior performance of PSO. The study underscores the significance of further research and practical implementation of PSO and GA to fully harness their potential in optimizing RAMD parameters across diverse domains.

RAMD optimization involves the application of various theories and concepts to enhance the performance and trustworthiness of systems. The theories include reliability analysis, which assesses the probability of failure-free operation; availability analysis, which measures the system's readiness to perform; maintainability analysis, which evaluates the ease and effectiveness of maintenance and repair; and dependability analysis, which encompasses overall system trustworthiness. These concepts are utilized to identify critical components, analyze failure modes, optimize maintenance strategies, and allocate resources effectively. RAMD optimization also incorporates techniques like FMEA and Fault Tree Analysis (FTA) to identify potential failure modes, assess their impact, and prioritize mitigation efforts [14]. Moreover, optimization algorithms such as GA and PSO are employed to find optimal solutions for maintenance scheduling, resource allocation, and system configurations [15]. Overall, RAMD optimization ensures systems operate reliably, remain available, can be effectively maintained, and deliver the intended performance, thus maximizing their overall dependability.

Current trends and developments in RAMD optimization are driven by advancements in technology and data-driven approaches. One prominent trend is the increasing utilization of machine learning and artificial intelligence techniques for RAMD analysis and optimization. These techniques enable predictive maintenance, fault diagnosis, and decision-making based on real-time

data and historical patterns. Another trend is the integration of RAMD optimization with other emerging fields such as the Internet of Things (IoT) and Industry 4.0 [16]. This integration allows for real-time monitoring, remote diagnostics, and proactive maintenance of interconnected systems [17]. Additionally, there is a growing focus on incorporating sustainability and resilience aspects into RAMD optimization, considering environmental impacts and the ability to withstand unforeseen disruptions. Overall, the current trends in RAMD optimization revolve around leveraging advanced technologies, data analytics, and interconnected systems to achieve higher levels of reliability, availability, maintainability, and dependability while considering sustainability and resilience as key factors [18].

The main methodologies used in conducting research on RAMD optimization encompass a combination of mathematical modeling, simulation, data analysis, statistical methods, and optimization algorithms. Researchers develop mathematical models and simulation frameworks to represent system behavior, analyze performance, and evaluate optimization strategies. Data analysis techniques, such as regression analysis and reliability data analysis, help identify patterns, correlations, and failure trends in large datasets. Optimization algorithms like GA, PSO, and other metaheuristic approaches are employed to find optimal solutions for RAMD optimization problems [19-21]. These methodologies enable researchers to gain insights into system performance, predict failures, optimize maintenance schedules, and enhance overall RAMD parameters.

In summary, the existing literature on RAMD optimization demonstrates significant progress in developing methodologies, models, and algorithms to improve system RAMD. However, addressing challenges related to realistic modeling, integration with emerging technologies, and holistic optimization remains an important area for future research.

The paper explores RAMD optimization in pharmaceutical plants within the context of Industry 4.0, utilizing Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). Through comparative analysis, PSO demonstrates superior performance, enhancing RAMD parameters like Reliability, Availability, Maintainability, and Dependability. The study emphasizes the importance of further research and practical implementation to fully leverage PSO and GA for optimizing RAMD across diverse domains.

II. Methods

I. Particle Swarm Optimization

Particle Swarm Optimization (PSO) is only one decade old in the optimization domain. It was introduced by Clerc, 2006 [22]. PSO is a stochastic algorithm that exhibits many similarities with the evolutionary algorithm for solving optimization problems. It essentially reproduces the food-searching behavior of society, such as a bird swarm or a school of fish [23]. In PSO, each member of the swarm is known as a particle. Every particle in the search space represents a potential solution. During the search process, every particle memorizes its current position, which is called personal best (best). All the particles explore the search space and the information collected by them is sorted to find the best particle in the swarm called the global best (gbest).

The position vector and the velocity vector of the i^{th}particle in the d-dimensional search space can be expressed as $x_i = (x_{i1}, x_{i2}, \ldots, x_{id})$ and $v_i = (v_{i1}, v_{i2}, \ldots, v_{id})$ respectively. According to a user defined fitness function, the best position of each particle (which corresponds to the best fitness value obtained by that particle at time (t) is $p_i = (p_{i1}, p_{i2}, \ldots, p_{id})$, denoted as pbest and the fittest particle found so far in the entire swarm is $p_g = (p_{g1}, p_{g2}, \ldots, p_{gd})$, denoted as gbest. Then the new velocities and the new positions of the particles for the next fitness evaluation are computed at time t +1 using Equation 1 and 2.

$$v_{\{id\}(t+1)} = w \cdot v_{\{id\}(t)} + c_1 \cdot rand_1 \cdot \left(p_{\{id\}(t)} - x_{\{id\}(t)}\right) + c_2 \cdot rand_2 \cdot \left(p_{\{gd\}(t)} - x_{\{id\}(t)}\right) \quad (1)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t) \quad (2)$$

$$R_{s(n)} = g\left(R_{1(n_1)}, \ldots, R_{n(n_n)}\right) = \prod_{\{i=1\}}^{7} R_{i(n)} \quad (3)$$

$$R_{i(n_i)} = \prod_{\{i=1\}}^{7} \left[1 - \left(Q_{i(n_i)}\right)^{\{n_i\}}\right] \quad (4)$$

Where, $rand_1$ and $rand_2$ are the random values that fall between (0, 1), w is the factor inertia weight which is used to direct the effect of the previous velocities on the current velocity of a particle, c_1 is a cognitive learning factor that represents that a particle moves toward its own success and c_2 is a social learning factor that represents that a particle moves toward the success of its neighbours. Some researchers have suggested the range of c_1 as (1.5 to 4) and c_2 as (2 to 2.5). Figure 1 shows a flow chart of particle swarm optimization.

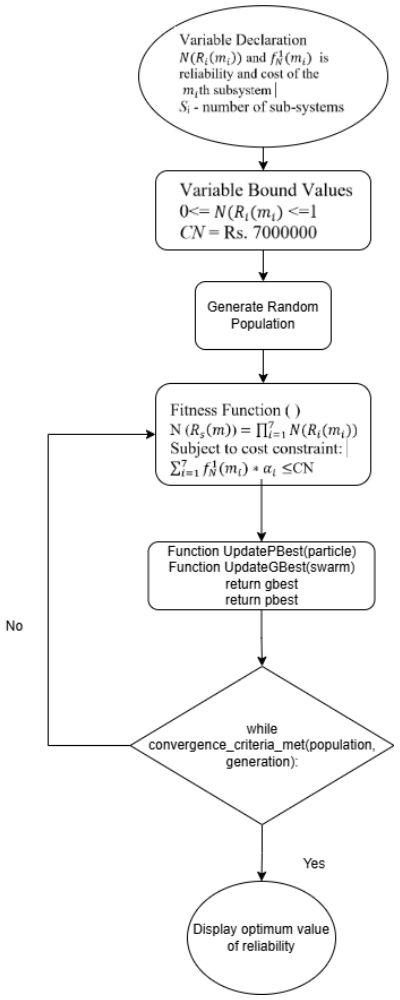


Figure 1: *Flow diagram of PSO for Redundancy allocation problem*

II. Genetic Algorithm

Genetic Algorithm (GA) is a powerful optimization technique that draws inspiration from the principles of natural selection and genetics.

What sets GA apart is its ability to effectively search through large solution spaces, handle nonlinear and non-differentiable objective functions, and explore multiple regions of the search space simultaneously. By employing genetic operators like crossover and mutation, GA mimics the evolutionary process to generate new candidate solutions and adaptively improve the population over generations. With its versatility, ability to handle complex problems, and capability to find near-optimal solutions, Genetic Algorithm remains a popular choice for a wide range of optimization problems in various domains. By using selection, genetic operators, and fitness evaluation, the GA explores the search space and evolves a population of individuals towards better solutions over generations. It's important to note that specific variations of GA can exist, such as elitism (preserving the best individuals), adaptive parameter control, or niching techniques to maintain diversity. These modifications can enhance the algorithm's performance in different problem domains. Additionally, the choice of encoding schemes, crossover and mutation operators, and fitness function design can also impact the effectiveness of the GA. In the present work, the GA methodology has been utilized [24,25]. Figure 2 shows the flow chart of the adapted GA approach for the present work.

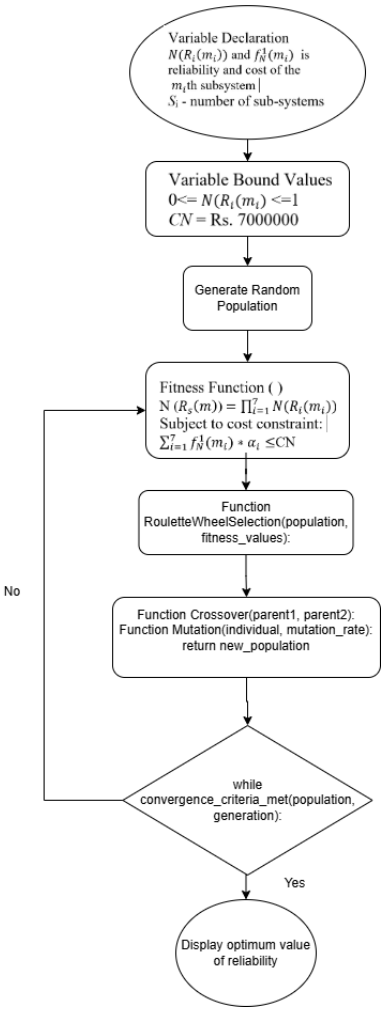


Figure 2: *Flow diagram of GA for Redundancy allocation problem*

III. Problem Definition

The 'ABC Machines Manufacturing Solutions Plant' consists of ten interconnected subsystems, each performing specific operations essential for manufacturing. The detailed descriptions of these machines and their roles in the production process are summarized in Table 1. All the machines are connected in series and perform specific operations.

Table 1: *Description of Manufacturing Plant*

SN	Machines	ABC M/C Mfg. Sol. products	Description
1	CNC Machine	Precision CNC	Performs precision cutting, drilling, and shaping of materials using computer numerical control technology
2	Injection Molding Machine	Pro Mold Master	produces plastic parts through the injection molding process, melting plastic granules and injecting them into molds.
3	Laser Cutting Machine	Laser Tech Cut	Utilizes a focused laser beam to accurately cut through various materials, such as metal, wood, acrylic, and fabric.
4	Assembly Machine	Swift Assembler	automates the assembly process by joining components together through screwing, fastening, gluing, or welding.
5	Packaging Machine	Smart Pack Pro	Automates the packaging process, including filling, sealing, labeling, and wrapping of products.
6	Press Machine	Power Form Press	Applies pressure to shape or form materials through compression, bending, or stamping.
7	Robotic Welding Machine	Weld Tech Bot	Uses industrial robots to perform precise and efficient welding operations
8	3D Printing Machine	Print Tech Pro	Creates three-dimensional objects layer by layer using additive manufacturing or 3D printing technology.
9	Automated Inspection Machine	Vision Inspect	Conducts automated inspections and quality control checks on manufactured products.
10	Conveyor System	Opti Flow Conveyor	Facilitates the smooth and efficient movement of materials and products throughout the manufacturing process.

I. Redundancy Allocation Problem

The redundancy allocation problem is formulated for the 'Prime Tech Manufacturing Solutions' manufacturing plant as follows.

$$N\left(R_{s(m)}\right) = \prod_{\{i=1\}}^{7} N\left(R_{i(m_i)}\right) \tag{5}$$

Subject to cost constraint:

$$\sum_{\{i=1\}}^{7} f_N^{1(m_i)} \cdot \alpha_i \leq C_N \tag{6}$$

Where,$N\left(R_{i(m_i)}\right)$ and $f_N^{1(m_i)}$is reliability and cost of the m_ith subsystem respectively. Here, considered cost constraint i.e. CN = Rs.7000000 as given in Equation (4) (since no. of constraint is one i.e. $\beta = 1$, therefore j=1) and system reliability given in Equation (3) is $T\left(R_{s(m)}\right) = 0.5237$. The failure rate, repair rate, and cost associated with each subsystem are crucial parameters for reliability analysis and optimization. A detailed breakdown of these parameters for all ten subsystems is provided in Table 2.

Table 2: *Parameters related to Subsystems S_i (1≤i≤10)*

Components	Failure rate f_i	Repair rate r_i	Cost of each component (in Rs) C_i S_i
S1	0.009	0.2	15000
S2	0.00017	0.1	500000
S3	0.122	0.15	90000
S4	0.10068	0.2	890000
S5	0.015	0.3	72000
S6	0.0023	0.25	20000
S7	0.017	0.2	770000
S8	0.021	0.15	70000
S9	0.133	0.2	190000
S10	0.217	0.25	195000

II. Measures of RAM Analysis

Considering that the system's components are all connected in series. As a result, if any component fails, the entire system will also fail. The following expressions are used to derive RAM and dependability for the components and the whole system [26].

IV. Reliability Analysis

The possibility that a system will work properly over the duration of the desired period and under the intended operating conditions is known as reliability [27]. Using the following formula, reliability can be determined for the components and the series system. Here RSTP stands the system's reliability, f_i denotes how frequently each subsystem fails for i ranging from 1 to 10.

$$R_{i(t)} = e^{-f_i t} \tag{7}$$

$$R_{\{STP\}} = \prod_{\{i=1\}}^{10} R_{i(t)} \tag{8}$$

I. Availability Analysis

The possibility that a system is functioning as intended under predetermined operating conditions at a specific moment is known as availability [28]. Following are the mathematical expressions of the availability function for the components and the whole system:

$$A_i = \frac{MTBF_i}{MTBF_i + MTTR_i} \tag{9}$$

And

$$A_{STP} = \prod_{i=1}^{10} A_i \tag{10}$$

Where Ai is the availability, $MTBF_i$is the mean-time between failures, and $MTTR_i$ is the mean-time to repair each subsystem for i ranging from 1 to 10, ASTP is the availability of the system.

II. Maintainability Analysis

It is the possibility that the failed system will be repaired and placed back into operation. The following expression is for computing the maintainability function for the components and the entire system [29]:

$$M_{i(t)} = \int_0^t r_i e^{-r_i t} dt = 1 - e^{-r_i t} \tag{11}$$

$$M_{STP(t)} = \prod_{i=1}^{10} M_{i(t)} \tag{12}$$

Where r_i is the component's repair rate, M_i(t) is the maintainability of each subsystem at time t for i ranging from 1 to 10 and MSTP (t) is the maintainability of the entire system.

III. Dependability Analysis

It evaluates how consistent the performance is. Dependability would be the ratio of failure to repair rates if they have exponential rates. It is mathematically expressed as follows [30].

$$d_i = \frac{r_i}{f_i} \tag{13}$$

Whenever availability is higher than 0.9, dependability will be highly valued. The following expression is used to determine the minimum value of dependability for both the system as a whole and each component individually:

$$D_{min,i} = 1 - \left(\frac{1}{1-d_i}\right)\left(e^{-\frac{\ln d_i}{d_i - 1}} - e^{-\frac{d_i + \ln d_i}{d_i - 1}}\right) \tag{14}$$

$$D_{(min,STP)} = \prod_{i=1}^{10} M_{(min,i)} \tag{15}$$

IV. N Mean-Time Between Failures (MTBF) Formulation

It symbolizes the mean failure time period. Mathematically, it can be calculated by following equation [31].

$$MTBF_i = \frac{1}{f_i} \tag{16}$$

V. Results and Discussion

RAMD (Reliability, Availability, Maintainability, and Durability) are interrelated attributes that are crucial for the performance, safety, and cost-effectiveness of various systems and processes. These factors are vital in ensuring that systems operate efficiently, remain operational for extended periods, and can be easily maintained with minimal disruptions. Organizations and industries that prioritize RAMD can benefit from increased operational efficiency, reduced downtime, enhanced safety, and improved customer satisfaction. Moreover, a strong focus on RAMD can lead to lower maintenance costs, fewer unexpected failures, and a longer system lifespan. Investing in RAMD considerations during the design, development, and operation phases can ultimately lead to better outcomes and competitive advantage in the marketplace.

In the current research, the RAMD problem is formulated for "The Prime Tech Manufacturing Solutions plant, having some systems that are in series. In RAMD, attributes like Reliability, Availability, Maintainability, and Dependability are considered for the analysis, and MTBFi and MTTR are calculated. The redundancy allocation problem (equations 3 and 4) is solved with the help of PSO and GA, and the results are mentioned in Table 3. The original cost was Rs. 2812000, but the cost obtained from PSO and GA is Rs. 6486000 and Rs. 5149000, respectively, as shown in Figure 3. The original reliability was 0.502298473, but the reliability obtained from PSO and GA is 0.912291618 and 0.842876326, respectively, given in Table 3. Reliability, Availability, Maintainability (Mi), Dependability (di), MTBFi (hrs.), MTTRi (hrs.), and Dmin I are also calculated for the individual systems in the original form given in Figure 4 and Table 4. Again, Availability, Maintainability (Mi), Dependability (di), MTBFi (hrs.), MTTRi (hrs.), and $D_{min\ i}$ are also calculated from the configuration obtained from PSO and then GA. Figure 4: Redundancy Levels for Components.

Table 3: *Redundancy level for* S_i ($\mathbf{1 \le i \le 10}$)

Techniques	Reliability	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
Original Form	0.5023	1	1	1	1	1	1	1	1	1	1
GA	0.8429	2	1	2	1	1	3	3	1	2	3
PSO	0.9123	2	2	2	2	3	3	3	2	2	2

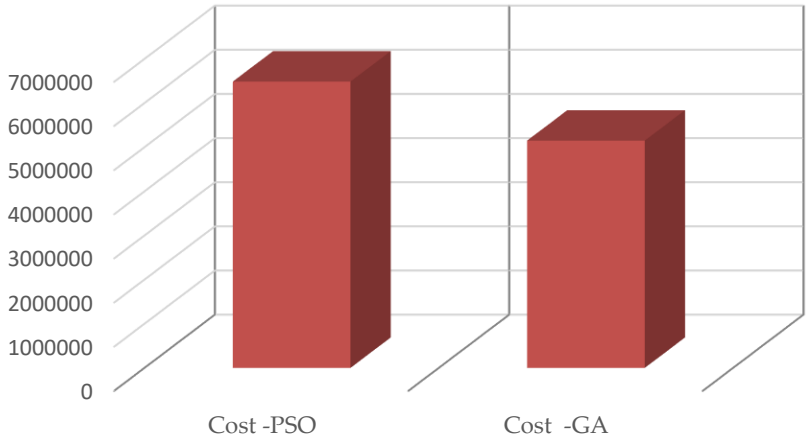


Figure 3: *Cost Analysis for PSO and GA.*

Table 4: *RAM Indices with Dependability, MTBF, MTTR and D_{min} for STP and its Components in the original form*

RAM Marices	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
Reliability (R_i)	0.9999	0.9998	0.9851	0.8993	0.9997	1	0.9999	0.979	0.9823	0.9897
Availability (A_i)	0.9997	0.9983	0.9527	0.6651	0.9996	1	0.9999	0.8771	0.9576	0.9865
Maintainability (M_i)	0.9816	0.6321	0.9502	0.8646	0.9975	0.9994	0.9975	0.7768	0.9816	0.9994
Dependability (d_i)	4938.2	588.23	20.155	1.9864	2666.6	61642	12212	7.1428	22.612	73.397
MTBF$_i$ (hrs.)	12345.3	5882.3	67.18	9.932	4444.4	8218952	203541	47.6190	56.5323	97.8635
MTTR$_i$ (hrs.)	2.5	10	3.3333	5	1.6666	1.3333	1.6666	6.6666	2.5	1.3333
$D_{min,i}$	1147.	136.08	4.8378	1.1623	619.05	14334518	28397.49	2.08476	5.378506	16.878

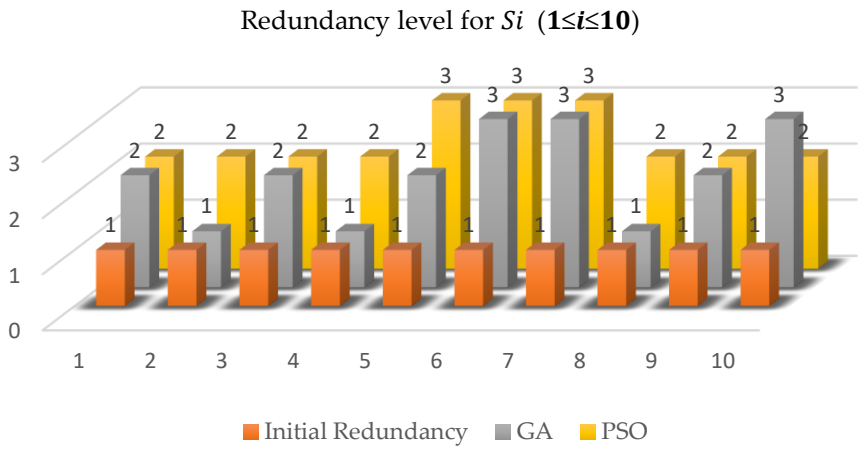


Figure 4: *Redundancy Levels for Components*

Table 5 presents RAM Indices with Dependability, MTBF, MTTR, and Dmin for STP and its Components after Technique 1. Further improvements in RAM indices, including Reliability, Availability, Maintainability, Dependability, Mean Time Between Failures (MTBF), and Mean Time To Repair (MTTR) after applying the second optimization technique, are detailed in Table 6, provides a comparative analysis with previous results.

Table 5: *RAM Indices with Dependability, MTBF, MTTR and Dmin i for STP and its Components after Technique 1*

RAM Marices	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
Reliability (R_i)	0.9999	0.9998	0.9851	0.8993	0.9998	1	1	0.979	0.9823	0.9898
Availability (A_i)	0.9998	0.9983	0.9527	0.6652	0.9996	1	1	0.8772	0.9577	0.9866
Maintainability (M_i)	0.9817	0.6321	0.9502	0.8647	0.9975	0.9994	0.9975	0.7769	0.9817	0.9994
Dependability (d_i)	4938.3	588.24	20.156	1.9865	2666.7	6E+07	122125	7.1429	22.613	73.398
$MTBF_i$ (hrs.)	12346	5882.4	67.186	9.9325	4444.4	8E+07	203542	47.619	56.532	97.864
$MTTR_i$ (hrs.)	2.5	10	3.3333	5	1.6667	1.3333	1.6667	6.6667	2.5	1.3333
$D_{min,i}$	1147.2	136.09	4.8379	1.1624	619.05	1E+07	28397	2.0848	5.3785	16.879

Table 6: *RAM Indices with Dependability, MTBF, MTTR and Dmin for STP and its Components after Technique 2*

RAM Marices	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
Reliability (R_i)	0.9999	0.9998	0.9851	0.8993	0.9998	1	1	0.979	0.9823	0.9898
Availability (A_i)	0.9998	0.9983	0.9527	0.6652	0.9996	1	1	0.8772	0.9577	0.9866
Maintainability (M_i)	0.9817	0.6321	0.9502	0.8647	0.9975	0.9994	0.9975	0.7769	0.9817	0.9994
Dependability (d_i)	4938.3	588.24	20.156	1.9865	2666.7	6E+07	122125	7.1429	22.613	73.398
$MTBF_i$ (hrs.)	12346	5882.4	67.186	9.9325	4444.4	8E+07	203542	47.619	56.532	97.864
$MTTR_i$ (hrs.)	2.5	10	3.3333	5	1.6667	1.3333	1.6667	6.6667	2.5	1.3333
$D_{min,i}$	1147.2	136.09	4.8379	1.1624	619.05	1E+07	28397	2.0848	5.3785	16.879

A graphical analysis of RAMD (Reliability, Availability, Maintainability, and Durability) is shown in Figures 5 to 8. These figures offer a comprehensive view of the system's overall performance across multiple attributes, emphasizing the advantages of using PSO over GA in certain areas. In Figure 5, Availability analysis shows good availability compared to the Configuration obtained from PSO, which has a higher availability level than GA, demonstrating the effectiveness of PSO in maintaining consistent performance. Figure 6 represents the optimized value of

dependability obtained from PSO, which supports better system performance and long-term reliability. Figure 7 shows the maintainability values obtained from both PSO and GA, where the maintainability achieved by PSO demonstrates better results, indicating that PSO contributes to a more sustainable and cost-effective system. D-min is a critical parameter in the design and development of reliable products and systems, setting the minimum acceptable level of reliability. Achieving or exceeding this threshold is essential to ensure that the product or system performs as intended and meets user expectations. Figure 8 shows that the results obtained by the PSO exhibit better value than GA, highlighting the superiority of PSO in providing more reliable and durable outcomes for system design.

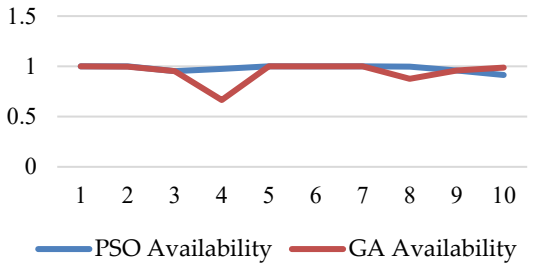


Figure 5: *Availability Analysis for PSO and GA*

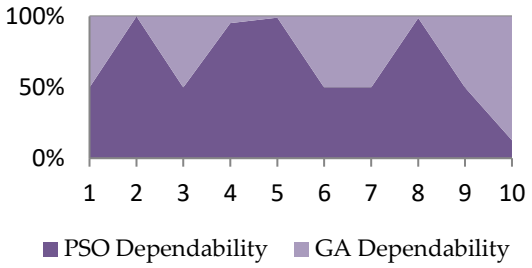


Figure 6: *Dependability Analysis for PSO and GA*

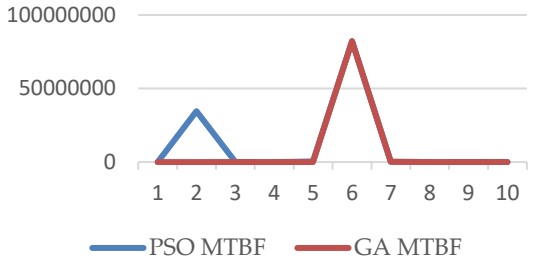


Figure 7: *Maintainability Analysis for PSO and GA*

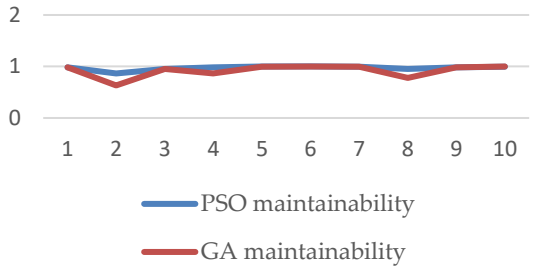


Figure 8: *Maintainability Analysis D_{min} Analysis*

Optimization techniques like Particle Swarm Optimization (PSO) play a crucial role in enhancing complex systems' performance over time. By refining key parameters such as availability, reliability, and maintainability, these methods ensure that products meet both technical standards and customer expectations. Effective optimization improves operational efficiency, minimizes failures, and enhances system longevity. By continuously adjusting variables, PSO helps achieve better resource utilization and overall system stability. This ensures that businesses can deliver high-quality, dependable solutions while reducing downtime and maintenance costs. Ultimately, integrating optimization techniques leads to more robust, efficient, and customer-satisfying products in various industries.

Mean Time Between Failures (MTBF) is a crucial statistical measure used to determine the average time a system or component operates before experiencing its first failure. It helps in assessing reliability and predicting system performance over time. The experimental results indicate that Genetic Algorithm (GA) performs better than Particle Swarm Optimization (PSO) in MTBF analysis, as illustrated in Figure 9. Mean Time to Repair (MTTR) is another key performance indicator that measures the average downtime a system experiences due to failures. It represents the mean duration, typically in hours, required to restore a system to its functional state after failure. Figure 10 presents the MTTR analysis derived from GA and PSO, showing that PSO performs better than GA in this aspect. For system S_2, the percentage increase in key metrics using PSO is as follows:

Reliability (0.01700289%), Availability (0.16998847%), Maintainability (36.78789346%), and Dependability. Additionally, PSO leads to an increase in Mean Time Between Failures (MTBFi) of 1,176,370.538 hours and Mean Time to Repair (MTTRi) of 588,135.2861 hours. However, Dependability (D_{min} i) decreases by 50%. This performance pattern is consistent for other systems such as S_4, S_5, and S_8, reinforcing PSO's effectiveness. These findings highlight the importance of selecting appropriate optimization techniques based on system performance requirements. While GA performs better in MTBF, PSO offers advantages in MTTR and other reliability parameters.

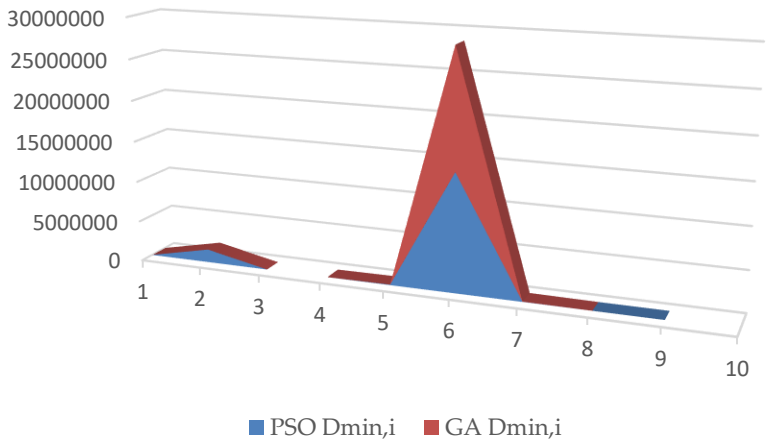


Figure 9: *MTBF Analysis PSO and GA*

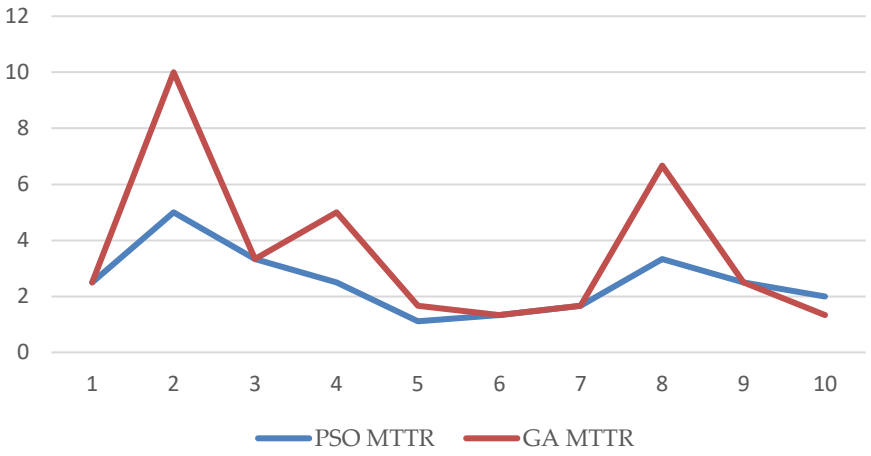


Figure 10: *MTTR Analysis PSO and GA.*

VI. Conclusion

The current research has delved into the critical aspects of Reliability, Availability, Maintainability, and Dependability (RAMD) within the realm of complex systems optimization. The study underscores the pivotal role of RAMD analysis and optimization in ensuring the smooth and uninterrupted operation of contemporary technological systems. By optimizing RAMD parameters, organizations stand to gain reduced downtime, more efficient resource allocation, decreased maintenance expenses, heightened system availability, improved safety measures and enhanced overall performance. The paper has focused on the utilization of two optimization techniques, Genetic Algorithms (GAs) and Particle Swarm Optimization (PSO), to enhance RAMD parameters.

While both techniques aim to optimize maintenance schedules, resource allocation, and system configurations, the comparative analysis demonstrates that PSO outperforms GA in achieving higher reliability, availability, maintainability, and dependability levels.

The study has examined the practical application of RAMD analysis and optimization by considering a manufacturing plant scenario with ten interconnected subsystems. The redundancy allocation problem has been formulated and solved using PSO and GA, resulting in improved RAMD attributes and a significant increase in cost. Additionally, individual RAMD indices, such as reliability, availability, maintainability, and dependability, have been computed for each subsystem, both in the original configuration and after optimization. The percentage increase in values for S_2's is 0.01700289, 0.16998847, 36.78789346, 1176370.538, 588135.2861, and -50 for Reliability, Availability, Maintainability, and Dependability MTBFi (hrs.), $MTTR_i$ (hrs.), $D_{min\ i}$ respectively by using PSO.

I. Future Research Directions

Future research should emphasis on development of comprehensive multi-objective optimization frameworks that incorporate sustainability and resilience metrics is essential. Further investigation into hybrid optimization algorithms that synergize various techniques will enhance robustness and efficacy. Extending the application of RAMD optimization to diverse domains, including healthcare, transportation, and energy, will provide sector-specific insights. Improving the scalability of optimization algorithms and their real-time adaptability is crucial. Additionally, leveraging big data analytics and machine learning to refine predictive maintenance models and enhance reliability predictions is imperative.

II. Implications

Advancements in these research areas will significantly enhance system performance, yielding higher reliability, availability, maintainability, and dependability, while reducing operational costs through predictive maintenance strategies. Improved safety and operational reliability will be achieved, particularly in critical sectors. Efficient resource utilization will be optimized, ensuring strategic allocation of maintenance personnel and financial investments. Organizations implementing these advanced RAMD techniques will gain a competitive advantage by delivering superior and dependable products and services, thereby increasing customer satisfaction and loyalty. Furthermore, these research directions will establish a robust foundation for future innovations in RAMD methodologies and applications, driving the field forward.

References

[1] A. Kumar, V. Kumar, V. Modgil, and A. Kumar, "Stochastic Petri Nets modelling for performance assessment of a manufacturing unit," Mater Today Proc, vol. 56, 2022, doi: 10.1016/j.matpr.2022.01.073.

[2] N. M. Usman and I. Yusuf, "Reliability, Availability, Maintainability, and Dependability of a Serial Rice Mill Plant (RMP) with the Incorporation of Coverage Factor," 2023. doi: 10.1007/978-981-19-9909-3_17.

[3] S. Devi, H. Garg, and D. Garg, "A review of redundancy allocation problem for two decades: bibliometrics and future directions," Artif Intell Rev, 2022.

[4] Y. Tang, S. Tan, and D. Zhou, "An Improved Failure Mode and Effects Analysis Method Using Belief Jensen–Shannon Divergence and Entropy Measure in the Evidence Theory," Arab J Sci Eng, vol. 48, no. 5, 2023, doi: 10.1007/s13369-022-07560-4.

[5] L. Carnevali, L. Ciani, A. Fantechi, G. Gori, and M. Papini, "An efficient library for reliability block diagram evaluation," Applied Sciences (Switzerland), vol. 11, no. 9, 2021..

[6] C. H. Mariano and A. F. Kuri-Morales, "Complex componential approach for redundancy allocation problem solved by simulation-optimization framework," J Intell Manuf, vol. 25, no. 4, 2014.

[7] H. Tadesse et al., "Investigation of production bottlenecks and productivity analysis in soft drink industry: a case study of East Africa Bottling Share Company," International Journal on Interactive Design and Manufacturing, 2024.

[8] A. S. Yadav, A. Kumar, K. K. Yadav, and S. Rathee, "Optimization of an inventory model for deteriorating items with both selling price and time-sensitive demand and carbon emission under green technology investment," International Journal on Interactive Design and Manufacturing, 2023.

[9] N. Kumar, N. Kumar, D. Kumar, and D. Kumar, "An Improved Grey Wolf Optimization-Based Learning Of Artificial Neural Network For Medical Data Classification," Journal of Information and Communication Technologies, 2021.

[10] P. Pourkarim Guilani, M. Sharifi, S. T. A. Niaki, and A. Zaretalab, "Redundancy allocation problem of a system with three-state components: A genetic algorithm," International Journal of Engineering, Transactions B: Applications, vol. 27, no. 11, 2014.

[11] M. Caserta and S. Voß, "An exact algorithm for the reliability redundancy allocation problem," in European Journal of Operational Research, 2015.

[12] P. P. Guilani, P. Azimi, S. T. A. Niaki, and S. A. A. Niaki, "Redundancy allocation problem of a system with increasing failure rates of components based on Weibull distribution: A simulation-based optimization approach," Reliab Eng Syst Saf, vol. 152, 2016, doi: 10.1016/j.ress.2016.03.010.

[13] A. Zaretalab, V. Hajipour, and M. Tavana, "Redundancy allocation problem with multi-state component systems and reliable supplier selection," Reliab Eng Syst Saf, vol. 193, 2020, doi: 10.1016/j.ress.2019.106629.

[14] M. Shafiee, E. Enjema, and A. Kolios, "An integrated FTA-FMEA model for risk analysis of engineering systems: A case study of subsea blowout preventers," Applied Sciences (Switzerland), vol. 9, no. 6, 2019.

[15] S. Devi and D. Garg, "Hybrid genetic and particle swarm algorithm: redundancy allocation problem," International Journal of System Assurance Engineering and Management, vol. 11, no. 2, 2020.

[16] M. A. Farsi and E. Zio, "Industry 4.0: Some Challenges and Opportunities for Reliability Engineering," International Journal of Reliability, Risk and Safety: Theory and Application, vol. 2, no. 1, 2019.

[17] D. Popli and M. Gupta, "Investigation of the Circularity and Conicity of Super Alloy During Rotary Ultrasonic Machining," Iranian Journal of Science and Technology, Transactions of Mechanical Engineering, 2018.

[18] D. Popli and M. Gupta, "Sequential procedure for selecting the ranges of process parameters in rotary ultrasonic machining," International Journal of Manufacturing Research, vol. 12, no. 3, pp. 364–378, 2017.

[19] D. Garg, T. Dahiya, and A. K. Shrivastava, "Developing a new heuristic algorithm for efficient reliability optimization," International Journal of Information Technology (Singapore), vol. 14, no. 5, 2022.

[20] A. Agrawal, D. Garg, R. Sethi, and A. K. Shrivastava, "Optimum redundancy allocation using spider monkey optimization," Soft comput, 2023.

[21] T. Dahiya, N. Vashishth, D. Garg, A. K. Shrivastava, and P. K. Kapur, "Novel Heuristic Algorithm & its Application for Reliability Optimization," International Journal of Mathematical, Engineering and Management Sciences, vol. 8, no. 4, 2023, doi: 10.33889/ijmems.2023.8.4.043.

[22] M. Clerc, Particle Swarm Optimization. 2006.

[23] D. Bratton and J. Kennedy, "Defining a standard for particle swarm optimization," Swarm Intelligence Symposium, 2007., 2007.

[24] D. Popli, U. Batra, V. Msomi, and S. Verma, "A systematic survey of RUM process parameter optimization and their influence on part characteristics of nickel 718," Sci Rep, vol. 13, no. 1, Dec. 2023.

[25] M. N. Ab Wahab, S. Nefti-Meziani, and A. Atyabi, "A comprehensive review of swarm optimization algorithms," PLoS One, vol. 10, no. 5, May 2015.

[26] M. V. Clavijo Mesa, C. E. Patino-Rodriguez, F. J. Guevara Carazas, I. Gunawan, and E. L. Droguett, "Asset management strategies using reliability, availability, and maintainability (RAM) analysis," Journal of the Brazilian Society of Mechanical Sciences and Engineering, vol. 43, no. 11, 2021.

[27] R. Teixeira, M. Nogal, and A. O'Connor, "Adaptive approaches in metamodel-based reliability analysis: A review," Structural Safety, vol. 89, 2021, doi: 10.1016/j.strusafe.2020.102019.

[28] J. Wang, Q. Zhang, S. Yoon, and Y. Yu, "Reliability and availability analysis of a hybrid cooling system with water-side economizer in data center," Build Environ, vol. 148, 2019, doi: 10.1016/j.buildenv.2018.11.021.

[29] M. Saini, A. Kumar, and D. Sinwar, "Parameter estimation, reliability and maintainability analysis of sugar manufacturing plant," International Journal of System Assurance Engineering and Management, vol. 13, no. 1, 2022.

[30] S. Kabir, "An overview of fault tree analysis and its application in model based dependability analysis," 2017.

[31] P. Zeng, W. Shao, and Y. Hao, "Study on preventive maintenance strategies of filling equipment based on reliability-cantered maintenance," Tehnicki Vjesnik, vol. 28, no. 2, 2021.

RT&A, No 2 (84)
Volume 20, June 2025

ON THE PROTECTION AGAINST NOISE FOR MEASUREMENT-BASED QUANTUM COMPUTATION

Valentin Vankov Iliev

•

Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia, Bulgaria
viliev@math.bas.bg

Abstract

In this paper we establish conditions for some pairs of quantum logic gates which operate on one qubit to be protected against crosstalk. We represent these logic gates by self-adjoint operators acting on an unitary plane, consider the resembling bipartite quantum systems, and extend the operators naturally to a commuting pair. Thus the corresponding two measurements can be considered as simultaneous and we can define them as events in a classical probability space. In particular, we are able to study their dependence, the information noise thus produced, and to find necessary and sufficient conditions for being informationally independent events.

Keywords: Quantum computation, Logical gates, Noise, Informational independence.

1. Introduction, Notation

1.1. Notation

The following notation will be frequently used in this paper:
$\delta_{k,\ell}$: Kronecker's delta;
$\mathbb{Z}_2 = \{0,1\}$: the additive group of two elements;
$\mathcal{H}$: 2-dimensional unitary space with inner product $\langle u|v\rangle$ of vectors $|u\rangle$ and $|v\rangle$;
$\{|0\rangle, |1\rangle\}$: orthonormal frame for $\mathcal{H}$, sometimes called the computational basis;
$\mathbb{I}_{\mathcal{H}}$: the identity linear operator on $\mathcal{H}$;
$\mathrm{Spec}(A)$: the spectre of a linear operator A on $\mathcal{H}$;

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}:$$

Pauli matrices and Hadamard matrix, respectively;
$\mathcal{E}$: the 3-dimensional real linear space of all self-adjoined operators on $\mathcal{H}$ with trace 0, which is furnished with orthonormal coordinates x, y, z;
$\mathbb{R}_+(A) = \{[(cx, cy, cz)|c \in \mathbb{R}, c \geq 0\}$: direction of the operator $A \in \mathcal{E}$ with coordinates (x, y, z);
$\mathcal{H}^{\otimes 2}$: the tensor square of $\mathcal{H}$ provided with the standard structure of a unitary space, the tensor product of vectors $|u\rangle$ and $|v\rangle$ being denoted $|uv\rangle$;
$\mathcal{U}^{(2)}$: the unit sphere of $\mathcal{H}^{\otimes 2}$;
$\psi_{s,t} = \frac{1}{\sqrt{2}}(|0t\rangle + (-1)^s|1(t+1)\rangle)$, $\psi_{s,t} \in \mathcal{U}^{(2)}$, $s, t \in \mathbb{Z}_2$: the four Bell's states;
We denote: $\mathrm{tr}_0\,\alpha = \cos\alpha$, $\mathrm{tr}_1\,\alpha = \sin\alpha$, $\alpha \in \mathbb{R}$.
We denote: $P(\alpha, \beta) = \mathrm{tr}_0\,\alpha\,\mathrm{tr}_0\,\beta\,\mathrm{tr}_1\,\alpha\,\mathrm{tr}_1\,\beta$, $\alpha, \beta \in \mathbb{R}$.

1.2. Introduction

Some of the logic gates in a quantum computational network can be represented by self-adjoined operators A, B, ... with spectre $\{1, -1\}$ on the unitary plane $\mathcal{H}$. Their set coincides with unit sphere of the linear space $\mathcal{E}$. After appropriate tensoring with the identity operator, any pair A, B of such logic operations generates a pair $\mathcal{A}$ and $\mathcal{B}$ of self-adjoined commuting operators with the same spectre, which are defined on the bipartite quantum system $\mathcal{H}^{\otimes 2}$. In particular, the measurements of the observables corresponding to these operators can be considered as simultaneous. In other words, we consider as an event "the outcome of measuring $\mathcal{A}$ is λ_k *and* the outcome of measuring $\mathcal{B}$ is λ_ℓ ", where $\lambda_k, \lambda_\ell \in \{1, -1\}$. We can suppose without loss of generality that $k, \ell \in \mathbb{Z}_2$, and that $\lambda_0 = 1$, $\lambda_1 = -1$. Born rule allows us to interpret the above conjunction as an intersection of two events in a sample space and this is done in Section 2. Namely, the tensor product of orthonormal frames of the unitary plane, consisting of eigenvectors of A and B constitutes an orthonormal frame of $\mathcal{H}^{\otimes 2}$ whose members are common eigenvectors of both $\mathcal{A}$ and $\mathcal{B}$. We chose this frame as the set of outcomes of a sample space $S(\psi; \mathcal{A}, \mathcal{B})$ with probability assignment $p_{k,\ell}$ created by quantum theory via fixing a state $\psi \in \mathcal{H}^{\otimes 2}$ and via Born rule. It turns out that $p_{k,\ell}$ is the probability of the intersection of the events $\mathcal{A} = \lambda_k$ and $\mathcal{B} = \lambda_\ell$ in this sample space. In order to find the probability $p_{k,\ell}$, we identify the self-adjoined operators with their matrices with respect to the computational basis for $\mathcal{H}$ and express the matrices A and B by using polar coordinates, $A = A_{\mu,\eta}$ and $B = A_{\nu,\zeta}$, see Theorem 2. Here μ, ν are polar angles and η, ζ are azimuthal angles.

Under the condition $\psi = \psi_{s,t}$ for some $s, t \in \mathbb{Z}_2$, (one of the four Bell's states) Theorem 3 yields that the probabilities of all events $\mathcal{A} = \lambda_k$ and $\mathcal{B} = \lambda_\ell$, $k, \ell \in \mathbb{Z}_2$, are equal and that the quantity $p_{k,\ell}$ is invariant with respect to the natural action of the group $\mathbb{Z}_2$ onto the Klein four-group $\mathbb{Z}_2 \times \mathbb{Z}_2$. In particular, the probability assignment of the sample space $S(\psi_{s,t}; \mathcal{A}, \mathcal{B})$ can be written in the form $(p_{k,k}, p_{k,k+1}, p_{k,k+1}, p_{k,k})$ for $k \in \mathbb{Z}_2$.

In Section 3 we consider the two binary trials $\mathfrak{A} = (\mathcal{A} = 1) \cup (\mathcal{A} = -1)$, $\mathfrak{B} = (\mathcal{B} = 1) \cup (\mathcal{B} = -1)$ and make use of the average quantity of information of one of the experiments $\mathfrak{A}$ and $\mathfrak{B}$ relative to the other (the *information flow*, or, *noise*, or, *crosstalk* between them), see [1, §1], defined in this particular case by the Shannon's formula [2, 5.3, (6)].

In accord with [2], the joint experiment of the above binary trials gives rise to the same probability distribution $(p_{k,k}, p_{k,k+1}, p_{k,k+1}, p_{k,k})$, $k \in \mathbb{Z}_2$. Via modification of its entropy, we bring out a function which measures the degree of dependence of the events $\mathcal{A} = \lambda_k$ and $\mathcal{B} = \lambda_\ell$. The above events are independent (the entropy is maximal) if and only if the crosstalk between $\mathfrak{A}$ and $\mathfrak{B}$ is zero. In this case we also say that the measurements of the observables corresponding to $\mathcal{A}$ and $\mathcal{B}$ are *informationally independent*. Thus, following [2, 5.2], we conclude that the above events are informationally independent exactly when the equation $p_{k,k} = \frac{1}{4}$ is satisfied.

Section 4 is devoted to three particular cases when the expression for $p_{k,k}$ from Theorem 3, (ii), (iii), has very simple form and the above equation can be solved explicitly in terms of sums or differences of polar or azimuthal angles. Those cases are defined by the property that both operators A, B have simultaneously directions $\mathbb{R}_+(A)$, $\mathbb{R}_+(B)$ laying in one of the three coordinate planes $x = 0$, $y = 0$, $z = 0$ of $\mathcal{E}$. In Subsection 4.1 we consider the set of operators with direction in coordinate plane $x = 0$ (a circle on the unit sphere of $\mathcal{E}$). Pauli matrices σ_2 and σ_3 belong to this set. The results of measurements performed on the observables $\mathcal{A}_{\mu,\eta}$ and $\mathcal{B}_{\nu,\zeta}$ are informationally independent if and only if $\mu + \nu = \frac{\pi}{2}$ or $\mu + \nu = \frac{3\pi}{2}$ in case $\psi = \frac{1}{\sqrt{2}}(|0t\rangle + |1(t+1)\rangle)$ for some $t \in \mathbb{Z}_2$, and if and only if $|\mu - \nu| = \frac{\pi}{2}$ in case $\psi = \frac{1}{\sqrt{2}}(|0t\rangle - |1(t+1)\rangle)$ for some $t \in \mathbb{Z}_2$.

We study the set of operators with direction in coordinate plane $y = 0$ in Subsection 4.2. Pauli matrices σ_1 and σ_3, as well as Hadamard matrix H, are members of this set. It turns out that the outcomes of measurement of observables $\mathcal{A}_{\mu,\eta}$ and $\mathcal{B}_{\nu,\zeta}$ are informationally independent exactly in case $\mu + \nu = \frac{\pi}{2}$ or $\mu + \nu = \frac{3\pi}{2}$, when $\psi = \frac{1}{\sqrt{2}}(|0(s+1)\rangle + (-1)^s|1s\rangle)$ for some $s \in \mathbb{Z}_2$, and in case $|\mu - \nu| = \frac{\pi}{2}$, when $\psi = \frac{1}{\sqrt{2}}(|0s\rangle + (-1)^s|1(s+1)\rangle)$ for some $s \in \mathbb{Z}_2$.

Finally, in Subsection 4.3 we examine the set of operators with direction in coordinate plane

$z = 0$. Pauli operators σ_1 and σ_2 belong there. In this case the measurements performed by the observables $\mathcal{A}_{\mu,\eta}$ and $\mathcal{B}_{\nu,\zeta}$ are informationally independent exactly in case $\eta + \zeta \in \{\frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \frac{7\pi}{2}\}$, when $\psi = \frac{1}{\sqrt{2}}(|0,0\rangle + (-1)^s|1,1\rangle)$ for some $s \in \mathbb{Z}_2$, and in case $|\eta - \zeta| = \frac{\pi}{2}$ or $|\eta - \zeta| = \frac{3\pi}{2}$, when $\psi = \frac{1}{\sqrt{2}}(|0,1\rangle + (-1)^s|1,0\rangle)$ for some $s \in \mathbb{Z}_2$.

When all was said and done, we hope that the above conditions for absence of crosstalk between two interacting logic gates can be checked experimentally.

2. Two Commuting Operators on $\mathcal{H}^{\otimes 2}$

The self-adjoined operators on $\mathcal{H}$ with spectre $\{1, -1\}$ have, in general, the form

$$A_{\mu,\eta} = \begin{pmatrix} \cos\mu & e^{-\mathrm{i}\eta}\sin\mu \\ e^{\mathrm{i}\eta}\sin\mu & -\cos\mu \end{pmatrix},$$

where $\mu \in [0, \pi]$ is the polar angle and $\eta \in [0, 2\pi)$ is the azimuthal angle. The polar coordinates $x = \sin\mu\cos\eta$, $y = \sin\mu\sin\eta$, $z = \cos\mu$, establish an isomorphism of the set of above operators and the unit sphere in $\mathcal{E}$.

We denote $B_{\nu,\zeta} = A_{\nu,\zeta}$, $\lambda_0 = 1$, $\lambda_1 = -1$. Corresponding (normalized) eigenvectors are

$$u^{(k)}_{\mu,\eta} = (-1)^k e^{-\mathrm{i}\eta}\,\mathrm{tr}_k\,\frac{\mu}{2}|0\rangle + \mathrm{tr}_{k+1}\,\frac{\mu}{2}|1\rangle, k \in \mathbb{Z}_2.$$

Moreover, $H^{(k)}_{\mu,\eta} = \mathbb{C}u^{(k)}_{\mu,\eta}$ is its λ_k-eigenspace. Note that

$$\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\} \text{ and } u_{\mu,\eta}u_{\nu,\zeta} = \{|u^{(0)}_{\mu,\eta}u^{(0)}_{\nu,\zeta}\rangle, |u^{(0)}_{\mu,\eta}u^{(1)}_{\nu,\zeta}\rangle, |u^{(1)}_{\mu,\eta}u^{(0)}_{\nu,\zeta}\rangle, |u^{(1)}_{\mu,\eta}u^{(1)}_{\nu,\zeta}\rangle\}$$

are orthonormal frames for $\mathcal{H}^{\otimes 2}$.

Let us set $\mathcal{A}_{\mu,\eta} = A_{\mu,\eta} \otimes \mathbb{I}_{\mathcal{H}}$, $\mathcal{B}_{\nu,\zeta} = \mathbb{I}_{\mathcal{H}} \otimes B_{\nu,\zeta}$. It is a straightforward check that the last two linear operators on $\mathcal{H}^{\otimes 2}$ are also self-adjoined with spectre $\{1, -1\}$. Moreover, the λ_k-eigenspace $\mathcal{H}^{(k)}_{\mathcal{A}_{\mu,\eta}} = H^{(k)}_{\mu,\eta} \otimes \mathcal{H}$ of the operator $\mathcal{A}_{\mu,\eta}$ has orthonormal frame $\{|u^{(k)}_{\mu,\eta}u^{(0)}_{\nu,\zeta}\rangle, |u^{(k)}_{\mu,\eta}u^{(1)}_{\nu,\zeta}\rangle\}$, and the λ_ℓ-eigenspace $\mathcal{H}^{(\ell)}_{\mathcal{B}_{\nu,\zeta}} = \mathcal{H} \otimes H^{(\ell)}_{\nu,\zeta}$ of the operator $\mathcal{B}_{\nu,\zeta}$ has orthonormal frame $\{|u^{(0)}_{\mu,\eta}u^{(\ell)}_{\nu,\zeta}\rangle, |u^{(1)}_{\mu,\eta}u^{(\ell)}_{\nu,\zeta}\rangle\}$, $k, \ell \in \mathbb{Z}_2$.

Since $u_{\mu,\eta}u_{\nu,\zeta}$ is an orthonornal frame of $\mathcal{H}^{\otimes 2}$ consisting of eigenvectors of both $\mathcal{A}_{\mu,\eta}$ and $\mathcal{B}_{\nu,\zeta}$, then the last two operators commute. In other words, the outcomes of measurements of these observables can be thought as simultaneous (the order is irrelevant) and the quantum theory predicts the probabilities of both outcomes by employing the frame $u_{\mu,\eta}u_{\nu,\zeta}$. Formally, this is done below.

2.1. A Sample Space and Two Random Variables

Let $\psi \in \mathcal{U}^{(2)}$ and let $S(\psi; \mathcal{A}, \mathcal{B})$ be the sample space with set of outcomes $u_{\mu,\eta}u_{\nu,\zeta}$ and probability assignment $p_{k,\ell} = |\langle|u^{(k)}_{\mu,\eta}u^{(\ell)}_{\nu,\zeta}\rangle|\psi\rangle|^2$, $k, \ell \in \mathbb{Z}_2$.

With an abuse of the language, we consider the observable $\mathcal{A}_{\mu,\eta}$ as a random variable

$$\mathcal{A}_{\mu,\eta}\colon u_{\mu,\eta}u_{\nu,\zeta} \to \mathbb{R}, \mathcal{A}_{\mu,\eta}(|u^{(0)}_{\mu,\eta}u^{(\ell)}_{\nu,\zeta}\rangle) = 1, \mathcal{A}_{\mu,\eta}(|u^{(1)}_{\mu,\eta}u^{(\ell)}_{\nu,\zeta}\rangle) = -1, \ell \in \mathbb{Z}_2,$$

on the sample space $S(\psi; \mathcal{A}, \mathcal{B})$ with probability distribution

$$p_{\mathcal{A}_{\mu,\eta}}(\lambda_k) = p_{k,0} + p_{k,1}, k \in \mathbb{Z}_2, p_{\mathcal{A}_{\mu,\eta}}(\lambda) = 0, \lambda \notin \mathrm{Spec}(\mathcal{A}_{\mu,\eta}).$$

Identifying the event $\{|u^{(k)}_{\mu,\eta}u^{(0)}_{\nu,\zeta}\rangle, |u^{(k)}_{\mu,\eta}u^{(1)}_{\nu,\zeta}\rangle\}$ with the "event" $\mathcal{A}_{\mu,\eta} = \lambda_k$ (the result of the measurement), we have $\mathrm{pr}(\mathcal{A}_{\mu,\eta} = \lambda_k) = p_{k,0} + p_{k,1}$, $k \in \mathbb{Z}_2$.

We also consider the observable $\mathcal{B}_{\nu,\zeta}$ as a random variable

$$\mathcal{B}_{\nu,\zeta}\colon u_{\mu,\eta}u_{\nu,\zeta} \to \mathbb{R}, \mathcal{B}_{\nu,\zeta}(|u^{(k)}_{\mu,\eta}u^{(0)}_{\nu,\zeta}\rangle) = 1, \mathcal{B}_{\nu,\zeta}(|u^{(k)}_{\mu,\eta}u^{(1)}_{\nu,\zeta}\rangle) = -1, k \in \mathbb{Z}_2,$$

on the above sample space with probability distribution

$$p_{\mathcal{B}_{\nu,\zeta}}(\lambda_\ell) = p_{0,\ell} + p_{1,\ell}, \ell \in \mathbb{Z}_2, p_{\mathcal{B}_{\nu,\zeta}}(\lambda) = 0, \lambda \notin \mathrm{Spec}(\mathcal{B}_{\nu,\zeta}).$$

Identifying the event $\{|u^{(0)}_{\mu,\eta}u^{(\ell)}_{\nu,\zeta}\rangle, |u^{(1)}_{\mu,\eta}u^{(\ell)}_{\nu,\zeta}\rangle\}$ with the "event" $\mathcal{B}_{\nu,\zeta} = \lambda_\ell$, we have $\mathrm{pr}(\mathcal{B}_{\nu,\zeta} = \lambda_\ell) = p_{0,\ell} + p_{1,\ell}$, $\ell \in \mathbb{Z}_2$. Moreover, the equality $(\mathcal{A}_{\mu,\eta} = \lambda_k) \cap (\mathcal{B}_{\nu,\zeta} = \lambda_\ell) = \{|u^{(k)}_{\mu,\eta}u^{(\ell)}_{\nu,\zeta}\rangle\}$ yields $\mathrm{pr}((\mathcal{A}_{\mu,\eta} = \lambda_k) \cap (\mathcal{B}_{\nu,\zeta} = \lambda_\ell)) = p_{k,\ell}$, $k, \ell \in \mathbb{Z}_2$.

We obtain immediately

Proposition 1. The following two statements are equivalent:
(i) One has $p_{0,0} = p_{1,1}$ and $p_{0,1} = p_{1,0}$.
(ii) For all $k, \ell \in \mathbb{Z}_2$ one has $\mathrm{pr}(\mathcal{A}_{\mu,\eta} = \lambda_k) = \mathrm{pr}(\mathcal{B}_{\nu,\zeta} = \lambda_\ell)$.

Theorem 2. Let $\psi = \psi_{s,t}$ where $s, t \in \mathbb{Z}_2$. Then for all $k, \ell \in \mathbb{Z}_2$ one has

$$p_{k,\ell} = \frac{1}{2}|(-1)^{k+\ell}e^{-\mathrm{i}(\eta+\zeta)}\,\mathrm{tr}_k\,\frac{\mu}{2}\,\mathrm{tr}_\ell\,\frac{\nu}{2}\delta_{0,t} + (-1)^k e^{-\mathrm{i}\eta}\,\mathrm{tr}_k\,\frac{\mu}{2}\,\mathrm{tr}_{\ell+1}\,\frac{\nu}{2}\delta_{1,t} +$$
$$(-1)^{s+\ell}e^{-\mathrm{i}\zeta}\,\mathrm{tr}_{k+1}\,\frac{\mu}{2}\,\mathrm{tr}_\ell\,\frac{\nu}{2}\delta_{0,t+1} + (-1)^s\,\mathrm{tr}_{k+1}\,\frac{\mu}{2}\,\mathrm{tr}_{\ell+1}\,\frac{\nu}{2}\delta_{1,t+1}|^2.$$

Proof. We have

$$p_{k,\ell} = |\langle u^{(k)}_{\mu,\eta}u^{(\ell)}_{\nu,\zeta}\rangle|\psi_{s,t}\rangle|^2 = \frac{1}{2}|\langle u^{(k)}_{\mu,\eta}|0\rangle\langle u^{(\ell)}_{\nu,\zeta}|t\rangle + (-1)^s\langle u^{(k)}_{\mu,\eta}|1\rangle\langle u^{(\ell)}_{\nu,\zeta}|t+1\rangle|^2 =$$
$$\frac{1}{2}|((-1)^k e^{-\mathrm{i}\eta}\,\mathrm{tr}_k\,\frac{\mu}{2}\delta_{0,0} + \mathrm{tr}_{k+1}\,\frac{\mu}{2}\delta_{1,0})((-1)^\ell e^{-\mathrm{i}\zeta}\,\mathrm{tr}_\ell\,\frac{\nu}{2}\delta_{0,t} + \mathrm{tr}_{\ell+1}\,\frac{\nu}{2}\delta_{1,t}) +$$
$$(-1)^s((-1)^k e^{-\mathrm{i}\eta}\,\mathrm{tr}_k\,\frac{\mu}{2}\delta_{0,1} + \mathrm{tr}_{k+1}\,\frac{\mu}{2}\delta_{1,1})((-1)^\ell e^{-\mathrm{i}\zeta}\,\mathrm{tr}_\ell\,\frac{\nu}{2}\delta_{0,t+1} + \mathrm{tr}_{\ell+1}\,\frac{\nu}{2}\delta_{1,t+1})|^2 =$$
$$\frac{1}{2}|(-1)^k e^{-\mathrm{i}\eta}\,\mathrm{tr}_k\,\frac{\mu}{2}((-1)^\ell e^{-\mathrm{i}\zeta}\,\mathrm{tr}_\ell\,\frac{\nu}{2}\delta_{0,t} + \mathrm{tr}_{\ell+1}\,\frac{\nu}{2}\delta_{1,t}) +$$
$$(-1)^s\,\mathrm{tr}_{k+1}\,\frac{\mu}{2}((-1)^\ell e^{-\mathrm{i}\zeta}\,\mathrm{tr}_\ell\,\frac{\nu}{2}\delta_{0,t+1} + \mathrm{tr}_{\ell+1}\,\frac{\nu}{2}\delta_{1,t+1})|^2 =$$
$$\frac{1}{2}|(-1)^{k+\ell}e^{-\mathrm{i}(\eta+\zeta)}\,\mathrm{tr}_k\,\frac{\mu}{2}\,\mathrm{tr}_\ell\,\frac{\nu}{2}\delta_{0,t} + (-1)^k e^{-\mathrm{i}\eta}\,\mathrm{tr}_k\,\frac{\mu}{2}\,\mathrm{tr}_{\ell+1}\,\frac{\nu}{2}\delta_{1,t} +$$
$$(-1)^{s+\ell}e^{-\mathrm{i}\zeta}\,\mathrm{tr}_{k+1}\,\frac{\mu}{2}\,\mathrm{tr}_\ell\,\frac{\nu}{2}\delta_{0,t+1} + (-1)^s\,\mathrm{tr}_{k+1}\,\frac{\mu}{2}\,\mathrm{tr}_{\ell+1}\,\frac{\nu}{2}\delta_{1,t+1}|^2.$$

■

2.2. The Probability Assignment of the Sample Space $S(\psi; \mathcal{A}, \mathcal{B})$

Below we express the probabilities $p_{k,\ell}$, $k, \ell \in \mathbb{Z}_2$, as functions in sums $\mu + (-1)^s\nu$ and $\eta + (-1)^t\zeta$, $s, t \in \mathbb{Z}_2$, by using Theorem 2.

Theorem 3. Let $\psi = \psi_{s,t}$ for some $s, t \in \mathbb{Z}_2$.
(i) One has $\mathrm{pr}(\mathcal{A}_{\mu,\eta} = \lambda_k) = \mathrm{pr}(\mathcal{B}_{\nu,\zeta} = \lambda_\ell) = \frac{1}{2}$ for any $k, \ell \in \mathbb{Z}_2$.
(ii) One has

$$p_{0,0} = p_{1,1} =$$
$$\frac{1}{2}\,\mathrm{tr}_t^2\,\frac{\mu + (-1)^s\nu}{2} + 2(-1)^{s+t}\,\mathrm{tr}_t^2\,\frac{\eta + (-1)^t\zeta}{2}P\left(\frac{\mu}{2}, \frac{\nu}{2}\right) =$$
$$\frac{1}{2}\,\mathrm{tr}_t^2\,\frac{\mu + (-1)^{s+1}\nu}{2} + 2(-1)^{s+t+1}\,\mathrm{tr}_{t+1}^2\,\frac{\eta + (-1)^t\zeta}{2}P\left(\frac{\mu}{2}, \frac{\nu}{2}\right)$$

for any $s, t \in \mathbb{Z}_2$.

(iii) One has

$$p_{0,1} = p_{1,0} =$$

$$\frac{1}{2}\operatorname{tr}_{t+1}^2 \frac{\mu + (-1)^s \nu}{2} + 2(-1)^{s+t+1} \operatorname{tr}_t^2 \frac{\eta + (-1)^t \zeta}{2} P\left(\frac{\mu}{2}, \frac{\nu}{2}\right) =$$

$$\frac{1}{2}\operatorname{tr}_{t+1}^2 \frac{\mu + (-1)^{s+1} \nu}{2} + 2(-1)^{s+t} \operatorname{tr}_{t+1}^2 \frac{\eta + (-1)^t \zeta}{2} P\left(\frac{\mu}{2}, \frac{\nu}{2}\right)$$

for any $s, t \in \mathbb{Z}_2$.

Proof. (i) In case $t = 0$ Theorem 2 yields

$$p_{0,0} = \frac{1}{2}|e^{-\mathrm{i}(\eta+\zeta)} \operatorname{tr}_0 \frac{\mu}{2} \operatorname{tr}_0 \frac{\nu}{2} + (-1)^s \operatorname{tr}_1 \frac{\mu}{2} \operatorname{tr}_1 \frac{\nu}{2}|^2,$$

$$p_{1,1} = \frac{1}{2}|e^{\mathrm{i}(\eta+\zeta)} \operatorname{tr}_0 \frac{\mu}{2} \operatorname{tr}_0 \frac{\nu}{2} + (-1)^s \operatorname{tr}_1 \frac{\mu}{2} \operatorname{tr}_1 \frac{\nu}{2}|^2,$$

and hence $p_{0,0} = p_{1,1}$. Similarly, we obtain

$$p_{0,1} = \frac{1}{2}|e^{-\mathrm{i}(\eta+\zeta)} \operatorname{tr}_0 \frac{\mu}{2} \operatorname{tr}_1 \frac{\nu}{2} + (-1)^{s+1} \operatorname{tr}_1 \frac{\mu}{2} \operatorname{tr}_0 \frac{\nu}{2}|^2,$$

$$p_{1,0} = \frac{1}{2}|e^{\mathrm{i}(\eta+\zeta)} \operatorname{tr}_0 \frac{\mu}{2} \operatorname{tr}_1 \frac{\nu}{2} + (-1)^{s+1} \operatorname{tr}_1 \frac{\mu}{2} \operatorname{tr}_0 \frac{\nu}{2}|^2,$$

and therefore $p_{0,1} = p_{1,0}$.

In case $t = 1$ Theorem 2 implies

$$p_{0,0} = \frac{1}{2}|e^{-\mathrm{i}(\eta-\zeta)} \operatorname{tr}_0 \frac{\mu}{2} \operatorname{tr}_1 \frac{\nu}{2} + (-1)^s \operatorname{tr}_1 \frac{\mu}{2} \operatorname{tr}_0 \frac{\nu}{2}|^2,$$

$$p_{1,1} = \frac{1}{2}|e^{\mathrm{i}(\eta-\zeta)} \operatorname{tr}_0 \frac{\mu}{2} \operatorname{tr}_1 \frac{\nu}{2} + (-1)^s \operatorname{tr}_1 \frac{\mu}{2} \operatorname{tr}_0 \frac{\nu}{2}|^2,$$

and therefore $p_{0,0} = p_{1,1}$. Similarly, we have

$$p_{0,1} = \frac{1}{2}|e^{-\mathrm{i}(\eta-\zeta)} \operatorname{tr}_0 \frac{\mu}{2} \operatorname{tr}_0 \frac{\nu}{2} + (-1)^{s+1} \operatorname{tr}_1 \frac{\mu}{2} \operatorname{tr}_1 \frac{\nu}{2}|^2,$$

$$p_{1,0} = \frac{1}{2}|e^{\mathrm{i}(\eta-\zeta)} \operatorname{tr}_0 \frac{\mu}{2} \operatorname{tr}_0 \frac{\nu}{2} + (-1)^{s+1} \operatorname{tr}_1 \frac{\mu}{2} \operatorname{tr}_1 \frac{\nu}{2}|^2,$$

so $p_{0,1} = p_{1,0}$.

In accord with Proposition 1, for any $k, \ell \in \mathbb{Z}_2$ we have $\operatorname{pr}(\mathcal{A}_{\mu,\eta} = \lambda_k) = \operatorname{pr}(\mathcal{B}_{\nu,\zeta} = \lambda_\ell)$ and part (i) is proved.

(ii) In case $t = 0$ we have

$$p_{0,0} = \frac{1}{2}|(\operatorname{tr}_0(\eta+\zeta) - \mathrm{i}\operatorname{tr}_1(\eta+\zeta)) \operatorname{tr}_0 \frac{\mu}{2} \operatorname{tr}_0 \frac{\nu}{2} + (-1)^s \operatorname{tr}_1 \frac{\mu}{2} \operatorname{tr}_1 \frac{\nu}{2}|^2 =$$

$$\frac{1}{2}\operatorname{tr}_0^2(\eta+\zeta) \operatorname{tr}_0^2 \frac{\mu}{2} \operatorname{tr}_0^2 \frac{\nu}{2} + (-1)^s \operatorname{tr}_0(\eta+\zeta) \operatorname{tr}_0 \frac{\mu}{2} \operatorname{tr}_0 \frac{\nu}{2} \operatorname{tr}_1 \frac{\mu}{2} \operatorname{tr}_1 \frac{\nu}{2} +$$

$$\frac{1}{2}\operatorname{tr}_1^2 \frac{\mu}{2} \operatorname{tr}_1^2 \frac{\nu}{2} + \frac{1}{2}\operatorname{tr}_1^2(\eta+\zeta) \operatorname{tr}_0^2 \frac{\mu}{2} \operatorname{tr}_0^2 \frac{\nu}{2} =$$

$$\frac{1}{2}(\operatorname{tr}_0^2 \frac{\mu}{2} \operatorname{tr}_0^2 \frac{\nu}{2} + 2(-1)^{s+1} \operatorname{tr}_0 \frac{\mu}{2} \operatorname{tr}_0 \frac{\nu}{2} \operatorname{tr}_1 \frac{\mu}{2} \operatorname{tr}_1 \frac{\nu}{2} + \operatorname{tr}_1^2 \frac{\mu}{2} \operatorname{tr}_1^2 \frac{\nu}{2}) +$$

$$(-1)^s(1 + \operatorname{tr}_0(\eta+\zeta)) \operatorname{tr}_0 \frac{\mu}{2} \operatorname{tr}_0 \frac{\nu}{2} \operatorname{tr}_1 \frac{\mu}{2} \operatorname{tr}_1 \frac{\nu}{2} =$$

$$\frac{1}{2}\operatorname{tr}_0^2 \frac{\mu + (-1)^s \nu}{2} + 2(-1)^s \operatorname{tr}_0^2 \frac{\eta+\zeta}{2} P\left(\frac{\mu}{2}, \frac{\nu}{2}\right).$$

In case $t = 1$ we obtain

$$p_{0,0} = \frac{1}{2}|(\operatorname{tr}_0(\eta-\zeta) - \mathrm{i}\operatorname{tr}_1(\eta-\zeta))\operatorname{tr}_0\frac{\mu}{2}\operatorname{tr}_1\frac{\nu}{2} + (-1)^s\operatorname{tr}_1\frac{\mu}{2}\operatorname{tr}_0\frac{\nu}{2}|^2 =$$

$$\frac{1}{2}\operatorname{tr}_0^2(\eta-\zeta)\operatorname{tr}_0^2\frac{\mu}{2}\operatorname{tr}_1^2\frac{\nu}{2} + (-1)^s\operatorname{tr}_0(\eta-\zeta)\operatorname{tr}_0\frac{\mu}{2}\operatorname{tr}_1\frac{\nu}{2}\operatorname{tr}_1\frac{\mu}{2}\operatorname{tr}_0\frac{\nu}{2} +$$

$$\frac{1}{2}\operatorname{tr}_1^2\frac{\mu}{2}\operatorname{tr}_0^2\frac{\nu}{2} + \frac{1}{2}\operatorname{tr}_1^2(\eta-\zeta)\operatorname{tr}_0^2\frac{\mu}{2}\operatorname{tr}_1^2\frac{\nu}{2} =$$

$$\frac{1}{2}\operatorname{tr}_0^2\frac{\mu}{2}\operatorname{tr}_1^2\frac{\nu}{2} + (-1)^s\operatorname{tr}_0\frac{\mu}{2}\operatorname{tr}_1\frac{\nu}{2}\operatorname{tr}_1\frac{\mu}{2}\operatorname{tr}_0\frac{\nu}{2} + \frac{1}{2}\operatorname{tr}_1^2\frac{\mu}{2}\operatorname{tr}_0^2\frac{\nu}{2} +$$

$$(-1)^{s+1}(1-\operatorname{tr}_0(\eta-\zeta))\operatorname{tr}_0\frac{\mu}{2}\operatorname{tr}_1\frac{\nu}{2}\operatorname{tr}_1\frac{\mu}{2}\operatorname{tr}_0\frac{\nu}{2} =$$

$$\frac{1}{2}\operatorname{tr}_1^2\frac{\mu+(-1)^s\nu}{2} + 2(-1)^{s+1}\operatorname{tr}_1^2\frac{\eta-\zeta}{2}P\left(\frac{\mu}{2},\frac{\nu}{2}\right).$$

(iii) When $t = 0$ we have

$$p_{0,1} = \frac{1}{2}|(\operatorname{tr}_0(\eta+\zeta) - \mathrm{i}\operatorname{tr}_1(\eta+\zeta))\operatorname{tr}_0\frac{\mu}{2}\operatorname{tr}_1\frac{\nu}{2} + (-1)^{s+1}\operatorname{tr}_1\frac{\mu}{2}\operatorname{tr}_0\frac{\nu}{2}|^2 =$$

$$\frac{1}{2}\operatorname{tr}_0^2(\eta+\zeta)\operatorname{tr}_0^2\frac{\mu}{2}\operatorname{tr}_1^2\frac{\nu}{2} + (-1)^{s+1}\operatorname{tr}_0(\eta+\zeta)\operatorname{tr}_0\frac{\mu}{2}\operatorname{tr}_1\frac{\nu}{2}\operatorname{tr}_1\frac{\mu}{2}\operatorname{tr}_0\frac{\nu}{2} +$$

$$\frac{1}{2}\operatorname{tr}_1^2\frac{\mu}{2}\operatorname{tr}_0^2\frac{\nu}{2} + \frac{1}{2}\operatorname{tr}_1^2(\eta+\zeta)\operatorname{tr}_0^2\frac{\mu}{2}\operatorname{tr}_1^2\frac{\nu}{2} =$$

$$\frac{1}{2}\operatorname{tr}_0^2\frac{\mu}{2}\operatorname{tr}_1^2\frac{\nu}{2} + (-1)^s\operatorname{tr}_0\frac{\mu}{2}\operatorname{tr}_1\frac{\nu}{2}\operatorname{tr}_1\frac{\mu}{2}\operatorname{tr}_0\frac{\nu}{2} + \frac{1}{2}\operatorname{tr}_1^2\frac{\mu}{2}\operatorname{tr}_0^2\frac{\nu}{2} +$$

$$(-1)^{s+1}(1+\operatorname{tr}_0(\eta+\zeta))\operatorname{tr}_0\frac{\mu}{2}\operatorname{tr}_1\frac{\nu}{2}\operatorname{tr}_1\frac{\mu}{2}\operatorname{tr}_0\frac{\nu}{2} =$$

$$\frac{1}{2}\operatorname{tr}_1^2\frac{\mu+(-1)^s\nu}{2} + 2(-1)^{s+1}\operatorname{tr}_0^2\frac{\eta+\zeta}{2}P\left(\frac{\mu}{2},\frac{\nu}{2}\right).$$

When $t = 1$ we obtain

$$p_{0,1} = \frac{1}{2}|(\operatorname{tr}_0(\eta-\zeta) - \mathrm{i}\operatorname{tr}_1(\eta-\zeta))\operatorname{tr}_0\frac{\mu}{2}\operatorname{tr}_0\frac{\nu}{2} + (-1)^{s+1}\operatorname{tr}_1\frac{\mu}{2}\operatorname{tr}_1\frac{\nu}{2}|^2 =$$

$$\frac{1}{2}\operatorname{tr}_0^2(\eta-\zeta)\operatorname{tr}_0^2\frac{\mu}{2}\operatorname{tr}_0^2\frac{\nu}{2} + (-1)^{s+1}\operatorname{tr}_0(\eta-\zeta)\operatorname{tr}_0\frac{\mu}{2}\operatorname{tr}_0\frac{\nu}{2}\operatorname{tr}_1\frac{\mu}{2}\operatorname{tr}_1\frac{\nu}{2} +$$

$$\frac{1}{2}\operatorname{tr}_1^2\frac{\mu}{2}\operatorname{tr}_1^2\frac{\nu}{2} + \frac{1}{2}\operatorname{tr}_1^2(\eta-\zeta)\operatorname{tr}_0^2\frac{\mu}{2}\operatorname{tr}_0^2\frac{\nu}{2} =$$

$$\frac{1}{2}\operatorname{tr}_0^2\frac{\mu}{2}\operatorname{tr}_0^2\frac{\nu}{2} + (-1)^{s+1}\operatorname{tr}_0\frac{\mu}{2}\operatorname{tr}_0\frac{\nu}{2}\operatorname{tr}_1\frac{\mu}{2}\operatorname{tr}_1\frac{\nu}{2} +$$

$$\frac{1}{2}\operatorname{tr}_1^2\frac{\mu}{2}\operatorname{tr}_1^2\frac{\nu}{2} + (-1)^s(1-\operatorname{tr}_0(\eta-\zeta))\operatorname{tr}_0\frac{\mu}{2}\operatorname{tr}_0\frac{\nu}{2}\operatorname{tr}_1\frac{\mu}{2}\operatorname{tr}_1\frac{\nu}{2} =$$

$$\frac{1}{2}\operatorname{tr}_0^2\frac{\mu+(-1)^s\nu}{2} + 2(-1)^s\operatorname{tr}_1^2\frac{\eta-\zeta}{2}P\left(\frac{\mu}{2},\frac{\nu}{2}\right).$$

Thus, we proved the third equalities from parts (ii) and (iii). The fourth equalities can be obtained by the trigonometric identity

$$P\left(\frac{\mu}{2},\frac{\nu}{2}\right) = \frac{1}{4}(-1)^{s+t+1}(\operatorname{tr}_t^2\frac{\mu+(-1)^s\nu}{2} - \operatorname{tr}_t^2\frac{\mu+(-1)^{s+1}\nu}{2}).$$

■

3. The Noise

Here we follow [2, Section 5] and [3, Section 3] with $\alpha = \beta = \frac{1}{2}$, $A = (\mathcal{A}_{\mu,\eta} = 1)$, $B = (\mathcal{B}_{\nu,\zeta} = 1)$.

The joint experiment of the binary trials $\mathfrak{A}_{\mu,\eta} = (\mathcal{A}_{\mu,\eta} = 1) \cup (\mathcal{A}_{\mu,\eta} = -1)$ and $\mathfrak{B}_{\nu,\zeta} = (\mathcal{B}_{\nu,\zeta} = 1) \cup (\mathcal{B}_{\nu,\zeta} = -1)$ (see [4, Part I, Section 6]) produces the probability distribution $\xi_1 = p_{k,k}, \xi_2 = p_{k,k+1}, \xi_3 = p_{k,k+1}, \xi_4 = p_{k,k}, k \in \mathbb{Z}_2$, (that is, the probability assignment of the sample space $S(\psi_{s,t}; \mathcal{A}, \mathcal{B})$), see Theorem 3). In turn, we obtain Boltzmann-Shannon entropy function $E(\theta) = -2\theta \ln \theta - 2(\frac{1}{2} - \theta) \ln(\frac{1}{2} - \theta)$, where $\theta = \xi_1 = p_{k,k}$, and the corresponding degree of dependence function $e(\theta)$. In accord with Section 3 of [3], the two events $\mathcal{A}_{\mu,\eta} = \lambda_k$ and $\mathcal{B}_{\nu,\zeta} = \lambda_\ell$ are independent (that is, $e(\theta) = 0$), exactly when the corresponding binary trials $\mathfrak{A}_{\mu,\eta}$ and $\mathfrak{B}_{\nu,\zeta}$ are informationally independent. In this case we also say that the above events are *informationally independent*. Since their probabilities are both $\frac{1}{2}$, we accomplish the equality $\theta = \frac{1}{4}$ as an equivalent condition for independence. In accord with Theorem 3, if $\psi = \psi_{s,t}$ for some $s, t \in \mathbb{Z}_2$, then the equation

$$\frac{1}{2} \operatorname{tr}_t^2 \frac{\mu + (-1)^s \nu}{2} + 2(-1)^{s+t} \operatorname{tr}_t^2 \frac{\eta + (-1)^t \zeta}{2} P\left(\frac{\mu}{2}, \frac{\nu}{2}\right) = \frac{1}{4},$$

or, equivalently, the equation

$$\frac{1}{2} \operatorname{tr}_t^2 \frac{\mu + (-1)^{s+1} \nu}{2} + 2(-1)^{s+t+1} \operatorname{tr}_{t+1}^2 \frac{\eta + (-1)^t \zeta}{2} P\left(\frac{\mu}{2}, \frac{\nu}{2}\right) = \frac{1}{4},$$

is a necessary and sufficient condition for the events $\mathcal{A}_{\mu,\eta} = \lambda_k$ and $\mathcal{B}_{\nu,\zeta} = \lambda_\ell$ to be informationally idependent.

4. Special Types of Self-Adjoined Operators with Spectre $\{1, -1\}$

In this Section we apply conditions of informational independence from Section 3 in some particular cases where they have simple form and explicit solutions.

4.1. Operators with Direction in Coordinate Plane $x = 0$

Here we consider self-adjoined operators of the form

$$A_{\mu,\frac{\pi}{2}} = \begin{pmatrix} \cos\mu & -\mathrm{i}\sin\mu \\ \mathrm{i}\sin\mu & -\cos\mu \end{pmatrix},$$

where $\mu \in [0, \pi]$. In particular, $A_{0,\frac{\pi}{2}} = \sigma_3$ and $A_{\frac{\pi}{2},\frac{\pi}{2}} = \sigma_2$. The two events $\mathcal{A}_{\mu,\eta} = \lambda_k$ and $\mathcal{B}_{\nu,\zeta} = \lambda_\ell$ are informationally independent if and only if $\frac{1}{2} \operatorname{tr}_t^2 \frac{\mu + (-1)^s \nu}{2} = \frac{1}{4}$. Equivalently, $|\mu + (-1)^s \nu| = \frac{\pi}{2}$ or $|\mu + (-1)^s \nu| = \frac{3\pi}{2}$. Hence the outcomes of measurement of observables $\mathcal{A}_{\mu,\eta}$ and $\mathcal{B}_{\nu,\zeta}$ are informationally independent precisely in case $\mu + \nu = \frac{\pi}{2}$ or $\mu + \nu = \frac{3\pi}{2}$ when $\psi = \psi_{0,t}$, $t \in \mathbb{Z}_2$, and precisely in case $|\mu - \nu| = \frac{\pi}{2}$ when $\psi = \psi_{1,t}$, $t \in \mathbb{Z}_2$.

4.2. Operators with Direction in Coordinate Plane $y = 0$

Now, we consider the self-adjoined operators of the form

$$A_{\mu,0} = \begin{pmatrix} \cos\mu & \sin\mu \\ \sin\mu & -\cos\mu \end{pmatrix}$$

where $\mu \in [0, \pi]$. We have $A_{0,0} = \sigma_3$, $A_{\frac{\pi}{2},0} = \sigma_1$, and $A_{\frac{\pi}{4},0} = H$.

The two events $\mathcal{A}_{\mu,\eta} = \lambda_k$ and $\mathcal{B}_{\nu,\zeta} = \lambda_\ell$ are informationally independent if and only if $\frac{1}{2} \operatorname{tr}_t^2 \frac{\mu + (-1)^{s+t+1} \nu}{2} = \frac{1}{4}$. Equivalently, $|\mu + (-1)^{s+t+1} \nu| = \frac{\pi}{2}$. Therefore the results of measurement

of observables $\mathcal{A}_{\mu,\eta}$ and $\mathcal{B}_{\nu,\zeta}$ are informationally independent exactly in case $\mu+\nu=\frac{\pi}{2}$ or $\mu+\nu=\frac{3\pi}{2}$ when $\psi=\psi_{s,s+1}$, and exactly in case $|\mu-\nu|=\frac{\pi}{2}$ when $\psi=\psi_{s,s}$, $s\in\mathbb{Z}_2$.

Remark 1. The case $t=s=1$ is discussed also in [3].

4.3. Operators with Direction in Coordinate Plane $z=0$

Finally, we consider self-adjoined operators of the form

$$A_{\frac{\pi}{2},\eta}=\begin{pmatrix}0 & e^{-\mathrm{i}\eta}\\ e^{\mathrm{i}\eta} & 0\end{pmatrix},$$

where $\eta\in[0,2\pi]$. We have $A_{\frac{\pi}{2},0}=\sigma_1$ and $A_{\frac{\pi}{2},\frac{\pi}{2}}=\sigma_2$.

The two events $\mathcal{A}_{\mu,\eta}=\lambda_k$ and $\mathcal{B}_{\nu,\zeta}=\lambda_\ell$ are informationally independent if and only if $\frac{1}{2}\operatorname{tr}_s^2\frac{\eta+(-1)^t\zeta}{2}=\frac{1}{4}$. Equivalently, $|\eta+(-1)^t\zeta|=\frac{\pi}{2}$. Therefore, the measurements performed by the observables $\mathcal{A}_{\mu,\eta}$ and $\mathcal{B}_{\nu,\zeta}$ are informationally independent precisely in case $\eta+\zeta\in\{\frac{\pi}{2},\frac{3\pi}{2},\frac{5\pi}{2},\frac{7\pi}{2}\}$, when $\psi=\psi_{s,0}$, and precisely in case $|\eta-\zeta|=\frac{\pi}{2}$ or $|\eta-\zeta|=\frac{3\pi}{2}$, when $\psi=\psi_{s,1}$, $s\in\mathbb{Z}_2$.

Acknowledgements

I would like to thank administration of the Institute of Mathematics and Informatics at the Bulgarian Academy of Sciences for creating safe working environment.

References

[1] Aspect A., Dalibard J., Roger G. (1982). Experimental Test of Bell's Inequalities Using Time-Varying Analysers. *Physical Review Letters*, 49 No 25:1804–1807.
[2] Bell J. (1964). On the Einstein Podolski Rosen Paradox. *Physics*, 1: 195–200.
[3] Gelfand I. M., Kolmogorov A. N., Yaglom A. M. Amount of Information and Entropy for Continuous Distributions. Mathematics and Its Applications, Selected Works of A. N. Kolmogorov, III: Information Theory and the Theory of Algorithms, 33–56, Springer Science+Business Media Dordrecht, 1993.
[4] Kolmogorov A. N. Foundations of the Theory of Probability, Chelsea Publishing Company New Yourk, 1956.
[5] A. Einstein, B. Podolsky, N. Rosen. (1935). Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review*, 47:777–780.
[6] V. V. Iliev. (2021). On the Use of Entropy as a Measure of Dependence of Two Events. *Reliability: Theory & Applications*, 16 No 3:237–248.
[7] C. E. Shannon. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27 No 3:379–423, 27 No 4:523–656.

HIDDEN MARKOV MODELS FOR ESTIMATING IMPORT DYNAMICS WITH PROBABILITY DISTRIBUTION ANALYSIS

Vyshnavi. M, Muthukumar. M

•

[1]Research Scholar, Department of Statistics, PSG College of Arts & Science, Coimbatore-641014
[2]Assistant Professor, Department of Statistics, PSG College of Arts & Science, Coimbatore-641014
E-mail: [1]vyshnavimp@gmail.com, [2]muthukumar@psgcas.ac.in

Abstract

This article examines the use of Hidden Markov Models and probability distributions to study agricultural import dynamics, with a focus on revealing hidden patterns in the data. The idea behind the study is to enhance our understanding of how transitions between different agricultural import states occur and to explore the most suitable probability distributions for modeling these hidden states. The purpose of the study is to identify the optimal distributional fit for hidden states by evaluating the Transition Probability Matrix, Emission Probability Matrix, and Initial Probability Vector (π) for each state. The research implements the Akaike Information Criterion and the Bayesian Information Criterion to select the best-fitting distribution for each scenario. Furthermore, the paper focuses on the practical implications of these discoveries, such as determining the most likely state sequence using the Viterbi path, which can help influence decision-making and forecasting. The analysis is carried out using R software, which provides information about the associations between probability distributions, stationarity tests, and the role of model selection criteria such as AIC and BIC. Graphical representations of AIC and BIC values over several probability distributions, and additionally a correlation matrix between selected distributions, help to highlight the findings. Overall, the paper enhances our understanding of probability distributions through HMM frameworks for agricultural import dynamics, providing recommendations for optimal model selection in various applications.

Keywords: Probability Distributions, Hidden Markov Model, Lomax Distribution, Viterbi Path, Correlation Matrix.

I. Introduction

Throughout the world, agriculture significantly impacts how countries develop in terms of their economies, societies, and environments. Indicators of a nation's reliance on foreign agricultural products and its ability to produce agricultural goods domestically include the percentage of agricultural imports to total national imports**.** Over the last few years, the agricultural industry has been using more sophisticated statistical techniques, such as HMM, to understand the complex dynamics present in agricultural systems and numerous variables that affect these systems, including market swings, weather conditions, and crop yields, cause them to analyze using conventional methods.

Hidden Markov Models and probability distributions are critical tools for studying and modelling agricultural import dynamics, providing advanced approaches for identifying

underlying trends and behaviours in import data. HMMs provide a strong foundation for capturing these dynamics by presenting import data as a series of hidden states and visible emissions. Hidden states in agricultural import contexts may represent specific import patterns or market conditions, whereas emissions relate to observed variables such as import amounts and prices. HMMs help to explain import dynamics, identify dominant patterns, and anticipate future import trends by estimating transition and emission probabilities. Finally, the investigation of probability distributions within the framework of Hidden Markov Models provides a reliable method for assessing agricultural import dynamics. This study improves our understanding of model selection and sheds light on optimal distribution choices by evaluating the suitability of various distributions using rigorous criteria such as AIC and BIC, thereby contributing to improved forecasting accuracy and decision-making in the agricultural sector. Forecast accuracy is greatly increased when agricultural import data is analyzed using Hidden Markov Models. By detecting underlying hidden states, HMMs can efficiently capture the dynamic elements that affect these imports, including market swings and seasonal variations. HMMs offer vital insights into the uncertainties that define agricultural markets by predicting the probability distributions connected to these states. This method provides a foundation for comparison with conventional forecasting techniques and enables in-depth analyses of changing import trends. This research paper was carried out using R software, which has robust statistical libraries and HMM programs. While R has several advantages, we ran into some computational issues during the investigation. Large datasets, in particular, increased calculation time, and memory management became an issue as the Baum-Welch algorithm was iterated several times. To address these issues, we used efficient coding techniques and R tools like depmixS4 and HMM to streamline calculations. We recommend that future users use parallel computing or data optimization approaches to improve the performance of their analysis, particularly with larger datasets or more sophisticated models.

Sebastian George and Ambily Jose [13] presented a Generalised Poisson Distribution (GPD) as a statistical tool for describing serial dependency in time series count data. The GP-HMM, paired with HMM, demonstrates good convergence properties in both simulated and actual data, indicating its superiority over the P-HMM. Joshni George and Seemon Thomas [5] presented a negative binomial hidden Markov model for Kerala AES situations, estimating parameters with the EM technique, getting hidden state sequences, and transition probabilities. Hao Zhang, Weidong Zhang, Ahmet Palazoglu, and Wei Sun [3] proposed HMM-Gamma model uses a Gamma distribution, pre-labeled monitoring days, and improved Expectation-Maximization approaches to forecast ozone exceedances in the Livermore Valley and Houston Metropolitan Area. Hui Zhang, Qing Ming Jonathan Wu, and Thanh Minh Nguyen [4] proposed a novel Student's t hidden Markov model (SHMM) that takes into account Markov states, latent components, and observations. Existing literature emphasizes the need to use proper probability distributions to improve model accuracy and forecasting capabilities. Several studies have shown that various HMM formulations and parameter estimation strategies are successful in capturing complicated patterns in dynamic datasets. The unique feature of this study is that it compares five different probability distributions in detail using Hidden Markov Models. The study aims to determine the best method for modeling sequential data by examining the effectiveness and suitability of each distribution. This paper also presents the creation of a Lomax HMM, which improves the modeling capabilities of conventional HMMs by incorporating the special characteristics of the Lomax distribution.

II. Methods

A Hidden Markov Model is a probabilistic model that is commonly used to represent time series data or observation sequences. It is made up of two major components: hidden states and observable states. Hidden States are unobservable or latent variables that represent the underlying states of the

system under consideration. Hidden states can change over time using a Markov process, which means that the likelihood of transitioning from one state to another is determined only by the present state and not by the previous states. Observable states are the data points or observations that can be directly observed and linked to each hidden state. The observations are derived from probability distributions that are conditional on the present concealed state. The type of probability distribution used depends on the nature of the data being modelled. The transition between concealed states and the creation of observations are probabilistic processes that are controlled by transition probabilities and emission probabilities, respectively. The parameters of these probability distributions are usually calculated from observed data using techniques like maximum likelihood estimation (MLE) or the Expectation-Maximization (EM) algorithm. The following are the distributions used for representing the observable outcomes.

An HMM is designed to model a system as it changes over discrete time steps. At each time step, the system is in one of a finite number of hidden states, and it emits an observable symbol based on the probability distribution associated with the state. Transition probabilities stimulate hidden state transitions, while emission probabilities govern observable symbol emission. An HMM is defined formally by:

- A set of hidden states S = $\{s_1, s_2, \dots, s_v\}$
- A set of observable symbols O = $\{o_1, o_2, \dots, o_u\}$
- A transition probability matrix A = $[a_{kl}]$, where a_{kl} represents the probability of transitioning from state a_k to a_l.
- An emission probability matrix B = $[b_{mn}]$, where b_{mn} represents the probability of emitting a symbol s_n when in state s_m.
- An initial probability distribution $\pi = [\pi_k]$ represents the probability of starting in the states$_k$.

In a HMM, the system under consideration is considered a Markov process with unobservable (hidden) states. The set S includes all conceivable hidden states the system can be in at any one time. Each state s_i (where i ranges from 1 to v) has attributes that may influence the system's behavior. The hidden states are not immediately observable, but they impact the observable outputs (symbols), allowing the states to be inferred indirectly from the observation sequence. All of the observable outputs or symbols that the system is capable of producing are contained in the set O. Although we are unable to directly witness the concealed states, we can observe the symbols that they produce. We deduce the hidden states from the succession of these observable symbols. One important part of the HMM is matrix A, which shows the chances of changing from one hidden state to another. In particular, the chance of changing from state s_k to state s_l is represented by a_{kl}. The transition probability matrix is of size v × v, where v is the number of hidden states. To guarantee the total probability of changing from a particular state to any other state, all of the entries in each row of this matrix must add up to 1.

The Markov property, which postulates that the probability of changing to the next state depends only on the present state and not on any previous states, is contained in this matrix. The emission probability matrix B represents the correspondence between the observable symbols and the hidden states. The probability of observing the symbol o_n is represented by the entry b_{mn}, assuming that the system is in the hidden state s_m. The size of this matrix is v×u, where u is the number of observable symbols and v is the number of concealed states. The entries in every row of the emission probability matrix must add up to 1, just like in the transition probability matrix. The model can provide a series of observations based on the underlying hidden state sequence because the emission matrix essentially transfers the hidden states to the observable symbols. The initial probability distribution π defines the probability of the system starting in each hidden state. The

vector π = [$\pi_1, \pi_2, \ldots, \pi_v$] is of length v, where π_k indicates the chance that the system starts in the state s_k.This distribution is crucial in determining the likelihood of different sequences of concealed states. It serves as the starting point for generating sequences in the HMM, and like other probability distributions, the entries must add to one.

The model parameters in the HMM must be constantly updated and improved in real-world applications, especially in scenarios involving temporal data, such agricultural imports over time. This is necessary to describe specific time points adequately. This is especially crucial when using the model to forecast or examine long-term patterns in agricultural data. Estimating the parameters governing the state transitions and the emission of visible symbols is a major difficulty in this context.

One popular method for estimating an HMM's parameters is the Baum-Welch algorithm, a variation of the Expectation-Maximization (EM) algorithm. The program looks for the values of A, B, and π in the model that maximize the probability of the observed sequence of symbols. Improved estimates are produced by iteratively adjusting the model parameters in light of the observed data. A continuous process [8] of developing model parameters in the transition state to explain the specific time point in importing agricultural data. A HMM is typically denoted by, $\mu = (A, B, \pi)$. This model provides us with the state transition probability, the observational probability, and the probability of starting in a specific state. The Baum-Welch algorithm, commonly known as the EM algorithm, focuses on parameter estimation via direct numerical maximum likelihood estimation. To maximize and determine the posterior analysis of the hidden variables.

I. Transition Probability Estimation

The likelihood of a transition from one state to another, such as from state k to state l, is represented by the symbol a_{kl}. The following formula can be used to estimate it [14], [15],

$$a_{kl} = \frac{Expected\ number\ of\ transitions\ from\ state\ k\ to\ l}{Expected\ number\ of\ transitions\ from\ state\ k} \tag{1}$$

The ratio indicates how frequently the system moves from state k to state l throughout the observation time. It can be mathematically represented as follows:

$$a_{kl} = \frac{\sum_{t=1}^{T} p_t(k,l)}{\sum_{t=1}^{T} \gamma_k(t)} \tag{2}$$

where, $p_t(k, l)$ represents the likelihood of being in state k at time t and moving to state l at time t+1 and $\gamma_k(t)$ represents the chance of being in state k at time t. An alternative expression for a_{kl} that integrates the forward-backward variables α and β, signifying the forward and backward probabilities is:

$$a_{kl} = \frac{\sum_{t=1}^{T} \alpha_k(t) a_{kl} b_k(O_{t+1}) \beta_l(t+1)}{\sum_{t=1}^{T} \alpha_k(t) \beta_k(t)} \tag{3}$$

The combined likelihood [16] of being in state k at time t and changing to state l at time t+1 is as follows:

$$Pr(k, l) = Pr(S_t = k, S_{t+1} = l \ / \ O, \mu)$$

$$= \frac{Pr(S_t = k, S_{t+1} = l / O, \mu}{P(O/\mu)}$$

$$= \frac{\alpha_k(t)a_{kl}b_l(O_{t+1})\beta_l(t+1)}{\sum_{m=1}^{v}\sum_{n=1}^{v}\alpha_m(t)a_{mn}b_n(O_{t+1})\beta_n(t+1)} \quad (4)$$

Given the observation sequence O and the model μ, the probability of being in state k at time t is represented by the symbol $\gamma_k(t)$ is defined as:

$$\gamma_k(t) = \Pr(S_t = k/O,\mu) = \sum_{l=1}^{v} \Pr(S_{t,} = k, S_{t+1} = l/\ O, \mu)$$

$$= \sum_{l=1}^{v} Pr(k,l) \quad (5)$$

II. Initial State Probability

The initial state probability represents the likelihood of beginning in state k at time t = 1 π_k. It can be stated as:

$$\pi_k = \gamma_k(t) \quad (6)$$

III. Emission Probability Estimation

When moving from state k to state l, the probability of determining symbols n is represented by the emission probability, b_{kln}.

$$b_{kln} = \frac{Expected\ number\ of\ transitions\ from\ k\ to\ l\ with\ symbol\ n\ observed}{Expected\ number\ of\ transitions\ from\ k\ to\ l} \quad (7)$$

This can be expressed as:

$$b_{kln} = \frac{\sum_{t:O_t=n,1\le t\le v} Pr(k,l)}{\sum_{t=1}^{v} Pr(k,l)} \quad (8)$$

Given a collection of observations, the Viterbi algorithm is a dynamic programming technique determining the most likely order of hidden states in an HMM. The process starts with an initialization step in which the first observation's emission probabilities and initial distribution are used to compute the likelihood of beginning in each stage. The algorithm iterates through each time step in the recursion step, utilizing back pointers for retaining the best path and the transition and emission probabilities to determine the most likely transitions between states. After analysing all observations, the termination step finds the state with the highest probability at the last time step. Ultimately, the algorithm uses the backpointers to trace the series of hidden states in the backtracking step, generating the most likely path that explains the observed data. MATLAB software is used in this investigation to determine the Viterbi path.

IV. HMM Models with Different Distributions

To determine whether different probability distributions are appropriate for representing the dynamics of agricultural imports, they are incorporated into the HMM framework. These distributions offer a variety of approaches for simulating the data's underlying patterns and state transitions. These are as follows:

Normal - HMM

"Normal-HMM" is a Hidden Markov Model (HMM) in which the emission probabilities are represented using a normal (Gaussian) distribution. In the case of HMM, each hidden state generates observations with a normal distribution. This means the chance of witnessing a specific value given a hidden state is determined using normal distribution characteristics, specifically the mean and

variance. To examine agricultural import dynamics in India, we built a Hidden Markov Model with a normal distribution with μ = 4.575 and σ = 1.53. We used the normal distribution's probability density function to model the emission probabilities within the HMM framework. The pdf of the Normal-HMM is expressed as follows:

$$P(x_t/y_t = k) = \frac{1}{\sqrt{2\pi}\sigma_k} e^{-(\frac{(x_t-\mu_k)^2}{2\sigma_k{}^2})} \quad (9)$$

N (μ, σ^2) as the Normal Distribution with mean μ and variance σ^2, x_t as the observed data at time t. y_t as the hidden state at time t. By incorporating this PDF into the HMM framework, we can successfully simulate and evaluate patterns and transitions in agricultural import dynamics in India, allowing for more informed decision-making and policy development in the agriculture sector. Given a sequence of observed import values $x_1, x_2, \dots, x_N$ and a sequence of hidden states $y_1, y_2, \dots, y_N$, the log-likelihood can be calculated as:

$$\text{Log-likelihood} = \sum_{t=1}^{N} log(P(x_n/y_n; \mu_{y_n}, \sigma_{y_n})) + \sum_{t=2}^{N} log(P(y_n/y_{n-1}; A)) \quad (10)$$

where, $P(x_n/y_n; \mu_{y_n}, \sigma_{y_n})$ is the probability of observing x_n given the hidden state y_n. $P(y_n/y_{n-1}; A)$ is the transition probability from state y_{n-1} to state y_n.

Gamma – HMM

In a Gamma-HMM, each hidden state generates observations with a Gamma distribution. The Gamma distribution, a continuous probability distribution, is commonly used to model positive-valued random variables. It's defined by two parameters: shape (α) and scale (β). The shape parameter indicates the shape of the distribution, whereas the scale parameter governs its spread. When we integrate a Gamma distribution within a Hidden Markov Model (HMM) to interpret agricultural import dynamics in India, with parameters α = 8.6192 and β = 1.8836, we use the Gamma distribution's probability density function to characterize emission probabilities within the HMM framework. The pdf of the Gamma-HMM is expressed as follows:

$$P(x_t/y_t = k) = \frac{\beta_k{}^{\alpha_k}}{\Gamma(\alpha_k)} x_t{}^{\alpha_k - 1} e^{-\beta_k x_k} \quad (11)$$

Γ(.) for the Gamma function. During the training phase of the Gamma HMM, the parameters α and β are commonly calculated using approaches such as maximum likelihood estimation (MLE) or EM algorithm, which are similar to other HMM versions. These algorithms seek to optimize the model parameters to maximize the likelihood of the observed data within the model. Overall, the Gamma HMM is a versatile framework for modelling sequential data with positive values, with emissions expected to follow a Gamma distribution. The forward algorithm, which is typically used to calculate the likelihood in HMMs, must be changed accordingly. Given a sequence of observations $X = \{x_1, x_2, \dots, x_N\}$ and a Gamma HMM with parameters ʎ = (A, B, π, α, β). The steps listed below are used to determine the log-likelihood [3]:

- Initialization

$$\alpha_1(k) = \pi_k P(x_1/y_{1=k}) \quad (12)$$

$$\alpha_1(k) = \pi_k \times \frac{\beta_k{}^{\alpha_k}}{\Gamma(\alpha_k)} x_1{}^{\alpha_k - 1} e^{-\beta_k x_1} \quad (13)$$

- Recursion

$$\alpha_{t+1}(l) = (\sum_{k=1}^{N} \alpha_t(k)\ A_{kl})\ P(\ x_{t+1}/y_{t+1=l}) \quad (14)$$

$$\alpha_{t+1}(l) = (\ \sum_{k=1}^{N} \alpha_t(k)\ A_{kl})\ \frac{\beta_l{}^{\alpha_l}}{\Gamma(\alpha_l)} x_{t+1}{}^{\alpha_l - 1} e^{-\beta_l x_{k+1}}$$

- Termination

$$log\ P\ (x/\Lambda) = log(\sum_{k=1}^{N} \alpha_T(k)) \quad (15)$$

Weibull – HMM

A Weibull Hidden Markov Model (Weibull HMM) is a variation of the classic Hidden Markov Model in which the emission distribution for each hidden state follows a Weibull distribution. Using a Weibull distribution with parameters α = 3.257 and β= 5.113. The emission probability for state k emitting observation x_t is given by the probability density function of the Weibull distribution:

$$P\ (x_t\ /y_t = k) = \frac{\alpha_k}{\beta_k} (\frac{x_t}{\beta_k})^{\alpha_k - 1} e^{-(\frac{x_t}{\beta_k})^{\alpha_k}} \quad (16)$$

where x_t is the observed value at time step t, α_k is the shape parameter and β_k is the scale parameter. We must modify the computation to take into account the emission probabilities modelled by Weibull distributions to determine the log-likelihood of a series of observations given Weibull HMM. We will apply the forward algorithm, which is comparable to the normal HMM log-likelihood computation, but using the probability density function for the Weibull distribution.

- Initialization: $\alpha_1(k) = \pi_k \frac{\alpha_k}{\beta_k} (\frac{x_t}{\beta_k})^{\alpha_k - 1} e^{-(\frac{x_t}{\beta_k})^{\alpha_k}}$ (17)
- Recursion: $\alpha_{t+1}(l) = (\ \sum_{k=1}^{N} \alpha_t(k)\ A_{kl})\ \frac{\alpha_k}{\beta_k} (\frac{x_t}{\beta_k})^{\alpha_k - 1} e^{-(\frac{x_t}{\beta_k})^{\alpha_k}}$ (18)
- Termination: $log\ P\ (x/\Lambda) = log(\sum_{k=1}^{N} \alpha_T(k))$ (19)

Lomax – HMM

A continuous probability distribution with two parameters, shape (α) and scale (λ), is the Lomax distribution. Using parameters α = 0.3520 and Λ = 0.0792, we utilize the probability density function of the Lomax distribution to model the emission probabilities within the HMM to peruse agricultural import dynamics in India. The Lomax HMM can be examined by,

$$P\ (x_t\ /y_t = k)\ = \frac{\alpha_k}{\Lambda_k} (1 + \frac{x_t}{\Lambda_k})^{-(\alpha_k + 1)} \quad (20)$$

The Lomax distribution, also known as the Pareto Type II distribution or the generalized Pareto distribution, is a probability distribution commonly used to simulate extreme value events. A Lomax-HMM is thus an HMM in which the emission distributions for each hidden state follow the Lomax distribution. This means that at each time step in the sequence, the observable data is created by a Lomax distribution whose parameters may differ based on the underlying hidden state. The log-likelihood can be calculated using the following modified algorithm:

- Initialization: $\alpha_1(k) = \pi_k\ \frac{\alpha_k}{\Lambda_k} (1 + \frac{x_t}{\Lambda_k})^{-(\alpha_k + 1)}$ (21)
- Recursion: $\alpha_{t+1}(l) = (\ \sum_{k=1}^{N} \alpha_t(k)\ A_{kl})\ \frac{\alpha_k}{\Lambda_k} (1 + \frac{x_t}{\Lambda_k})^{-(\alpha_k + 1)}$ (22)

- Termination: log P (x / ʎ) =log($\sum_{k=1}^{N} \alpha_T(k)$) (23)

log P(x/λ) calculates the log-likelihood of the observed sequence based on Lomax HMM parameters. During the model's training or parameter estimation phase, it is typical to aim to maximize this log-likelihood about the model parameters.

Log-Normal – HMM

A Lognormal-HMM is a probabilistic model that is widely employed in time-series research, particularly when working with data with a log-normal distribution. HMMs are a sort of probabilistic graphical model that is used to represent sequences of observations. They are made up of a series of hidden states that are not directly observable, as well as a series of seen emissions that are affected by the hidden states. With parameters mean of the logarithm (mean log) = 1.4616 and standard deviation of the logarithm (standard deviation log) = 0.3503. The following is the pdf,

$$P\ (x_t\ /\ y_t = k)\ = \frac{1}{x_k \sigma_k \sqrt{2\pi}}\, e^{-\frac{(ln\, x_t - \mu_k)^2}{2{\sigma_k}^2}} \tag{24}$$

Log-normal HMM training involves estimating parameters μ_k and σ_k^2 using approaches such as MLE, EM, or Baum-Welch algorithm, as with other HMM versions. These algorithms seek to optimize the model parameters to maximize the likelihood of the training data given the model. To determine the log-likelihood of a sequence of observations in a Log-Normal Hidden Markov Model, we must modify the method to account for the emission probabilities represented by Log-Normal distributions. We will compute the likelihood using the forward technique:

- Initialization: $\alpha_1(k) = \pi_k \frac{1}{x_k \sigma_k \sqrt{2\pi}}\, e^{-\frac{(\ln x_t - \mu_k)^2}{2{\sigma_k}^2}}$ (25)
- Recursion: $\alpha_{t+1}(l) = (\sum_{k=1}^{N} \alpha_t(k)\ A_{kl})\ \frac{1}{x_k \sigma_k \sqrt{2\pi}}\, e^{-\frac{(\ln x_t - \mu_k)^2}{2{\sigma_k}^2}}$ (26)
- Termination: log P (x / ʎ) =log($\sum_{k=1}^{N} \alpha_T(k)$) (27)

V. Correlation Matrix

A correlation matrix is a square matrix that shows the correlation coefficients between pairs of variables. In statistics, correlation assesses the degree and direction of a linear relationship between two variables. The correlation coefficient, abbreviated ϱ, varies from -1 to 1. To create a correlation matrix with various probability distributions such as Normal, Gamma, Weibull, Lomax, and Lognormal, the correlation between pairs of variables must be specified. Because these distributions reflect continuous random variables, we can use the correlation coefficient between any two variables to determine the strength and direction of their linear link. Here, Statistical R software is used to construct the correlation coefficient between variables X_i and X_j. The correlation matrix serves as a valuable tool for understanding the relationships between different probability distributions, offering insights into their interdependencies and guiding the selection of the most appropriate distributions for modelling agricultural import dynamics within the HMM framework.

VI. Algorithm

The algorithm for the study involves several key steps:

1. Collecting import data for agricultural products from reliable sources covering specific years.
2. Select five statistically appropriate probability distributions for modelling the data.
3. Using R software to determine the parameters of each selected probability distribution.
4. Reconstructing the import data using the values of the parameters.
5. Determine the correlation matrix among the chosen probability distributions to investigate their connections.
6. Use the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test to conduct stationarity tests to determine whether the data is stationary.
7. Use Hidden Markov Models to model the data. Determine the HMM's parameters, such as the initial probability vector, the emission probability matrix, and the transition probability matrix.
8. When evaluating various models, choose the best-fitting HMM using information metrics such as the Akaike Information Criterion and the Bayesian Information Criterion.
9. In a Hidden Markov Model, figuring out the Viterbi path provides the best estimate of the hidden states that produced the observed data.

These steps are crucial for analyzing agricultural import dynamics and selecting the most appropriate probability distributions within the Hidden Markov Model framework. They provide a systematic approach to understanding the relationships between distributions and their suitability for modelling complex systems.

VII. Data Source

The "Economics & Statistics Division within the Ministry of Agriculture & Farmers Welfare" website, may be accessed at http://desagri.gov.in.

III. Results and Discussion

Table 1: Correlation Coefficient for Different Probability Distributions.

Distribution	Normal	Gamma	Weibull	Lomax	Lognormal
Normal	1.00	-0.5157	0.9975	-0.2504	0.7913
Gamma	-0.5157	1.000	-0.5646	-0.6641	-0.7316
Weibull	0.9975	-0.5646	1.00	-0.2057	0.81337
Lomax	-0.25048	-0.6641	-0.2057	1.00	0.1569
Lognormal	0.7913	-0.7316	0.8133	0.1569	1.00

The correlation table shows notable relationships between probability distributions, such as strong negative correlations between Normal-Gamma and Gamma-Lomax distributions and a strong positive correlation between Weibull and Lognormal distributions, which allows for more informed model selection for dependent variables.

Table 2: Stationarity of KPSS Test Results in Different Distributions

Test For	Test Statistics	Lag Parameter	P-Value	Result
Normal	0.3096	3	0.1	Stationary
Gamma	0.1808	3	0.1	Stationary
Weibull	0.3096	3	0.1	Stationary
Lomax	0.1075	3	0.1	Stationary
Lognormal	0.2245	3	0.1	Stationary

According to the KPSS test results, all distributions exhibit stationarity at a significance level of 0.1. The KPSS test assesses if a time series is stationary around a mean or has a linear trend; in this case, all distributions pass the test, implying that they meet the stationarity assumption. As a result, the data is appropriate for researching HMM.

The parameter ranges for 2, 3, and 4 states are shown in the table below for various HMM. Each model is specifically designed to capture various data distributions, with parameter ranges reflecting import-level characteristics. By comparing these ranges, the table emphasizes the strengths and limitations of each HMM in predicting import dynamics, delivering useful insights into enhancing forecast accuracy.

Table 3: Parameter ranges for various hidden Markov models with different distributions and numbers of states.

States		Normal-HMM	Gamma-HMM	Weibull-HMM	Lomax-HMM	Log-Normal HMM
2	Low Export (S_1)	[0.016, 0.138]	[0.000021, 0.013]	[0.018, 0.132]	[0.008, 0.029]	[0.026, 0.153]
	High Export (S_2)	[0.138, 0.260]	[0.013, 0.027]	[0.132, 0.247]	[0.029, 0.05]	[0.153, 0.280]
3	Low Export (S_1)	[0.016, 0.097]	[0.000021, 0.009]	[0.018, 0.085]	[0.008, 0.021]	[0.026, 0.102]
	Medium Export (S_2)	[0.097, 0.179]	[0.009, 0.018]	[0.085, 0.151]	[0.021, 0.035]	[0.102, 0.177]
	High Export (S_3)	[0.179, 0.260]	[0.018, 0.027]	[0.151, 0.247]	[0.035, 0.050]	[0.177, 0.280]
4	Very low Export (S_1)	[0.016, 0.077]	[0.00002, 0.006]	[0.018, 0.075]	[0.008, 0.018]	[0.028, 0.090]
	Low Export (S_2)	[0.077, 0.138]	[0.006, 0.0125]	[0.075, 0.132]	[0.018, 0.029]	[0.090, 0.153]
	High Export (S_3)	[0.138, 0.199]	[0.012, 0.020]	[0.132, 0.189]	[0.029, 0.040]	[0.153, 0.217]
	Very High Export (S_4)	[0.199, 0.260]	[0.020, 0.027]	[0.189, 0.247]	[0.040, 0.050]	[0.217, 0.280]

This table highlights the delicate impact of distribution and state count variations on export-level modelling in HMM. From Normal to Log-Normal distributions, and from two to four states, each configuration provides unique probabilistic insights on export behaviour, underscoring the need for careful model selection in effectively representing export dynamics.

In a 2-State Hidden Markov Model, the framework is divided into two states: "Low" and "High" with observations showing whether import levels are "Decreasing (D)" or "Increasing (I)". This setup provides a broad view of trends. The 3-State Model adds a "Moderate" state, allowing for a more nuanced analysis where observations can be "Decreasing", "Stable (S)," or "Increasing" capturing a range of fluctuations between low and high import levels. The 4-State Model introduces even more detail with states like "Very Low", "Low," "High," and "Very High," and observations of "Very Decreasing (VD)", "Decreasing", "Increasing", and "Very Increasing (VI)" offering a comprehensive look at significant changes and trends in import levels.

TPM, EPM, and π for Normal-HMM (2,3 & 4 States)

$$\text{TPM} = \begin{matrix} & s_1 & s_2 \\ s_1 \\ s_2 \end{matrix}\begin{bmatrix} 0.4 & 0.6 \\ 0.238 & 0.727 \end{bmatrix},\ \text{EPM} = \begin{matrix} & D & I \\ s_1 \\ s_2 \end{matrix}\begin{bmatrix} 0.1 & 0.9 \\ 0.59 & 0.40 \end{bmatrix},\quad \pi_0 = [\,0.312\,, 0.687]$$

$$\text{TPM} = \begin{matrix} & s_1 & s_2 & s_3 \\ s_1 \\ s_2 \\ s_3 \end{matrix}\begin{bmatrix} 0.5 & 0 & 0.5 \\ 0.076 & 0.461 & 0.461 \\ 0.071 & 0.428 & 0.5 \end{bmatrix},\text{EPM} = \begin{matrix} & D & S & I \\ s_1 \\ s_2 \\ s_3 \end{matrix}\begin{bmatrix} 0.25 & 0 & 0.75 \\ 0.15 & 0.07 & 0.76 \\ 0.2 & 0.53 & 0.26 \end{bmatrix}, \pi_0 = [\,0.125, 0.406, 0.468]$$

$$\text{TPM} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{matrix}\begin{bmatrix} 0.33 & 0 & 0 & 0.66 \\ 0.16 & 0.33 & 0.16 & 0.33 \\ 0.11 & 0.11 & 0.33 & 0.44 \\ 0 & 0.15 & 0.38 & 0.46 \end{bmatrix}, \text{EPM} = \begin{matrix} & VD & D & I & VI \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{matrix}\begin{bmatrix} 0 & 0 & 0.33 & 0.66 \\ 0 & 0 & 0.66 & 0.33 \\ 0 & 0.2 & 0.55 & 0.22 \\ 0.5 & 0.2 & 0.28 & 0 \end{bmatrix}, \pi_0 = [0.09,\ 0.18,\ 0.28,\ 0.43]$$

TPM, EPM, and π for Gamma-HMM (2,3 & 4 States)

$$\text{TPM} = \begin{matrix} & s_1 & s_2 \\ s_1 \\ s_2 \end{matrix}\begin{bmatrix} 0.93 & 0.03 \\ 0.5 & 0.5 \end{bmatrix},\ \text{EPM} = \begin{matrix} & D & I \\ s_1 \\ s_2 \end{matrix}\begin{bmatrix} 0.43 & 0.56 \\ 0.5 & 0.5 \end{bmatrix},\quad \pi_0 = [\,0.937\,, 0.062]$$

$$\text{TPM} = \begin{matrix} & s_1 & s_2 & s_3 \\ s_1 \\ s_2 \\ s_3 \end{matrix}\begin{bmatrix} 0.846 & 0.115 & 0.03 \\ 1 & 0 & 0 \\ 0.5 & 0 & 0.5 \end{bmatrix}, \text{EPM} = \begin{matrix} & D & S & I \\ s_1 \\ s_2 \\ s_3 \end{matrix}\begin{bmatrix} 0.18 & 0.18 & 0.62 \\ 0.66 & 0.33 & 0 \\ 0.5 & 0 & 0.5 \end{bmatrix}, \pi_0 = [\,0.843, 0.093, 0.062]$$

$$\text{TPM} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{matrix}\begin{bmatrix} 0.82 & 0.09 & 0.09 & 0 \\ 0.4 & 0.4 & 0 & 0.2 \\ 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0 & 0 & 0.5 \end{bmatrix},\text{EPM} = \begin{matrix} & VD & D & I & VI \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{matrix}\begin{bmatrix} 0.17 & 0.13 & 0.35 & 0.35 \\ 0.2 & 0.6 & 0.2 & 0 \\ 1 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \end{bmatrix}, \pi_0 = [0.72,\ 0.16,\ 0.63,\ 0.06]$$

TPM, EPM, and π for Weibull-HMM (2,3 & 4 States)

$$\text{TPM} = \begin{matrix} & s_1 & s_2 \\ s_1 \\ s_2 \end{matrix}\begin{bmatrix} 0.33 & 0.66 \\ 0.16 & 0.84 \end{bmatrix},\ \text{EPM} = \begin{matrix} & D & I \\ s_1 \\ s_2 \end{matrix}\begin{bmatrix} 0.16 & 0.83 \\ 0.5 & 0.5 \end{bmatrix},\quad \pi_0 = [\,0.187, 0.812]$$

$$\text{TPM} = \begin{matrix} & s_1 & s_2 & s_3 \\ s_1 \\ s_2 \\ s_3 \end{matrix}\begin{bmatrix} 0.33 & 0 & 0.66 \\ 0.14 & 0.28 & 0.57 \\ 0.04 & 0.19 & 0.76 \end{bmatrix}, \text{EPM} = \begin{matrix} & D & S & I \\ s_1 \\ s_2 \\ s_3 \end{matrix}\begin{bmatrix} 0 & 0.33 & 0.66 \\ 0 & 0 & 1 \\ 0.41 & 0.18 & 0.41 \end{bmatrix}, \pi_0 = [\,0.09, 0.218, 0.68]$$

$$\text{TPM} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{matrix}\begin{bmatrix} 0.5 & 0 & 0 & 0.5 \\ 0 & 0.25 & 0.25 & 0.5 \\ 0 & 0.18 & 0.36 & 0.45 \\ 0.07 & 0.07 & 0.35 & 0.5 \end{bmatrix},\text{EPM} = \begin{matrix} & VD & D & I & VI \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{matrix}\begin{bmatrix} 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0.75 & 0.25 \\ 0 & 0.18 & 0.45 & 0.36 \\ 0.46 & 0.26 & 0.26 & 0 \end{bmatrix}, \pi_0 = [0.06, 0.13,\ 0.34,\ 0.46]$$

TPM, EPM, and π for Lomax-HMM (2,3 & 4 States)

$$\text{TPM} = \begin{matrix} & s_1 & s_2 \\ s_1 \\ s_2 \end{matrix}\begin{bmatrix} 0.92 & 0.08 \\ 0.37 & 0.62 \end{bmatrix},\ \text{EPM} = \begin{matrix} & D & I \\ s_1 \\ s_2 \end{matrix}\begin{bmatrix} 0.45 & 0.54 \\ 0.87 & 0.13 \end{bmatrix},\quad \pi_0 = [\,0.75, 0.25]$$

$$\text{TPM} = \begin{matrix} \\ s_1 \\ s_2 \\ s_3 \end{matrix}\begin{matrix} \begin{matrix} s_1 & s_2 & s_3 \end{matrix} \\ \begin{bmatrix} 0.78 & 0.21 & 0 \\ 0.4 & 0.5 & 0.1 \\ 0.5 & 0 & 0.5 \end{bmatrix} \end{matrix}, \text{EPM} = \begin{matrix} \\ s_1 \\ s_2 \\ s_3 \end{matrix}\begin{matrix} \begin{matrix} \text{D} & \text{S} & \text{I} \end{matrix} \\ \begin{bmatrix} 0.05 & 0.35 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.5 & 0.5 & 0 \end{bmatrix} \end{matrix}, \pi_0 = [\,0.63, 0.315, 0.062]$$

$$\text{TPM} = \begin{matrix} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{matrix}\begin{matrix} \begin{matrix} s_1 & s_2 & s_3 & s_4 \end{matrix} \\ \begin{bmatrix} 0.81 & 0.13 & 0.06 & 0 \\ 0.43 & 0.43 & 0.14 & 0 \\ 0.17 & 0.17 & 0.5 & 0.17 \\ 0 & 0.5 & 0 & 0.5 \end{bmatrix} \end{matrix}, \text{EPM} = \begin{matrix} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{matrix}\begin{matrix} \begin{matrix} \text{VD} & \text{D} & \text{I} & \text{VI} \end{matrix} \\ \begin{bmatrix} 0.05 & 0.35 & 0.41 & 0.17 \\ 0.14 & 0.28 & 0.57 & 0 \\ 0.17 & 0.66 & 0.17 & 0 \\ 0.5 & 0.5 & 0 & 0 \end{bmatrix} \end{matrix}, \pi_0 = [0.53, 0.22,\ 0.18,\ 0.06]$$

TPM, EPM, and π for Lognormal-HMM (2,3 & 4 States)

$$\text{TPM} = \begin{matrix} \\ s_1 \\ s_2 \end{matrix}\begin{matrix} \begin{matrix} s_1 & s_2 \end{matrix} \\ \begin{bmatrix} 0.54 & 0.45 \\ 0.25 & 0.75 \end{bmatrix} \end{matrix}, \text{EPM} = \begin{matrix} \\ s_1 \\ s_2 \end{matrix}\begin{matrix} \begin{matrix} \text{D} & \text{I} \end{matrix} \\ \begin{bmatrix} 0.18 & 0.82 \\ 0.52 & 0.47 \end{bmatrix} \end{matrix}, \quad \pi_0 = [\,0.343, 0.656]$$

$$\text{TPM} = \begin{matrix} \\ s_1 \\ s_2 \\ s_3 \end{matrix}\begin{matrix} \begin{matrix} s_1 & s_2 & s_3 \end{matrix} \\ \begin{bmatrix} 0.28 & 0.14 & 0.57 \\ 0.5 & 0.25 & 0.25 \\ 0.15 & 0.1 & 0.75 \end{bmatrix} \end{matrix}, \text{EPM} = \begin{matrix} \\ s_1 \\ s_2 \\ s_3 \end{matrix}\begin{matrix} \begin{matrix} \text{D} & \text{S} & \text{I} \end{matrix} \\ \begin{bmatrix} 0 & 0.14 & 0.85 \\ 0.25 & 0 & 0.75 \\ 0.38 & 0.14 & 0.47 \end{bmatrix} \end{matrix}, \pi_0 = [\,0.22, 0.12, 0.65]$$

$$\text{TPM} = \begin{matrix} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{matrix}\begin{matrix} \begin{matrix} s_1 & s_2 & s_3 & s_4 \end{matrix} \\ \begin{bmatrix} 0.33 & 0.16 & 0.16 & 0.33 \\ 0.2 & 0.4 & 0.2 & 0.2 \\ 0.33 & 0 & 0.16 & 0.5 \\ 0.07 & 0.14 & 0.14 & 0.64 \end{bmatrix} \end{matrix}, \text{EPM} = \begin{matrix} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{matrix}\begin{matrix} \begin{matrix} \text{VD} & \text{D} & \text{I} & \text{VI} \end{matrix} \\ \begin{bmatrix} 0 & 0.16 & 0.16 & 0.66 \\ 0 & 0.2 & 0.2 & 0.6 \\ 0.33 & 0 & 0.5 & 0.16 \\ 0.4 & 0.26 & 0.33 & 0 \end{bmatrix} \end{matrix}, \pi_0 = [0.18, 0.15,\ 0.18,\ 0.46]$$

Table 4: Model Comparison for Various Distributions Based on States and Information Criteria.

Distribution	States	Log-Likelihood	AIC	BIC
Normal	2	-21.621	63.243	77.902
	3	-32.522	107.04	137.82
	4	-46.865	165.73	218.49
Gamma	2	-26.886	73.772	88.429
	3	-43.243	128.48	159.26
	4	-43.828	159.65	212.42
Weibull	2	-19.714	59.428	74.085
	3	-29.521	101.04	131.82
	4	-45.444	162.88	215.65
Lomax	**2**	**-19.136**	**58.272**	**72.930**
	3	-43.226	128.45	159.23
	4	-43.240	158.48	211.24
Lognormal	2	-20.448	60.896	75.553
	3	-27.879	97.759	128.54
	4	-42.204	156.40	209.17

Based on the AIC and BIC values provided, the Lomax distribution appears to be the best fit among the distributions listed. This conclusion is based on the consistently lower AIC and BIC values across a variety of states when compared to other distributions.

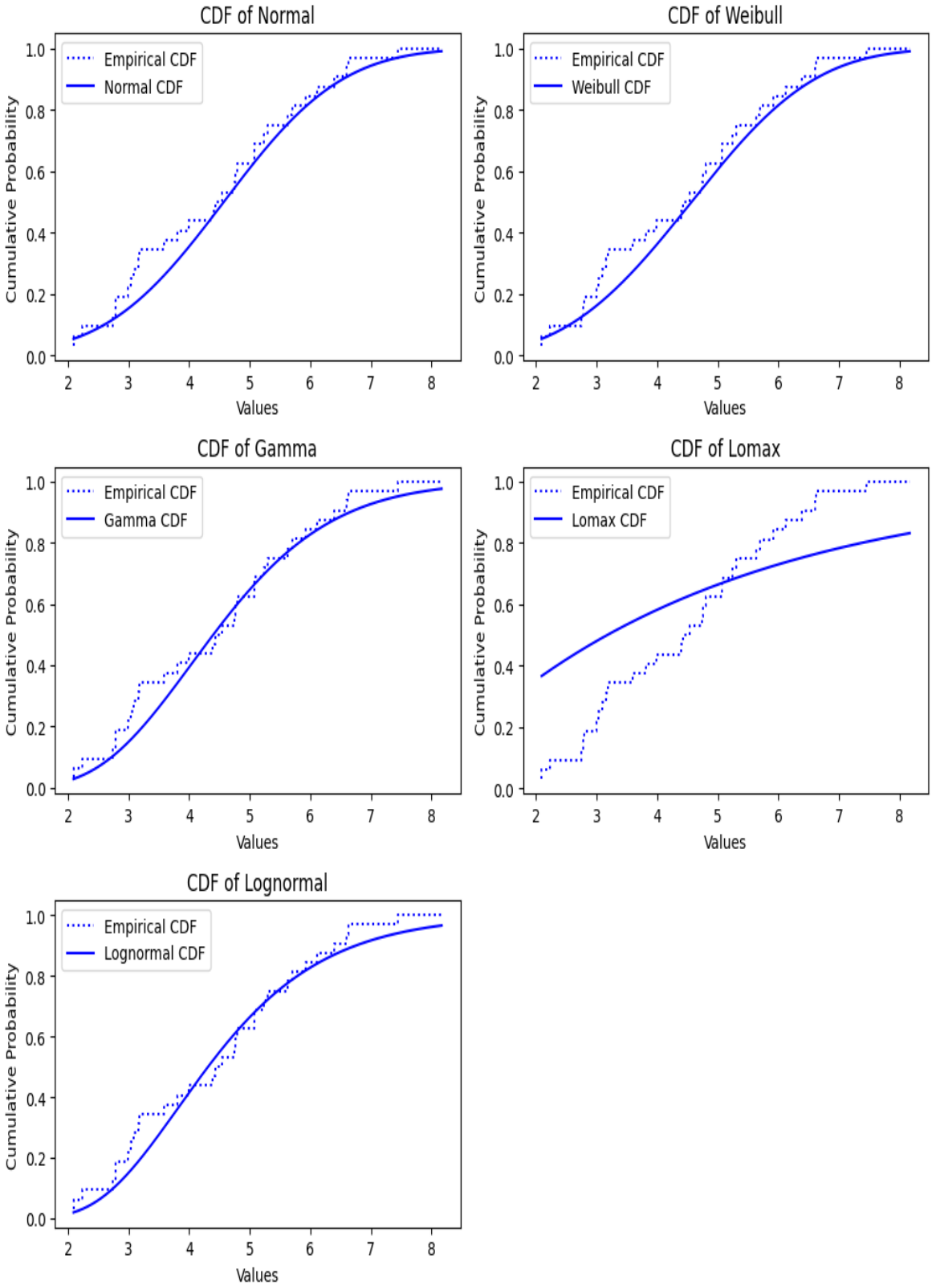


Figure 1: The graphs show the comparative evaluation of CDFs using HMM in different distributions.

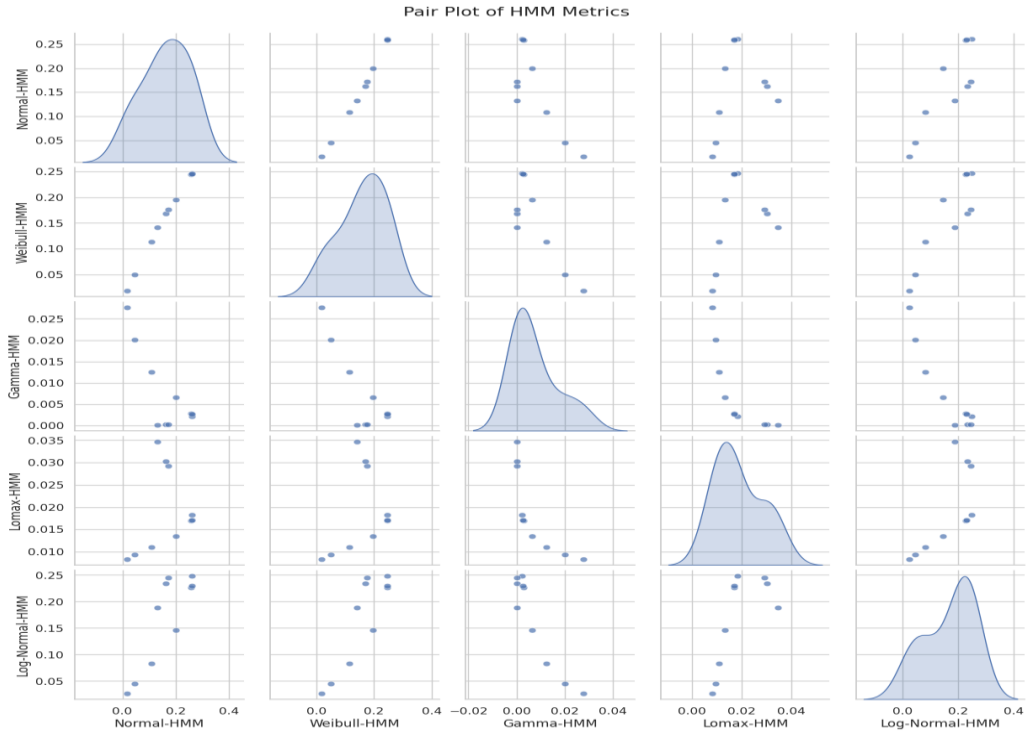


Figure 2: This graph shows the pair plot of different HMM distributions for comparison

The pair plot is a helpful visualization that shows the relationships between the performance metrics of various Hidden Markov Models. The scatter plots reveal how different models are correlated, while the diagonal plots indicate the distribution of each model's performance. This enables a comparison of trends, correlations, and distribution patterns, assisting in identifying the model that performs the best.

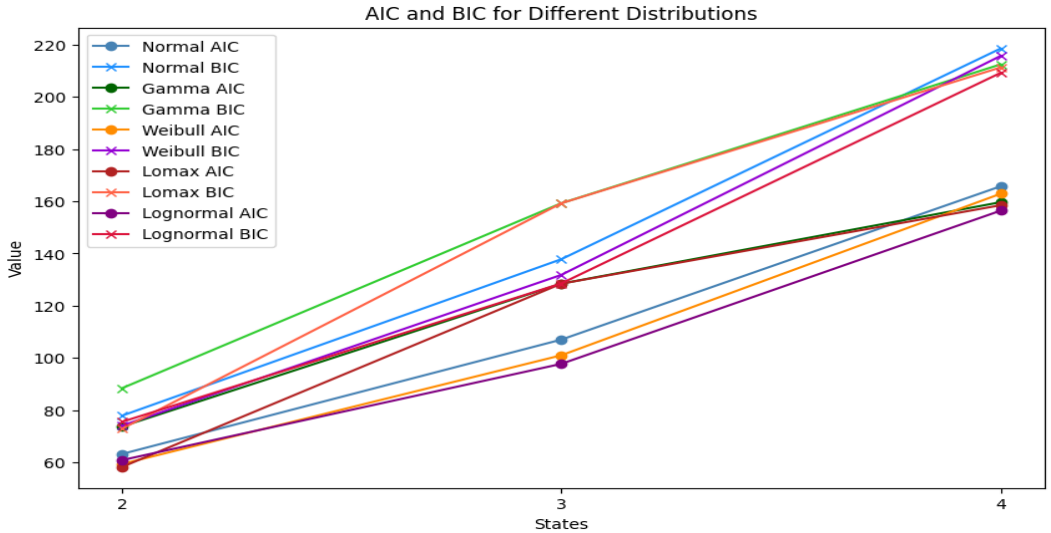


Figure 4: Graphical representation of comparative analysis of AIC and BIC values for different probability distributions.

IV. Conclusion

In conclusion, our findings shed light on the selection of probability distributions for Hidden Markov Models. We recognized the Lomax distribution as extremely beneficial, particularly within

the HMM's state 2. This conclusion emphasizes the need to consider multiple distributions when modelling complex systems. Furthermore, our investigation of the Viterbi path demonstrates the practical relevance of our findings in determining the most likely state sequence. The Viterbi path is 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. Overall, this study helps advance our understanding of probability distributions in HMMs.

References

[1] C.D. Mitchell and L.H. Jamieson. (1993). Modeling duration in a hidden Markov model with the exponential family. IEEE International Conference on Acoustics, Speech, and Signal Processing, Minneapolis, MN, USA, 1993, pp. 331-334 vol.2.

[2] Cota, Napat, Teerasit Kasetkasem, Preesan Rakwatin, Thitiporn Chanwimaluang, and Itsuo Kumazawa. (2015). Rice phenology estimation based on statistical models for time-series SAR data. In 2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, pp. 1-6. IEEE.

[3] Hao Zhang, Weidong Zhang, Ahmet Palazoglu and Wei Sun. (2012). Prediction of ozone levels using a Hidden Markov Model (HMM) with Gamma distribution. Atmospheric Environment, Volume 62, Pages 64-73, ISSN 1352-2310.

[4] Hui Zhang, Qing Ming Jonathan Wu and Thanh Minh Nguyen. (2013). Modified student's t-hidden Markov model for pattern recognition and classification. IET Signal Processing, Vol. 7, Iss. 3, pp. 219–227.

[5] Joshni George and Seemon Thomas. (2019). Negative Binomial Hidden Markov Model for AES count data. Think India Journal, Vol-22- Issue-14, ISSN:0971-1260.

[6] Lethanh, Nam, Kiyoyuki Kaito, and Kiyoshi Kobayashi. (2015). Infrastructure deterioration prediction with a Poisson hidden Markov model on time series data. Journal of Infrastructure Systems 21.3, 04014051.

[7] MacDonald I. L, and Bhamani F. (2018). A time-series model for underdispersed or overdispersed counts. The American Statistician, Taylor & Francis.

[8] Mahalakshmi Rajendran, Sentamarai Kannan Kaliyaperumal and Balasubramaniam Ramakrishnan. (2021). Hidden Markov Model of Evaluation of Break-Even Point of HIV patients: A Simulation Study. International Journal of Medical Sciences and Nursing Research,1(2):19-22.

[9] Marisa, Ukur Arianto Sembiring and Helena Margaretha. (2019). Earthquake Probability Prediction in Sumatra Island Using Poisson Hidden Markov Model (HMM). AIP Conf. Proc. 2192, 090006-1–090006-12.

[10] Mingyuan Zhou and Lawrence Carin. (2013). Negative binomial process count and mixture modelling. IEEE Transactions on Pattern Analysis and Machine Intelligence ,37.2, 307-320.

[11] Sarvi F, Nadali A, Khodadost M, Moghaddam M. K, and Sadeghifar M. (2017). Application of Poisson Hidden Markov model to predict the number of PM2. 5 exceedance days in Tehran during 2016-2017. Avicenna Journal of Environmental Health Engineering, 4(1), 58031-58031.

[12] Satu Helske and Jouni Helske. (2019). Mixture Hidden Markov Models for Sequence Data: The seqHMM Package in R, Journal of Statistical Software, 88.

[13] Sebastian George and Ambily Jose. (2020). Generalized Poisson Hidden Markov Model for Overdispersed or Underdispersed Count Data. Revista Colombiana de Estadística, Volume 43, Issue 1, pp. 71 to 82.

[14] Sebastian T, Jeyaseelan V, Jeyaseelan L, Anandan S, George S, and Bangdiwala S.I. [2019]. Decoding and modelling of time series count data using Poisson hidden Markov model and Markov ordinal logistic regression models. Statistical Methods in Medical Research, 28(5):1552-1563.

[15] Takashi Kaburagi, and Takashi Matsumoto. (2008). A generalized hidden Markov model approach to transmembrane region prediction with Poisson distribution as state duration probabilities. IPSJ Digital Courier, 4: 193-206.

OPTIMIZED SCHEDULING FOR THREE-STAGE FLOW SHOP WITH PARALLEL MACHINES: APPLICATION IN WASTE MANAGEMENT AND RECYCLING FACILITIES

Sonia[1], Pooja kaushik[2] Deepak Gupta[3], Sonia Goel[*4]

•

[1,2,3,4]Department of Mathematics, Maharishi Markandeshwar Engineering College, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala
[1]dhimansonia13@gmail.com, [2]drtobekaushik2024@gmail.com [3]guptadeepak2003@yahoo.com, [*4]sonia.mangla14@gmail.com
[*]Corresponding author

Abstract

Scheduling is one of the most important aspects of manufacturing sectors. Various factors in terms of performance evaluation, make business more competitive. Scheduling is the process of allocating scarce resources to tasks throughout time in order to accomplish one or more optimization goals. In fact, the methodical process of organizing, managing and maximizing work while ensuring the greatest possible use of both time and resources is the main aim of scheduling. This paper indicates the desirable and necessary steps to discover the optimum solution by efficient scheduling in waste recycling facilities which is essential for optimizing resources utilization and reducing operational delay. In this study, we have taken three stage flow shop scheduling with multiple machines at first two stages and single machine at last stage. This type of model has a number of applications not only in several industries but also in saving natural resources like water, electricity and reducing waste, idle time and unnecessary energy consumption. As, this research specially focused on the scenario of waste management and recycling facility where incoming waste undergoes three primary stages sorting, cleaning, and recycling. Here the scheduling evaluation is done by comparing the Branch and Bound method with Palmer's and Dannenbring method to determine an optimal job sequence that minimize the total processing time. The processing time is represented as a trapezoidal fuzzy number making an impact. The goal of this activity is to determine the optimal job plan in order to reduce the total amount of time that has processed and allocate the lead in the best feasible way to reduce the machine idle time, minimizes the water usage and energy consumption. This comparison demonstrated how well our suggested approach handled this challenging scheduling issue the findings contribute to the efficient management of recycling facilities, ensuring improved throughput. Because better scheduling strategies must be implemented in order to the growing demand for waste processing and recycling. So, the proposed approach aids in streamlining waste management processes, ultimately supporting environmental conservation efforts. Here, Branch and Bound is the best approach as compared to Palmer and Dannenbring approach. Future work can be extending this methodology to the three-phase flow shop scheduling with parallel machines at all the stages.

Keywords: Scheduling, Equipotential machines, trapezoidal fuzzy numbers

I. Introduction

Customer satisfaction in a global supply chain is crucial to the long-term viability of the company. Given the abundance of items on the market, the goal of any manufacturing sector is to satisfy customer demands on schedule, with the appropriate quality, and most important at the lowest possible cost. One of the most important choices for maximizing the use of tools and resources is scheduling.

It is intimately related to how well the company performs in terms of cost, quality, speed, dependability, flexibility, and time. Johnson [1] conducted the crucial research on flow shop scheduling for the first time for two and three stages. Ignall and Schrage [2] worked on the B&B method to systematically explore possible job sequences. Maggu et.al [4] extended the classical flow shop scheduling by transportation and job block concepts. Chung Yee Lee [5] investigates the scheduling problems with the concept of parallel machines and proposed Polynomial time algorithm. MasriKhan et.al [11] explore the hybrid scheduling algorithms palmer-NEH, Gupta-NEH and Dannenbring-NEH to reduce energy costs in production environment. Narain [6] addressed the scheduling problems under no idle constraints and demonstrate the application and effectiveness of the proposed algorithm. Sharma et.al [8] studied probabilistic processing times and machine breakdown intervals into the scheduling model and included breakdowns where machine might stop working. Palmer [3] provided priority index to rank the jobs served as a foundation for later algorithms for flow shop scheduling. T.P. Singh [7] focused on 3-stage flow shop scheduling problems with parallel machines. Gupta [9] studied on heuristic optimization technique for total flow time of three sequential processing stages of equipotential flow shop scheduling. Gupta et.al [12] presented an exact branch and bound algorithm for two stage flow shop environment and offered practical value for industries. Utama et.al [14] provided a comprehensive review of energy efficient scheduling approaches in hybrid flow shop environment. Malhotra et.al [13] worked on the scheduling problem with three identical parallel machines based on branch and bound. Rahmani et.al [10] worked on Dannenbring method to find the optimal schedule including work weights and transportation time. Kaushik et.al [15] investigated scheduling strategies in a bi-stage flow shop by comparing B&B with two heuristic algorithms HEH and CDS.

II. Statistical Development

In this study we are presenting the ideas of multiple machines at level 1st and 2nd with single machine at level 3rd . This kind of model may produce many things on a huge scale in a variety of industrial businesses. There was another approach to solving the challenges provided by the idea of trapezoidal fuzzy numbers accepted as time. The description of a mathematical problem is assuming three processors, $\mathcal{H}$, $\mathcal{I}$ and $\mathcal{R}$ to complete n jobs (i = 1, 2, 3, 4, ... n) referred to as the first, second, and third stages respectively. At phase one and two, there are parallel machines of type $\mathcal{H}_j$ and $\mathcal{I}_k$ (j = 1, 2, 3, 4, ..., m_1) (k= 1, 2, 3, 4, ..., m_2) respectively, where as in the third stages, there is only one machine. In the initial step, each parallel machine's unit operational cost for ith job on jth machine is also provided in the form of $\mathcal{H}_{ij}$ and $\mathcal{I}_{ij}$ respectively. Here the processing time is represented as trapezoidal fuzzy numbers of processors $\mathcal{H}$, $\mathcal{I}$ and $\mathcal{R}$ in the form of ($\hat{a}_i$, $\hat{e}_i$, $\hat{\imath}_i$, $\hat{o}_i$), ($\hat{u}_i$, $\hat{r}_i$, $\hat{s}_i$, $\hat{t}_i$) and ($\hat{l}_i$, $\hat{n}_i$, $\hat{t}_i$, $\hat{\jmath}_i$) respectively. The total available time is denoted by t_{rs} of like machines of type $\mathcal{H}_j$ and $\mathcal{I}_k$ (j = 1, 2, 3, 4, ..., m_1), (k = 1, 2, 3, 4, ..., m_2) respectively. The statistical form of model is shown as follow,

Table1: *Statistical form of model*

Jobs (i)	Stage1($\mathcal{H}$)					Stage2($\mathcal{J}$)					Stage3($\mathcal{R}$)
	$\mathcal{H}_1$	$\mathcal{H}_2$	$\mathcal{H}_3$	$\mathcal{H}_{m1}$	Processing time of $\mathcal{H}$ $\alpha_i = (\hat{a}_i, \hat{e}_i, \hat{\imath}_i, \hat{o}_i)$	$\mathcal{J}_1$	$\mathcal{J}_2$	$\mathcal{J}_3$	$\mathcal{J}_{m2}$	Processing time of $\mathcal{J}$ $\beta_i = (\hat{u}_i, \hat{r}_i, \hat{s}_i, \hat{t}_i)$	Processing time of $\mathcal{R}$ $\gamma_i = (\hat{l}_i, \hat{n}_i, \hat{t}_i, \jmath_i)$
1	$\mathcal{H}_{11}$	$\mathcal{H}_{12}$	$\mathcal{H}_{13}$	$\mathcal{H}_{1m1}$	$(\hat{a}_1, \hat{e}_1, \hat{\imath}_1, \hat{o}_1)$	$\mathcal{J}_{11}$	$\mathcal{J}_{12}$	$\mathcal{J}_{13}$	$\mathcal{J}_{1m2}$	$(\hat{u}_1, \hat{r}_1, \hat{s}_1, \hat{t}_1)$	$(\hat{l}_1, \hat{n}_1, \hat{t}_1, \jmath_1)$
2	$\mathcal{H}_{21}$	$\mathcal{H}_{22}$	$\mathcal{H}_{23}$	$\mathcal{H}_{2m1}$	$(\hat{a}_2, \hat{e}_2, \hat{\imath}_2, \hat{o}_2)$	$\mathcal{J}_{21}$	$\mathcal{J}_{22}$	$\mathcal{J}_{23}$	$\mathcal{J}_{2m2}$	$(\hat{u}_2, \hat{r}_2, \hat{s}_2, \hat{t}_2)$	$(\hat{l}_2, \hat{n}_2, \hat{t}_2, \jmath_2)$
3	$\mathcal{H}_{31}$	$\mathcal{H}_{32}$	$\mathcal{H}_{33}$	$\mathcal{H}_{3m1}$	$(\hat{a}_3, \hat{e}_3, \hat{\imath}_3, \hat{o}_3)$	$\mathcal{J}_{31}$	$\mathcal{J}_{32}$	$\mathcal{J}_{33}$	$\mathcal{J}_{3m2}$	$(\hat{u}_3, \hat{r}_3, \hat{s}_3, \hat{t}_3)$	$(\hat{l}_3, \hat{n}_3, \hat{t}_3, \jmath_3)$
4	$\mathcal{H}_{41}$	$\mathcal{H}_{42}$	$\mathcal{H}_{43}$	$\mathcal{H}_{4m1}$	$(\hat{a}_4, \hat{e}_4, \hat{\imath}_4, \hat{o}_4)$	$\mathcal{J}_{41}$	$\mathcal{J}_{42}$	$\mathcal{J}_{43}$	$\mathcal{J}_{4m2}$	$(\hat{u}_4, \hat{r}_4, \hat{s}_4, \hat{t}_4)$	$(\hat{l}_4, \hat{n}_4, \hat{t}_4, \jmath_4)$
n	$\mathcal{H}_{n1}$	$\mathcal{H}_{n2}$	$\mathcal{H}_{n3}$	$\mathcal{H}_{nm1}$	$(\hat{a}_n, \hat{e}_n, \hat{\imath}_n, \hat{o}_n)$	$\mathcal{J}_{n1}$	$\mathcal{J}_{n2}$	$\mathcal{J}_{n3}$	$\mathcal{J}_{nm2}$	$(\hat{u}_n, \hat{r}_n, \hat{s}_n, \hat{t}_n)$	$(\hat{l}_n, \hat{n}_n, \hat{t}_n, \jmath_n)$
t_{rs}	t_{11}	t_{12}	t_{13}	t_{1m1}		t_{21}	t_{22}	t_{23}	t_{2m2}		

I. Assumptions

- Each task has its own autonomy.
- The setup time is not taken into account.
- It is not necessary to process every task on every possible machine.
- Each task has a unique operational cost.
- All of the equipotential parallel processors may begin working at the same time.
- The availability of processors is constant.
- Here, pre-emption is not taken into consideration.

II. Methodology to solve problem

Step (1): Initially, we will use crisp(A)= 1/2[(b + c)- 4/5(b-a) +2/3(d-c)] to apply Yager's Formula of fuzzy dispensation time of different jobs, where (a, b, c, d) represents the fuzzy processing time of ith job of corresponding machine. The processing time as trapezoidal fuzzy will be reduced to a single number using this procedure.

Step (2): Check the condition $\sum_{r=1}^{m1} t_{1r} = \sum_{i=1}^{n} \alpha_i$ for first stage and $\sum_{r=1}^{m2} t_{1r} = \sum_{i=1}^{n} \beta_i$ for second stage. Then apply Vogal Approximation Method (VAM) and Modified Distribution Method. If condition is not satisfied then apply Modified Distribution Method for unbalanced problem.

Step (3): Use the Branch and Bound approach. The procedure is summarized in the following steps.

- Apply the formula
  - $g' = \max\left(\sum_{i=1}^{n} \mathcal{H}_{ij}\right) + \min_{i \epsilon j_{r'}} (\max \mathcal{J}_{ij} + \mathcal{R}_{ij})$
  - $g'' = \max_{i \epsilon j_r} (\mathcal{H}_{ij}) + \max\left(\sum_{i=1}^{n} \mathcal{J}_{ij}\right) + \min_{i \epsilon j_{r'}} \mathcal{R}_{ij}$

- $g''' = \max_{i \epsilon j_r}(\mathcal{H}_{ij}) + \max_{i \epsilon j_r}(\mathcal{I}_{ij}) + \sum_{i \epsilon j_r} \mathcal{R}_{ij}$

- Calculate $\mathcal{M}$= max (g′, g″, g‴)
- Calculate $\mathcal{M}$ for all corresponding jobs.
- Determine the lowest value of $\mathcal{M}$ obtained in step(c).
- Then begin the optimal sequence by the job corresponding to minimum value $\mathcal{M}$ obtained in step (d).
- Now repeat the above process for (n-1) subsequence, then for (n-2) and so on. Finally, the process gives the optimal or near optimal schedule.
- Formulate in-out table for the optimal sequence obtained in step (f).

III. Practical Situation

In our model the first two stages challenge with parallel machines and third stage with single machines. This kind of scheduling can be effectively used in waste management which involve multiple steps in different stages. Here the case study involved the three stages.

The first stage is sorting stage. In this stage the waste is sorted into different categories like plastics, metals and papers using multiple parallel machines i.e. $\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{H}_3$. The second stage is cleaning stage with multiple parallel machines $\mathcal{I}_1$, $\mathcal{I}_2$ and $\mathcal{I}_3$. where the sorted recyclables are washed and cleaned to remove all kind of impurities. The third stage is processing or recycling stage with a single machine, where the cleaned sorted recyclables are melted and compacted into reusable materials. The difficulties lie in using the machineries as efficiently as possible while reducing the makespan i.e. by scheduling we can minimize the sorting time, the water and electricity usage by balancing the load on cleaning machines. Overall, we can save money by reducing the cost of waste disposal and protect the environment. So, the scheduling has a crucial role in this type of scenario also.

I. Process Stages

Let us assume we want to process 5 jobs i.e. five types of waste, that can be plastic waste, metal waste, glass waste, paper waste and electronic waste etc. We want to schedule all these jobs across three stages. This model worked in the following way,

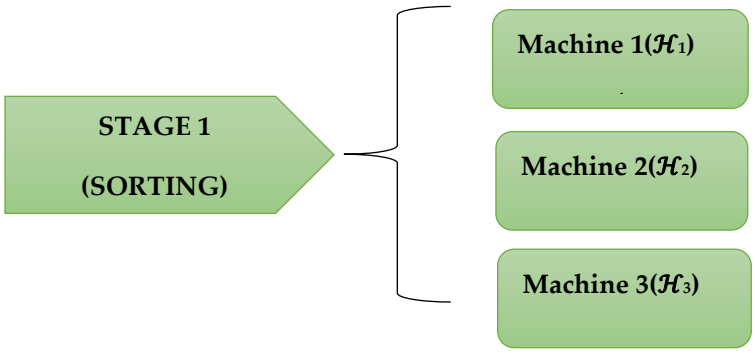


Figure I(a): *Three parallel machines at 1st stage*

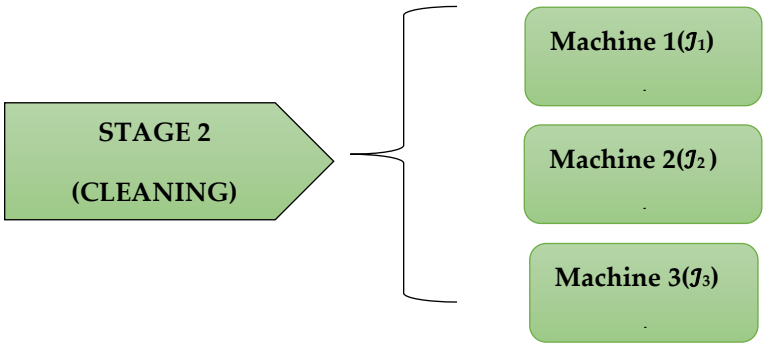


Figure I(b): *Three parallel machines at 2nd stage*

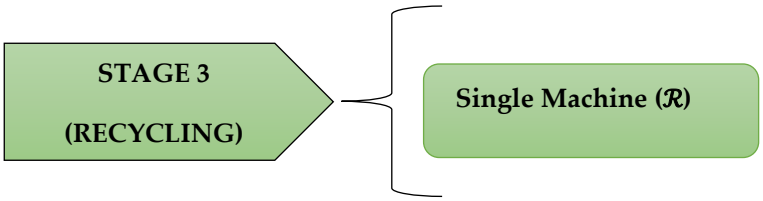


Figure I(c): *Single machines at 3rd stage*

IV. Numerical Example

Consider the three phase Flow Shop Scheduling with three parallel machines at first two stages each and single machine at last stage. The statement of numerical problem is given in table 2.

Table2: *Numerical Example*

Jobs	Stage1($\mathcal{H}$)				Stage2($\mathcal{J}$)				Stage3($\mathcal{R}$)
(i)	$\mathcal{H}_1$	$\mathcal{H}_2$	$\mathcal{H}_3$	Processing time of $\mathcal{H}$	$\mathcal{J}_1$	$\mathcal{J}_2$	$\mathcal{J}_3$	Processing time of $\mathcal{J}$	Processing time of $\mathcal{R}$
1	1	5	4	(1,3,4,5)	7	2	10	(3,5,6,8)	(2,3,5,7)
2	7	1	9	(2,4,5,7)	8	3	4	(4,7,8,9)	(1,2,3,4)
3	6	7	3	(6,7,8,9)	5	1	6	(3,5,4,7)	(1,2,4,9)
4	4	8	5	(3,5,7,8)	3	7	2	(4,5,6,8)	(3,5,7,8)
5	10	2	8	(2,3,4,5)	1	6	8	(2,5,8,9)	(4,5,6,8)
t_{rs}	10	4.3	9		10	10	7.95		

Solution:

Step 1: Apply Yager's ranking formula, the simplified form of problem is represented in table3.

Table3: *Yager's ranking formula of fuzzy dispensation time*

Jobs	Stage1($\mathcal{H}$)				Stage2($\mathcal{J}$)				Stage3($\mathcal{R}$)
(i)	$\mathcal{H}_1$	$\mathcal{H}_2$	$\mathcal{H}_3$	Processing time of $\mathcal{H}$	$\mathcal{J}_1$	$\mathcal{J}_2$	$\mathcal{J}_3$	Processing time of $\mathcal{J}$	Processing time of $\mathcal{R}$
1	1	5	4	3.0	7	2	10	5.35	4.25
2	7	1	9	4.0	8	3	4	6.6	2.4
3	6	7	3	7.4	5	1	6	4.7	4.25
4	4	8	5	5.5	3	7	2	5.7	5.5
5	10	2	8	3.4	1	6	8	5.6	5.75
t_{rs}	10	4.3	9		10	10	7.95		

Step 2: Since our problem is balanced as it satisfies the condition of step 2 of algorithm. Now, to find the optimal allocations of processing time on machine $\mathcal{H}$, $\mathcal{I}$ and $\mathcal{R}$ by using VAM method is given in table 4 below;

Table4: *Optimal allocations of processing time*

Jobs	Stage1($\mathcal{H}$)			Stage2($\mathcal{I}$)			Stage3($\mathcal{R}$)
(i)	$\mathcal{H}_1$	$\mathcal{H}_2$	$\mathcal{H}_3$	$\mathcal{I}_1$	$\mathcal{I}_2$	$\mathcal{I}_3$	$\mathcal{R}_{ij}$
1	3.0	0	0	0	5.35	0	4.25
2	0	4	0	0	0	6.6	2.4
3	0	0	7.4	0.05	4.65	0	4.25
4	5.5	0	0	4.35	0	1.35	5.5
5	1.5	0.3	1.6	5.6	0	0	5.75

Step 3: Now to apply B&B algorithm to find out the lower bound of jobs for the optimal sequence. The calculations are shown in table 5.

Table5: *Branch and Bound procedure*

Task (i)	g′= max $(\sum_{i=1}^{n} \mathcal{H}_{ij})$ + $\min_{i\epsilon j_{r'}}(\max \mathcal{I}_{ij} + \mathcal{R}_{ij})$	g″ = $\max_{i\epsilon j_r}(\mathcal{H}_{ij})$ + max ($\sum_{i=1}^{n} \mathcal{I}_{ij}$) + $\min_{i\epsilon j_{r'}} \mathcal{R}_{ij}$	g‴= $\max_{i\epsilon j_r}(\mathcal{H}_{ij})$ + $\max_{i\epsilon j_r}(\mathcal{I}_{ij})$ + $\sum_{i\epsilon j_r} \mathcal{R}_{ij}$	$\mathcal{M}$= max (g′, g″, g‴)
1	10 + 8.9 = 18.9	3.0 +10 +2.4 = 15.4	3.0 + 5.35 + 22.15 = 30.5	30.5
2	10 + 8.9 = 18.9	4+10+4.25 = 18.25	4 + 6.6+22.15 = 32.75	32.75
3	10 + 9 = 19	7.4+10+2.4 = 19.8	7.4 + 4.65 + 22.15 = 34.2	34.2
4	10 + 8.9 = 18.9	5.5 + 10 +2.4 = 17.9	5.5+4.35+22.15 = 32	32
5	10 + 8.9 = 18.9	1.6 +10 +2.4 = 14	1.6 + 5.6 + 22.15 = 29.35	29.35

In table 5, min {30.5, 32.75, 34.2, 32, 29.35} = 29.35, which corresponds to job5. So, job5 will be at the first position of the optimal sequence. Continue with the same process of branch and bound for the subsequence {5, 1}, {5,2}, {5,3} and {5, 4}. The results shown in table 6.

Table6: *Branch and Bound procedure for 2nd position in optimal schedule*

Task (i)	g′= max $(\sum_{i=1}^{n} \mathcal{H}_{ij})$ + $\min_{i\epsilon j_{r'}}(\max \mathcal{I}_{ij} + \mathcal{R}_{ij})$	g″ = $\max_{i\epsilon j_r}(\mathcal{H}_{ij})$ + max ($\sum_{i=1}^{n} \mathcal{I}_{ij}$) + $\min_{i\epsilon j_{r'}} \mathcal{R}_{ij}$	g‴= $\max_{i\epsilon j_r}(\mathcal{H}_{ij})$ + $\max_{i\epsilon j_r}(\mathcal{I}_{ij})$ + $\sum_{i\epsilon j_r} \mathcal{R}_{ij}$	$\mathcal{M}$= max (g′, g″, g‴)
{5,1}	10 + 8.9 = 18.9	9.85 +7.95 +2.4 =20.2	9.85 + 22.15 = 32	32
{5,2}	10 + 8.9 = 18.9	10.9 + 4.65 +4.25 =19.8	10.9 + 22.15 =33.05	33.05
{5,3}	10 + 9 = 19	9.05 +7.95+2.4 = 19.4	9.05 + 22.15 = 31.2	31.2
{5,4}	10 + 8.9 = 18.9	12.25 + 10 +2.4 = 24.65	12.25 + 22.15 = 34.4	34.4

From table 6, we observe that the min {32, 33.05, 31.2, 34.4} =31.2, which corresponds to the subsequence {5,3}. Therefore job 3 will be at second position. Continue the process for the 3rd and 4th position in the optimal schedule. Table 7 and table 8 show the outcomes.

Table7: *Branch and Bound procedure for the 3rd position in optimal schedule*

Task (i)	$g' = \max(\sum_{i=1}^{n} \mathcal{H}_{ij}) + \min_{i \epsilon j_{r'}}(\max \mathcal{I}_{ij} + \mathcal{R}_{ij})$	$g'' = \max_{i \epsilon j_r}(\mathcal{H}_{ij}) + \max(\sum_{i=1}^{n} \mathcal{I}_{ij}) + \min_{i \epsilon j_{r'}} \mathcal{R}_{ij}$	$g''' = \max_{i \epsilon j_r}(\mathcal{H}_{ij}) + \max_{i \epsilon j_r}(\mathcal{I}_{ij}) + \sum_{i \epsilon j_r} \mathcal{R}_{ij}$	$\mathcal{M} = \max(g', g'', g''')$
{5,3,1}	10 + 9 =19	10 + 7.95 +2.4 = 20.35	10 + 22.15 =32.15	32.15
{5,3,2}	10 + 9.6 =19.6	6.6 + 5.35 + 4.25 =16.2	6.6 + 22.15 = 28.75	28.75
{5,3,4}	10 + 9 = 19	13.4 + 6.6 + 2.4 = 22.4	13.4 +22.15 =35.55	35.55

Here min {32.15, 28.75, 35.55} = 28.75, which corresponds to the subsequence {5,3,2}. Which show that job 2 will be at 3rd position.

Table8: *Branch and Bound procedure for the 4th position in optimal schedule*

Task (i)	$g' = \max(\sum_{i=1}^{n} \mathcal{H}_{ij}) + \min_{i \epsilon j_{r'}}(\max \mathcal{I}_{ij} + \mathcal{R}_{ij})$	$g'' = \max_{i \epsilon j_r}(\mathcal{H}_{ij}) + \max(\sum_{i=1}^{n} \mathcal{I}_{ij}) + \min_{i \epsilon j_{r'}} \mathcal{R}_{ij}$	$g''' = \max_{i \epsilon j_r}(\mathcal{H}_{ij}) + \max_{i \epsilon j_r}(\mathcal{I}_{ij}) + \sum_{i \epsilon j_r} \mathcal{R}_{ij}$	$\mathcal{M} = \max(g', g'', g''')$
{5,3,2,1}	10 + 9.85 = 19.85	10 + 4.35 +5.5 =19.85	10 + 22.15 = 32.15	32.15
{5,3,2,4}	10 + 9.6 =19.6	13.4 + 5.35 +4.25 = 23	13.4 + 22.15 =35.55	35.55

Clearly, min {32.15, 35.55} = 32.15, which corresponds to {5,3,2,1}. So, job 1 will be at the 4th place and job 4 at the 5th place in the optimal schedule.
Finally, {5, 3, 2, 1,4} is the optimal schedule by B&B and table 9 show the minimum makespan.

Table 9: *In/Out table for the minimum makespan*

Jobs(i)	$\mathcal{H}_1$	$\mathcal{H}_2$	$\mathcal{H}_3$	$\mathcal{I}_1$	$\mathcal{I}_2$	$\mathcal{I}_3$	$\mathcal{R}$
5	0 - 1.5	0 - 0.3	0 – 1.6	1.6 – 7.2	-	-	7.2 – 12.95
3	-	-	1.6 – 9	9 – 9.05	0 – 4.65	-	12.95 – 17.2
2	-	0.3 - 4.3	-	-	-	0 – 6.6	17.2 – 19.6
1	1.5 - 4.5	-	-	-	4.65 – 10	-	19.6 – 23.85
4	4.5 - 10	-	-	10-14.35	-	6.6 –7.95	23.85 - **29.35**

The minimum makespan by B&B is 29.35

V. Solution by Dannenbring Algorithm

Dannenbring Method is a heuristic approach designed for solving the scheduling issues involving more than two machines. The Johnson method served as the basis for the Dannenbring approach. Here the complete work with the nth machine processing time is transformed into a recapitulation table with two values says Vi_1 and Vi_2. This method ranks jobs using a composite priority function based on the sum of processing time. It improved the Johnsons rule by considering more than first two machines.

I. Working procedure

Step1:

Calculate aggregate processing time of machines at level first and two i.e. $\frac{\sum pij}{m1}$ where m_1 is number of parallel machines and p_{ij} is the processing time of i^{th} job on j^{th} machine at first stage. Similarly calculate the same for the second stage also.

Step2:

Find the processing time for first machine by using the formula $T_{j1} = \sum_{i=1}^{m}(m - i + 1)^* t_{ij}$ and the processing time for the second machine will be calculated by $T_{j2} = \sum_{i=1}^{m} i^* t_{ij}$.

Where;

m : Total number stages

i : The stage index

t_{ij} : Processing time for job 'i' on machine 'j'

Step3: Calculate the value Vi_1 and Vi_2 by using the formula in step 2.

Step4: Find out the optimal schedule by ordering the either of the value Vi_1 and Vi_2 in descending order.

Step5: find the in – out schedule for calculating makespan.

Solution:

Find the aggregate processing time of machine at first and second stage using table 4.

We get the result in table 10.

Table10: *Aggregate processing time*

Jobs(i)	$\mathcal{H}$	$\mathcal{I}$	$\mathcal{R}$
1	1	1.78	4.25
2	1.33	2.2	2.4
3	2.46	1.56	4.25
4	1.8	1.9	5.5
5	1.13	1.86	5.75

Now, Calculate the value Vi_1 and Vi_2 by using the formula in step 2. Table 11 demonstrate the results below;

Table 11: *The value of Vi_1 and Vi_2*

Jobs	Vi_1	Vi_2
1	3(1 + 1.78 + 4.25) = 21.09	1 + 3.56 + 12.75 =17.31
2	3(1.33 + 2.2 + 2.4) = 17.79	1.33 + 4.4 + 7.2 = 12.93
3	3(2.46 + 1.56 + 4.25) = 24.81	2.46 + 3.12 + 12.75 = 18.33
4	3(1.8 + 1.9 +5.5) = 27.6	1.8 + 3.8 + 16.5 = 22.1
5	3(1.13 +1.86 +5.75) 26.22	1.13 + 3.72 + 17.25 =22.1

From step 4 the optimal order is {4, 5, 3, 1,2}. Now, find the makespan corresponding to the obtained schedule. Result is shown in table 12 below.

Table 12: *In-out table for makespan*

Jobs(i)	$\mathcal{H}_1$	$\mathcal{H}_2$	$\mathcal{H}_3$	$\mathcal{I}_1$	$\mathcal{I}_2$	$\mathcal{I}_3$	$\mathcal{R}$
4	0 – 5.5	-	-	5.5 – 9.85	-	5.5 – 6.85	9.85 –15.35
5	5.5 – 7	0 – 0.3	0 – 1.6	9.85 – 15.45	-	-	15.45 – 21.2
3	-	-	1.6 – 9	15.45 – 15.5	9 – 13.65	-	21.2 – 25.45
1	7 – 10	-	-	-	13.65 – 19	-	25.45 – 29.7
2	-	0.3 – 4.3	-	-	-	6.85 -13.45	29.7 – **32.1**

VI. Solution with the help of Palmer's Heuristic Algorithm

Palmer Heuristic [3] is well-known guideline for solving scheduling issues in flow shop. This is quick solving approach specifically designed for flow shop scheduling to find near optimal solution. Palmer's heuristic reduces the total makespan by implicitly attempting to minimize the peak load on any one machine by giving the middle machines larger weights. This method gives a job index to every job according to its processing time after that jobs are ordered according to given index. The algorithm of Palmer's heuristic for the three-stage flow shop scheduling with parallel machines at first two stages is given below;

Step1 Calculate aggregate processing time at level 1 and level 2:

For the step 1, find the aggregate processing time of machines at the first and second stage i.e. $\frac{\sum pij}{m1}$ where m_1 is number of parallel machines and p_{ij} is the processing time of i^{th} job on j^{th} machine at first two stages.

Step2 Calculate weighted Slop Index (WSI):

First of all, give some specific weights to each machine by using the formula

S(i) = [$\sum_{i=2}^{m}(m-2i+1)$]* T_{ij} where

T_{ij} = processing time of job 'i' on machine 'j 'at all stages

m = number of stages

Step3 Job Rank

Sort the job in descending order according to the job indices S(i) calculated in step 2.

Step4 Makespan

Calculate the total completion time of job for the job sequence obtained in step 3.

Solution: To find the aggregate of processing time at stage first and two. We proceed from table 10, where the results already shown. Now, Calculate weighted Slop Index (WSI) as follow;

S(1) = (2*1) + (0*1.78) + (-2*4.25) = 2+0-8.5 = -6.5

S(2) = (2*1.3) + (0*2.2) + (-2*2.4) = 2.6+0-4.8= -2.2

S(3) = (2*2.4) + (0*1.56) + (-2*4.25) = 4.8+0- 8.5= -3.7

S(4) = (2*1.8) + (0*1.9) + (-2*5.5) = 3.6+0-11=-7.4

S(5) = (2*1.13) + (0*1.86) + (-2*5.7) = 2.26+0-11.4=-9.24

As mention above, sort the job in descending order according to the job indices S(i), we get the job sequence {J_2, J_3, J_1, J_4, J_5}. The makespan for the obtained schedule is shown in table13.

Table13: *Makespan for the optimal sequence by Palmer's heuristics*

Jobs(i)	$\mathcal{H}_1$	$\mathcal{H}_2$	$\mathcal{H}_3$	$\mathcal{I}_1$	$\mathcal{I}_2$	$\mathcal{I}_3$	$\mathcal{R}$
2	-	0 - 4	-	-	-	4 – 10.6	10.6 - 13
3	-	-	0 - 7.4	7.4 – 7.45	7.4 -12.05	-	13 – 17.25
1	0 -3.0	-	-	-	12.05-17.4	-	17.25-21.5
4	3.0-8.5	-	-	8.5-12.85	-	12.85-14.2	21.5-28
5	8.5-10	4- 4.3	7.4-9	12.85-18.45	-	-	28 - **33.75**

Conclusion

We observe from the solution using all the three methods, we found that the B&B technique yield a lower minimum makespan as compared to Palmer and Dannenbring heuristic. It systematically evaluates all possible sequences, ensuring the best possible schedule. Dannenbring heuristic improves upon Palmer's heuristic by introducing a weighted priority index. It provides better approximation than Palmer but does not guarantee an optimal solution. Thus, we may conclude that B&B is the precise approach when compared to Dannenbring and Palmer approach.

References

[1] Johnson, S. M. (1954). Optimal two-and three-stage production schedules with setup times included. *Naval research logistics quarterly*, *1*:61-68.

[2] Ignall, E., & Schrage, L. (1965). Application of the branch and bound technique to some flow-shop scheduling problems. *Operations research*, *13*:400-412.

[3] Palmer, D. S. (1965). Sequencing jobs through a multi-stage process in the minimum total time—a quick method of obtaining a near optimum. *Journal of the Operational Research Society*, *16*:101-107.

[4] Maggu, P. L., Das, G., & Kumar, R. (1981). On equivalent-job for job-block in 2× n sequencing problem with transportation-times. *Journal of the operations research society of japan*, *24*:136-146.

[5] Lee, C. Y. (1991). Parallel machines scheduling with nonsimultaneous machine available time. *Discrete Applied Mathematics*, *30*:53-61.

[6] Narain, L., & Bagga, P. C. (2005). Flowshop/no-idle scheduling to minimize total elapsed time. *Journal of Global Optimization*, *33*:349-367.

[7] Singh, T. P., & Gupta, S. (2008). Minimization of rental cost in 3-stage flow shop with parallel machines including transportation time. *Acta Cinencia Indica*, *34*:629-634.

[8] Sharma, S., & Gupta, D. (2011). Minimizing rental cost under specified rental policy in two stage flow shop, the processing time associated with probabilities including break-down interval and job block criteria. *European Journal of Business and Management*, *3*:85-103.

[9] Gupta, D., & Goel, S. (2018). Three stage flow shop scheduling model with m-equipotential machines. *International Journal on Future Revolution in Computer Science & Communication Engineering*, *4*:269-274.

[10] Rahmani, U., Wulan, E. R., Huda, A. F., Supian, S., & Subiyanto, S. (2019). Scheduling of Flowshop Type Production with Dannenbring Method to Obtain Optimal Results with Observing Waiting Time and Weight of Jobs, *127*:212-224.

[11] Masrikhan, M., & Kurniawati, D. A. (2021). Flow Shop scheduling based on Palmer-neh, Gupta-neh and Dannenbring-NEH algorithms to minimize the energy cost. *Sinergi*, *25*:111-118.

[12] Gupta, D., & Goel, S. (2022). Branch and bound technique for two stage flow shop scheduling model with equipotential machines at every stage. *International Journal of Operational Research*, *44*:462-472.

[13] Malhotra, K., Gupta, D., Goel, S., & Tripathi, A. K. (2022). Bi-objective flow shop scheduling with equipotential parallel machines. *Malaysian Journal of Mathematical Sciences*, *16*:451-470.

[14] Utama, D. M., Primayesti, M. D., Umamy, S. Z., Kholifa, B. M. N., & Yasa, A. D. (2023). A systematic literature review on energy-efficient hybrid flow shop scheduling. *Cogent Engineering*, *10*:2206074.

[15] Kaushik, P., Gupta, D., & Goel, S. (2025). COMPARATIVE STUDY OF B&B WITH HEURISTICS NEH AND CDS FOR BI STAGE FLOW SHOP SCHEDULING PROBLEM UNDER FUZZY ENVIRONMENT. *Journal of Mechanics of Continua and Mathematical Sciences*, 20:73-89.

RT&A, No 2 (84)
Volume 20, June 2025

RELIABILITY MEASURES AND CLASSICAL AND BAYESIAN PARAMETRIZATION OF TWO NON-IDENTICAL UNITS SYSTEM MODEL WITH ON-LINE/OFF-LINE REPAIRS OF REPAIR MACHINE

Poonam Sharma[1], Pawan Kumar*[2], Shallu Sharma[2]

•

[1,*2,2]Department of Statistics, University of Jammu, J&K.
sharmapoonam1038@gmail.com[1], pkk_skumar@yahoo.co.in*[2], shallusharma28udh@gmail.com[2]

Abstract

The present paper deals with the reliability analysis of a system model which consists of two non identical units and a repair machine. One of the units initially, is in operative mode and other is kept in standby mode. Repair machine is initially in good condition, it may fails while repairing the failed units. In case of failure of repair machine, firstly it undergoes on-line repair and then to off-line repair if not repaired online. The failure time distributions of both the units and repair machine are assumed to be exponential while the repair time distributions are taken as general. By using regenerative point techniques, various measures of system effectiveness such as transition probabilities, MTSF, reliability, availability, busy period etc. have been obtained. Also, some of them have been studied graphically. A simulation study is also carried out to analyse the considered system model both in Classical and Bayesian setups. From the graphs and tables, various important conclusions are drawn.

Keywords: Reliability, Mean Time to System Failure, Busy Period Analysis, Regenerative Point Technique, Repair Machine, Bayesian Estimation, Maximum Likelihood Estimation.

1. Introduction

Reliability theory is a fundamental framework that plays a pivotal role in various fields, ranging from engineering and economics to healthcare and social sciences. The essence of reliability theory lies in assessing and ensuring the dependability and consistency of systems, processes, or components over time.At its core, reliability theory addresses the inherent uncertainty and variability in real-world systems, acknowledging that components may fail or degrade over time. The primary objective is to quantify and enhance the reliability of these systems, ultimately aiming for optimal functionality and durability. One key aspect of reliability theory is the study of failure mechanisms and their impact on overall system performance.In the real-world, the situations arise in many cases when the repair machine fails itself. Kumar et al. [5] studied reliability measures of a two unit system with Inspection and On-Line/Off-Line Repairs Using the Regenerative Point Technique. In this study they analysed a system model having two identical units with the inspection of each unit also with single repair facility which is always available with the system. Tyagi and Chaudhary [1] analysed a two non-identical unit parallel system subject to two types of failure and correlated life times. They used the concept of two non identical unit system with two types of repair. Also the units in this model are arranged in parallel configuration. Kumar and Gupta [3] studied performance measures of a two non-identical unit system model with repair and replacement policies. Kishan and Jain [4] carried a study on two non-identical unit standby system model with repair, inspection and post-repair from Classical

and Bayesian viewpoints, considering both failure and repair time distribution as weibull. Also a Monte carlo simulation study is carried out. Lin et al.[6] adhered to the reliability analysis and preventive maintenance using Classical and Bayesian semi-parametric degradation approaches of a locomotive wheel-sets as a case study. Yue et al. [8] accomplished performance analysis and optimization of a machine repair problem with warm spares and two heterogeneous repairmen. Here first repairman is always available for serving the failed units, while the second repairman leaves for a vacation of random length when the number of failed units is less than N. Goyal et al.[2] studied reliability measures and profit exploration of windmill water-pumping systems incorporating warranty and two types of repair. Malik et al.[7] studied reliability and economic measures of a system with inspection for online repair and no repair activity in abnormal weather. They considered two reliability models in which unit fails completely via partial failure. In the present study, we analyze the reliability characteristics and estimate in Classical and Bayesian paradigms. The parameters of two non identical units system model with a repair machine. Here, one of the units initially is in operative mode and other is kept in standby mode. Repair machine is initially in good condition, it may fail while repairing the failed units. In case of failure of repair machine, firstly it undergoes on-line repair and then to off-line repair. The failure time distributions of both the units and repair machine are assumed to be exponential while the repair time distributions are taken as general. By using regenerative point techniques, various measures of system effectiveness such as transition probabilities, MTSF, reliability, availability, busy period etc. have been obtained. Some of them have been studied graphically. Also a simulation study is carried out to analyse the considered system model both in Classical and Bayesian setups. From the graphs and tables, various important conclusions are drawn.

2. System Description and Assumptions

The system comprises of two non identical units A and B and one Repair machine. Initially one of the units is in operative mode and other is kept in standby mode. Repair machine is used to repair the failed units. Initially, repair machine is assumed to be in good condition. There might be online and offline repair of repair machine. As soon as repair machine fails firstly it goes to online repair and then to offline repair, if online repair fails to fix problem. The failure time distributions of both the units and Repair machine are assumed to be exponential while the repair time distributions are taken general in nature. Once a component is repaired it is as good as new.

3. Notations and Symbols

α_1 :Failure rate of unit A
α_2 :Failure rate of unit B
α_3 :Failure rate of repair machine
β : Online to offline repair rate.
A_o/B_o :Units are in operative mode
A_r/B_r :Units are under repair
$F_1(.)$: Cdf of repair time of unit A
$F_2(.)$ Cdf of repair time of unit B
$F_3(.)$: Cdf of online repair time of repair machine
$F_4(.)$: Cdf of offline repair time of repair machine
A_s/B_s :Units are in standby mode
A_{wr}/B_{wr}:Units are waiting for repair
${}_{on}RM_r/{}_{off}RM_r$: Online/Offline repair of Repair machine.

3.1. Symbols for the states of the system

$S_0 = [A_o, B_s, RM]$
$S_1 = [A_r, B_o, RM]$
$S_2 = [A_r, B_{wr}, RM]$
$S_7 = [A_o, B_r, RM]$
$S_8 = [A_o, B_{wr}, {}_{on}RM_r]$
$S_3 = [A_{wr}, B_o, {}_{on}RM_r]$
$S_4 = [A_{wr}, B_o, {}_{off}RM_r]$
$S_5 = [A_{wr}, B_{wr}, {}_{on}RM_r]$
$S_6 = [A_{wr}, B_{wr}, {}_{off}RM_r]$
$S_9 = [A_o, B_{wr}, {}_{off}RM_r]$

The transition diagram along with all transitions is shown in Figure 1

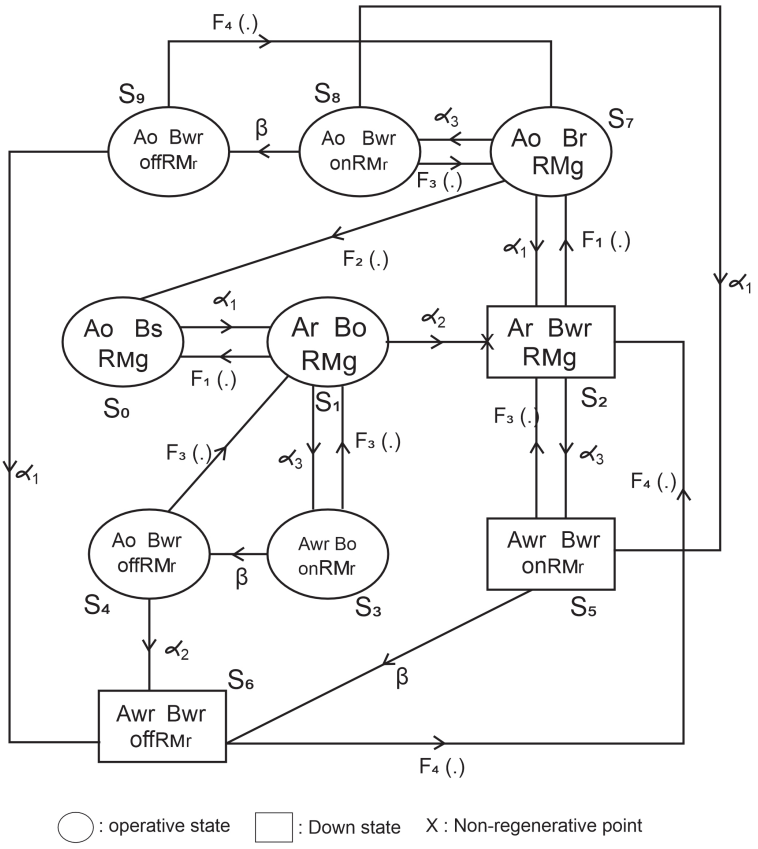


Figure 1: *Transition Diagram*

4. Transition Probabilities and Sojourn Times

The steady state transition probabilities are given by,

$$p_{ij} = \lim_{t\to\infty} Q_{ij}(t) = \int q_{ij}(t)dt \quad p_{ij}^{(k)} = \lim_{t\to\infty} Q_{ij}^{(k)}(t) \quad and \quad p_{ij}^{(k,l)} = \lim_{t\to\infty} Q_{ij}^{k,l}(t)$$

In particular we have

$$p_{10}(t) = \int e^{-\alpha_3 t} e^{-\alpha_2 t} dF_1(t) \quad = \quad \widetilde{F_1}(\alpha_2+\alpha_3)$$

Similarly,

$p_{13} = \frac{\alpha_3}{\alpha_2+\alpha_3}[1-\widetilde{F_1}(\alpha_2+\alpha_3)]$ $\quad$ $p_{15}^{(2)} = [1-\widetilde{F_1}(\alpha_3)] - \frac{\alpha_3}{\alpha_2+\alpha_3}[1-\widetilde{F_1}(\alpha_2+\alpha_3)]$

$p_{17}^{(2)} = \widetilde{F_1}(\alpha_3) - \widetilde{F_1}(\alpha_2+\alpha_3)$ $\quad$ $p_{25} = 1-\widetilde{F_1}\ (\alpha_3)$

$p_{27} = \widetilde{F_1}\ (\alpha_3)$ $\quad$ $p_{31} = \widetilde{F_3}(\alpha_2+\beta)$

$p_{34} = \frac{\beta}{\alpha_2+\beta}[1-\widetilde{F_3}(\alpha_2+\beta)]$ $\quad$ $p_{35} = \frac{\alpha_2}{\alpha_2+\beta}[1-\widetilde{F_3}(\alpha_2+\beta)]$

$p_{41} = \widetilde{F_4}(\alpha_2)$ $\quad$ $p_{46} = [1-\widetilde{F_4}(\alpha_2)]$

$p_{52} = \widetilde{F_3}\ (\beta)]$ $\quad$ $p_{56} = 1-\widetilde{F_3}\ (\beta)$

$p_{70} = \widetilde{F_2}(\alpha_1+\alpha_3)$ $\quad$ $p_{71} = \frac{\alpha_1}{\alpha_1+\alpha_3}[1-\widetilde{F_2}(\alpha_1+\alpha_3)]$

$p_{78} = \frac{\alpha_3}{\alpha_1+\alpha_3}[1-\widetilde{F_2}(\alpha_1+\alpha_3)]$ $\quad$ $p_{85} = \frac{\alpha_1}{\alpha_1+\beta}[1-\widetilde{F_3}(\alpha_1+\beta)]$

$p_{87} = \widetilde{F_3}(\alpha_1+\beta)$ $\quad$ $p_{89} = \frac{\beta}{\alpha_1+\beta}[1-\widetilde{F_3}(\alpha_1+\beta)]$

$p_{96} = [1 - \widetilde{F_4}(\alpha_1)]$ $p_{97} = \widetilde{F_4}(\alpha_1)$

Thus, we observe the following relations:

$p_{10} + p_{13} + p_{15}^{(2)} + p_{17}^{(2)} = 1$ $p_{25} + p_{27} = 1$

$p_{52} + p_{56} = 1$ $p_{31} + p_{34} + p_{35} = 1$

$p_{41} + p_{46} = 1$ $p_{70} + p_{72} + p_{78} = 1$

$p_{85} + p_{87} + p_{89} = 1$ $p_{96} + p_{97} = 1$

$p_{01} = p_{62} = 1$

4.1. Mean Sojourn times

In reliability, Mean Sojourn time ψ_i, refers to the expected amount of time a system spends in a specific state before making a transition to another state. To find ψ_i for state S_i, we observe that there is no transition from S_i to any otherstate as long as the system is in state S_i. Let T_i denotes the sojourn time in state S_i then mean sojourn time ψ_i in state S_i is:

$$\psi_i = E[T_i] = \int P(T_i > t)dt$$

Hence, using the above formula following values for mean sojourn time are obtained:

$\psi_0 = \frac{1}{\alpha_1}$ $\psi_1 = \frac{1}{(\alpha_2+\alpha_3)}[1 - \widetilde{F_1}(\alpha_2 + \alpha_3)]$

$\psi_2 = \frac{1}{\alpha_3}[1 - \widetilde{F_1}(\alpha_3)]$ $\psi_3 = \frac{1}{(\alpha_2+\beta)}[1 - \widetilde{F_2}(\alpha_2 + \beta)]$

$\psi_4 = \frac{1}{\alpha_2}[1 - \widetilde{F_4}(\alpha_2)]$ $\psi_5 = \frac{1}{\beta}[1 - \widetilde{F_3}(\beta)]$

$\psi_6 = \int \bar{F_4}(t)dt$ $\psi_7 = \frac{1}{(\alpha_1+\alpha_3)}[1 - \widetilde{F_2}(\alpha_1 + \alpha_3)]$

$\psi_8 = \frac{1}{(\alpha_1+\beta)}[1 - \widetilde{F_3}(\alpha_1 + \beta)]$ $\psi_9 = \frac{1}{\alpha_1}[1 - \widetilde{F_4}(\alpha_1)]$

5. Analysis of Reliability and MTSF

Let T_i be any random variable which denotes the time to failiure of the system when system starts up from state $S_i \in E_i$, then the reliability of the system is given by

$$R_i(t) = P[T_i > t]$$

To obtain $R_i(t)$, we consider failed states as absorbing states. By referring to the state transition diagram, the recursive relations among $R_i(t)$ can be formulated on the basis of probabilistic arguments. Taking Laplace transform and solving the resultant set of equations for$R_0^*(s)$, we get

$$R_0^*(s) = \frac{N_1(s)}{D_1(s)} \tag{1}$$

where,

$$N_1(s) = Z_0^*[1 - q_{13}^*(q_{31}^* + q_{34}^*(q_{41}^* + Z_4^*))] + q_{01}^*[Z_1^* + q_{13}^*(Z_3^* + Z_4^* q_{34}^*)]$$

$$D_1(s) = 1 - q_{13}^*(q_{31}^* + q_{34}^* q_{41}^*) - q_{01}^* q_{10}^*$$

By taking inverse Laplace Transform of (1), we get reliability of the system.To get MTSF, we use the well known formula

$$E(T_0) = \int R_0(t)dt = \lim_{s\to 0} R_0^*(s) = \frac{N_1(0)}{D_1(0)} \tag{2}$$

where,

$$N_1(0) = \psi_0[1 - p_{13}(p_{31} + p_{34}(p_{41} + \psi_4))] + p_{01}[\psi_1 + p_{13}(\psi_3 + \psi_4 p_{34})]$$

and

$$D_1(0) = 1 - p_{13}(p_{31} + p_{34}p_{41}) - p_{01}p_{10}$$

Since we have $q_{ij}^*(0) = p_{ij}$ and $\lim_{s\to 0} Z_i^*(s) = \int Z_i(t)dt = \psi_i$ [1]

6. Availability Analysis

Availability refers to the probability that a system is available at time 't' given that initially it started from $S_i \in E_i$. Point wise availability refers to the availability of a system at a specific point in time. It is a measure of system performance and indicates whether a system is operational and providing its intended service at a particular moment. By using stochastic arguments, the recurrence relations among different point wise availabilities are obtained and taking the Laplace transforms and solving the resultant set of equations for$A_0^*(s)$, we have

$$A_0^*(s) = \frac{N_2(s)}{D_2(s)}$$

where,

$$\begin{aligned} N_2(s) =& Z_0^*[1 - q_{13}^* A(q_{31}^* + q_{34}^* q_{41}^*)] + A(Z_1^* + Z_3^* q_{13}^* + Z_4^* q_{13}^* q_{34}^*) + Z_5^* \\ & [B(q_{15}^{(2)*} + q_{13}^* q_{35}^*) + q_{17}^{(2)*} C + q_{13}^* q_{34}^* D] + q_{27}^*(q_{15}^{(2)*} + q_{17}^{(2)*})[Z_7^* \\ & + Z_8^* + q_{78}^* + Z_9^* q_{78}^* q_{89}^*] + q_{13}^* q_{27}^* E[Z_7^* + Z_8^* q_{78}^* + Z_9^* q_{78}^* q_{89}^*] \end{aligned}$$

Here,

$$\text{A} = q_{27}^*(1 - q_{78}^*)$$

$$B = 1 - q_{27}^* q_{72}^* - q_{78}^* q_{87}^* - q_{78}^* q_{89}^* q_{97}^* - q_{27}^* q_{78}^* q_{89}^* q_{96}^*$$

$$C = q_{25}^* q_{72}^* + q_{78}^* q_{87}^* - q_{78}^* q_{89}^* q_{97}^* - q_{27}^* q_{78}^* q_{89}^* q_{96}^*$$

$$D = q_{25}^* q_{46}^* q_{q46}^* q_{78}^* q_{87}^* (1 - q_{27}^* q_{97}^*)$$

$$E = q_{34}^* q_{46}^* + q_{35}^*$$

and,

$$\begin{aligned} D_2(s) =& q_{27}^*[1 - q_{72}^* - q_{78}^* q_{87}^* - q_{70}^* q_{17}^{(2)*} - q_{78}^* q_{89}^* q_{97}^* - q_{78}^* q_{85}^* + q_{70}^* q_{15}^{(2)*} - q_{13}^* q_{70}^* \\ & (q_{35}^* + q_{34}^* q_{46}^*)] - F[q_{10}^* + q_{13}^*(q_{31}^* + q_{34}^* q_{41}^*)] \end{aligned} \tag{3}$$

where,

$$F = q_{27}^*[1 - q_{72}^* - q_{78}^*(q_{87}^* - q_{85}^* + q_{89}^* q_{97}^* - q_{89}^* q_{96}^*)]$$

The steady state availability is given by

$$A_0 = \lim_{t\to\infty} A_0(t) = \lim_{s\to 0} s A_0^*(s) = \frac{N_2(0)}{D_2(0)}$$

[1]Limits of integeration whenever they are 0 to ∞ are not mentioned.

now we know that, $q_{ij}(t)$ is the pdf of the transition time from state Si to Sj and also $q_{ij}(t)$ is the probability of transition from state S_i to S_j during the interval (t, t + dt), thus

$$q_{ij}^*(s)|s = 0 = q_{ij}^*(0) = p_{ij}$$

We also know that

$$\lim_{s\to 0} Z_i^*(s) = \int Z_i(t)dt = \psi_i$$

Therefore,

$$\begin{aligned} N_2(0) =& \psi_0[1 - p_{13}A(p_{31} + p_{34}p_{41})] + A(\psi_1 + \psi_3 p_{13} + \psi_4 p_{13}p_{34}) + \psi_5[B(p_{15}^{(2)} + \\ & p_{13}p_{35}) + p_{17}^{(2)}C + p_{13}p_{34}D] + p_{27}(p_{15}^{(2)} + p_{17}^{(2)})[\psi_7 + \psi_8 + p_{78} + \psi_9 p_{78}p_{89}] \\ & + p_{13}p_{27}E[\psi_7 + Z\psi_8 p_{78} + Z\psi_9 q p_{78}p_{89}] \end{aligned}$$

Here,

$$A = p_{27}(1 - p_{78})$$

$$B = 1 - p_{27}p_{72} - p_{78}p_{87} - p_{78}p_{89}p_{97} - p_{27}p_{78}p_{89}p_{96}$$

$$C = p_{25}p_{72} + p_{78}p_{87} - p_{78}p_{89}p_{97} - p_{27}p_{78}p_{89}p_{96}$$

$$D = p_{25}p_{46} - p_{46}p_{78}p_{87}(1 - p_{27}p_{97})$$

$$E = p_{34}p_{46} + p_{35}$$

and

$$\begin{aligned} D_2(0) =& p_{27}[1 - p_{72} - p_{78}p_{87} - p_{70}p_{17}^{(2)} - p_{78}p_{89}p_{97} - p_{78}p_{85} + p_{70}p_{15}^{(2)} - p_{13}p_{70} \\ & (p_{35} + p_{34}p_{46})] - F[p_{10} + p_{13}(p_{31} + p_{34}p_{41})] \end{aligned}$$

where,

$$F = p_{27}[1 - p_{72} - p_{78}(p_{87} - p_{85} + p_{89}p_{97} - p_{89}p_{96})]$$

The steady state probability that the system will be up in long run is given by

$$A_0 = \lim_{t\to\infty} A_0(t) = \lim_{s\to 0} sA_0^*(s)$$

$$\lim_{s\to 0} \frac{sN_2(s)}{D_2(s)} = \lim_{s\to 0} N_2(s) \lim_{s\to 0} \frac{s}{D_2(s)}$$

Since as s $\to$ 0, $D_2(s)$ becomes zero. Hence, on using L'Hospital's rule, A_0 becomes

$$A_0 = \frac{N_2(0)}{D_2^{'}(0)} \tag{4}$$

where,

$$\begin{aligned} D_2^{'}(0) =& \psi_0 A + p_{27}p_{70}(\psi_1 + \psi_3 p_{13}) + \psi_2 B + p_{27}(\psi_4 p_{13}p_{34}p_{78} + \psi_7 F) + \psi_5 C + \psi_6 D \\ & + p_{27}p_{78}(\psi_8 E + \psi_9 p_{89}F) \end{aligned} \tag{5}$$

here,

$$A = p_{10}[(p_{27}(1-p_{72}) - p_{78}p_{87} - p_{78}p_{89}p_{97}) - p_{78}(p_{25}p_{87} + p_{27}p_{85}) + p_{78}p_{89}(p_{25}p_{97} - p_{27}p_{96})] - p_{70}(p_{25}p_{17}^{(2)} - p_{27}p_{15}^{(2)}) + p_{13}p_{27}p_{70}(p_{35} + p_{34}p_{46})$$

$$B = 1 - p_{10}(1 - p_{78}p_{89}p_{97}) - p_{13}p_{34}p_{41}(p_{52} - p_{78}p_{89}p_{97}) - p_{13}p_{31}p_{56}(1 - p_{78}p_{87}) - p_{25}p_{70}p_{17}^{(2)} - p_{78}p_{87}(1 - p_{10}) - p_{78}p_{89}p_{97}(1 - p_{13}p_{31})$$

$$C = p_{25}[1 - E(p_{13}p_{31} + p_{10} + p_{13}p_{34}p_{41}) - p_{70}p_{17}^{(2)} - p_{78}(p_{87} + p_{89}p_{97})] + p_{27}[p_{78}p_{85} F(p_{70}(p_{15}^{(2)} + p_{13}p_{35}))]$$

$$D = p_{25}p_{56}[1 - E(p_{13}p_{31} + p_{10} + p_{13}p_{34}p_{41}) - p_{70}p_{17}^{(2)} - p_{78}(p_{87} + p_{89}p_{97})] + p_{27}p_{56}[p_{78}p_{85}F + p_{78}p_{89} + p_{70}(p_{15}^{(2)} + p_{13}p_{35})] + p_{13}p_{27}[p_{34}(p_{46}p_{70} - p_{41}p_{78}p_{89}p_{96}) - p_{13}p_{78}p_{89}p_{96}] - p_{10}p_{27}p_{78}p_{89}p_{96}$$

$$E = 1 - p_{78}p_{87} - p_{78}p_{89}p_{97}$$

$$F = 1 - p_{10} - p_{13}p_{31} - p_{13}p_{34}p_{41}$$

Using $N_2(0)$ and $D_2'(0)$ in equation [4], we get the expression for A_0. The expected up time of the system during (0,t] is given by

$$\mu_{up}(t) = \int_0^t A_0(u)du$$

So that,

$$\mu_{up}^* = \frac{A_0^*(s)}{s}$$

7. Busy Period Analysis

$B_i(t)$ refers to the probability that the system having started from regenerative state Si∈ E, at time t=0, is under repair at time t due to failure of the unit. To find these probabilities, we use simple probabilistic arguments and further on taking Laplace transform and solving the resultant set of equation for $B_0^*(s)$, we have

$$B_0^*(s) = \frac{N_3(s)}{D_2(s)}$$

where,

$$N_3(s) = Z_1q_{27}^*q_{70}^* + Z_2[q_{15}^{(2)*}A + q_{17}^{(2)*}B + Aq_{13}^*C]Z_7[q_{27}^*(q_{15}^{(2)*} + q_{17}^{(2)*}) + q_{13}^*q_{27}^*C]$$

here,

$$A = 1 - q_{78}^*q_{87}^* - q_{78}^*q_{89}^*q_{97}^*$$

$$B = q_{27}^* + q_{78}^*q_{85}^* + q_{78}^*q_{89}^*q_{96}^*$$

$C = q_{35}^{*} + q_{34}^{*} q_{46}^{*}$

and, D_2(s) is same as given by[3]. In the steady state, the probability that the repairman will be busy is given by

$$B_0 = \lim_{t\to\infty} B_0(t) = \lim_{s\to 0} B_0^{*}(s) = \frac{N_3(0)}{D_2^{'}(0)}$$

where,

$N_3(0) = Z_1 p_{27} p_{70} + Z_2[p_{15}^{(2)} A + p_{17}^{(2)} B + A p_{13} C] Z_7[p_{27}(p_{15}^{(2)} + p_{17}^{(2)}) + p_{13} p_{27} C]$

here,

$A = 1 - p_{78} p_{87} - p_{78} p_{89} p_{97}$

$B = p_{27} + p_{78} p_{85} + p_{78} p_{89} p_{96}$

$C = p_{35} + p_{34} p_{46}$

and $D_2^{'}(0)$ is same as obtained in [5]. The expected busy period of the repairman during (0,t] is given by

$$\mu_b(t) = \int_0^t B_0(u) du$$

So that,

$$\mu_b^{*}(s) = \frac{B_0^{*}(s)}{s}$$

8. Expected Number of Repairs

$V_i(t)$ is defined as the expeted number of repairs during the time interval (0,t] of the failed units, when the system initially starts from regenerative state S_i. Further, by the definition of $V_i(t)$ the recurrence relations are easily obtained and on taking their Laplace- Stieltjes transforms and solving the resultant set of equations for $\widetilde{V}_0(s)$, we get

$$\widetilde{V}_0(s) = \frac{N_4(s)}{D_3(s)}$$

where,

$N_4(s) = \widetilde{Q}_{27}[\widetilde{Q}_{70}(\widetilde{Q}_{10} - \widetilde{Q}_{17}^{(2)}) + \widetilde{Q}_{15}^{(2)} A + \widetilde{Q}_{13} B(\widetilde{Q}_{35} + \widetilde{Q}_{34}\widetilde{Q}_{46})]$

here,

$A = 1 + \widetilde{Q}_{56}\widetilde{Q}_{70} - \widetilde{Q}_{78}\widetilde{Q}_{87} - \widetilde{Q}_{78}\widetilde{Q}_{89}\widetilde{Q}_{97}$

$B = 1 + \widetilde{Q}_{70} - \widetilde{Q}_{78}\widetilde{Q}_{87} - \widetilde{Q}_{78}\widetilde{Q}_{89}\widetilde{Q}_{97}$

and $D_3(s)$ is written by replacing q_{ij}^{*} and $q_{ij}^{(k)*}$ by $\widetilde{Q}_{ij}$ and $\widetilde{Q}_{ij}^{(k)}$ respectively in the equation (7). In the steady state, the expected number of repairs per unit time is given by

$$V_0 = \lim_{t\to\infty} V_0(t) = \lim_{s\to 0} s\widetilde{V}_0(s) = \frac{N_4(0)}{D_2^{'}(0)}$$

where,

$$N_4(0) = p_{27}[p_{70}(p_{10} - p_{17}^{(2)}) + p_{15}^{(2)}A + p_{13}B(p_{35} + p_{34}p_{46})]$$

here,

$$A = 1 + p_{56}p_{70} - p_{78}p_{87} - p_{78}p_{89}p_{97}$$

$$B = 1 + p_{70} - p_{78}p_{87} - p_{78}p_{89}p_{97}$$

9. Profit Function Analysis

As the reliability charateristics are obtained, the profit function P(t) for the system can be obtained. Profit is defined as excess of revenue over the cost, therefore the expected total profits generated during(0,t] is given as :
P(t) = Expected total revenue in(0,t] - Expected total expenditure in(0,t]

$$= K_0\mu_{up}(t) - K_1\mu_b(t) - K_2V_0(t)$$

where,
K_0 = Revenue per unit up time of the system.
K_1 = Cost per unit time for which repair man is busy in repairing the failed unit.
K_2 = Cost of repair per unit.
In steady state, the expected total profits per unit time, is given by:

$$P = \lim_{t\to\infty}\frac{P(t)}{t} = \lim_{s\to 0} s^2\overset{*}{P}(s)$$

Therefore, we have

$$P = K_0A_0 - K_1B_0 - K_2V_0 \tag{6}$$

10. Estimation of the Parameters, MTSF, And Profit Function

10.1. Classical Estimation

10.1.1 ML Estimation

Let

$$\underset{\sim}{X_1} = (x_{11}, x_{12}, ..., x_{1n_1}), \quad \underset{\sim}{X_2} = (x_{21}, x_{22}, ...x_{2n_2}), \quad \underset{\sim}{X_3} = (x_{31}, x_{32}, ..., x_{3n_3}), \underset{\sim}{X_4} = (x_{41}, x_{42}, ..., x_{4n_4}),$$

$$\underset{\sim}{X_5} = (x_{51}, x_{52}, ..., x_{5n_5}, \quad \underset{\sim}{X_6} = (x_{61}, x_{62}, ..., x_{6n_6}), \underset{\sim}{X_7} = (x_{71}, x_{72}, ..., x_{7n_7}, \quad \underset{\sim}{X_8} = (x_{81}, x_{82}, ..., x_{8n_8})$$

Therefore, Likelihood function of combined sample is :

$$L = (X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8|\alpha_1, \alpha_2, \alpha_3, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \beta)$$

Now pdf of exponential distribution is :$\lambda \exp(-\lambda x)$

$$L = {\alpha_1}^{n_1}{\alpha_2}^{n_2}{\alpha_3}^{n_3}{\lambda_1}^{n_4}{\lambda_2}^{n_5}{\lambda_3}^{n_6}{\gamma_1}^{n_7}{\gamma_2}^{n_8} \exp -(\alpha_1W_1 + \alpha_2W_2 + \alpha_3W_3 + \lambda_1W_4 + \lambda_2W_5 + \lambda_3W_6 + \lambda_4W_7 + \beta W_8)$$

Here,

$$Wi = \sum_{n=1}^{n_i} x_{i_j}; 1 = 1, 2, 3, 4, 5, 6, 7, 8$$

Taking log b/s, we get

$$\begin{aligned} logL =& n_1 log\alpha_1 + n_2 log\alpha_2 + n_3 log\alpha_3 + n_4 log\lambda_1 + n_5 log\lambda_2 + n_6 log\lambda_3 + n_7 log\lambda_4 + n_8 log\beta_1 \\ & - (\alpha_1 W_1 + \alpha_2 W_2 + \alpha_3 W_3 + \lambda_1 W_4 + \lambda_2 W_5 + \lambda_3 W_5 + \lambda_4 W_6 + \beta W_8) \end{aligned} \tag{7}$$

Now we obtain, MLE $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3\ , \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4\ , \hat{\beta})$of the parameters $(\alpha_1, \alpha_2, \alpha_3, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \beta)$ as

$\hat{\alpha}_1 = \frac{n_1}{W_1}$, $\hat{\alpha}_2 = \frac{n_2}{W_2}$
$\hat{\alpha}_3 = \frac{n_3}{W_3}$, $\hat{\lambda}_1 = \frac{n_4}{W_4}$
$\hat{\lambda}_2 = \frac{n_5}{W_5}$, $\hat{\lambda}_3 = \frac{n_6}{W_6}$
$\hat{\lambda}_4 = \frac{n_7}{W_7}$, $\hat{\beta} = \frac{n_8}{W_8}$

The asymptotic distribution of$(\hat{\alpha_1} - \alpha_1, \hat{\alpha_2} - \alpha_2, \hat{\alpha_3} - \alpha_3, \hat{\lambda_1} - \lambda_1, \hat{\lambda_2} - \lambda_2, \hat{\lambda_3} - \lambda_3, \hat{\lambda_4} - \lambda_4, \hat{\beta} - \beta) \sim N_8(0, I^{-1})$, where I is the Fisher Information matrix with diagonal elements as

$I_{11} = \frac{n_1}{\alpha_1^2}$, $I_{22} = \frac{n_2}{\alpha_2^2}$, $I_{33} = \frac{n_3}{\alpha_3^2}$, $I_{44} = \frac{n_4}{\lambda_1^2}$, $I_{55} = \frac{n_5}{\lambda_2^2}$, $I_{66} = \frac{n_6}{\lambda_3^2}$, $I_{77} = \frac{n_7}{\lambda_4^2}$, $I_{88} = \frac{n_8}{\beta^2}$

and all non-diagonal elements are zero. The MLE $\hat{M}$ & $\hat{P}$ of MTSF and Profit function can be obtained using the invariance property of MLE. Also the asymptotic distribution of $(\hat{M} - M) is N(0, A' I^{-1} A)$ & that of $(\hat{P} - P) is N(0, B' I^{-1} B)$, where

$A' = (\frac{\delta M}{\delta \alpha_1}, \frac{\delta M}{\delta \alpha_2}, \frac{\delta M}{\delta \alpha_3}, \frac{\delta M}{\delta \alpha_2}, \frac{\delta M}{\delta \lambda_2}, \frac{\delta M}{\delta \lambda_3}, \frac{\delta M}{\delta \lambda_4}, \frac{\delta M}{\delta \beta})$

$B' = (\frac{\delta P}{\delta \alpha_1}, \frac{\delta P}{\delta \alpha_2}, \frac{\delta P}{\delta \alpha_3}, \frac{\delta P}{\delta \lambda_1}, \frac{\delta P}{\delta \lambda_2}, \frac{\delta P}{\delta \lambda_3}, \frac{\delta P}{\delta \lambda_4}, \frac{\delta P}{\delta \beta})$

10.2. Bayesian Estimation

Bayesian estimation is a statistical approach used to determine the impact of prior knowledge as well as the sample information.The Bayesian method of estimation is taken into consideration in this part for estimating the model parameters.The parameters involved in the model are random variables having independent Gamma prior distribution.Here, we estimate the unknown parameters while taking into account the gamma prior distribution and the corresponding PDFs.

$$\alpha_1 \sim Gamma(a_1, b_1) \qquad (\alpha_1, a_1, b_1) > 0, \tag{8}$$

$$\alpha_2 \sim Gamma(a_2, b_2) \qquad (\alpha_2, a_2, b_2) > 0, \tag{9}$$

$$\alpha_3 \sim Gamma(a_3, b_3) \qquad (\alpha_3, a_3, b_3) > 0, \tag{10}$$

$$\lambda_1 \sim Gamma(a_4, b_4) \qquad (\lambda_1, a_4, b_4) > 0, \tag{11}$$

$$\lambda_2 \sim Gamma(a_5, b_5) \qquad (\lambda_2, a_5, b_5) > 0, \tag{12}$$

$$\lambda_3 \sim Gamma(a_6, b_6) \qquad (\lambda_3, a_6, b_6) > 0, \tag{13}$$

$$\lambda_4 \sim Gamma(a_7, b_7) \qquad (\lambda_4, a_7, b_7) > 0, \tag{14}$$

$$\beta \sim Gamma(a_8, b_8) \qquad (\beta, a_8, b_8) > 0, \tag{15}$$

Here,a_i *and* b_i (i = 1,2,3,4,5,6,7,8) denotes the shape and scale parameters. Now using likelihood function and taking prior distributions, the posterior distributions of these parameters is obtained as follows:

$$\alpha_1 | \underset{\sim}{X_1} \sim Gamma(n_1 + a_1, b_1 + W_1) \tag{16}$$

$$\alpha_2 | \underset{\sim}{X_2} \sim Gamma(n_2 + a_2, b_2 + W_2) \tag{17}$$

$$\alpha_3 | \underset{\sim}{X_3} \sim Gamma(n_3 + a_3, b_3 + W_3) \tag{18}$$

$$\lambda_1|\underset{\sim}{X_4} \sim Gamma(n_4 + a_4, b_4 + W_4) \tag{19}$$

$$\lambda_2|\underset{\sim}{X_5} \sim Gamma(n_5 + a_5, b_5 + W_5) \tag{20}$$

$$\lambda_3|\underset{\sim}{X_6} \sim Gamma(n_6 + a_6, b_6 + W_6) \tag{21}$$

$$\lambda_4|\underset{\sim}{X_7} \sim Gamma(n_7 + a_7, b_7 + W_7) \tag{22}$$

$$\beta|\underset{\sim}{X_8} \sim Gamma(n_8 + a_8, b_8 + W_8) \tag{23}$$

To obtain Bayes estimates and width of HPD intervals of the parameters,we generated observations from the above posterior distributions. For attaining Bayesian estimation of MTSF and profit function we substituted the above draws directly into the equations [2] & [6]. Assuming square error loss function, the sample means of respective draws are taken as Bayes estimates of parameters and reliability characteristics.

11. Simulation Study

A simulation research is carried out to investigate the behaviour of parameters estimates and reliability features. The Standard Error(SE)/Posterior Standard Error(PSE) and width of confidence/HPD intervals are computed and are given in Table 1-6. Samples of sizes $n_1 = n_2 = n_3 = n_4 = n_5 = n_6 = n_7 = n_8 = 100$ have been drawn from the six considered distributions by assuming various values of the parameters as given in Tables 1-6. The number of iterations taken are 10000. All calculations are performed in R software.

Table 1: *The values of MTSF for fixed* $\lambda_1 = 0.15$ *and varying* α_1

α_1	True MTSF	MLE.MTSF	SE	C.I	Bayes MTSF	PSE	HPD Interval
0.1	15.729	10.11	0.0042	0.068	10.05	0.00098	0.00076
0.2	9.114	5.03	0.0032	0.034	5.043	0.00075	0.00062
0.3	6.909	3.39	0.0029	0.024	3.369	0.00050	0.00049
0.4	5.807	2.55	0.0027	0.021	2.536	0.00044	0.00037
0.5	5.145	2.06	0.0025	0.019	2.032	0.00047	0.00028
0.6	4.704	1.70	0.0025	0.018	1.697	0.00042	0.00038
0.7	4.389	1.47	0.010	0.018	1.456	0.00044	0.00034
0.8	4.1536	1.28	0.0024	0.017	1.282	0.00049	0.00030
0.9	3.969	1.14	0.0024	0.017	1.138	0.00037	0.00025
1	3.822	1.02	0.0024	0.017	1.029	0.00047	0.00038

Table 2: *The values of MTSF for fixed* λ_1*=0.35 and varying* α_1

α_1	True MTSF	MLE.MTSF	SE	C.I	Bayes MTSF	PSE	HPD Interval
0.1	20.034	10.15	0.0068	0.069	10.109	0.0022	0.0012
0.2	11.267	5.10	0.0045	0.034	5.065	0.0011	0.00069
0.3	8.344	3.42	0.0037	0.025	3.387	0.00073	0.00055
0.4	6.883	2.55	0.0032	0.021	2.545	0.00061	0.00063
0.5	6.006	2.10	0.0031	0.019	2.038	0.00061	0.00056
0.6	5.422	1.70	0.0029	0.018	1.707	0.00060	0.00043
0.7	5.005	1.48	0.0028	0.018	1.464	0.00080	0.00049
0.8	4.691	1.32	0.0027	0.018	1.283	0.00061	0.00038
0.9	4.448	1.16	0.0026	0.017	1.145	0.00049	0.00044
1	4.253	1.08	0.0026	0.017	1.034	0.00059	0.00030

Table 3: *The values of MTSF for fixed λ_1=0.65 and varying α_1*

α_1	True MTSF	MlE.MTSF	SE	C.I	Bayes MTSF	PSE	HPD Interval
0.1	26.493	10.18	0.011	0.070	10.1892	0.0024	0.0019
0.2	14.496	5.15	0.0066	0.035	5.114	0.0016	0.0014
0.3	10.497	3.47	0.0050	0.025	3.418	0.0010	0.00083
0.4	8.498	2.57	0.0043	0.021	2.563	0.0010	0.00088
0.5	7.298	2.10	0.0038	0.019	2.061	0.0010	0.00056
0.6	6.498	1.73	0.0036	0.018	1.719	0.0010	0.00063
0.7	5.927	1.48	0.0034	0.018	1.478	0.00085	0.00046
0.8	5.499	1.31	0.0032	0.018	1.297	0.00063	0.00060
0.9	5.165	1.18	0.0031	0.017	1.152	0.00057	0.00052
1	4.899	1.08	0.0030	0.017	1.040	0.00053	0.00052

Table 4: *The values of Profit for fixed λ_1=0.15 and varying α_1*

α_1	True profit	MlE.Profit	SE	C.I	Bayes Profit	PSE	HPD Interval
0.1	565.66	504.78	498.53	575.20	479.28	397.45	291.21
0.2	402.36	383.26	397.36	810.27	350.27	323.22	159.32
0.3	311.70	317.39	635.72	636.85	265.39	63.56	46.57
0.4	254.26	275.92	526.51	951.11	204.74	45.93	33.66
0.5	214.62	247.42	877.84	464.26	159.05	743.24	544.48
0.6	185.61	226.48	376.34	961.33	123.31	25.80	18.91
0.7	163.44	210.38	631.29	451.36	94.53	24.65	18.06
0.8	145.92	197.58	662.10	739.46	70.84	20.19	14.79
0.9	131.71	187.11	718.50	516.67	50.96	81.40	59.64
1	119.95	178.39	643.89	545.35	34.03	15.60	10.70

Table 5: *The values of Profit for fixed λ_1=0.35 and varying α_1*

α_1	True profit	MlE.Profit	SE	C.I	Bayes Profit	PSE	HPD Interval
0.1	821.06	543.02	602.82	626.98	629.57	319.08	233.79
0.2	620.12	419.43	689.35	691.46	547.35	275.89	161.56
0.3	496.74	345.93	678.57	543.09	483.00	141.41	103.61
0.4	414.10	298.73	653.09	802.17	430.65	195.75	143.42
0.5	355.20	266.05	632.30	583.35	387.01	89.59	65.64
0.6	311.25	242.08	652.66	1043.69	249.95	139.17	101.97
0.7	277.25	223.72	750.20	1033.47	318.03	28.31	20.74
0.8	250.25	209.17	527.56	876.81	290.21	156.78	114.87
0.9	228.19	3197.35	712.67	607.05	265.74	51.75	37.92
1	209.93	187.52	3884.21	993.43	244.01	89.38	65.49

Table 6: *The values of Profit for fixed λ_1=0.65 and varying α_1*

α_1	True profit	MlE.Profit	SE	C.I	Bayes Profit	PSE	HPD Interval
0.1	1021.3	547.18	893.57	364.97	718.30	249.01	182.45
0.2	813.1	441.50	531.40	436.79	667.84	526.05	110.66
0.3	672.6	369.25	580.49	481.96	624.46	405.31	296.97
0.4	572.1	319.94	405.80	313.23	586.36	290.30	212.70
0.5	497.1	284.78	613.66	522.86	552.48	191.82	140.54
0.6	439.2	258.59	624.87	678.83	522.07	67.79	49.67
0.7	393.4	238.37	554.60	720.39	494.59	71.09	52.09
0.8	356.2	222.28	766.52	416.82	469.60	89.92	65.89
0.9	325.5	209.17	733.96	552.96	446.77	52.28	72.69
1	299.8	198.27	788.98	788.43	425.81	30.95	22.68

12. Graphical Study

A graphical analysis of the system model provides a more insightful and vivid representation of system behaviour, allowing for a clear visualization of failure transitions, repair processes, and overall system performance. So for more concrete study, we plot MTSF and Profit function wrt α_1 failure rate of unit A for different values of λ_1 repair rate of unit A as 0.15, 0.35 and 0.65. Here all other parameters are fixed α_2= 0.007, α_3= 0.009,λ_2=0.09,λ_3=0.06,λ_4 = 0.07 β=0.7,K_0=1500,K_1=300 and K_2=250.

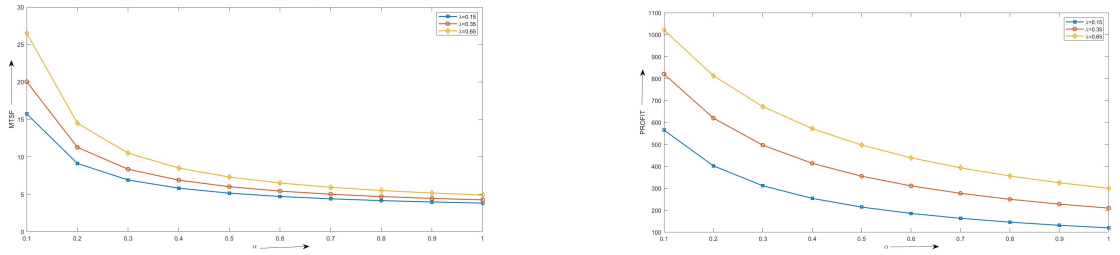


Figure 2: *(a)Behaviour of MTSF wrt to α_1 for different values of λ_1 and (b)Behaviour of Profit wrt to α_1 for different values of λ_1*

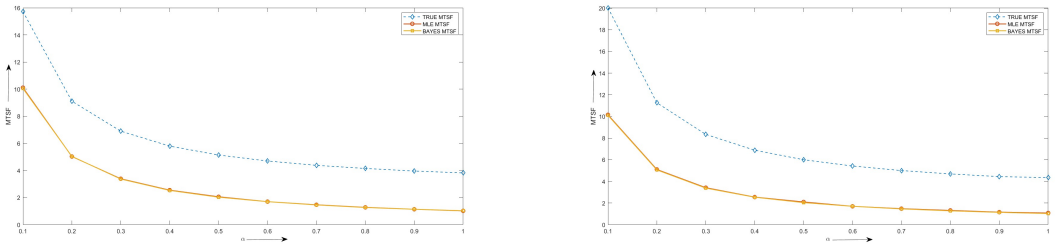


Figure 3: *(a)Behaviour of True MTSF, MLE MTSF & Bayes MTSF wrt to α_1 for λ_1= 0.15 and (b) Behaviour of True MTSF, MLE MTSF & Bayes MTSF wrt to α_1 for λ_1= 0.35*

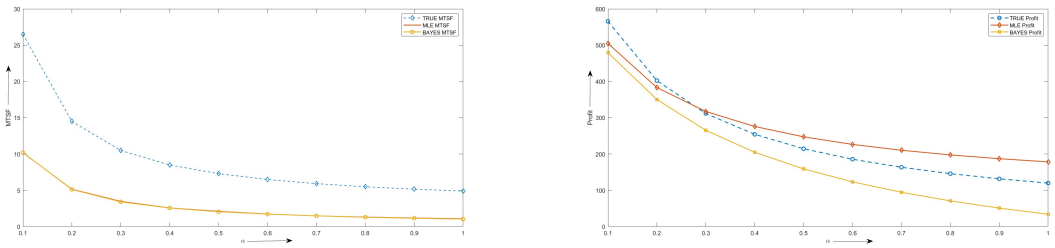


Figure 4: *(a) Behaviour of True MTSF, MLE MTSF & Bayes MTSF wrt to α_1 for λ_1= 0.65 and (b) Behaviour of True Profit, MLE Profit & Bayes Profit wrt to α_1 for λ_1= 0.15*

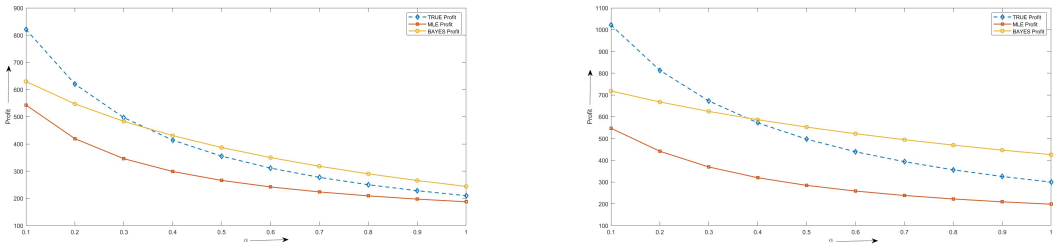


Figure 5: *(a)Behaviour of True Profit, MLE Profit & Bayes Profit wrt to α_1 for λ_1= 0.35 and (b) Behaviour of True Profit, MLE Profit & Bayes Profit wrt to α_1 for λ_1= 0.65*

13. Discussion and Conclusion

From tables and figures we may conclude that MTSF decreases as the failure rate α_1 increases whereas it increases as the repair rate λ_1 increases. Same trends follows for profit function. Additionally, Tables 1-6 show that for fixed and variable parameters, Bayes estimates of the MTSF and profit function are better than MLEs because they have lower PSE, and the width of HPD intervals is smaller than the width of confidence intervals. Therefore based on this analysis we conclude that the Bayes approach is superior to the Classical approach for estimating the MTSF and profit function for the examined model.

14. Acknowledgement

First author is thankful to Department of Science and Technology, Govt. of India for providing financial assitance in the form of INSPIRE fellowship.

References

[1] Chaudhary, P. and Tyagi, L. (2021). A two non-identical unit parallel system subject to two types of failure and correlated life times *Reliability: Theory & Applications*, 162 (62):247–258.

[2] Goyal, N., Ram, M., Kumar, A., B, Soni. and Klochkov, Yury(2021). Reliability measures and profit exploration of windmill water-pumping systems incorporating warranty and two types of repair *Mathematics*, 9(8):822.

[3] Gupta, U. and Kumar, P. Performance measures of a two non-identical unit system model with repair and replacement policies(2019) *Reliability: Theory & Applications*, 14(1):53-62.

[4] Kishan, R. and Jain, D. (2012). A two non-identical unit standby system model with repair, inspection and post-repair under classical and Bayesian viewpoints *Journal of reliability and statistical studies*,85–103.

[5] Kumar, K., Kumar, B. and Ravi, V.(2021). Study of Reliability Measures of A Two Unit System with Inspection and On-Line/Off-Line Repairs Using the Regenerative Point Technique *International Journal of Operations Research*, 18(3):57-66.

[6] Lin, J., Pulido, J. and Asplund, M.(2015). Reliability analysis for preventive maintenance based on classical and Bayesian semi-parametric degradation approaches using locomotive wheel-sets as a case study *Reliability Engineering & System Safety*, 134:143-156.

[7] Malik, SC and Pawar, D.(2010). Reliability and economic measures of a system with inspection for on-line repair and no repair activity in abnormal weather *Bulletin of Pure and Applied Sciences*, 29(2)355–368.

[8] Yue, D., Yue, W. and Qi, H.(2012). Performance analysis and optimization of a machine repair problem with warm spares and two heterogeneous repairmen *Optimization and Engineering*, 13:545-562.

RT&A, No 2 (84)
Volume 20, June 2025

DEVELOPMENT OF AN ATTRIBUTE CONTROL CHART BASED ON THE INVERSE KUMARASWAMY DISTRIBUTION

Nagaraju, R.[1], Palanivel, M.[2], Sriramachandran, G. V.[3*]

[1]Department of Statistics, Chikkanna Government Arts College, Tiruppur, India.
Email: rmnagarajustat@gmail.com
[2]Department of Statistics, Government Arts College, Udumalpet, India.
Email: palanivel6060@gmail.com
[3]*Department of Mathematics, Dr. N. G. P. Institute of Technology, Coimbatore, India.
Corresponding author Email: gvsriramachandran72@gmail.com

Abstract

This article presents the development of an attribute control chart designed for products with lifespans following the inverse Kumaraswamy distribution, using the distribution's mean as the quality metric. The study assumes one distribution parameter is known while varying the other to design the control chart. The chart's performance is evaluated using Average Run Length (ARL), with adjustments to parameter values determining ARL in out-of-control conditions. The proposed control chart is assessed through simulations. A comparative analysis with other similar charts is conducted based on ARL values, and the findings are discussed in detail.

Keywords: Attribute control chart, variable control chart, Inverse Kumaraswamy distribution, average run length, control limits.

I. Introduction

A control chart is an essential tool in statistical process control used to improve process quality. It is used in manufacturing and non-manufacturing settings to reduce variability by identifying special causes of variation. This ultimately leads to improved products and services and increased profit margins. Control charts (CCs) monitor process quality and identify deviations. The visual representation of data collected by CCs makes it easy to detect process variability and take necessary measures to eliminate the source of variation. A process is considered under control when all data points fall between the upper control line (UCL) and lower control line (LCL) within the CCs. Conversely, an out-of-control state is identified when any data point falls outside the control limits.

There are two main types of CCs for process monitoring: attribute and variable CCs. A process is considered variable if it relies on continuous data. There are several variable CCs designed to monitor the state of the process. Among these, Shewhart x bar CCs are the most popular due to their simplicity. Page [18] proposed the cumulative sum CCs, a graphical technique showing the cumulative sums of deviations of sample values related to a quality characteristic. In a study that compared the effectiveness of Shewhart x bar CCs and cumulative sum CCs in identifying slight changes in process averages, Hawkins and Olwell [10] concluded that the former was more effective. Abbas et al. [1] investigated process location monitoring using the nonparametric adaptive cumulative sum CCs

method. Robert [21] also introduced exponentially weighted moving average CCs with average run lengths (ARL), studied the characteristics of CCs tests based on moving averages, and described a process for creating geometric moving averages. Jing-Er Chiu and Tsen-I Kuo [13] developed the MP chart, a multivariate Poisson count data control chart. A new control chart, the np S2 control chart, was developed by Linda and Quinino [16] to monitor process variance instead of using S2 or R control charts. Based on the findings of an attribute inspection, this control chart's primary contribution is the assessment of process variability status.

Conversely, CCs are categorized as attribute CCs when the product's compliance with specific requirements determines the quality rating. Processes with non-normal distributions are best supervised by attribute CCs. Aslam and Jun [6] demonstrated that attribute CCs can successfully monitor the lifespan of high-quality products in the literature on single sampling schemes. To accommodate the Weibull distribution, they created attribute CCs. When the process is stable, the test's duration and the parameters of the recommended CCs are set up to guarantee that the average run length stays near the intended goal. The CCs were found to be highly flexible and appropriate for monitoring the lifespans of high-quality products. For processes that generate count data, Ahangar and Chimka [4] introduced attribute control charts with optimal limits. The chart's objective is to lower the overall cost related to its errors, which are represented linearly by Type I and Type II errors. Poisson, geometric, and negative binomial distributions can all be represented by the proposed chart, indicating that it can detect both quality improvement and degradation. Baklizi and Ghannam [7] explored attribute CCS based on the inverse Weibull distribution during their discussion of the use of CCS in industry. A quality indicator for the product was the quantity of failures that occurred during the life testing. The Average Run Length (ARL) values when the process deviates from control for different shift coefficient values are computed to evaluate the CCs. The truncation coefficient and control limit coefficients were also calculated for various sample sizes and target average run lengths. An example was provided to demonstrate how they explained the application of this CCS. Rao [19] created characteristic CCs for the exponentiated half Logistic distribution calculated dispersion with either a known or vague shape parameter, noticing a drift in ARL values. They concluded that the proposed control outline is appropriate for covering non-conforming items. Rao et al. [20] moreover proposed CCs for the Dagum distribution. These CCs can be utilized to track the continuances of quality items. The chart execution is detailed in terms of ARLs, showing that ARL values increase with bigger sample sizes and drop with progressed dissemination of shape parameters. Adeoti and Ogundipe [2] evaluated the execution of CCs utilizing a generalized exponential distribution. They studied the plan based on the assumption that the shape parameter is known and the scale parameter is varied. Srinivasa Rao et al. [23] created an attribute control chart by presuming the item's lifetime adheres to the Exponentiated Inverse Kumaraswamy distribution within the framework of a time-truncated life test. The operation of the formulated control chart is analyzed using the average run length (ARL) values. The devised chart is subsequently evaluated by the ARLs obtained from simulation studies for various sample sizes. Sriramachandran et. al., [24] constructed a CC for a repetitive sampling plan if the product life follows the Lindley distribution.

Many authors studied the properties of different forms of the Kumaraswamy distribution such as log-Kumaraswamy distribution, Ishaq et. al., [12], Transmuted Exponentiated Kumaraswamy Distribution, Joseph and Ravindran [14], exponentiated Kumaraswamy distribution, Lemontea et. al., [15], Odd Lindley-Kumaraswamy Distribution, Samson et. al., [22], Type II Half Logistic Kumaraswamy Distribution, El-Sayed et. al., [9], Generalized Odd Maxwell-Kumaraswamy Distribution, Ishaq et. al., [11] and many more. Based on our review of the literature, it appears that no research has been done on the attribute control chart for the life span of items that follow the Inverse Kumaraswamy distribution. This study's remaining sections are arranged in the following deliberate manner. Section 2 defines the Inverse Kumaraswamy distribution as studied by Daghistani et al. [8] Both the CC design and its implementation strategy are explained in Section 3. Section 4, compares the

efficacy of the suggested CC to other CC that are available in the literature. Section 5 studies the proposed CC using simulated. Section 6 studies the proposed CC using a real application. In Section 7, conclusions are provided.

II. Inverse Kumaraswamy Distribution

One of the most significant lifetime distributions, the Kumaraswamy distribution (KD) was introduced in Kumaraswamy and has a range of [0, 1]. Because of its properties such as unimodal, increasing, decreasing, or constant—it makes a good substitute for beta dissemination. AL-Fattah, et al.,[3] studied some of the properties of the inverse Kumaraswamy distribution (IKD). This distribution has several applications in various fields, such as medical research, life-testing problems, and stress-strength analysis. In addition, the IKD can be used to model failure rates in reliability and biological studies. Areej and AL-Zaydi [5] and Daghistani et al. [8] proposed a new version of the IKD based on the transformation $X = Z^{-1}$. The PDF and CDF of their proposed IKD are

$$f(x) = \theta\beta x^{-(\theta+1)}\left(1 - x^{-\theta}\right)^{\beta-1}, x > 1, \theta, \beta > 0 \quad (1)$$

$$F(x) = \left(1 - x^{-\theta}\right)^{\beta}, x > 1, \theta, \beta > 0 \quad (2)$$

The mean of the IKD is defined as

$$\mu = \beta B\left(1 - \frac{1}{\theta}, \beta\right) \quad (3)$$

where B(..) is the beta function.

III. Design of The Proposed Control Chart

The design of an np-Control Chart (np-CC) for the Inverse Kumaraswamy Distribution (IKD) under a time-truncated life test based on the failure count in each subgroup can be summarized as follows:

- A random sample of size n is selected and subjected to the life test.
- The test is terminated at a pre-specified time $t_0 = a.\mu_0$, where μ_0 is the mean of the distribution, and a is the control chart coefficient determining the termination time.
- Count the number of failures, D, observed during the test.
- If the failure count satisfies LCL ≤ D ≤ UCL, the process is in control.
- if D > UCL or D < LCL, the process is declared out-of-control.
- The number of failures, D, is modelled by a binomial distribution with parameters n (sample size) and with p_0 (probability of failure within the time t_0).
- The Upper Control Limit (UCL), Central Line (CL), and Lower Control Limit (LCL) for the proposed control chart are calculated using the binomial distribution and expressed as:

$$UCL/LCL = CL \pm K\sqrt{np_0q_0} \quad (4)$$

where n- sample size, p_0 - failure probability before time t_0, $q_0 = 1 - p_0$, $CL = np_0$ and K- coefficient of control limit. For unknown failure probability p_0, the control limits for the mean number of failures (MNF) is

$$UCL/LCL = MNF \pm K\sqrt{MNF\left(1 - \frac{MNF}{n}\right)} \quad (5)$$

As the number of items that failed in a production process may vary, implementing *CC* is complicated due to varying sample sizes. Hence the CC is designed based on *ARL*. Here *ARL* for the in-control process is denoted as ARL_0 and ARL for the out-of-control process is denoted as ARL_1. Let r_0 be the specified ARL_0. Then the in-control *ARL* is given as

$$ARL_0 = \frac{1}{1-P_{in}^0} \tag{6}$$

where p_{in}^0 is the probability that a production process is under control and can be obtained

$$P_{in}^0 = P(LCL \le f \le UCL) = \sum_{f=LCL+1}^{UCL} \binom{n}{f} {p_0}^f (1-p_0)^{n-f}, \text{where } p_0 = \left(1-(a.\mu_0)^{-\theta}\right)^{\beta} \tag{7}$$

If the process is declared out of control when the scale parameter is shifted to $z_1 = cz_0$ then

$$P_{in}^1 = P(LCL \le f \le UCL) = \sum_{f=LCL+1}^{UCL} \binom{n}{d} {p_1}^f (1-p_1)^{n-f}, \text{ where } p_1 = \left(1-(c.\mu_0)^{-\theta}\right)^{\beta} \tag{8}$$

The out-of-control - ARL_1 is given as

$$ARL_1 = \frac{1}{1-P_{in}^1} \tag{9}$$

For the target $ARL_0 = 200, 250, 300 \text{ and } 370$, shape parameter $\beta_0 = 0.5, 0.7, 0.9, 1.0, 1.2$ the sample size n=20, a, k, UCL, and LCL are determined using eqn (5). By substituting the values in the eqn (9), the ARL_1 for shifts c= 0.9, 0.8, 0.7... 0.1 are obtained and the values obtained are tabulated in Tables 1 to 4.

Table 1: *CC Parameters for ARL_0=200 and ARL_1 for different shifts when n=20 and $\theta = 3$*

β	0.5	0.7	0.9	1.0	1.2
UCL	16	17	16	18	18
LCL	3	4	3	6	5
a	0.8371	0.8518	0.8706	0.9242	0.9498
L	2.9132	2.9214	2.9354	2.9502	2.9754
Shift c			ARL_1		
1	201.97	200.64	200.52	200.31	201.32
0.9	23.32	60.29	75.45	83.09	92.82
0.8	3.43	5.35	9.3	14.04	21.68
0.7	1.51	2.59	2.75	4.29	5.58
0.6	1.11	1.42	1.53	2.05	2.35
0.5	1.02	1.1	1.24	1.34	1.43
0.4	1	1.02	1.03	1.09	1.12
0.3	1	1	1	1.02	1.02
0.2	1	1	1	1	1
0.1	1	1	1	1	1

From Table 1 to 4, the following trends are noticed: As β increases from 0.5 to 1.2, the ARL_1 values increase, i.e. from Table 4, for the shift c=0.9, the ARL_1 values are 16.29, 32.08, 64.35, 98.08, and 104.95 when $\beta = 0.5, 0.7, 0.9, 1.0 \text{ and } 1.2$ respectively.

- The control chart coefficient a and also the coefficient of control limit K increase as β changes from 0.5 to 1.2
- As shift c moves from 1 to 0.1, the ARL_1 decreases which is represented in Figure 1.

Table 2: *CC Parameters for ARL_0=250 and ARL_1 for different shifts when n=20 and $\theta = 3$*

β	0.5	0.7	0.9	1.0	1.2
UCL	17	17	16	18	16
LCL	4	4	3	5	3
a	0.8612	0.8768	0.8958	0.9292	0.9512
L	2.9031	2.9035	2.9058	2.9065	2.9139
Shift c			ARL_1		
1	251.5	250.22	250.61	250.82	252.87
0.9	39.39	40.78	67.15	88.11	93.99
0.8	5.55	7.77	13.83	15.87	18.56
0.7	2.06	2.5	2.98	5.67	7.29
0.6	1.29	1.4	1.52	2.36	3.3
0.5	1.07	1.09	1.18	1.43	1.5
0.4	1.01	1.02	1.08	1.12	1.2
0.3	1	1	1.04	1.08	1.14
0.2	1	1	1.01	1.03	1.05
0.1	1	1	1	1	1.01

Table 3: *CC Parameters for ARL_0=300 and ARL_1 for different shifts when n=20 and $\theta = 3$*

β	0.5	0.7	0.9	1.0	1.2
UCL	16	17	16	16	18
LCL	3	4	3	3	5
a	0.8437	0.8617	0.8807	0.9286	0.9562
L	2.9039	2.9053	2.9055	2.9061	2.9065
Shift c			ARL_1		
1	300.08	302.99	301.87	302.33	302.02
0.9	19.2	43.01	60.74	92.27	87.44
0.8	3.19	5.26	7.44	18.57	12.87
0.7	1.47	2.42	3.92	6.28	8.75
0.6	1.1	1.38	1.45	3.52	5.39
0.5	1.02	1.09	1.21	1.56	2.44
0.4	1	1.02	1.12	1.15	1.18
0.3	1	1	1.01	1.1	1.09
0.2	1	1	1	1.05	1.02
0.1	1	1	1	1.01	1

IV. Comparative Study

The out-of-control Average Run Length (ARL) values are utilized to evaluate and compare the performance of the proposed control charts against other similar control charts. Specifically, Table 5 at ARL_0=370 and ARL_0=300 illustrates the comparative analysis of out-of-control ARL values for the

proposed control charts with those based on the Exponentiated Inverse Kumaraswamy Distribution by Srinivasa Rao et. al., [23], Inverse Weibull Distribution by Baklizi and Ghannam [7], and Inverse Rayleigh Distribution by Nanthakumar, and Kavitha [17]. The analysis reveals that the proposed control chart demonstrates superior performance, consistently achieving smaller out-of-control ARL

Table 4: *CC Parameters for ARL_0=370 and ARL_1 for different shifts when n=20 and* $\theta = 3$

β	0.5	0.7	0.9	1.0	1.2
UCL	16	17	16	15	16
LCL	3	4	3	2	3
a	0.8497	0.8684	0.8392	0.8024	0.8361
L	2.9045	2.904	2.9032	2.8023	2.8067
Shift c			ARL 1		
1	372.27	372.68	371.03	372.08	370.11
0.9	16.29	32.08	64.35	98.08	104.95
0.8	2.99	8.85	16.64	22.48	32.08
0.7	1.44	1.83	4.32	8.62	10.08
0.6	1.09	1.19	1.36	4.12	6.25
0.5	1.01	1.03	1.08	1.84	2.25
0.4	1	1	1.01	1.2	1.14
0.3	1	1	1	1.14	1.18
0.2	1	1	1	1.08	1.10
0.1	1	1	1	1.02	1.04

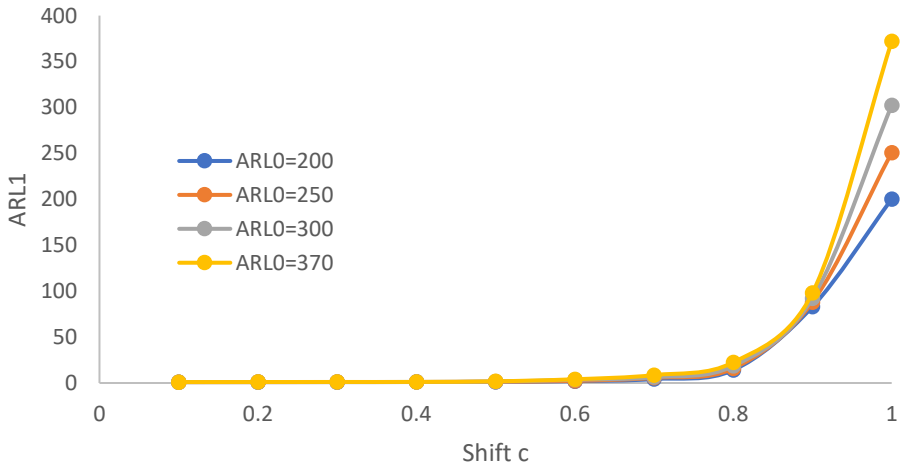


Figure 1: *ARL1 values for different shift c when* $\beta = 1, \theta = 3$ *when n=20*

values than the other charts. This is significant because control charts with lower out-of-control ARL values are more effective at quickly detecting shifts in the process, making them better suited for maintaining quality control.

Table 5: *Comparative analysis of out-of-control ARL values of similar CC*

Shift c	ARL_0=370			ARL_0=300		
	Exponentiated Inverse Kumaraswamy CC	Inverse Weibull CC	Proposed CC	Inverse Rayleigh CC	Exponentiated Inverse Kumaraswamy CC	Proposed CC
0	370.89	370.11	372.27	300.04	300.37	300.08
0.9	178.75	261.72	16.29	152.82	119.55	19.2
0.8	71.21	10.36	2.99	45.04	46.68	3.19
0.7	28.76	1.89	1.44	14.34	19.16	1.47
0.6	12.29	1.08	1.09	5.37	8.45	1.1
0.5	5.66	1	1.01	2.46	4.09	1.02
0.4	2.89	1	1	1.44	2.23	1
0.3	1.69	1	1	1.09	1.42	1
0.2	1.76	1	1	1.01	1.09	1
0.1	1.01	1	1	1	1	1

V. Study on Simulated Data

A simulated dataset was used to demonstrate the effectiveness of the proposed control charts. Specifically, two sets of 20 samples each, with a subgroup size of n=40, were randomly generated from an IKD to showcase the in-control and out-of-control performance of the control charts. These samples were generated using a target ARL_0=370, with a truncated life test time of t_0 = 0.8024(1.5) = 72 sec (approximately). The number of failed items was recorded in each subgroup and denoted as D and it is tabulated in Table 6. Based on Equation (5) from the study, the calculated control limits for the control charts are LCL=2 and UCL=15. The observed results are illustrated in Figure 2, which plots the number of failed items (D) against the subgroup index. The first 20 subgroups are in control, whereas the subsequent 20 subgroups reflect an out-of-control process due to a shift in the underlying distribution. The proposed control charts successfully detect the shift, with out-of-control signals observed at the 24th and 25th subgroups. These findings highlight the effectiveness of the proposed control charts in detecting process shifts, showcasing their reliability in identifying changes in quality within a timely manner.

VI. Real-Life Application

Suppose that a producer wants to enhance the quality of a product. It is known that the failure time of the product follows the IKD with shape parameter 3 and the scale parameter $\beta = 1.2$ and the target mean life is 1 hour and ARL_0=370, with the failure probability p_0 = 0.3254 from Eq. (7). Also, from Table 4, the control charts parameters such as sample size n = 20, LCL = 3, UCL = 16, a=0.8361 and L=2.8067 are obtained. Thus, the testing time t_0 =50 min.

Therefore, the proposed charts work as follows:

1. Randomly select 20 samples from each group.
2. Test the samples for t_0 =50 min.
3. Observe the MNF.
4. If $3 \leq MNF \leq 16$, declare the process as under control, otherwise, declare it as out of control.

Table 6: *Simulated Data*

Sample No	No. of Defectives (D)	Sample No	No. of Defectives (D)	Sample No	No. of Defectives (D)	Sample No	No. of Defectives (D)
1	11	11	11	21	7	31	5
2	8	12	5	22	6	32	14
3	4	13	12	23	5	33	8
4	6	14	5	24	17	34	13
5	12	15	11	25	16	35	9
6	14	16	9	26	11	36	7
7	4	17	5	27	10	37	8
8	2	18	15	28	13	38	9
9	6	19	5	29	13	39	11
10	3	20	5	30	14	40	12

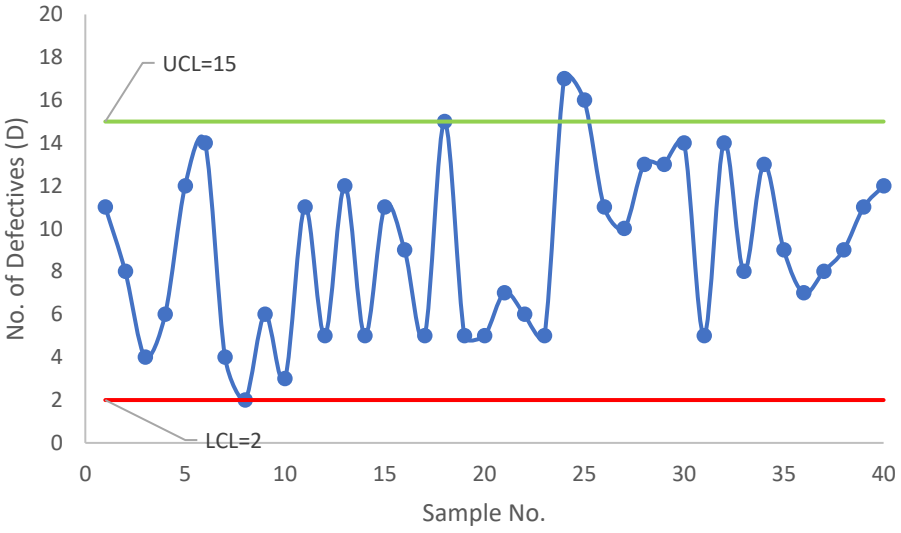


Figure 2: *Control chart for simulated data*

VII. Conclusions

This article proposes a new attribute control chart based on the Inverse Kumaraswamy distribution (IKD). The mean value of the IKD is used as the quality parameter to monitor the process. The scale parameter of the IKD is adjusted to accommodate various values, enabling the design of flexible control charts for different scenarios. The failure probability is calculated for the process shifts (c), and the corresponding out-of-control Average Run Length (ARL_1) is determined. Results for different process conditions are tabulated and compared to assess the chart's performance. Simulated data is used to evaluate the performance of the proposed control chart. The results demonstrate the chart's effectiveness in detecting shifts, as evidenced by the data and its ability to signal out-of-control conditions promptly. The applicability of the control chart is explained in real-life applications. This showcases its ability to monitor the lifespan of products that follow the IKD. The new control charts are suitable for industries that need to track product lifespans and ensure quality control, particularly for processes characterized by the IKD. The study suggests potential extensions, such as: employing the proposed control charts in repetitive sampling schemes to improve efficiency in monitoring

processes, incorporating neutrosophic statistics to handle uncertainty and imprecision in data, and further enhancing the applicability of the control charts in complex real-world scenarios. The proposed IKD-based control charts provide a robust tool for process monitoring, with proven efficacy in detecting shifts through both simulated and real-world data. This study lays the groundwork for future research in applying these charts to advanced sampling schemes and integrating statistical methods to handle uncertain environments. These advancements can enhance their utility in diverse industrial applications.

References

[1] Abbas, Z, Hafiz Zafar Nazir, Dongdong Xiang & Jianqing Shi, (2024). Nonparametric adaptive cumulative sum charting scheme for monitoring process location, *Quality and Reliability Engineering*, 40(5).

[2] Adeoti, O. A., & Ogundipe, P. (2020). A control chart for the generalized exponential distribution under time truncated life test, *Life Cycle Reliability and Safety Engineering*, doi.org/10.1007/s41872-020-00146-9.

[3] Al- Fattah, EL-Helbawy, & Al-Dayian, (2017). Inverted Kumaraswamy distribution: Properties and estimation, *Pakistan Journal of Statistics*, 33(1), 37-61.

[4] Ahangar, N.E., & Chimka, J. (2016). Attribute Control Charts with Optimal Limits, *Quality and Reliability Engineering International*, DOI:10.1002/qre.1839.

[5] Areej, M. & AL-Zaydi., (2024). On Moments of Inverse Kumaraswamy Distribution Based on Progressive Type-II Censored Order Statistics, *Journal of Statistical Theory and Applications*, https://doi.org/10.1007/s44199-024-00099-3.

[6] Aslam, M., & Jun, C.H (2015). Attribute control charts for the Weibull distribution under truncated life tests, *Quality Engineering*, 27(3), 283–288.

[7] Baklizi, A & Ghannam, S.A., (2022). An attribute control chart for the inverse Weibull distribution under truncated life tests, *Heliyon*, 8(11976).

[8] Daghistani, A.M., Al-Zahrani, B., and Shahbaz, M. Q., (2023). A New Inverse Kumaraswamy Family of Distributions: Properties and Applications, *Pakistan Journal of Statistics and Operation Research*, 19(2), 313-326.

[9] El-Sayed, A., El-Sherpieny & Mahmoud M. Elsehetry., (2019). Type II Kumaraswamy Half Logistic Family of Distributions with Applications to Exponential Model, *Annals of Data Science*, 6(1).

[10] Hawkins, D. M. and Olwell, D. H. (1998). Cumulative Sum Charts and Charting for Quality Improvement, Springer, New York.

[11] Ishaq, Panitanarak, U., Abiodun, A. A., Suleiman, A. A & Daud, H., (2024). The Generalized Odd Maxwell-Kumaraswamy Distribution: Its Properties and Applications, *Contemp. Math*. DOI: doi:10.37256/cm.5120242888.

[12] Ishaq, A. I., Suleiman, A. A., Daud, H., & Abba, S.I., (2023). Log-Kumaraswamy distribution: its features and applications, *Frontiers in Applied Mathematics and Statistics*, doi: 10.3389/fams.2023.1258961.

[13] Jing-Er Chiu & Tsen-I Kuo (2007). Attribute Control Chart for Multivariate Poisson Distribution, *Communications in Statistics - Theory and Methods*, 37(1),146-158.

[14] Joseph, J & Ravindran, M (2023). Transmuted Exponentiated Kumaraswamy Distribution, *Reliability: Theory and Applications*, 1(72).

[15] Lemontea, A.J., Wagner, B.S & Cordeiro, G.M., (2013). The exponentiated Kumaraswamy distribution and its log-transform, *Brazilian Journal of Probability and Statistics*, 27(1), 31-53.

[16] Linda Lee Ho & Quinino, RC (2013). An attribute control chart for monitoring the variability of a process, *International Journal of Production Economics*, 145(1), 263-267.

[17] Nanthakumar, C. & Kavitha, K. (2017). Design of Attribute control chart based on Inverse

Rayleigh distribution under Type- I censoring, *International Journal of Scientific Research in Mathematical and Statistical Sciences*, 4(6), 17-22.

[18] Page, E. S. (1954). Continuous Inspection Scheme, *Biometrika*, 41(1), 100–115.

[19] Rao, G. S. (2019). A control chart for time truncated life tests using exponentiated half logistic distribution, *Applied Mathematics & Information Sciences: An International Journal*, 12(1), 125–131.

[20] Rao, G.S., Fulment, A.K., & Josephat, P.K., (2019). Attribute control charts for the Dagum distribution under truncated life tests, *Life Cycle Reliability and Safety Engineering*, 8(4), 329–335.

[21] Robert.S (1959). Control chart tests based on geometric moving averages, *Technometrics*, 1(3), 239–50.

[22] Samson, K., Abubakar, M.A., & Asongo, Alhaji, (2020). Odd Lindley-Kumaraswamy Distribution: Model, Properties and Application, *Journal of Advances in Mathematics and Computer Science*, 35(6).

[23] Srinivasa Rao, B., Rami Reddy, M., & Rosaiah, K., (2023). An Attribute Control Chart for Time Truncated Life Tests Using Exponentiated Inverse Kumaraswamy Distribution, *Reliability: Theory and Applications*, 3(74).

[24] Sriramachandran, G V, Gadde, SR & Aslam, M. (2023). Utilizing Repetitive Sampling in the Construction of a Control Chart for Lindley Distribution, *Journal of Statistical Theory and Applications*, 23(4), 224-239.

ASSESSMENT OF POSSIBLE SEISMIC HAZARD AND RISK USING THE EXAMPLE OF ALMATY

Gennadii Nigmetov[1,a], Andrey Savinov[1,b], Temir Nigmetov[1,c], Syrym Gabbasov[2,d]

1) All-Russian Research Institute for Civil Defense and Emergencies, 7, Davidkovskaja Street, Moscow, 121352, Russian Federation, 2) Kazakh Agrotechnical Research University named after Saken Seifullin, Zhenis, avenue 62, Astana, 010011, Kazakhstan
a) tagirmaks@mail.ru b) savandr198@mail.ru c) t.nigmetov@yandex.ru d) taee@mail.ru

Abstract

In January 2024, residents of Almaty felt the seismic impact of an earthquake that occurred on the border of China and Kyrgyzstan. The city was not prepared for such an event. How could the city residents prepare for this devastating earthquake? In order to properly prepare the population for a destructive earthquake, it is necessary to have a system for assessing the individual seismic risk of the city, which uses data for risk assessment: on predicted possible earthquake sources (PES) dangerous for the city for the next 10 years, on the seismic resistance of buildings and the city's population. Based on these data, the possible individual risk for the city population is assessed. In order to assess the individual risk as accurately as possible, in addition to data on the PES and seismic resistance of buildings, data on the macroseismic field from the PES forecast are needed. At present, there are all the necessary scientific and technical capabilities for creating monitoring systems for seismic protection of cities.

Keywords: Dynamic-geophysical method, seismic hazard, possible earthquake source, building seismic resistance, individual seismic risk

I. Introduction

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nam dui ligula, fringilla a, euismod Since 1999, for several years, the Federal State Budgetary Institution All-Russian Research Institute of Civil Defense and Emergencies has been on duty 24 hours a day using the Extremum geographic information system (GIS). The GIS used an original probabilistic model for assessing the consequences of strong earthquakes. The model used a database of seismic resistance of buildings in all countries of the world, including Kazakhstan, data on tectonic faults and data on geology for its calculations. These data are used to construct a macroseismic field [1,2]. The model is based on the probability laws of destruction of buildings of various design depending on seismic impact [3].

Analysis of data on the rescue of people after the impact of a catastrophic earthquake, who were in the rubble of buildings and structures, shows that after 24 hours up to 47% of people in the rubble die, and after 3 days the percentage of fatalities increases to 60%, after 6 days the percentage of fatalities reaches 85%. The graph of the deaths of those trapped in the rubble is shown in Figure 2. Experience in assessing the consequences of catastrophic earthquakes shows that data on the real consequences in populated areas arrive with a significant delay measured in days, while it is important to know as accurately as possible the distribution of victims in the six-point macroseismic field zone of a catastrophic earthquake. The correct distribution of forces and resources provides a gain in time and the opportunity to save the maximum number of survivors.

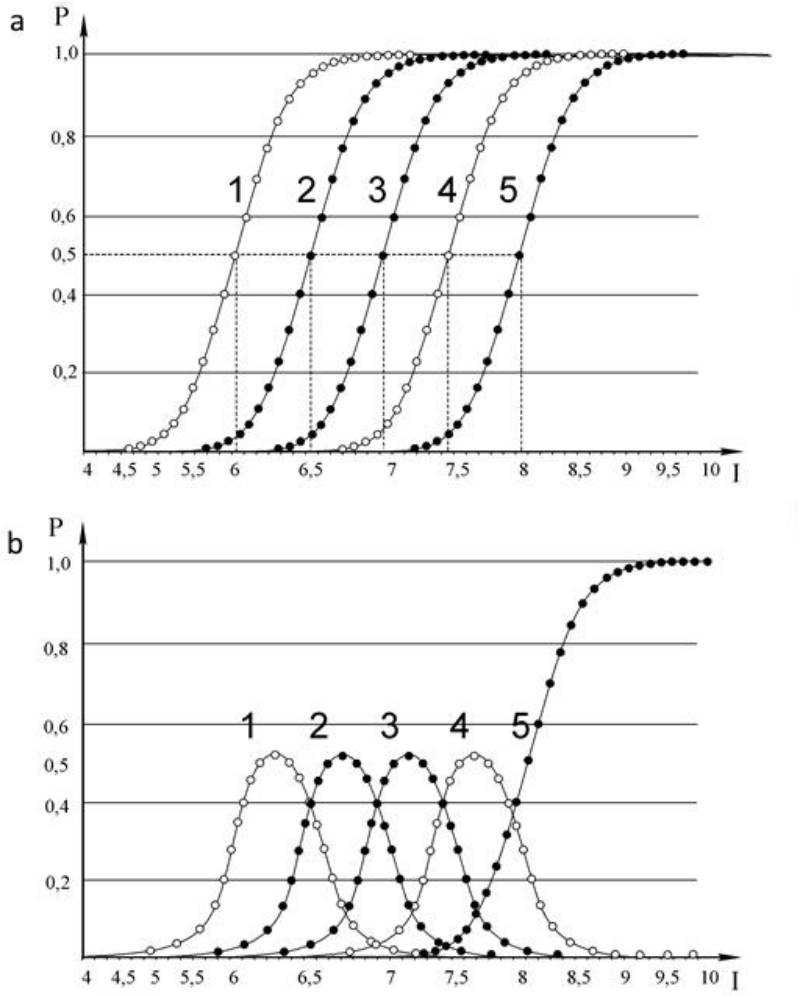


Figure 1: *Example of destruction laws during earthquakes for class A buildings (made of local materials) a — not less than certain degrees of damage; b — certain degrees of damage; 1, 2, 3, 4, 5 — degrees of damage to buildings.*

Analysis of data on the rescue of people after the impact of a catastrophic earthquake, who were in the rubble of buildings and structures, shows that after 24 hours up to 47% of people in the rubble die, and after 3 days the percentage of fatalities increases to 60%, after 6 days the percentage of fatalities reaches 85%. The graph of the deaths of those trapped in the rubble is shown in Figure 2. Experience in assessing the consequences of catastrophic earthquakes shows that data on the real consequences in populated areas arrive with a significant delay measured in days, while it is important to know as accurately as possible the distribution of victims in the six-point macroseismic field zone of a catastrophic earthquake. The correct distribution of forces and resources provides a gain in time and the opportunity to save the maximum number of survivors.

Using the GIS "Extremum" it is possible to quickly assess the possible consequences both according to the PES and according to promptly received data on the parameters of seismic events. The calculation of consequences is performed within a time not exceeding 0.1-0.5 hours. Based on the data received, an assessment of the individual risk for the population that fell into the 6-point zone of the territory in the epicentral zone of a catastrophic earthquake is performed.

As can be seen from Table 1, when calculating the consequences for the territories of Asia using the GIS "Extremum", out of 13 cases considered, in 7 cases the calculations of the consequences were 50% or more accurate, in the remaining 5 cases the calculations were erroneous. Errors arose due to incomplete data on development, population, scatter of data on the earthquake source, macroseismic field.

Table 1: *Examples of operational assessment of the consequences of strong earthquakes, carried out using the GIS "Extremum" in 2008.*

№	Date	Time	Longitude	Latitude	Magnitude	Depth km	Place	Irrecoverable losses	Sanitary losses	Fact Irrecoverable losses	Fact Sanitary losses	Convergence, %
1	18.06.2008	08:12:12.5	92,1	33,28	5,5	10	Qinghai, China	0-2	0-3	1	2	100
2	05.08.2008	09:49:15.9	105,59	32,86	6	10	Sichuan, China	12-34	0-38	4	35	100
3	25.08.2008	13:22:00.7	83,59	31,03	6,4	33	Xizan, China	2-11	0-15	> 42	> 30	50
4	27.08.2008	01:35:29.3	104,19	51,63	6,1	10	Lake Baikal, Russia	2139-4528	816-2533	0	0	0
5	10.09.2008	11:00:34.3	55,86	27,01	5,9	15	Southern Iran	16-49	0-61	7	30	50
6	11.09.2008	17:38:59.0	105,72	32,93	5,6	10	Sichuan, China	2-7	0-8	> 3	> 10	80
7	05.10.2008	15:52:49.2	73,75	39,49	6,6	40	Tajikistan-Xinjiang border region	19-75	1-82	> 70	> 50	100
8	05.10.2008	22:56:27.9	69,59	33,97	6	10	Southeast Afghanistan	1590-3396	577-2107	> 60	> 50	10
9	06.10.2008	08:30:43.4	90,48	29,76	6,2	10	Xizan, China	26-63	9-55	0	0	0
10	08.10.2008	14:07:15.2	90,4	29,77	5,6	10	Xizan, China	0-2	0-2	0	0	100
11	13.10.2008	17:16:11.8	70,28	38,67	5,6	33	Afghanistan - Tajikistan, border area	0-1	0-1	>72	>60	1,67
12	28.10.2008	23:09:56.3	67,47	30,74	6,5	15	Pakistan	6093-12589	2529 - 7185	>1000	>350	16,4
13	29.10.2008	11:32:42.6	67,66	30,68	6,5	15	Pakistan	4838-10088	1958 - 5802	> 150	> 82	3,1

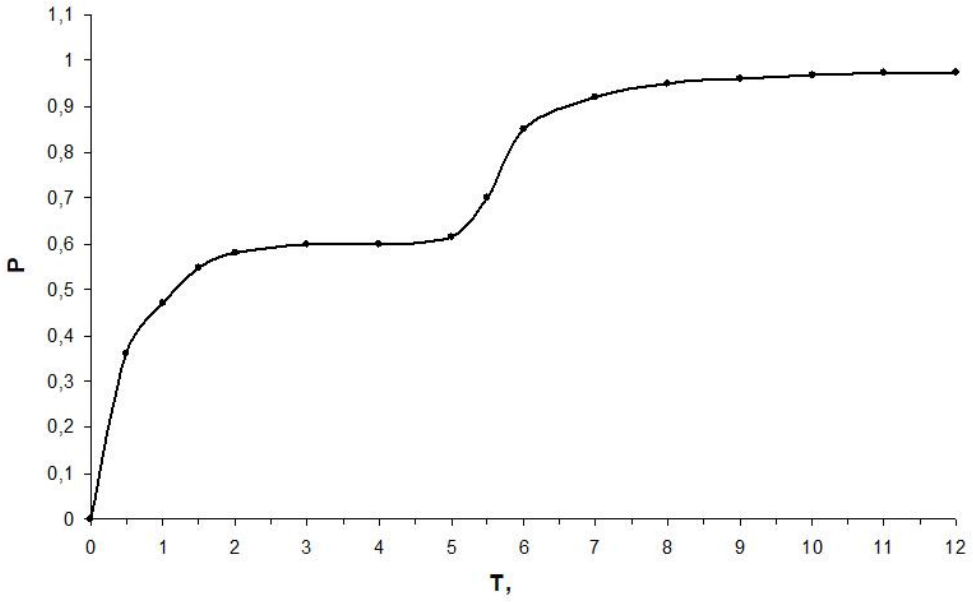


Figure 2: *Dynamics of fatalities among victims trapped in rubble.*

As can be seen from Table 1, when calculating the consequences for the territories of Asia using the GIS "Extremum", out of 13 cases considered, in 7 cases the calculations of the consequences were 50% or more accurate, in the remaining 5 cases the calculations were erroneous. Errors arose due to incomplete data on development, population, scatter of data on the earthquake source, macroseismic field.

To clarify data on seismic resistance of buildings, it is proposed to use mobile diagnostic complexes "Strela-Struna", "Kub", "Tolkyn-1". [2]

For effective planning of measures to protect the population, it is proposed to assess the individual seismic risk at various stages of seismic hazard:

1) when there is a threat of a possible earthquake;
2) when an earthquake occurs (main shock stage);
3) after the earthquake occurs (aftershock stage).

II. Materials and methods

In case of a possible earthquake threat, to calculate the individual risk for the population, it is necessary to have:

1) data on the coordinates and power of possible earthquake sources and the time of their possible triggering;
2) data on tectonics, geology and terrain in the considered possible epicentral 6-point zone;
3) data on the seismic resistance of buildings and structures in the considered zone;
4) macroseismic field from the PES;
5) the population that fell into the 6-point zone of the macroseismic field;
6) the mathematical expectation of population losses in the 6-point zone of the macroseismic field.

How can the parameters of the PES be determined? In the works [5,6,7] it was shown that the PES manifests itself much earlier than the main shock and it manifests itself in a complex

manner through mechanical, electromagnetic and hydrometeorological parameters.

The probability of a PES outbreak being triggered can be expressed through the following relationship:

$$P_z = 1 - \prod_{i=1}^{i=n} (\mathbf{1} - \boldsymbol{P_i}) \quad (1)$$

where Pz is the probability of the PES being triggered;
Pi – the probability that the i-th precursor will trigger the PES.

Table 2: *Parameters of precursors and their manifestation in time before earthquakes.*

№	Harbinger type	Time of manifestation before the main shock*	Characteristic types of manifestation of precursors*	Probability of triggering with possible parameters of PES
1	A sharp drop in atmospheric pressure (sharp gusts of wind)	1-20 day	A sharp 72-hour drop and rise in atmospheric pressure to 20 gPa	0,5-5,5 M
		1-7 day	A sharp 12-hour drop and rise in atmospheric pressure to 25 gPa	0,5-5,5 M
2	Manifestation of lightning discharges in a certain sequence	1-7 day	The seismically active region is being fought off	0,3-5,5 M
		4 day	Activity on tectonic faults	0,5-5,5 M
4	Manifestation of a special cloud portrait	1-3 day	Manifestation of linear prorans	0,5-5,5 M
		0,5-7 day	Tectonic faults and the epicenter are reflected linearly	0,7-5,5 M
5	Manifestation of foreshock activity or lull	1-20 day	Tectonic faults or epicenter were repulsed	0,5-5,5 M
		1-20 day May not manifest	Tectonic faults or epicenter were repulsed	0,5-5,5 M
6	Changes in the spectral composition of seismic signals	1-20 day	High frequency noise	0,7-5,5 M
		1-20 day	High and low frequency noises	0,7-5,5 M

*During the process of analysis and observation, the criteria may be refined.

Individual seismic risk is a complex value defined as the quotient of the value of the mathematical expectation of losses in the considered 6-point zone of a possible seismic event for the time during which an earthquake is expected and the number of people in the considered zone.

$$Re_i = P\frac{m_6}{T \times N_6} = [\frac{1}{year}] \quad (2)$$

where: m6 – mathematical expectation of losses in the considered 6-point zone of a possible earthquake, people;

T – the time during which a possible seismic event is predicted, year;

N_6 – the number of people located in the 6-point earthquake zone under consideration;

P - the probability of the PES being triggered during which an earthquake is predicted.

When exposed to an earthquake, it is necessary to immediately after receiving information about the parameters of the seismic event calculate the possible consequences and determine the possible individual risk.

After the impact of the main shock, using forecast data on possible aftershocks and data on the vulnerability of buildings and structures taking into account their damage from repeated impact of seismic events, it is necessary to clarify the possible individual risk. Thus, to calculate the value of individual risk, data on the possible seismic hazard, the seismic resistance of buildings and data on the possible number of people in the zone of dangerous seismic impact are needed.

To assess the activity of the PES, it is proposed to create monitoring networks in the area of probable PES, consisting of sensors, acceleration, speed, displacement, water level sensors in wells, and atmospheric pressure sensors.

The technical parameters of accelerometers installed at seismic stations have a frequency range from 0.5 Hz to 20 Hz and are intended only for timely detection of coordinates of earthquake hypocenters and their power and are not intended for monitoring the technical condition of the soil-structure system. Seismometric soil-structure monitoring systems must include at least four adjacent sensors: one sensor must be located on the ground next to the structure, the second at the base of the structure, the third in the middle part of the structure, and the fourth at the top of the structure. The frequency characteristics of soil-structure monitoring systems must have a range of 0.1-1000 Hz with a sensitivity of 1 V*m/s2. The task of the soil-structure monitoring system is not only to record the parameters of the impact of seismodynamic loads, but also to continuously determine the technical condition of the soil-structure system, its stability and seismic resistance. The studies conducted under the scientific supervision of the author in conditions of real impact of dynamic and seismic loads using multichannel networks provide grounds to conclude that it is possible to create monitoring technologies that ensure timely determination of the technical condition of soil-structure systems, their stability and seismic resistance.

III. Results

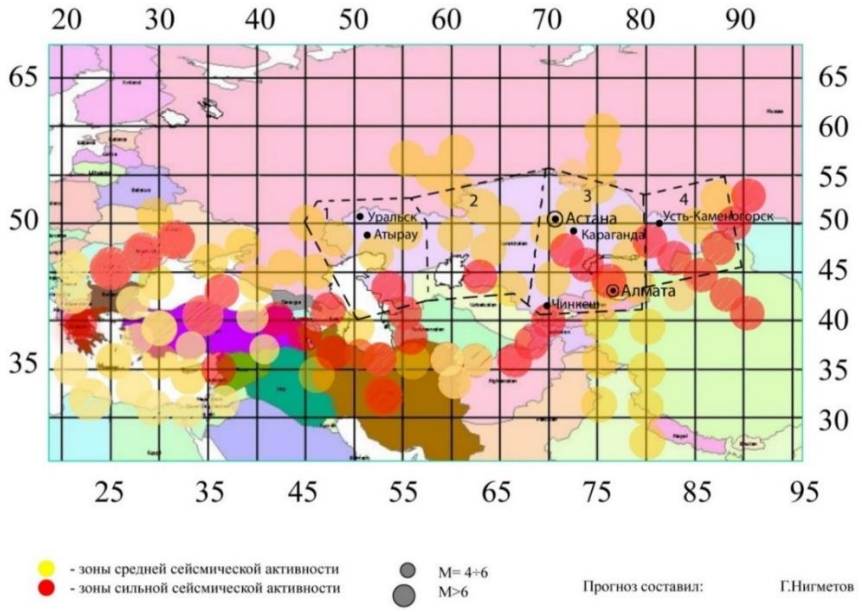


Figure 3: *PES map of the Almaty region, obtained in 2013-2018 as a result of the analysis of the manifestation of complex precursors.*

Let us consider one example of the application of a geographic information system for forecasting possible consequences of a possible earthquake with the parameters specified in Table 2. A PES map was obtained in the Almaty region.

To assess possible risks, the following parameters of the predicted earthquake hypocenter were used: coordinates 77 N, 42.85 E, M=7, depth 10 km (see Fig. 4 and 5), which coincided with the predicted zones. The risk forecast was completed in 2015 and reported at a conference in Almaty.

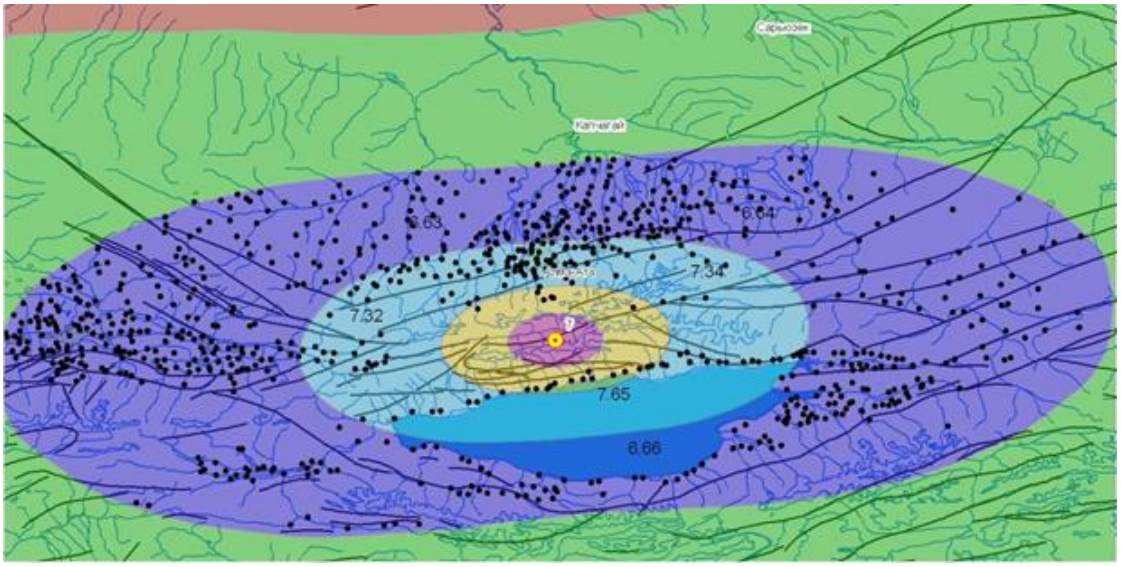


Figure 4: *Predicted field of vibration intensity for a possible earthquake 40 km south of Almaty.*

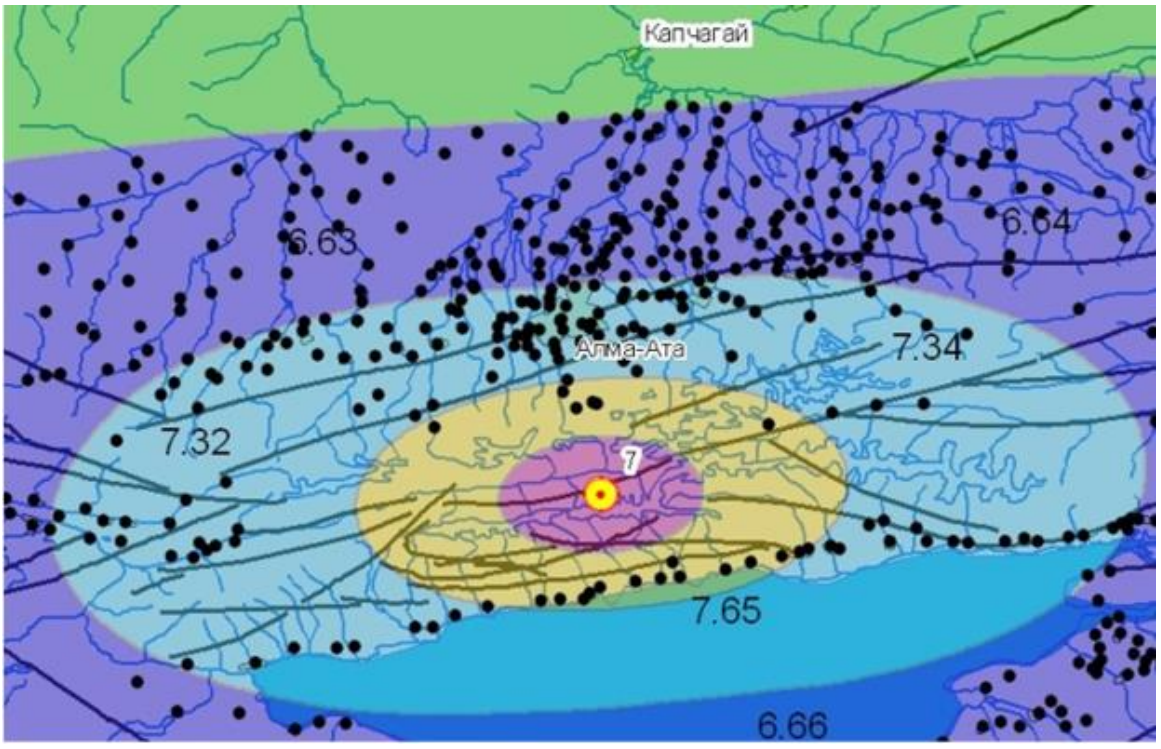


Figure 5: *Predicted field of oscillation intensity for a possible earthquake 40 km south of Almaty, with a close-up view of the oscillation intensity zones.*

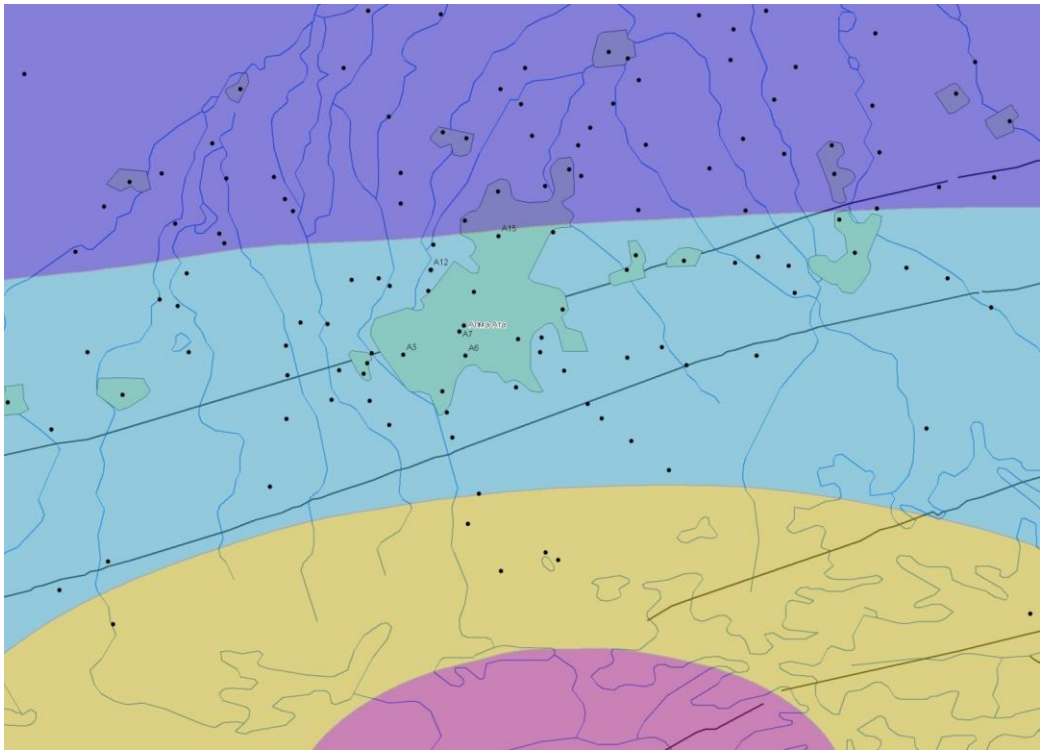


Figure 6: *Predicted field of oscillation intensity for a possible earthquake 40 km south of Almaty, with a close-up view of the oscillation intensity zones.*

Based on the results of the calculations performed for a possible earthquake source with the predicted parameters, the individual seismic risk will be:

$$Re_i = \frac{m}{T \times N} = \frac{300000}{10 \times 1500000}[\frac{1}{year}] = 2 \times 10^{-2}[\frac{1}{year}] \quad (3)$$

where: m is the mathematical expectation of losses in the considered zone of a possible earthquake, people;

T – time during which a possible seismic event is predicted, (1-10 years) year.;

N is the number of people (1,500,000 people) located in the 6-point earthquake zone.

Thus, the risk under this scenario is 2000 times higher than the norm.

For a timely response to a possible seismic hazard in the Almaty region, it is necessary to create a system for monitoring and forecasting the territory in the areas predicted by the PES.

IV. Discussion

According to the medium-term forecast of seismic activity of the territory of Eurasia for the period from 2013 onwards, including 2026, developed by Nigmetov G.M. and presented at the Russian Expert Council on Forecasting Seismic Hazard and Risk (see 2.1 and 3.1). This forecast was first reported to the RES in 2008 as a forecast until 2009, then at regular meetings of the RES it was extended. The forecast came true for all destructive earthquakes: China (Sichuan), Kyrgyzstan, Kazakhstan, Chechnya, Tajikistan, Italy, Japan.

As can be seen from the forecast based on data from Nigmetov G.M., the most seismically active part of the territory of Kazakhstan is the southeastern part of the territory of the Republic of Kazakhstan.

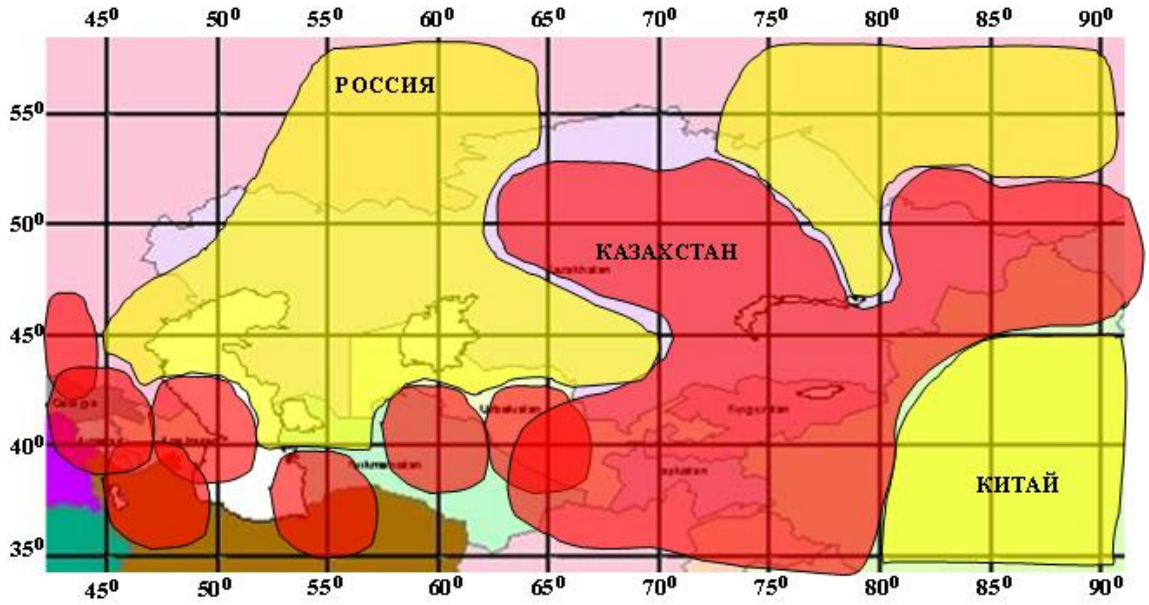


Figure 7: *Forecast of zones of increased seismic activity in the territory of the Republic of Kazakhstan and adjacent territories, which was presented by G.M. Nigmetov at a scientific and practical conference in 2009.*

Fig. 8-10 shows the manifestation of precursors before the earthquake of 04.03.24, which occurred on the border of Kazakhstan and Kyrgyzstan. As can be seen from the figures, the parameters of the precursors correspond to the criteria presented in Table № 2

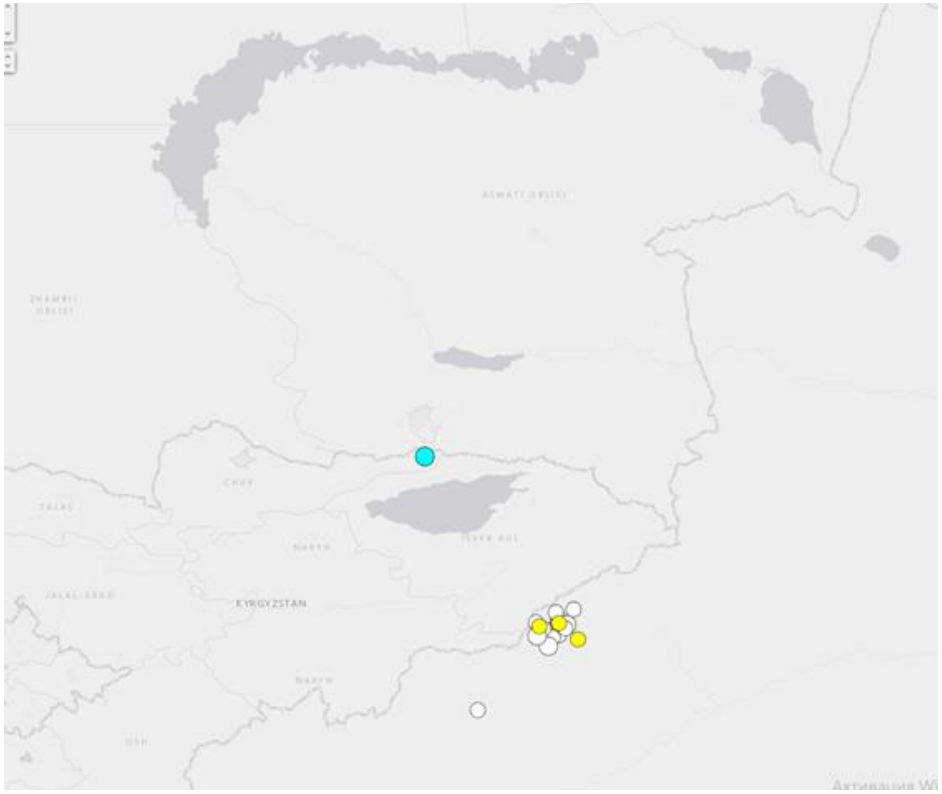


Figure 8: *Earthquake of magnitude 5.3 that occurred on 03/04/24 in the forecast zone on the border of Kazakhstan and Kyrgyzstan (according to the US Geological Survey).*

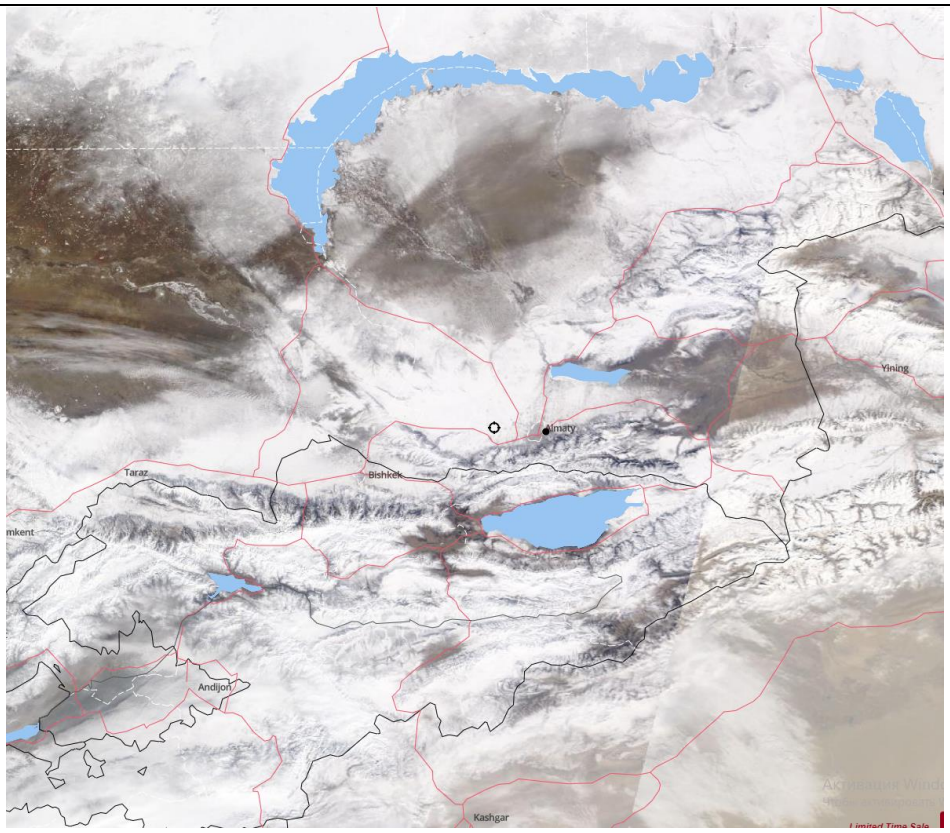


Figure 9: *Manifestation of a precursor in the cloudiness portrait before the earthquake that occurred on 04.03.24 on the border of Kazakhstan and Kyrgyzstan (according to the data of the image of 29.02.24 of the Terra satellite). A characteristic gap in the cloudiness in the area of the expected epicenter is visible.*

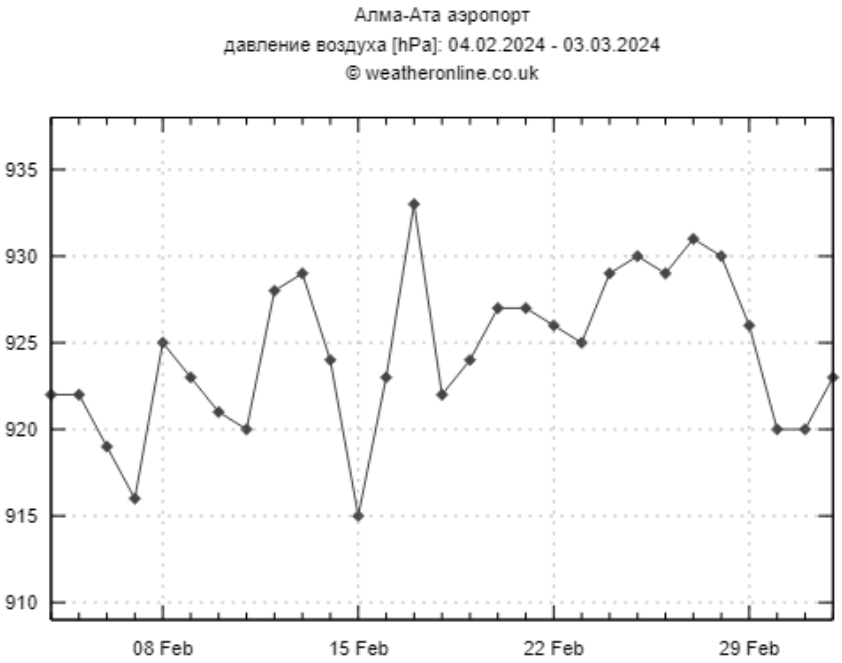


Figure 10: *Manifestation of a sharp drop in atmospheric pressure at the Almaty airport before the earthquake that occurred on 04.03.24 on the border of Kazakhstan and Kyrgyzstan (data obtained from Internet sources).*

To monitor seismic hazard and individual seismic risk for the population of the Almaty region, it is proposed to use the above-described end-to-end approach from PES forecasting, seismic resistance assessment to calculations of individual seismic risk using GIS.

At the first stage of the work, it is proposed:

1) to assess the seismic hazard of the territory in the vicinity of Almaty; for this, use the adapted models of Nigmetov G.M. and other authors for the analysis of geodynamic, seismic, geophysical, electromagnetic, meteorological, satellite data and determination of the coordinates of possible earthquake sources 2-7 days in advance with a magnitude of M less than 5 and M greater than 5; Determine possible earthquake sources;

2) perform work on the prompt assessment of seismic resistance of buildings and structures using the dynamic testing method, clarify the database on seismic resistance of buildings and objects in the territory of Eurasia;

3) adapt the model for assessing the consequences of strong earthquakes for the territory of Almaty and assessing individual seismic risk taking into account geological features and buildings;

4) develop a geoinformation system for the territory of Almaty and surrounding areas for predicting possible sources of strong earthquakes, predicting possible consequences of strong earthquakes, monitoring assessment of individual seismic risk and displaying monitoring information on seismic hazard and individual seismic risk taking into account individual blocks and buildings.

At the second stage of the work, it is proposed to modernize the technical and software equipment of the monitoring system. The modernization will be carried out based on the results of the trial operation of the system, clarification of the parameters of possible earthquake sources, a refined database of seismicity and seismic resistance, and refined criteria for assessing seismic hazard and risk for the region in question.

The results of the conducted research and the considered examples show that the accuracy of the results of forecasting the consequences of strong earthquakes depends on:

1) the accuracy of constructing the intensity field or macroseismic fields of earthquakes;

2) the degree of reliability of data on development and its seismic resistance;

3) the reliability of data on the population, tectonics and geology of the territories.

In the course of the conducted research and analysis, it was revealed that the terrain has a great influence on the intensity of the earthquake, as well as tectonic fault lines.

Long-term data calculated on GIS show that the geographic information system, with the help of which it is possible to construct and study in detail the macroseismic fields of earthquakes, constantly needs to update the data, in more expanded information about various territories of the world and its population, about their geological structure, as well as the latest data on tectonic faults, terrain and types of design solutions for development and its seismic resistance.

V. Conclusion

From higher research results it is clear that the individual risk for the population within 10 years. To introduce the Almaty norm in 2000, when urgent measures are required to reduce the risk, therefore it is necessary:

1) to specify the coordinates, depths and magnitudes of the earthquake zone at a distance of 100-200 km from the city boundaries;

2) to assess the seismic resistance of typical city buildings located in the i-th elements of the city districts and to update the database, presenting it in the form of the following table:

Table 3: *An example of the proposed form of a database on development for seismic resistance of buildings (for the updated GIS "Extremum" database).*

Name of the city district	1 element	2 element	3 element	i
X	7,2	6	7,1	
Y				

3) perform a risk assessment for i-th elements of city districts and identify elements with increased risk exceeding the risk standard by 2 or more times;

4) develop risk reduction measures (development of seismic hazard warning systems, increasing the seismic resistance of buildings, reducing seismicity at the base of buildings, demolishing emergency buildings and relocating the population to earthquake-resistant buildings)

5) after completing the measures, perform a re-assessment of the individual risk

6) if the individual risk is not exceeded in any of the i-th elements, the measures to prepare for a catastrophic earthquake are considered complete

References

[1] Larionov V.I., Nigmetov G.M., Nikolaev A.V., Sushchev S.P., Ugarov A.N., Kozlov M.A., Frolova N.I. et al. Global (world) geographic information system "Extremum" for decision-making on risk management and rapid response. In the collection: 25 years - from ideas to technologies Collection of scientific and technical works. Moscow, 2001. pp. 87-101.

[2] Nigmetov G.M. Problems of monitoring buildings and structures // Monitoring. Science and safety.-M.2011.-№21, pp.36-42.

[3] "Methodology for assessment and certification of engineering safety of buildings and structures", EMERCOM of Russia. Federal Center for Science and High Technologies "VNII GOChS" - M., 2003.

[4] Nigmetov G.M. Some issues of short-term forecasting of seismic hazard and risk. CATALOQUE of seismoforecasting research carried out in Azerbaijan territory in 2009, pp.147-154.

[5] Nigmetov G.M. Earth crust oscillations before destructive earthquakes. CATALOQUE of seismoforecasting research carried out in Azerbaijan territory in 2010, pp. 99-104.

[6] Nigmetov G.M., Akimov V.A., Maklakov A.S., Savinov A.M., Nigmetov T.G., Egorova A.A. Assessment of medium-term and long-term risks for the population and territories of the northwestern and central Caucasus from natural hazards Natural and man-made risks. Safety of structures. 2021. No. 3 (52). P. 34-37.

[7] Larionov V.I., Nigmetov G.M., Sushchev S.P., Ugarov A.N., Frolova N.I. Vulnerability and seismic risk assessment using GIS technologies from the occurrence of instability of building soil foundations during earthquakes. Journal "Seismic Resistant Construction", No. 2, 1999.

[8] Nigmetov G. M., Rybakov A. V., Savinov A. M., Nigmetov T. G. Modern approaches to assessing the risk of collapse of structures. Civil safety technologies. Vol. 15 2018. No. 2 (56) P. 26-29.

[9] GOST R22.10.02-2016 Safety in emergency situations, the procedure for substantiating and taking into account civil defense measures, measures to prevent natural and man-made emergencies when developing territorial planning documents

RT&A, No 2 (84)
Volume 20, June 2025

AN APPLICATIONS-BASED STOCHASTIC MODELING OF MIXTURE DISTRIBUTION FOR CANCER DATA

Poovitha R[1] and Pandiyan P[2]

•

[1]Research Scholar, Department of Statistics,
Annamalai University, Tamilnadu, India
[2]Professor, Department of Statistics,
Annamalai University, Tamilnadu, India
Corresponding Authour: [1]pranaveesai29@gamil.com, [2]pandiyanau@gmail.com

Abstract

This study proposed a new stochastic model by the mixture of Gamma and Zeghdoudi distribution (MGZ Distribution). The Reliability Analysis and statistical features such as moments, Moment generating function, order statistics, stochastic ordering, entropy, and the Maximum Likelihood Estimation of the model parameters were derived. Two real-time datasets of cancer patients' survival times were used to demonstrate the model's goodness-of-fit using the AIC, BIC, and AICC model selection methods. The new MGZ model has been compared with the other classical models like Exponential, Lindley and Shanker. The result shows that the proposed model is more flexible than the other models with the lowest values of AIC, BIC and AICC.

Keywords: Mixture distribution, Moments, Order Statistics, Maximum likelihood Estimation, Goodness of fit Criteria.

1. Introduction

The statistical distributions have been extensively utilized for analyzing time-to-event data, also referred to as survival or reliability data, in different areas of applicability, including medical science. In recent years, an impressive set of new statistical distributions has been explored by statisticians. The development of an expanded class of classical distribution is required in analysis, biomedicine, reliability, insurance, and finance. Recently, many researchers have been working in this area and have proposed new methods to develop improved probability distributions with utility. A statistical analysis is often employed, which heavily relies on the assumed probability model or distributions.

A Zeghdoudi distribution [8] random variable X with a parameter $\theta > 0$ is described by its probability density function (pdf)

$$f_1(x;\theta) = \frac{\theta^3}{2+\theta} x(1+x)e^{-\theta x}, \quad x > 0, \quad \theta > 0 \tag{1}$$

Considering the gamma distribution [18] with parameters $\beta = 3, \theta$, the pdf can be defined as:

$$f_2(x;\theta) = \frac{1}{2}\theta^3 x^2 e^{-\theta x}, \quad x > 0, \quad \theta > 0 \tag{2}$$

Lindley [5] introduced the fiducial distribution and Bayes theorem. Rama Shanker [17] proposed a mixture of exponential (θ) and gamma $(2, \theta)$ distributions, leading to the Shanker

distribution. Shanker [14] also combined exponential (θ), gamma $(2,\theta)$, and gamma $(4,\theta)$ distributions with respective mixing proportions $\frac{\theta^3}{\theta^3+\theta^2+6}$, $\frac{\theta^2}{\theta^3+\theta^2+6}$, and $\frac{6}{\theta^3+\theta^2+6}$ This distribution is known as the Uma distribution. The Akash distribution is a two-component mixture of an exponential and a gamma distribution with mixing proportions $\frac{\theta^2}{\theta^2+2}$, and $\frac{2}{\theta^2+2}$.

Shanker [16] later proposed the Komal distribution with applications in survival analysis. Rama Shanker [13] also worked on combining exponential (θ) and gamma $(2,\theta)$ distributions with mixing proportions $\frac{\theta(\theta+1)}{\theta^2+\theta+1}$, and $\frac{1}{\theta^2+\theta+1}$ This article introduces the (MGZD), which was developed using a blend of Gamma and Zeghdoudi distributions. The probability density function (pdf) and cumulative distribution function (cdf) of the proposed distribution, along with several statistical characteristics and a reliability analysis, are discussed in the following sections. The maximum likelihood estimation (MLE) method is used to estimate model parameters. Cancer remains one of the leading causes of morbidity and mortality worldwide, impacting millions of lives annually. Despite advancements in cancer detection, treatment, and prevention, the global cancer burden continues to rise due to population aging, expansion, and shifts in risk factors. According to Sung et al. [19], the most recent global cancer data in 2020 showed approximately 10 million cancer-related deaths and 19.3 million new cases. The most common diagnoses were breast, lung, colon, prostate, and stomach cancers. The American Cancer Society [3] reported that advancements in early detection and treatment have improved survival rates, with breast cancer 5-year survival rates exceeding 85% in high-income countries. The World Health Organization [20] estimated the global economic impact of cancer at over $1 trillion annually, covering direct treatment costs and indirect losses such as productivity declines. Additionally, regional disparities in cancer incidence and mortality exist due to differences in healthcare access, risk factors, and early detection services. Approximately 70% of cancer deaths occur in low- and middle-income countries.

Statistical models are essential for studying cancer due to its complexity and heterogeneity. These models estimate survival rates, assess treatment effectiveness, and predict cancer risk. By incorporating new data and trends, these models improve predictions, enhance patient outcomes, and guide public health policies. Lastly, the results of fitting various well-known distributions to cancer survival data using (MGZD) are presented. All computations in this research were performed using the statistical programming language R.

2. New Mixture Distribution

This section introduces the MGZ distribution, a new distribution created by combining two existing distributions. Let X be a random variable with a mixed distribution. Its probability density function (p.d.f), $f(x)$, is expressed as follows

$$f(x) = \sum_{i=1}^{k} \omega_i f_i(x)$$

where $f_i(x)$ is the probability density function for all i, and ω_i, $i = 1,\ldots,k$, denote mixing proportions that are non-negative and satisfy

$$\sum_{i=1}^{k} \omega_i = 1$$

The components of the mixture are - $f_1(x) \sim$ Gamma$(3,\theta)$ with parameters $\beta = 3,\theta$ - $f_2(x) \sim$ Zeghdoudi(θ) with parameter θ. These two independent random variables have mixing proportions $\frac{\beta}{\beta+1}$ and $\frac{1}{\beta+1}$, respectively. Thus, the density function of X is given by

$$f(x;\theta,\beta) = \frac{\theta^3}{\beta+1}\left(\frac{\beta}{2}x^2 + \frac{x}{2+\theta}(1+x)\right)e^{-\theta x} \tag{3}$$

The function defined in Equation (3) represents a valid probability density function $f(x;\theta,\beta)$ for all $x > 0$

$$\int_0^\infty f(x;\theta,\beta)\,dx = \int_0^\infty \frac{\theta^3}{\beta+1}\left(\frac{\beta}{2}x^2 + \frac{x}{2+\theta}(1+x)\right)e^{-\theta x}\,dx$$

after simplification, we get

$$\int_0^\infty f(x;\theta,\beta)\,dx = 1$$

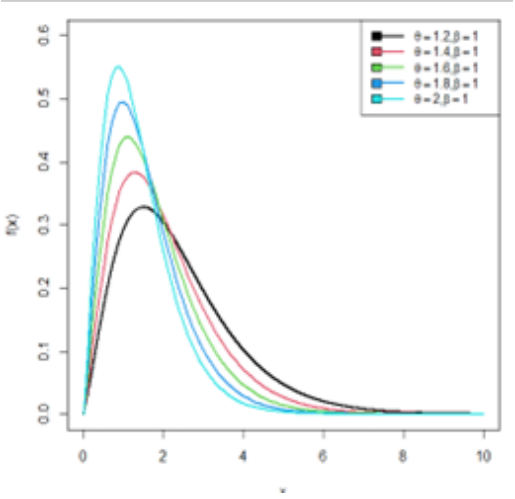


Figure 1: *PDF plot of MGZD different θ and β.*

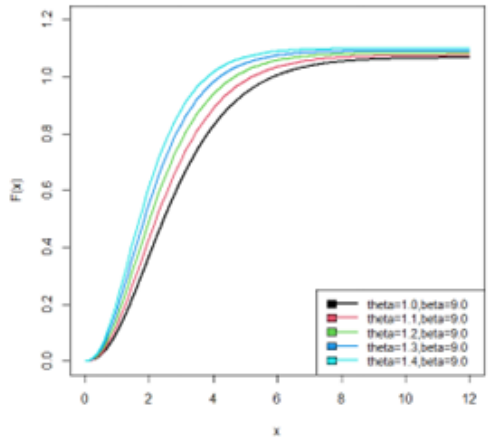


Figure 2: *CDF plot of MGZD different θ and β.*

The cumulative distribution function (c.d.f) of MGZD is defined as

$$F(x;\theta,\beta) = \int_0^x \frac{\theta^3}{\beta+1}\left(\frac{\beta}{2}x^2 + \frac{x}{2+\theta}(1+x)\right)e^{-\theta x}\,dx$$

after simplification, we obtain

$$F(x;\theta,\beta) = \frac{1}{\beta+1}\left[\frac{\beta\gamma(3,\theta x)}{2} + \frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{2+\theta}\right], \quad x>0,\theta,\beta>0 \tag{4}$$

3. Reliability Analysis

This section provides the reliability function, hazard function, reverse hazard function, cumulative hazard function, odds rate, Mills ratio, and mean residual function for the MGZ distribution.

3.1. Survival Function

The survival function of the MGZ distribution is defined as

$$S(x) = 1 - F(x;\theta,\beta)$$

$$S(x) = 1 - \frac{1}{\beta+1}\left[\frac{\beta\gamma(3,\theta x)}{2} + \frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{2+\theta}\right]$$

3.2. Hazard Rate Function

An important metric for describing life phenomena is the hazard rate function of the MGZ distribution, which is defined as

$$h(x) = \frac{f(x;\theta,\beta)}{1-F(x;\theta,\beta)}$$

$$h(x) = \left(\frac{\theta^3 \left(\frac{\beta}{2} x^2 + \frac{x}{(2+\theta)} (1+x) \right) e^{-\theta x}}{(\beta+1) - \left[\frac{\beta\gamma(3,\theta x)}{2} + \frac{\theta\gamma(2,\theta x) + \gamma(3,\theta x)}{2+\theta} \right]} \right)$$

3.3. Reverse Hazard Function

The reverse hazard rate function of the MGZ distribution is defined as

$$h_r(x) = \frac{f(x;\theta,\beta)}{F(x;\theta,\beta)}$$

$$h_r(x) = \left(\frac{\theta^3 \left(\frac{\beta}{2} x^2 + \frac{x}{(2+\theta)} (1+x) \right) e^{-\theta x}}{\left[\frac{\beta\gamma(3,\theta x)}{2} + \frac{\theta\gamma(2,\theta x) + \gamma(3,\theta x)}{2+\theta} \right]} \right)$$

3.4. Cumulative Hazard Function

The cumulative hazard function of the MGZ distribution is defined as

$$H(x) = -\ln(1 - F(x;\theta,\beta))$$

$$H(x) = \ln \left(\frac{1}{\beta+1} \left[\frac{\beta\gamma(3,\theta x)}{2} + \frac{\theta\gamma(2,\theta x) + \gamma(3,\theta x)}{2+\theta} \right] - 1 \right)$$

3.5. Odds Rate Function

The odds rate function of the MGZ distribution is defined as

$$O(x) = \frac{F(x;\theta,\beta)}{1 - F(x;\theta,\beta)}$$

$$O(x) = \left(\frac{\left[\frac{\beta\gamma(3,\theta x)}{2} + \frac{\theta\gamma(2,\theta x) + \gamma(3,\theta x)}{2+\theta} \right]}{(\beta+1) - \left[\frac{\beta\gamma(3,\theta x)}{2} + \frac{\theta\gamma(2,\theta x) + \gamma(3,\theta x)}{2+\theta} \right]} \right)$$

3.6. Mean Residual Function

The mean residual function of the MGZ distribution is defined as

$$M(x) = \frac{1}{S(x;\theta,\beta)} \int_x^\infty t f(t;\theta,\beta) dt - x$$

$$M(x) = \frac{1}{1 - \frac{1}{\beta+1} \left[\frac{\beta\gamma(3,\theta x)}{2} + \frac{\theta\gamma(2,\theta x) + \gamma(3,\theta x)}{2+\theta} \right]} \int_x^\infty t \frac{\theta^3}{\beta+1} \left(\frac{\beta}{2} x^2 + \frac{x}{(2+\theta)} (1+x) \right) e^{-\theta x} \, dt - x$$

using the substitution method and

$$\Gamma(s,x) = \int_x^\infty t^{s-1} e^{-t} dt,$$

we obtain

$$M(x) = \left[\frac{\frac{\beta\Gamma(4,\theta x)}{2} + \frac{\theta\Gamma(3,\theta x) + \Gamma(4,\theta x)}{2+\theta}}{\theta(\beta+1) - \left(\frac{\beta\gamma(3,\theta x)}{2} + \frac{\theta\gamma(2,\theta x) + \gamma(3,\theta x)}{2+\theta} \right)} \right] - x$$

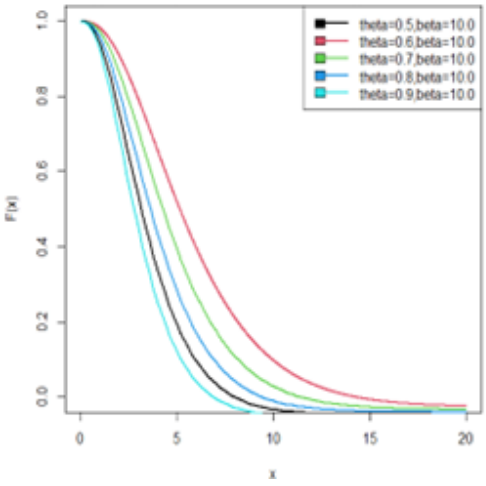


Figure 3: *Survival function of MGZD.*

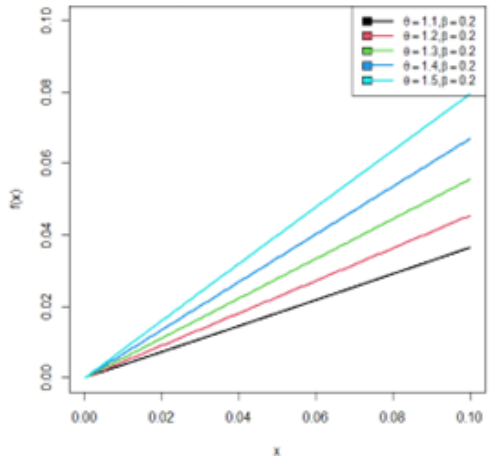


Figure 4: *Hazard function of the MGZD.*

4. Statistical Properties

In this section, we derive the structural properties, moments, moment-generating function, characteristic function, and the r^{th} moment for the MGZ distribution. We also investigate the mean, variance, coefficient of variation, skewness, kurtosis, and dispersion.

4.1. Moments

The r^{th} moment of a random variable X is defined as

$$E(X^r) = \mu'_r = \int_0^\infty x^r f(x;\theta,\beta)\, dx$$

$$\mu'_r = \int_0^\infty x^r \frac{\theta^3}{\beta+1}\left(\frac{\beta}{2}x^2 + \frac{x}{(2+\theta)}(1+x)\right) e^{-\theta x}\, dx$$

on simplifying the integral and using, $\frac{\Gamma(t)}{a^t} = \int_0^\infty x^{t-1}e^{-ax}dx$, we obtain

$$\mu'_r = \frac{1}{\theta^r(\beta+1)}\left[\frac{\beta\Gamma(r+3)}{2} + \frac{\theta\Gamma(r+2)+\Gamma(r+3)}{2+\theta}\right] \tag{5}$$

where $\Gamma(.)$ is the gamma function. By substituting $r = 1, 2, 3$, and 4 in Equation (5), we obtain

$$E(X) = \text{Mean} = \left(\frac{6(\beta+1)+\theta(3\beta+2)}{\theta(\beta+1)(2+\theta)}\right)$$

$$E(X^2) = \left(\frac{24(\beta+1)+6\theta(2\beta+1)}{\theta^2(\beta+1)(2+\theta)}\right)$$

$$E(X^3) = \left(\frac{120(\beta+1)+6\theta(10\beta+4)}{\theta^3(\beta+1)(2+\theta)}\right)$$

$$E(X^4) = \left(\frac{720(\beta+1)+10\theta(36\beta+12)}{\theta^4(\beta+1)(2+\theta)}\right)$$

4.2. Variance

The variance is given by $\sigma^2 = E(X^2) - [E(X)]^2$. On substituting the values, we get

$$\sigma^2 = \left[\frac{24(\beta+1)+6\theta(2\beta+1)}{\theta^2(\beta+1)(2+\theta)}\right] - \left[\frac{6(\beta+1)+\theta(3\beta+2)}{\theta(\beta+1)(2+\theta)}\right]^2$$

after simplification

$$\sigma^2 = \left(\frac{3\beta^2(\theta^2+4\theta+4) + 6\beta(\theta^2+4\theta+4) + 2(\theta^2+6\theta+6)}{\theta^2(\beta+1)^2(2+\theta)^2} \right)$$

4.3. Coefficient of Variation

The coefficient of variation (C.V.) is defined as

$$C.V. = \frac{\sigma}{\mu}$$

$$C.V. = \frac{\sqrt{3\beta^2(\theta^2+4\theta+4) + 6\beta(\theta^2+4\theta+4) + 2(\theta^2+6\theta+6)}}{6(\beta+1) + \theta(3\beta+2)}$$

4.4. Skewness

The skewness is given by

$$\text{Skewness} = \sqrt{\beta_1} = \frac{E(X^3)}{\sigma^3}$$

after simplification

$$\text{Skewness} = \left(\frac{(120(\beta+1) + 6\theta(10\beta+4))}{\left[(3\beta^2(\theta^2+4\theta+4) + 6\beta(\theta^2+4\theta+4) + 2(\theta^2+6\theta+6)\right]^{3/2}} \right)$$

4.5. Kurtosis

The kurtosis is defined as

$$\text{Kurtosis} = \beta_2 = \frac{E(X^4)}{\sigma^4}$$

$$\text{Kurtosis} = \left(\frac{[720(\beta+1) + 10\theta(36\beta+12)]}{\left[(3\beta^2(\theta^2+4\theta+4) + 6\beta(\theta^2+4\theta+4) + 2(\theta^2+6\theta+6)\right]^2} \right)$$

4.6. Moment Generating Function

The moment-generating function (MGF) of a random variable X is denoted by $M_X(t)$ and is defined as

$$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx} f(x;\theta,\beta)\, dx, \quad t \in \mathbb{R}$$

$$M_X(t) = \frac{\theta^3}{\beta+1} \int_0^\infty e^{tx} \left(\frac{\beta}{2}x^2 + \frac{x}{2+\theta}(1+x) \right) e^{-\theta x}\, dx$$

simplifying and by applying the Taylor series $e^x = \sum_{j=0}^{\infty} \frac{t^j}{j!}\ \forall\ \infty < x < \infty$, we obtain

$$M_X(t) = \int_0^\infty \left(1 + tx + \frac{(tx)^2}{2!} + \dots \right) \left(\frac{\beta}{2}x^2 + \frac{x}{2+\theta}(1+x) \right) e^{-\theta x}\, dx$$

$$M_X(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!} \int_0^\infty x^j f(x;\theta,\beta)\, dx$$

$$M_X(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!} E(x^j)$$

$$M_X(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!\theta^j(\beta+1)} \left[\frac{\beta\Gamma(j+3)}{2} + \frac{\theta\Gamma(j+2) + \Gamma(j+3)}{2+\theta} \right]$$

4.7. Characteristic Function

The characteristic function of a random variable X is denoted by $\phi_X(t)$ and is defined as

$$\phi_X(t) = E(e^{itX}) = \int_0^\infty e^{itx} f(x;\theta,\beta)\, dx$$

using the definition of the moment-generating function

$$\phi_X(t) = M_X(it)$$

$$\phi_X(t) = \sum_{j=0}^{\infty} \frac{(it)^j}{j!\theta^j(\beta+1)} \left[\frac{\beta\Gamma(j+3)}{2} + \frac{\theta\Gamma(j+2)+\Gamma(j+3)}{2+\theta}\right]$$

5. Order Statistics

Let $X_{(1)}, X_{(2)}, ..., X_{(n)}$ be the order statistics of a random sample $X_1, X_2, ..., X_n$ from a continuous population with probability density function $f_X(x)$ and cumulative density function $F_X(x)$. Then, the pdf of the r^{th} order statistic $X_{(r)}$ can be written as

$$f_{r:n}(x) = \frac{n!}{(r-1)!(n-r)!}[F(x;\theta)]^{r-1}[1-F(x;\theta)]^{n-r} f(x;\theta), \quad x > 0. \tag{6}$$

simplifying Equation (6) and substituting Equations (3) and (4), we obtain

$$f_{r:n}(x) = \frac{n!}{(r-1)!(n-r)!} \sum_{j=0}^{\infty} (-1)^j \binom{n-r}{j} \left(\frac{1}{\beta+1}\right)^{r+j-1}$$

$$\times \left[\frac{\beta\gamma(3,\theta x)}{2} + \frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{2+\theta}\right]^{r+j-1}$$

$$\times \frac{\theta^3}{\beta+1}\left(\frac{\beta}{2}x^2 + \frac{x}{2+\theta}(1+x)\right) e^{-\theta x}$$

First order statistic ($X_{(1)}$)

$$f_{1:n}(x) = \frac{n!}{(n-1)!} \sum_{j=0}^{\infty} (-1)^j \binom{n-1}{j} \left(\frac{1}{\beta+1}\right)^{j}$$

$$\times \left[\frac{\beta\gamma(3,\theta x)}{2} + \frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{2+\theta}\right]^{j}$$

$$\times \frac{\theta^3}{\beta+1}\left(\frac{\beta}{2}x^2 + \frac{x}{2+\theta}(1+x)\right) e^{-\theta x}$$

n^{th} Order statistic ($X_{(n)}$)

$$f_{n:n}(x) = \frac{n!}{(n-1)!} \sum_{j=0}^{\infty} (-1)^j \binom{n-1}{j} \left(\frac{1}{\beta+1}\right)^{n+j-1}$$

$$\times \left[\frac{\beta\gamma(3,\theta x)}{2} + \frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{2+\theta}\right]^{n+j-1}$$

$$\times \frac{\theta^3}{\beta+1}\left(\frac{\beta}{2}x^2 + \frac{x}{2+\theta}(1+x)\right) e^{-\theta x}$$

6. Stochastic Ordering

A crucial technique in reliability and finance for evaluating the relative performance of models is stochastic ordering. Let X and Y be two random variables with probability density functions $f(x)$ and $f(y)$, cumulative distribution functions $F(x)$ and $F(y)$, and survival functions $S(x) = 1 - F(x)$ and $S(y) = 1 - F(y)$. Then,

- Mean residual life order (MRLO) is denoted as

$$X \leq_{MRLO} Y, \quad \text{if} \quad m_X(x) \leq m_Y(y), \quad \forall x$$

- Hazard rate order (HRO) is denoted as

$$X \leq_{HRO} Y, \quad \text{if} \quad \frac{S_X(x)}{S_Y(y)} \quad \text{is decreasing for } x \geq 0$$

- Stochastic order (SO) is denoted as

$$X \leq_{SO} Y, \quad \text{if} \quad S_X(x) \leq S_Y(y), \quad \forall x$$

Likelihood ratio ordering proof, assume that X and Y are two independent random variables with probability density functions $f_X(x;\theta,\beta)$ and $f_Y(x;\lambda,\alpha)$. If $\theta < \lambda$ and $\beta < \alpha$, then the likelihood ratio is

$$\Lambda = \frac{f_X(x;\theta,\beta)}{f_Y(x;\lambda,\alpha)}$$

substituting the density functions

$$\Lambda = \frac{\left[\frac{\theta^3}{\beta+1}\left(\frac{\beta}{2}x^2 + \frac{x}{2+\theta}(1+x)\right)e^{-\theta x}\right]}{\left[\frac{\lambda^3}{\alpha+1}\left(\frac{\alpha}{2}x^2 + \frac{x}{2+\lambda}(1+x)\right)e^{-\lambda x}\right]}$$

therefore,

$$\log(\Lambda) = \log\left[\frac{\theta^3(\alpha+1)}{\lambda^3(\beta+1)}\right] + \log\left[\frac{\beta}{2}x^2 + \frac{x}{2+\theta}(1+x)\right] - \log\left[\frac{\alpha}{2}x^2 + \frac{x}{2+\lambda}(1+x)\right] - (\theta-\lambda)x$$

differentiating with respect to x

$$\frac{\partial \log(\Lambda)}{\partial x} = \left(\frac{\left[\beta x + \frac{2x+1}{2+\theta}\right]}{\left[\frac{\beta}{2}x^2 + \frac{x}{2+\theta}(1+x)\right]} - \frac{\left[\alpha x + \frac{2x+1}{2+\lambda}\right]}{\left[\frac{\alpha}{2}x^2 + \frac{x}{2+\lambda}(1+x)\right]} + (\lambda-\theta)\right)$$

hence,

$$\frac{\partial \log(\Lambda)}{\partial x} < 0$$

if $\theta < \lambda$ and $\beta < \alpha$,

Thus, the likelihood ratio is decreasing, proving that the MGZ distribution follows likelihood ratio ordering.

7. Bonferroni and Lorenz Curves

In this section, the Bonferroni and Lorenz curves are derived using the MGZ distribution. These curves are powerful tools for analysing distributions and have diverse applications in fields such as economics, insurance, income inequality, reliability, and medicine. For a random variable X

and its probability density function $f(x)$, the Bonferroni and Lorenz curves are constructed. Here, $f(x)dx$ represents the probability that a randomly selected unit falls within a given interval.

$$B(p) = \frac{1}{p\mu_1'} \int_0^q xf(x;\theta,\beta)\, dx \quad \text{and}$$

$$L(p) = \frac{1}{\mu_1'} \int_0^q xf(x;\theta,\beta)\, dx,$$

where $q = F^{-1}(p)$, with $q \in [0,1]$ and

$$\mu_1' = \left(\frac{6(\beta+1)+\theta(3\beta+2)}{\theta(\beta+1)(2+\theta)} \right)$$

Thus, the Bonferroni and Lorenz curves for our distribution are given by

$$B(p) = \frac{1}{p\mu_1'} \int_0^q x \frac{\theta^3}{\beta+1} \left(\frac{\beta}{2}x^2 + \frac{x}{(2+\theta)}(1+x) \right) e^{-\theta x}\, dx.$$

integrating by the substitution method, we obtain

$$B(p) = \left(\frac{\left[\frac{\beta\gamma(3,\theta q)}{2} + \frac{\theta\gamma(2,\theta q)+\gamma(3,\theta q)}{2+\theta} \right]}{p\left[\frac{6\beta(\theta+1)+3\theta\beta+2\theta}{2+\theta} \right]} \right) \tag{7}$$

The Lorenz curve is given by

$$L(p) = pB(p)$$

$$L(p) = \left(\frac{\left[\frac{\beta\gamma(3,\theta q)}{2} + \frac{\theta\gamma(2,\theta q)+\gamma(3,\theta q)}{2+\theta} \right]}{\left[\frac{6\beta(\theta+1)+3\theta\beta+2\theta}{2+\theta} \right]} \right)$$

8. Entropies

In this section, we derived the Renyi entropy, and Tsallis entropy from the MGZ distribution. Entropy and information are widely recognized as measures of uncertainty or randomness within a probability distribution. The entropy functionals of probability distributions are rooted in a variational definition of uncertainty measures, providing a robust framework for analyzing probabilistic systems.

8.1. Renyi Entropy

Entropy is defined as a random variable X is a measure of the variation of the uncertainty. It is used in many fields, such as engineering, statistical mechanics, finance, information theory, biomedicine, and economics. The Renyi entropy of order λ is defined as

$$R_\lambda = \frac{1}{1-\lambda} \log \int_0^\infty [f(x;\theta,\beta)]^\lambda\, dx, \quad \lambda > 0, \lambda \neq 1.$$

$$R_\lambda = \frac{1}{1-\lambda} \log \int_0^\infty \left[\frac{\theta^3}{\beta+1} \left(\frac{\beta}{2}x^2 + \frac{x}{2+\theta}(1+x) \right) e^{-\theta x} \right]^\lambda\, dx.$$

using the gamma function $\frac{\Gamma(t)}{a^t} = \int_0^\infty x^{t-1}e^{-ax}dx$, and the binomial theroem $(a+b)^z = \sum_{j=0}^\infty \binom{z}{j} a^j b^{z-j}$, and simplifying the integral, we obtain

$$R_\lambda = \frac{1}{1-\lambda} \log \left(\frac{\theta^3}{\beta+1} \right)^\lambda \sum_{j=0}^\infty \sum_{k=0}^\infty \binom{\lambda}{j} \binom{\lambda-j}{k} \left(\frac{\beta}{2} \right)^j \left(\frac{1}{2+\theta} \right)^{\lambda-j} \frac{\Gamma(2\lambda-k+1)}{(\lambda\theta)^{2\lambda-k+1}}. \tag{8}$$

8.2. Tsallis Entropy

The Boltzmann-Gibbs (B-G) statistical properties introduced by Tsallis have received significant attention. This generalization of B-G statistics was first proposed by Tsallis (1988) through the introduction of Tsallis entropy. For continuous random variables, it is defined as

$$T_\lambda = \frac{1}{\lambda - 1}\left[1 - \int_0^\infty [f(x;\theta,\beta)]^\lambda \, dx\right], \quad \lambda > 0, \lambda \neq 1.$$

$$T_\lambda = \frac{1}{\lambda - 1}\left[1 - \int_0^\infty \left(\frac{\theta^3}{\beta+1}\right)^\lambda \left[\left(\frac{\beta}{2}x^2 + \frac{x}{2+\theta}(1+x)\right)e^{-\theta x}\right]^\lambda \, dx\right].$$

simplifying the integral using the gamma function $\frac{\Gamma(t)}{a^t} = \int_0^\infty x^{t-1}e^{-ax}dx$, and the Taylor series $e^x = \sum_{j=0}^\infty \frac{t^j}{j!} \; \forall \; \infty < x < \infty$, we obtain

$$T_\lambda = \frac{1}{\lambda - 1}\left[1 - \left(\frac{\theta^3}{\beta+1}\right)^\lambda \sum_{j=0}^\infty \sum_{k=0}^\infty \binom{\lambda}{j}\binom{\lambda - j}{k}\left(\frac{\beta}{2}\right)^j \left(\frac{1}{2+\theta}\right)^{\lambda - j} \frac{\Gamma(2\lambda - k + 1)}{(\lambda\theta)^{2\lambda - k + 1}}\right] \tag{9}$$

9. Estimation of Parameter

The MGZ distribution parameter's maximum likelihood estimates and Fisher's information matrix are provided in this section.

9.1. Maximum likelihood Estimation (MLE) and Fisher's Information Matrix

Consider $x_1, x_2, x_3, , x_n$ be a random sample of size n from the MGZ distribution with parameter θ, β the likelihood function, which is defined as

$$L(x;\theta,\beta) = \prod_{i=1}^n f(x_i;\theta,\beta)$$

$$L = \prod_{i=1}^n \frac{\theta^3}{\beta+1}\left(\frac{\beta}{2}x_i^2 + \frac{x_i}{2+\theta}(1+x_i)\right)e^{-\theta x_i}$$

then, the log-likelihood function is

$$\log L = 3n\log\theta - n\log(\beta+1) + \sum_{i=1}^n \log\left(\frac{\beta}{2}x_i^2 + \frac{x_i}{2+\theta}(1+x_i)\right) - \theta\sum_{i=1}^n x_i$$

differentiating with respect to θ and β, we obtain

$$\frac{\partial \log L}{\partial\theta} = \left(\frac{3n}{\theta}\right) + \sum_{i=1}^n \left(\frac{\left[\frac{x_i + x_i^2}{(2+\theta)^2}\right]}{\left(\frac{\beta}{2}x_i^2 + \frac{x_i}{2+\theta}(1+x_i)\right)}\right) - \sum_{i=1}^n x_i = 0 \tag{10}$$

$$\frac{\partial \log L}{\partial\beta} = -\left(\frac{n}{\beta+1}\right) + \sum_{i=1}^n \left(\frac{\left(\frac{x_i^2}{2}\right)}{\left(\frac{\beta}{2}x_i^2 + \frac{x_i}{2+\theta}(1+x_i)\right)}\right) = 0 \tag{11}$$

The maximum likelihood estimate of the parameters for the MGZ distribution is provided by equations (10) and (11). The equation, however, cannot be solved analytically, so we used R programming and a data set to solve it numerically. The asymptotic normality results are used

to derive the confidence interval. Given that if $\lambda = (\theta, \beta)$ represents the MLE of $\lambda = (\theta, \beta)$, the results can be expressed as follows

$$\sqrt{n}(\hat{\lambda} - \lambda) \to N_2(0, I^{-1}(\lambda))$$

in this case, $I(\lambda)$ represents Fisher's Information Matrix,

$$I(\lambda) = -\frac{1}{n}\begin{bmatrix} E\left[\frac{\partial^2 \log L}{\partial \theta^2}\right] & E\left[\frac{\partial^2 \log L}{\partial \theta \partial \beta}\right] \\ E\left[\frac{\partial^2 \log L}{\partial \beta \partial \theta}\right] & E\left[\frac{\partial^2 \log L}{\partial \beta^2}\right] \end{bmatrix}$$

The second derivatives of the log-likelihood function are

$$\frac{\partial^2 \log L}{\partial \theta^2} = -\left(\frac{3n}{\theta^2}\right) + \sum_{i=1}^{n}\left(\frac{\left(\frac{\beta}{2}x_i^2 + \frac{x_i}{2+\theta}(1+x_i)\right)\left(\frac{2\theta(x_i+x_i^2)+4(x_i+x_i^2)}{(2+\theta)^4}\right) - \left(\frac{x_i+x_i^2}{(2+\theta)^2}\right)^2}{\left(\frac{\beta}{2}x_i^2 + \frac{x_i}{2+\theta}(1+x_i)\right)^2}\right)$$

$$\frac{\partial^2 \log L}{\partial \beta^2} = \left(\frac{n}{(\beta+1)^2}\right) - \sum_{i=1}^{n}\left(\frac{\left(\frac{x_i^2}{2}\right)^2}{2\left(\frac{\beta}{2}x_i^2 + \frac{x_i}{2+\theta}(1+x_i)\right)^2}\right)$$

$$\frac{\partial^2 \log L}{\partial \theta \partial \beta} = \sum_{i=1}^{n}\left(\frac{x_i^3(1+x_i)}{2\left(\theta^2+4(\theta+1)\right)}\right)$$

10. Applications

In this section, the proposed MGZ distribution is applied to the real data sets.

Dataset 1: Survival time of breast cancer patients (in months) of patients. The secondary data has been taken from a private tertiary care hospital in Puducherry. 63, 61, 58, 57, 55, 26, 5, 20, 54, 58, 30, 24, 35, 27, 47, 67, 70, 78, 39, 26, 65, 77, 49, 50, 27, 41, 65, 52, 48, 62, 38, 26.

Dataset 2: We have used the data representing the size of the Tumour (in mm) of 210 patients with node-positive breast cancer, as reported by Royston and Altman (2013). The entire data set contains patient records from a 1984-1989 trial conducted by the German Breast Cancer Study Group (GBSG) of 720 patients with node-positive breast cancer; it retains the 686 patients with complete data for the predictive variables. We used the initial 210 sample values for the analysis. 18, 20, 40, 25, 30, 52, 21, 20, 20, 30, 12, 30, 21, 25, 57, 25, 21, 27, 12, 45, 18, 18, 30, 30, 25, 14, 21, 30, 23, 30, 55, 11, 40, 25, 15, 29, 15, 45, 40, 30, 60, 21, 20, 58, 30, 28, 25, 18, 40, 35, 30, 21, 45, 40, 20, 19, 52, 80, 40, 20, 80, 30, 33,16, 25, 39, 52, 50, 39, 42, 65, 28, 25, 23, 32, 15, 40, 35, 25, 30, 22, 45, 20, 27, 29, 22, 24, 25, 23, 30, 50, 45, 7, 8, 18, 16, 28, 21, 15, 20, 12, 20, 30, 50, 25, 30, 25, 32, 35, 20, 28, 26, 35, 23, 80, 20, 20, 34, 30, 23, 35,70, 20, 45, 24, 28, 35, 25, 60, 35, 40, 23, 40, 21, 35, 55, 120, 23, 50, 55, 16, 22, 21, 12, 14, 60, 55, 120, 23, 50, 55, 16, 22, 21, 12, 14, 60, 30, 25, 40, 48, 35, 25, 28, 35, 30, 35, 22, 12, 21, 35, 52, 36, 25, 31, 20, 30, 25, 44, 21, 30, 32, 50, 21, 20, 60, 35, 30, 20, 23, 49, 55, 20, 25, 22, 18, 30, 20, 40, 35, 13, 70, 27, 21, 25, 80, 36, 32, 25, 19.

Dataset 3: Age of the patients with advanced lung cancer in years (J. F. Lawless (2003)). 69, 64, 38, 63, 65, 49, 69, 68, 43, 70, 81, 63, 63, 52, 48, 61, 42, 35, 63, 56, 55, 67, 63, 65, 46, 53, 69, 68, 43, 55, 42, 64, 65, 65, 55, 66, 60, 67, 53, 62, 67, 72, 48, 68, 67, 61, 60, 62, 38, 50, 63, 64, 43, 34, 66, 62, 52, 47, 63, 68, 45, 41, 66, 62, 60, 66, 38, 53, 37, 54, 60, 48, 52, 70, 50, 62, 65, 58, 62, 64, 63, 58, 64, 52, 35, 63, 70, 51, 40, 69, 36, 71, 62, 60, 44, 54, 66, 49, 72, 68, 62, 71, 70, 61, 71, 59, 67, 60, 69, 57, 39, 62, 50, 43, 70, 66, 61, 81, 58, 63, 60, 62, 42, 69, 63, 45, 68, 39, 66, 63, 49, 64, 65, 64, 67, 65, 37.

We observe its flexibility over some well-known existing distributions. The results for the analysis in this present study are obtained using R software. We have also calculated the Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), Akaike Information Criteria Corrected (AICC), and -2logL for the considered distributions to observe their fit. The distribution with the value of lowest AIC and BIC and AICC is considered as the best. The Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), Akaike Information Criteria Corrected (AICC), and -2logL are used to compare the goodness of fit of the fitted distribution. The following formula can be used to determine AIC, BIC, AICC, and -2logL.

$$\text{AIC} = 2k - 2\log L, \quad \text{BIC} = k\log n - 2\log L, \quad \text{and AICC} = \text{AIC} + \frac{2k(k+1)}{(n-k-1)}$$

Where, k= number of parameters, n sample size, and -2logL is the maximized value of the log-likelihood function.

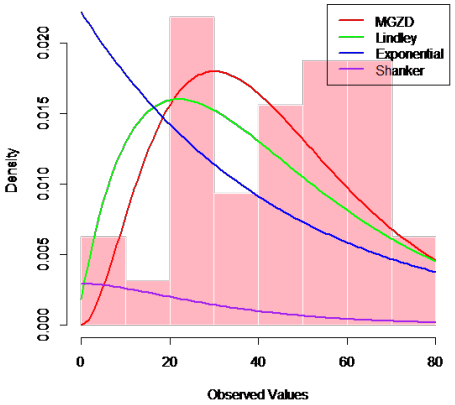


Figure 5: *Fitting density curve for dataset 1*

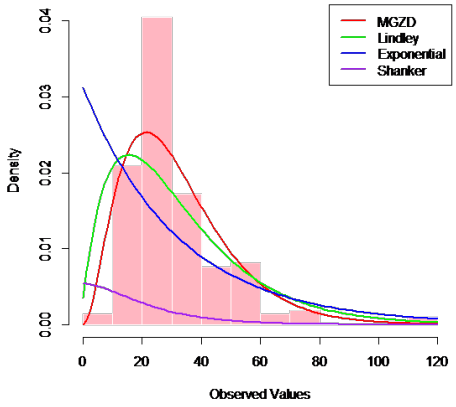


Figure 6: *Fitting density curve for dataset 2*

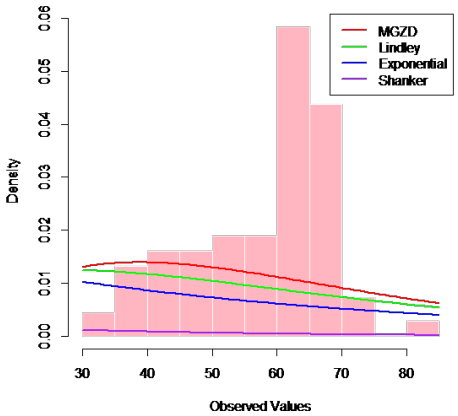


Figure 7: *Fitting density curve for dataset 3*

Table 1: *The MLEs of MGZD parameters and AIC, BIC, AICC, and -2logL for Data Set 1.*

Distributions	ML Estimates	-2logL	AIC	BIC	AICC
Mixture of Gamma and Zeghdoudi	$\hat{\theta} = 0.05542\,(0.0067)$	203.2229	207.2229	209.4939	207.8229
	$\hat{\beta} = 0.00403\,(0.00017)$				
Exponential	$\hat{\theta} = 0.0185\,(0.0038)$	229.6043	231.6043	232.3970	231.8043
Lindley	$\hat{\theta} = 0.0363\,(0.00535)$	213.2929	215.2929	216.4284	215.4929
Shanker	$\hat{\theta} = 0.0369\,(0.00543)$	212.536	214.536	215.6727	214.736

Table 2: *The MLEs of MGZD parameters and AIC, BIC, AICC, and -2logL for Data Set 2.*

Distributions	ML Estimates	-2logL	AIC	BIC	AICC
Mixture of Gamma and Zeghdoudi	$\hat{\theta} = 0.09355\,(0.00372)$	1717.616	1721.616	1728.310	1721.674
	$\hat{\beta} = 0.00382\,(0.00593)$				
Exponential	$\hat{\theta} = 0.03119\,(0.00215)$	1876.483	1878.483	1881.830	1878.502
Lindley	$\hat{\theta} = 0.06059\,(0.00295)$	1770.519	1772.519	1775.866	1772.539
Shanker	$\hat{\theta} = 0.07918\,(0.00348)$	2753.765	2755.765	2759.112	2755.784

Table 3: *The MLEs of MGZD parameters and AIC, BIC, AICC, and -2logL for Data Set 3.*

Distributions	ML Estimates	-2logL	AIC	BIC	AICC
Mixture of Gamma and Zeghdoudi	$\hat{\theta} = 0.05143\,(0.00253)$	1233.039	1237.039	1242.878	1237.128
	$\hat{\beta} = 14.3464\,(41.4323)$				
Exponential	$\hat{\theta} = 0.01716\,(0.00146)$	1388.006	1390.006	1392.926	1390.036
Lindley	$\hat{\theta} = 0.03374\,(0.00203)$	1291.502	1293.502	1296.422	1293.532
Shanker	$\hat{\theta} = 0.04516\,(0.00243)$	2092.173	2094.173	2097.093	2094.203

From table 1, 2 and table 3, the (MGZD) mixture gamma and zeghdoudi distributions has the lowest AIC, BIC and AICC values, thus making it to be fitted better than the (Exponential, Lindley and Shanker) distributions. Figure 5, 6 and 7 also explicates the better model.

11. Conclusion

This study introduces a novel two-parameter distribution known as the Modified Gamma-Zeghdoudi Distribution (MGZD). This distribution is a compound mixture of the Zeghdoudi and Gamma distributions, two well-known statistical models. Essential features of the MGZD, such as survival and hazard functions, have been analyzed. We have explored some of the important statistical properties, such as entropies, Bonferroni and Lorenz curves, moments (mean, variance, skewness, and kurtosis), order statistics, and stochastic ordering. We have estimated the parameters of the MGZD are estimated using the maximum likelihood estimation method.

To assess the effectiveness, the MGZD is applied to three real-world data sets and compared with well-known distributions. The results illustrate that the MGZD gives a better fit to the data, proving its superiority over other known classical models in enclosing the above data structure.

References

[1] Alam, F., Alrajhi, S., Nassar, M., and Afify, A. Z. (2021). Modeling liver cancer and leukemia data using arcsine-Gaussian distribution. *Computers, Materials and Continua (CMC)*, 67(2), 2185–2202.

[2] Al-Kadim, K. A., and Mahdi, A. A. (2018). Exponentiated transmuted exponential distribution. *Journal of University of Babylon for Pure and Applied Sciences*, 26(2), 78–90.

[3] American Cancer Society. (2021). Breast Cancer Survival Rates. Available at: `https://www.cancer.org/cancer/breast-cancer/understanding-a-breast-cancer-diagnosis/breast-cancer-survival-rates.html`

[4] David, H. A. (1970). *Order Statistics*. Wiley and Sons, New York.

[5] Ghitany, M. E., Atieh, B., and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78(4), 493–506.

[6] Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data* (2nd ed.). Hoboken, NJ: Wiley-Interscience.

[7] Messaadia, H., and Zeghdoudi, H. (2018). Zeghdoudi distribution and its applications. *International Journal of Computing Science and Mathematics*, 9(1), 58–65.

[8] Macdonald, P. (1971). An estimation procedure for mixtures of distributions. *Journal of the Royal Statistical Society*, 20, 102–107.

[9] R©nyi, A. (1961). On measures of entropy and information. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, Vol. 4, 547–562. University of California Press.

[10] R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[11] Royston, P., and Altman, D. G. (2013). External validation of a Cox prognostic model: Principles and methods. *BMC Medical Research Methodology*, 13(1), 33. `https://doi.org/10.1186/1471-2288-13-33`

[12] Sakthivel, M.,and Pandiyan, P. (2024). Stochastic Modelling on New Mixture Distribution with Properties and their Application. *International Journal of Mathematics, Statistics, and Computer Science*, 2, 259–275. `https://doi.org/10.59543/ijmscs.v2i.9015`

[13] Shanker, R. (2016). Aradhana distribution and its applications. *International Journal of Statistics and Applications*, 6(1), 23–34.

[14] Shanker, R. (2022). Uma distribution properties and application. *Biometrics Biostatistics International Journal*, 11(5), 165–169.

[15] Shanker R. (2015). Shanker distribution and its applications. *International Journal of Statistics and Applications*, a; 5(6):338–348, doi: org/10.5923/j.statistics.20150506.08.

[16] Shanker R, Hagos F, Sujatha S. (2015).On modeling of lifetimes data using exponential and Lindley distributions. *Biom Biostat Int J*; 2(5):140–147, doi.org/10.15406/bbij.2015.02.00042.

[17] Shanker R. (2015). Akash distribution and its applications. *International Journal of Probability and Statistics*. b;4(3):65–75, doi: 10.5923/j.statistics.20150506.08.

[18] Shanker R. (2016) Garima distribution and its application to model behavioral science data. *Biometrics and Biostatistics International Journal*, 4(7):1–9, doi.org/10.15406/bbij.2016.04.00116

[19] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* 2021 May;71(3):209-249. doi: 10.3322/caac.21660. Epub 2021 Feb 4. PMID: 33538338.

[20] World Health Organization. (2011). *Global Status Report on Noncommunicable Diseases 2010*. Available at: `https://www.who.int/nmh/publications/ncd_report_full_en.pdf`

HIGH-TEMPERATURE HEAT-INSULATING MATERIALS: A COMPROMISE BETWEEN THERMAL CONDUCTIVITY AND RELIABILITY

N.M. Piriyeva[1], R.K. Karimova[2], G.K.Abdullayeva[3]

•

Azerbaijan State Oil and Industry University, Baku, Azerbaijan
[1]naciba.piriyeva@asoiu.edu.az, [2] rashida.karimova@asoiu.edu.az,
[3]gulshan.abdullayeva@asoiu.edu.az

Abstract

This article examines heat transfer mechanisms in high-temperature thermal insulation materials, addressing key operational challenges such as thermal degradation, mechanical stability, and chemical resistance. The novelty of the study lies in an integrated approach that includes optimizing the micro- and nanostructure to minimize radiative and conductive heat transfer, developing advanced multilayer and composite systems, applying predictive computational modeling for property assessment, and exploring self-healing coatings and intelligent materials. These innovations ensure reduced thermal conductivity without compromising mechanical strength, enhance resistance to thermal cycling and aggressive chemical exposure, and significantly extend the service life of insulation systems, making them highly promising for aerospace, energy, metallurgy, and other industries requiring reliable thermal protection. Additionally, the proposed solutions contribute to increased energy efficiency, reduced material consumption, and improved environmental sustainability, paving the way for next-generation thermal insulation technologies with broader industrial applicability.

Keywords: High-temperature insulation materials, thermal conductivity, mechanical strength, reliability.

I. Introduction

Modern industrial and energy systems operating at extremely high temperatures require reliable thermal insulation materials capable of simultaneously reducing heat loss and maintaining mechanical strength. In industries such as aviation and space technology, metallurgy and energy, the use of traditional thermal insulation materials faces problems with degradation, shrinkage and loss of structural integrity during long-term operation.

The main task of developers is to find the optimal balance between low thermal conductivity, providing effective thermal insulation, and high mechanical resistance, preventing material destruction under thermal loads. Modern solutions include the use of porous ceramic materials, aerogels, multilayer structures and nanocomposites, which demonstrate improved characteristics compared to traditional insulators.

The purpose of this article is to consider the key heat transfer mechanisms in high-temperature thermal insulation materials, identify the main problems associated with their operation and offer promising solutions to improve their reliability.

II. Problem setting

High-temperature thermal insulation materials play a key role in metallurgy, energy, aviation, space and chemical industries, protecting structures from overheating and reducing heat loss. However, their development faces serious scientific and technical challenges associated with the need to achieve a compromise between thermal conductivity, mechanical strength and resistance to aggressive operating conditions.

Table 1.: *Main difficulties in the development and application of high-temperature thermal insulation materials*

Difficulty	Problem Description	Consequences
Compromise between thermal conductivity and mechanical strength	Materials with low thermal conductivity have high porosity, which reduces their mechanical strength.	Brittleness, tendency to fracture under load, reduced operational lifespan.
Degradation of structure at high temperatures	Oxidation, sintering, thermochemical erosion, and structural breakdown reduce insulation properties.	Loss of thermal insulation efficiency, frequent material replacement.
Thermal cycling loads	Frequent temperature fluctuations cause thermal expansion and formation of microcracks.	Reduced strength, accelerated deterioration, and shortened lifespan.
Need to reduce density while maintaining strength	In the aerospace industry, it's crucial to minimize the structure's weight without compromising reliability.	Limited material options, difficulty in achieving the required balance of properties.
Limited temperature ranges of traditional materials	Most commercial thermal insulators lose their properties at temperatures above 1500°C.	Limited applicability, risk of overheating and structural failure in extreme conditions.

One of the main factors affecting the reliability of insulating materials is their structural stability under high temperatures [1-3]. Traditional thermal insulators, such as ceramic fibers, foam glass, and refractory bricks, are susceptible to thermal aging, sintering, and oxidation, which leads to a deterioration in their insulation properties and mechanical strength. For example, porous ceramics, which are effective at reducing thermal conductivity, lose their mechanical load resistance over time due to grain growth and structural degradation.

To illustrate the key issues with thermal insulation materials, Table 2 is presented, showing the main operational limitations of various types of insulators.

Another serious issue is the impact of thermal cycling loads. Materials subjected to rapid temperature fluctuations experience repeated expansion and contraction, leading to the formation of microcracks and gradual destruction. This is particularly critical in aerospace and aviation technology, where temperatures can change from extreme cold to very high values in a short time.

Additional challenges arise when there is a need to reduce material density without losing strength, especially in the aerospace industry, where the minimum weight of a structure is a key requirement [4]. Ultra-lightweight thermal insulation materials, such as aerogels, have unique properties, but their mechanical fragility limits their field of application.

Thermal conductivity of thermal insulation materials is determined by various heat transfer mechanisms, including heat conduction through the solid skeleton, convection in porous structures, and radiative heat exchange [5-7]. Under conditions of extremely high temperatures, the contribution of radiative heat transfer increases, requiring the development of special methods for its suppression. Optimizing the material structure to reduce thermal conductivity without compromising mechanical properties remains one of the key research challenges.

Table 2: *Problems and limitations of various thermal insulation materials*

Material Type	Main Advantages	Main Problems and Limitations
Ceramic Fibers	Lightness, high temperature resistance	Brittleness, tendency to degrade, limited mechanical strength
Aerogels	Low thermal conductivity, high thermal resistance	High cost, mechanical fragility, production complexity
Foam Glass	Resistance to chemical environments, good mechanical strength	High density, limited flexibility, high production cost
Refractory Bricks	Durability, resistance to thermal cycling	High thermal conductivity, significant weight, structural sintering
Multi-layer Coatings	High adaptability, property regulation capability	Application difficulty, potential delamination under thermal cycling

In addition, in a number of industries, the chemical resistance of insulation materials is an important factor. In the operation of gas turbine engines, reactor installations, and high-temperature furnaces, materials are exposed to aggressive components such as oxygen, water vapor, and sulfur compounds [8-10].

Oxidation and chemical erosion can significantly reduce the service life of thermal insulation, requiring the creation of new materials with high chemical inertness. Thus, solving the problem of the reliability of thermal insulation materials requires a comprehensive approach, including optimizing their microstructure, using nanotechnology, developing multilayer structures, and applying modern computer modeling methods to predict operational characteristics.

III. Examples of practical application

The development of high-temperature thermal insulation materials is focused on achieving an optimal balance between low thermal conductivity, mechanical strength, and thermal stability. Modern technologies allow for the creation of materials with enhanced properties by utilizing nanostructures, multilayer systems, optimizing porosity, and phase transition processes (Table 3). Computer modeling also plays a key role in accelerating the development and predicting the performance characteristics of new thermal insulators.

To demonstrate the impact of various structural modifications on thermal conductivity, the graph below (Figure 1) shows the relationship between thermal conductivity and density for different classes of thermal insulation materials based on Table 4.

Based on Figure 1, it can be concluded that the lower the material density, the lower its thermal conductivity [11]. Aerogels and porous ceramics demonstrate the best performance; however, as the density decreases, mechanical strength diminishes, which requires additional solutions.

Approaches to Improving the Reliability of Thermal Insulation Materials

Modern technologies in thermal insulation not only focus on reducing thermal conductivity but also on enhancing the operational reliability of materials [13-17]. This is especially important in conditions of high temperatures, mechanical stresses and aggressive environments, such as in the aerospace industry, energy sector, and metallurgy.

Developers use combined methods that integrate various technologies to enhance thermal stability, mechanical strength, and durability (Table 5). For example, integrating aerogels with reinforcing fibers results in materials with low density but high strength, which is critical for aviation and space applications.

Table 3. *Modern methods for improving thermal insulation materials*

Method	Principle of Action	Achievable Results
Nanomaterials and aerogels	Reducing the size of structural elements to the nanoscale to suppress heat transfer.	Ultra-low thermal conductivity (~0.02 W/m·K), high thermal stability, reduced material density.
Multilayer systems	Combination of materials with different thermal conductivity and mechanical strength.	Improved balance between strength and insulation, high resistance to thermal cycling.
Porous ceramics and glass ceramics	Optimization of porosity to reduce thermal conductivity and increase thermal stability.	Reduced heat loss, resistance to oxidation and high temperatures.
Phase change materials (PCM)	Use of phase transition energy for temperature regulation.	Stabilization of temperature fluctuations, protection from overheating.
Computer modeling	Numerical methods (finite element method, molecular dynamics) for optimizing material structure and composition.	Reduced development costs, accurate prediction of properties.

Table 4. *Dependence of thermal conductivity on density for various insulation materials*

Material	Density (kg/m³)	Thermal Conductivity (W/m·K)
Silicon aerogel	150	0.02
Porous ceramics	500	0.1
Traditional refractories	1800	1.2
Metal ceramics	3000	3.5

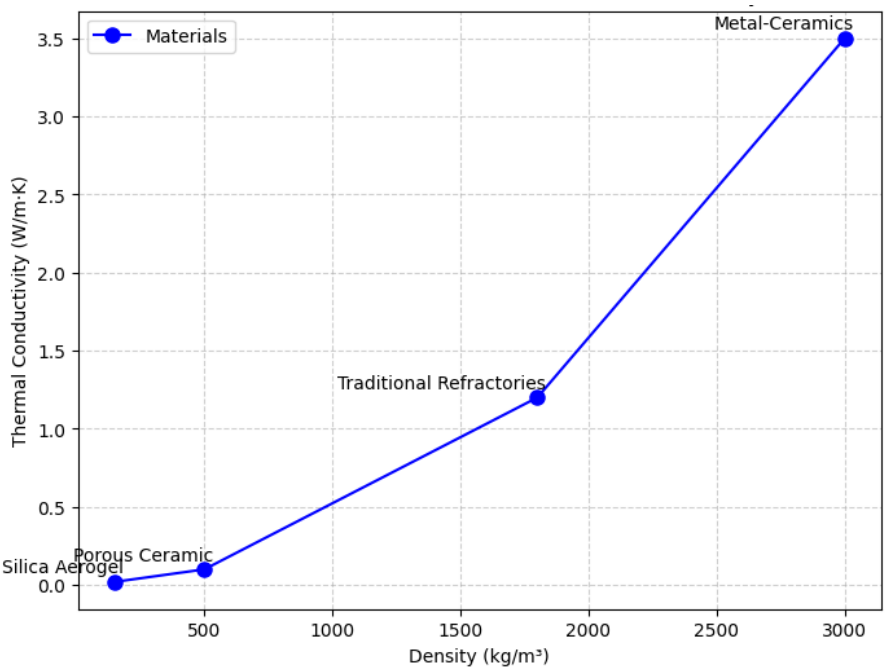


Figure 1.: *Effect of material structure on thermal conductivity*

Table 5. *Combined Methods for Improving the Reliability of Thermal Insulation Materials*

Combined Method	Application	Advantages
Aerogels + Reinforcing Fibers	Aviation, Space, Metallurgy	Resistance to mechanical stresses, reduced weight
Porous Ceramics + Protective Coatings	Gas Turbines, Industrial Furnaces	Enhanced resistance to oxidation and thermal cycles
Phase-Change Materials + Multi-Layered Structures	Energy Sector, Thermal Protection Systems	Effective temperature regulation, reduced heat loss

An example of combined technology application is the use of porous ceramics with protective coatings in gas turbines. In gas turbines, operating temperatures can exceed 1400°C, which requires the use of heat-resistant materials. Ordinary metal components cannot withstand such conditions without additional protection. One solution is the use of porous ceramics (based on ZrO_2 — zirconium dioxide) with a protective coating of Al_2O_3 (aluminum oxide). This results in reduced heat transfer to the turbine's metal parts, increasing the lifespan of the components.

To illustrate the effect of multi-layer thermal insulation coatings, let's consider a graph showing the temperature change across the thickness of the material (Figure 2).

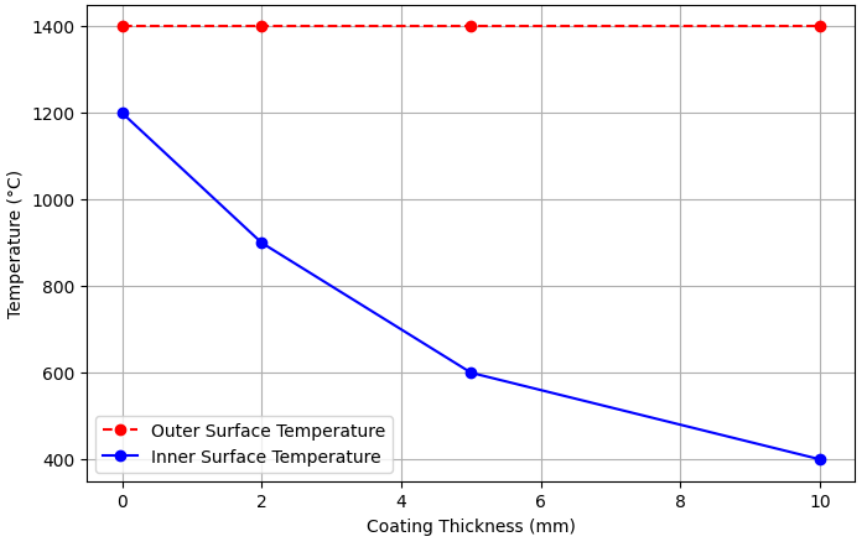


Figure 2.: *Influence of multilayer structures on the temperature distribution.*

From the graph shown in Figure 2, it can be concluded that adding a multilayer thermal insulation coating reduces the temperature on the inner surface by more than three times, significantly decreasing the thermal impact on the structure.

The future of thermal insulation materials lies in the development of innovative technologies aimed at improving their reliability and efficiency. One promising direction is the creation of self-healing coatings, which, due to oxide layers, can regenerate during oxidation, extending the insulation's lifespan. Research is also being conducted in the field of smart materials, which have the ability to change their properties depending on temperature, allowing them to adapt to various operating conditions.

Additionally, the active development of 3D printing for ceramic materials is enabling the design of complex thermal insulation structures with optimized geometries, improving their operational characteristics. These technologies open new horizons in the field of thermal protection, providing not only a reduction in heat losses but also increased resistance to thermal, mechanical, and chemical impacts. Future research will focus on creating adaptive and self-regulating materials that can function in extreme conditions without losing effectiveness.

References

[1] Ahmedov, E. N., Safiyev, E. S., & Rzayeva, S. V. (2025). Innovative methods of electrodischarge purification of chemical liquids for the synthesis of polymeric materials with predefined properties. *Alternative Energy and Ecology (ISJAEE)*, (1), 204–217. (In Russ.) https://doi.org/10.15518/isjaee.2025.01.204-217.

[2] Safiyev, E., Rzayeva, S., & Karimova, R. (2024). The importance of diagnostics of electrical equipment at thermal power plants for ensuring the reliability of power systems. *Przegląd Elektrotechniczny, 2024*(9), 127-131

[3] Ahmedov, E., Rzayeva, S., & Mammadov, N. (2023). The mechanism of electric discharge effect on the modification process of linear low-density polyethylene. *Przegląd Elektrotechniczny, 99*(6).

[4] Karimova, R. K., & Rzayeva, S. V. (2023). Comparison of thermal conductivity of aqueous and formamide solutions ($BeCl_2$) at high temperatures. *International Journal of Engineering Science*, *15*(3), 156–159.

[5] Piriyeva, N. M. (2022). Design of electric devices with induction levitation elements. *International Journal on Technical and Physical Problems of Engineering (IJTPE)*, *14*(1), 124–129.

[6] Safiyev, E. S., & Piriyeva, N. M. (2022). On the issue of assessing the temperature index and the range of heat resistance of polymeric electrical insulating materials. *News of Azerbaijan Higher Technical Schools, (1)*, 49–51.

[7] Safiyev, E. S., Karimova, S. M., & Ganiyeva, N. A. (2023). Investigation of high-temperature superconductors. *Vestnik Nauki, 2*(3), 269–278.

[8] Ahmedov, E., Safiyev, E., Rzayeva, S., Mammadov, N., Ganiyeva, N., & Mukhtarova, K. (2023). Obtaining a graft copolymer of polyethylene by electrodischarge synthesis. *Przegląd Elektrotechniczny, 2023*(11)., 100-103

[9] Piriyeva, N. M., & Kerimzade, G. S. (2024). Systematization of levitation equations for electrical devices with levitation elements. *Przegląd Elektrotechniczny, 100*(8), 175–177.

[10] Rzayeva, S. V., Piriyeva, N. M., & Guseynova, I. A. (2024). Analysis of reliability of typical power supply circuits. *Reliability: Theory and Applications (RTA), 19*(3), 173–178.

[11] Rzayeva, S. V., Ganiyeva, N. A., & Piriyeva, N. M. (2024). Modern approaches to electrical equipment diagnostics. *International Journal on Technical and Physical Problems of Engineering (IJTPE), 16*(1), 182–189.

[12] Piriyeva, N. M., Abdullayeva, G. K., & Bakhtiyarov, A. L. (2025). Engineering approaches to minimizing the environmental impact of thermal power plants. *International Journal on Technical and Physical Problems of Engineering (IJTPE), 16*(4), 231–243.

[13] Mamedova, G. V., Kerimzade, G. S., & Piriyeva, N. M. (2022). Electromagnetic calculation of tension devices for winding wires of small cross-sections. *International Journal on Technical and Physical Problems of Engineering (IJTPE), 14*(4), 80–85.

[14] Piriyeva, N. M., & Kerimzade, G. S. (2024). Methods for increasing electromagnetic efficiency in induction levitator. *Przegląd Elektrotechniczny, 100*(10), 192–196.

[15] Piriyeva, N. M., Ganiyeva, N. A., & Karimova, R. K. (2025). Innovative approaches to selective automatic reclosing on overhead cable lines. *International Journal on Technical and Physical Problems of Engineering (IJTPE), 17*(1), 32–44.

[16] Rzayeva, S. V., Piriyeva, N. M., & Ismayilova, S. I. (2025). High and low voltage coordination in electrical power systems. *International Journal on Technical and Physical Problems of Engineering (IJTPE), 17*(1), 19–31.

[17] Ahmedov, E. N., Rzayeva, S. V., Mukhtarova, K. M., & Aliyev, N. A. (2025). Innovative enhancements in ground wire configurations for superior lightning protection of overhead power lines. *International Journal on Technical and Physical Problems of Engineering (IJTPE), 17*(1), 45–58.

EVALUATING MULTI SERVER QUEUING SYSTEM EFFICIENCY: A COMPARATIVE STUDY BETWEEN FUZZY QUEUING MODEL AND INTUITIONISTIC FUZZY QUEUING MODEL WITH INFINITE CAPACITY CONSTRAINTS

S. Aarthi
•
Department of Mathematics & Statistics, Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, Chennai 603203, Tamil Nadu, India.
aarthis4@srmist.edu.in

Abstract

In this study, we suggest an innovative analytical approach employing triangular fuzzy and triangular intuitionistic fuzzy numbers to ascertain the membership functions governing significant service-execution proportions within multiserver queuing models. Departing from conventional methodologies, both the inter-entry and completion rates are characterized by fuzzy natures, introducing a novel dimension to the analysis. Through numerical validation, our model demonstrates viability in the multiserver queuing context, bolstering its credibility. Furthermore, we conduct a contextual investigation, juxtaposing individual fuzzy metrics, thereby revealing the superior categorization potential of intuitionistic fuzzy queuing models over their fuzzy counterparts. Expanding the horizons of traditional fuzzy queuing theory, our integration of intuitionistic fuzzy environments promises enhanced queuing model implementation efficacy. Our primary objective lies in evaluating the productivity of a multi-server queuing model with limitless capacity, leveraging both fuzzy queuing and intuitionistic fuzzy queuing theories. Embracing a framework where arrival and service rates are characterized by triangular and intuitionistic triangular fuzzy numbers, we undertake a thorough assessment to establish evaluation criteria, adhering to a design protocol that preserves fuzzy values without converting them into crisp values. Additionally, we address two statistical challenges to ascertain the method's validity.

Keywords: queuing model, fuzzy numbers, multi-server, execution proportions

I. Introduction

Queuing models have found extensive application in various domains, aiding in the formulation of routine procedures, workplace configurations, and labor regulations. However, real-world scenarios often present challenges where certain variables within these queuing models may be inaccurately known due to various factors. As a result, there arises a need to integrate fuzzy logic into queuing systems, enabling adaptability to diverse real-life situations. Nevertheless, in numerous practical instances, precise data is crucial for effective decision-making by system designers and researchers.

To address the inherent vagueness in such situations, many scholars have proposed methodologies. One such approach involves constructing a membership function for queue performance measures, thereby providing a means to navigate ambiguity. Additionally, the notion of intuitionistic fuzzy sets [2] has emerged as a viable method for defining fuzzy sets, particularly when data accessibility is limited, making it challenging to determine ambiguous conceptions accurately.

Acknowledging the quandary of operationalizing fuzzy set hypotheses, given that fuzzy numbers deviate from a traditional linear approach like real numbers, the paper propounds a methodology for resolving a multiple transmission queuing model in both fuzzy and intuitionistic fuzzy environments without undermining their essence. In comparison to antecedent methodologies, this approach is advantageous due to its succinctness, adaptability, and relevance. The analysis outcomes suggest that the performance metrics of the fuzzy queuing model align with the estimated spectrum of the intuitionistic fuzzy queuing model.

Building upon foundational research by Lie and Lee [10] and Buckley [4], subsequent progressions by Negi and Lee [12] and Chen [6] expanded upon these models, introducing fuzzified exponential time based on queuing theory. Srinivasan [17] introduced a fuzzy queuing model using the DSW algorithm, while Thamotharan [18] investigated a multi-server fuzzy queuing model using triangular and trapezoidal fuzzy numbers with α cuts. Shanmugasundaram et al. [15] developed a predictive fuzzy multi-server queuing model utilizing the DSW algorithm. Madhuri et al. (2017) conducted an analysis of fuzzy queues employing Zadeh's [19] augmentation rule, and Selvam et al. [14] proposed a ranking method for pentagonal fuzzy numbers applied to the incentre of centroids. Samouylov et al. [13] assessed a multi-server queuing system catering to two correlated flows of requests. Bin Sun et al. [3] investigated a multi-server queuing system with two input flows. Dudin et al. [8] interpreted a multi-server vacation queuing model. Kuaban et al. [9] scrutinized a multi-server queuing model with balking and correlated reneging, with implications for healthcare management. Anupama Kumar [1] examined a $M/M/3$ queuing system with a queue-dependent multi-server. Diletta Olliaro et al. [7] studied a multiserver queue where jobs request a variable number of servers for random service time. Selvam et al. [14] also explored a ranking technique for handling fuzzy numbers. S.P. Chen [5] used the α-cut approach to convert a fuzzy bulk-service queue into a series of traditional crisp bulk-service queues. Anggraito et al. [20] determined the stability condition for a multi-server queuing model with both large and small jobs, assuming an infinite number of servers and considering various service time distributions for the small jobs. Golovin et al. [21] studied about the multiresource queuing system with multiserver customers. Anupama [22] considered $M/M/3$ Queuing system with a Queue-dependent multi-server. Fuzzy set theory, an extension of traditional set theory, addresses the limitations of crisp numbers in handling uncertainty. Crisp numbers, commonly used in applications, may not be suitable for scenarios with uncertainties at multiple points. In such cases, intuitionistic fuzzy numbers prove to be a more robust tool.

The central objective of this work is to juxtapose fuzzy queuing models against intuitionistic fuzzy queuing models while also exploring methods to tackle fuzzy queuing challenges without fundamentally altering their core tenets. Existing endeavors typically involve translating fuzzy queuing problems into classical equivalents and then employing various techniques such as the α-cut approach, the L-R approach, the interval arithmetic approach, and the DSW algorithm for resolution. However, the outcomes derived from these methods produce definite results, thereby altering the intrinsic nature of the original problem. This manuscript, titled " Evaluating Multi-Server Queuing System Efficiency: A Comparative Study between Fuzzy Queuing Model and Intuitionistic Fuzzy Queuing Model with Infinite Capacity Constraints," endeavors to circumvent this issue. We propose a methodology to address queuing problems without the need for their conversion into classical equivalents, thus facilitating a direct comparison between fuzzy and intuitionistic fuzzy

queuing models. The effectiveness of our proposed approach is demonstrated through the resolution of numerical problems and subsequent comparative analysis of outcomes.

The structural framework of the article can be delineated as follows: Module 2 elucidates essential definitions, while Module 3 delves into the intricate details of the system model. Manuals and preconceptions are scrutinized in Module 4, followed by the exploration of fundamental theorems in Module 5. Module 6 presents the envisaged queuing model. Module 7 offers a mathematical illustration, providing a tangible demonstration of the proposed methodology's efficacy. Module 8 serving as the culminating segment of the framework.

II. Preliminaries

The motive of this division is to give some basic definitions, annotations, and outcomes that are used in our subsequent calculations.

Definition 2.1. [11] A fuzzy number $\tilde{A}$ is defined on R, the set of real numbers is said to be a triangular fuzzy number (TFN) if its membership function $\mu_{\tilde{A}}: R \to [0,1]$ which satisfy the following conditions:

$$\mu_{\tilde{A}}(x)= \begin{cases} \frac{x-\tilde{a}_1}{\tilde{a}_2-\tilde{a}_1} \text{ for } \tilde{a}_1 \le x \le \tilde{a}_2 \\ 1 \text{ for } x = \tilde{a}_2 \\ \frac{\tilde{a}_3-x}{\tilde{a}_3-\tilde{a}_2} \text{ for } \tilde{a}_2 \le x \le \tilde{a}_3 \\ 0 \text{ otherwise} \end{cases}$$

Definition 2.2. [11] Let the two triangular fuzzy numbers be $\tilde{P} \approx (\tilde{a}_1, \tilde{a}_2, \tilde{a}_3) = (\tilde{m}_1, \tilde{\alpha}_1, \tilde{\beta}_1)$ and $\tilde{Q} \approx (\tilde{b}_1, \tilde{b}_2, \tilde{b}_3) = (\tilde{m}_2, \tilde{\alpha}_2, \tilde{\beta}_2)$ and then the arithmetic operations on TFN be given as follows:

Addition

$$\tilde{P} + \tilde{Q} \approx (\tilde{m}_1 + \tilde{m}_2, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}) \quad (1)$$

Subtraction

$$\tilde{P} - \tilde{Q} \approx (\tilde{m}_1 - \tilde{m}_2, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}) \quad (2)$$

Multiplication

$$\tilde{P}.\tilde{Q} \approx (\tilde{m}_1.\tilde{m}_2, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}) \quad (3)$$

Division

$$\frac{\tilde{P}}{\tilde{Q}} \approx \left(\frac{\tilde{m}_1}{\tilde{m}_2}, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}\right) \quad (4)$$

Scalar Multiplication

$$\begin{aligned} k\tilde{P} &\approx (ka_2, \alpha_1, \beta_1), k \ge 0 \\ k\tilde{P} &\approx (-ka_2, \alpha_1, \beta_1), k < 0 \end{aligned} \quad (5)$$

A triangular fuzzy number $\tilde{P} \approx (\tilde{a}_1, \tilde{a}_2, \tilde{a}_3) \in F(R)$ is known to be positive if $\Re(\tilde{P}) > 0$ and defined by $\tilde{P} \succ 0$.

Definition 2.3. [16] A fuzzy number $\tilde{A}'$ on R is said to be a triangular intuitionistic fuzzy number (TIFN) if its membership function $\mu_{\tilde{A}'}: R \to [0,1]$ and non – membership function $\gamma_{\tilde{A}'}: R \to [0,1]$ have the following conditions:

$$\mu_{\tilde{A}'}(x) = \begin{cases} \dfrac{x-\tilde{a}_1}{\tilde{a}_2-\tilde{a}_1} \text{ for } \tilde{a}_1 \le x \le \tilde{a}_2 \\ 1 \text{ for } x = \tilde{a}_2 \\ \dfrac{\tilde{a}_3-x}{\tilde{a}_3-\tilde{a}_2} \text{ for } \tilde{a}_2 \le x \le \tilde{a}_3 \\ 0 \text{ otherwise} \end{cases}$$

and

$$\gamma_{\tilde{A}'}(x) = \begin{cases} 1 \text{ for } x < \tilde{a}_1', x > \tilde{a}_3' \\ \dfrac{\tilde{a}_2-x}{\tilde{a}_2-\tilde{a}_1'} \text{ for } \tilde{a}_1' \le x \le \tilde{a}_2 \\ 0 \text{ for } x = \tilde{a}_2 \\ \dfrac{x-\tilde{a}_2}{\tilde{a}_3-\tilde{a}_2} \text{ for } \tilde{a}_2 \le x \le \tilde{a}_3' \end{cases}$$

and is given by $\tilde{A}' = (a_1, a_2, a_3; a_1', a_2, a_3')$ where $\tilde{a}_1' \le \tilde{a}_1 \le \tilde{a}_2 \le \tilde{a}_3 \le \tilde{a}_3'$

Definition 2.4. The extension of fuzzy arithmetic operations of Ming Ma et al. [11] to the set of TIFN based upon both location indices and functions of fuzziness indices. The location indices number is taken in the regular arithmetic while the functions of fuzziness indices are assumed to follow the lattice rule, which is the least upper bound in the lattice $\tilde{I}'$. For any two arbitrary TIFN $\tilde{P}' \approx (\tilde{m}_1, \tilde{\alpha}_1, \tilde{\beta}_1; \tilde{m}_1, \tilde{\alpha}_1', \tilde{\beta}_1')$ and $\tilde{Q}' \approx (\tilde{m}_2, \tilde{\alpha}_2, \tilde{\beta}_2; \tilde{m}_2, \tilde{\alpha}_2', \tilde{\beta}_2')$ and $* \in \{+, -, \times, \div\}$, then the arithmetic operations on TIFN are defined by $\tilde{P}' * \tilde{Q}' = (\tilde{m}_1 * \tilde{m}_2, \tilde{\alpha}_1 \vee \tilde{\alpha}_2, \tilde{\beta}_1 \vee \tilde{\beta}_2; \tilde{m}_1 * \tilde{m}_2, \tilde{\alpha}_1' \vee \tilde{\alpha}_2', \tilde{\beta}_1' \vee \tilde{\beta}_2')$.

In particular, for any two TIFNs $\tilde{P}' \approx (\tilde{m}_1, \tilde{\alpha}_1, \tilde{\beta}_1; \tilde{m}_1, \tilde{\alpha}_1', \tilde{\beta}_1')$ and $\tilde{Q}' \approx (\tilde{m}_2, \tilde{\alpha}_2, \tilde{\beta}_2; \tilde{m}_2, \tilde{\alpha}_2', \tilde{\beta}_2')$ the arithmetic operations are defined as

$$\tilde{P}' * \tilde{Q}' = (\tilde{m}_1, \tilde{\alpha}_1, \tilde{\beta}_1; \tilde{m}_1, \tilde{\alpha}_1', \tilde{\beta}_1') * (\tilde{m}_2, \tilde{\alpha}_2, \tilde{\beta}_2; \tilde{m}_2, \tilde{\alpha}_2', \tilde{\beta}_2')$$

$$\tilde{P}' * \tilde{Q}' = (\tilde{m}_1 * \tilde{m}_2, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}; \tilde{m}_1 * \tilde{m}_2, max\{\tilde{\alpha}_1', \tilde{\alpha}_2'\}, max\{\tilde{\beta}_1', \tilde{\beta}_2'\})$$

$$\tilde{P}' * \tilde{Q}' = (\tilde{m}_1 * \tilde{m}_2, \tilde{\alpha}_1 \vee \tilde{\alpha}_2, \tilde{\beta}_1 \vee \tilde{\beta}_2; \tilde{m}_1 * \tilde{m}_2, \tilde{\alpha}_1' \vee \tilde{\alpha}_2', \tilde{\beta}_1' \vee \tilde{\beta}_2')$$

In particular, for any two TIFN $\tilde{P}' \approx (\tilde{a}_1, \tilde{a}_2, \tilde{a}_3; \tilde{a}_1', \tilde{a}_2', \tilde{a}_3') \approx (\tilde{m}_1, \tilde{\alpha}_1, \tilde{\beta}_1; \tilde{m}_1, \tilde{\alpha}_1', \tilde{\beta}_1')$, $\tilde{Q}' \approx (\tilde{b}_1, \tilde{b}_2, \tilde{b}_3; \tilde{b}_1', \tilde{b}_2', \tilde{b}_3') \approx (\tilde{m}_2, \tilde{\alpha}_2, \tilde{\beta}_2; \tilde{m}_2, \tilde{\alpha}_2', \tilde{\beta}_2')$we define:

Addition

$$\tilde{P}' + \tilde{Q}' = (m_1 + m_2, max\{\alpha_1, \alpha_2\}, max\{\beta_1, \beta_2\}; m_1 + m_2, max\{\alpha_1', \alpha_2'\}, max\{\beta_1', \beta_2'\}) \quad (6)$$

Subtraction

$$\tilde{P}' - \tilde{Q}' = (m_1 - m_2, max\{\alpha_1, \alpha_2\}, max\{\beta_1, \beta_2\}; m_1 - m_2, max\{\alpha_1', \alpha_2'\}, max\{\beta_1', \beta_2'\}) \quad (7)$$

Multiplication

$$\tilde{P}' \times \tilde{Q}' = (m_1 \times m_2, max\{\alpha_1, \alpha_2\}, max\{\beta_1, \beta_2\}; m_1 \times m_2, max\{\alpha_1', \alpha_2'\}, max\{\beta_1', \beta_2'\}) \quad (8)$$

Division

$$\tilde{P}' \div \tilde{Q}' = (m_1 \div m_2, max\{\alpha_1, \alpha_2\}, max\{\beta_1, \beta_2\}; m_1 \div m_2, max\{\alpha_1', \alpha_2'\}, max\{\beta_1', \beta_2'\}) \quad (9)$$

Scalar Multiplication

$$k\tilde{P}' = (ka_2, \alpha_1, \beta_1; ka_2, \alpha_1', \beta_1'), \text{ for } k \ge 0 \quad (10)$$

$$k\tilde{P}' = (-ka_2, \alpha_1, \beta_1; -ka_2, \alpha_1', \beta_1'), \text{ for } k < 0$$

The triangular fuzzy number and triangular intuitionistic fuzzy number are illustrated in Figures 1 and 2.

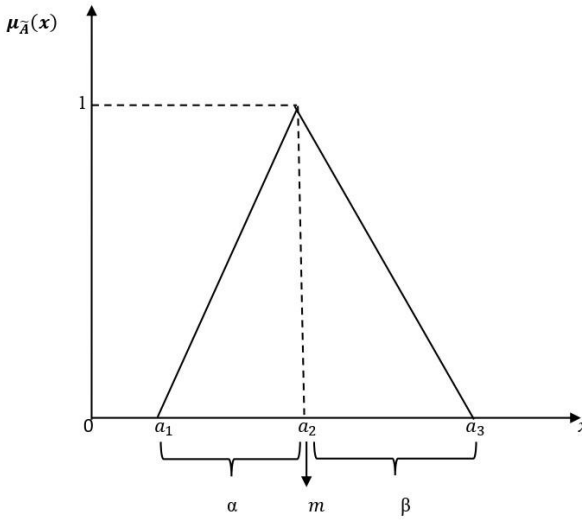


Figure 1: *Triangular fuzzy number*

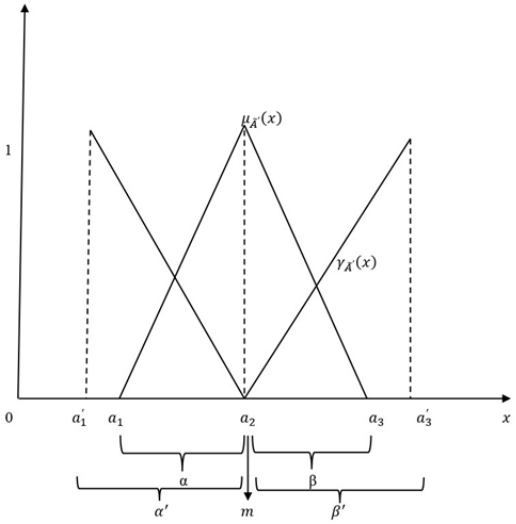


Figure 2: *Triangular intuitionistic fuzzy number*

III. Model Description

We suggest a queuing model with an infinite population and multiple parallel service facilities, following a First-Come-First-Served (FCFS) rule denoted as $(FM/FM/C)$: $(\infty/FCFS)$. In this model, entry times and service durations adhere to Poisson and Exponential distributions, respectively, with fuzzy parameters λ and μ.

Both TFN and TIFN are used for determining the entry rate and processing rate. By analyzing both triangular fuzzy numbers and intuitionistic fuzzy numbers, performance measures are determined, and models are compared according to the mean number of customers in the sequence and infrastructure as well as the duration of their time in the queue and system. As long as we maintain the fuzziness values until the end of the process, the problems can be solved without changing them from fuzzy to crisp. By doing so, it becomes easier to apply to real-life situations. Figure 3 depicts this situation.

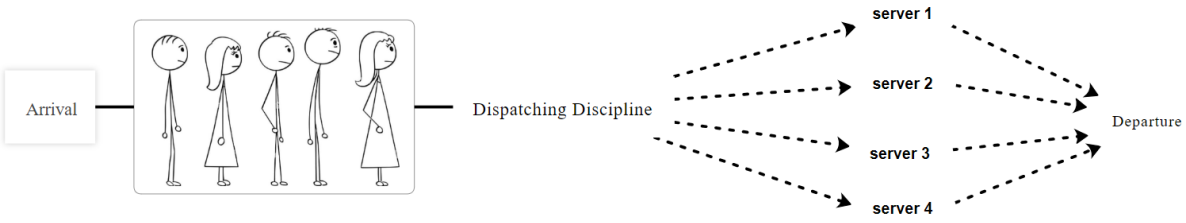


Figure 3: *Multiserver queuing system*

IV. Concerns and Engravings

I. Concerns

i) Consider the endless capacity $FM/FM/c$ fuzzy queuing model under the FCFS discipline.

ii) The system is served by a multiple server.

iii) The time between arrivals follows a Poisson distribution.

iv) Service time is distributed exponentially.

v) Both arrival rate and service rate are described using triangular fuzzy numbers and intuitionistic triangular fuzzy numbers.

II. Syntaxes

Here we are using the following notations:

$\tilde{\lambda}, \tilde{\lambda}' \rightarrow$ The average number of patrons who arrive in a predetermined period of time.

$\tilde{\mu}, \tilde{\mu}' \rightarrow$ The average number of patrons being serviced per unit of time.

$\tilde{\rho} \rightarrow$Traffic intensity.

$\widetilde{N}_q, \widetilde{N}_q' \rightarrow$ The estimated number of patrons in the line.

$\widetilde{N}_s, \widetilde{N}_s' \rightarrow$The estimated number of patrons in the system.

$\tilde{T}_q, \tilde{T}_q' \rightarrow$ The estimated sojourn time of the patrons in the queue.

$\tilde{T}_s, \tilde{T}_s' \rightarrow$ The estimated sojourn time of the patrons in the system.

$c \rightarrow$ Number of servers

$\tilde{P}, \tilde{P}' \rightarrow$ Interarrival rate.

$\tilde{Q}, \tilde{Q}' \rightarrow$ Service rate.

$X \rightarrow$ Arrival time schedule.

$Y \rightarrow$ Service time schedule.

V. Multiserver queuing model with infinite capacity

The Probability on n units in the $(FM/FM/c)$: $(\infty/FIFO)$ is given as

The birth and death process in steady state,

$$\tilde{P}_n' = \left(\frac{\tilde{\lambda}_0'\tilde{\lambda}_1'\tilde{\lambda}_2'\tilde{\lambda}_3'\cdots\tilde{\lambda}_{n-1}'}{\tilde{\mu}_1'\tilde{\mu}_2'\tilde{\mu}_3'\tilde{\mu}_4'\cdots\tilde{\mu}_n'}\right)\tilde{P}_0' \tag{11}$$

For this model,

$$\tilde{\lambda}_n' = \tilde{\lambda}' \quad \text{for all } n$$

$$\tilde{\mu}_n' = \begin{cases} n\tilde{\mu}' & \text{for all } n < c \\ c\tilde{\mu}' & \text{for all } n \geq c \end{cases}$$

$$\tilde{P}_n' = \begin{cases} \dfrac{1}{n!}\left(\dfrac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n \tilde{P}_0' & \text{for all } n < c \\ \dfrac{\tilde{\lambda}'^n}{(1\tilde{\mu}')(2\tilde{\mu}')(3\tilde{\mu}')(4\tilde{\mu}')\cdots(c-1)\tilde{\mu}'(c\tilde{\mu}')(c\tilde{\mu}')(c\tilde{\mu}')\cdots(n-(c-1))times}\tilde{P}_0' & \text{for all } n \geq c \end{cases}$$

$$\tilde{P}_n' = \begin{cases} \frac{1}{n!}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n \tilde{P}_0' & \text{for all } n < c \\ \frac{\tilde{\lambda}'^n}{(c-1)!\,\tilde{\mu}'^{c-1}\;\; c^{n-c+1}\tilde{\mu}'^{n-c+1}} \tilde{P}_0' & \text{for all } n \geq c \end{cases}$$

$$\tilde{P}_n' = \begin{cases} \frac{1}{n!}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n \tilde{P}_0' & \text{for all } n < c \\ \frac{\tilde{\lambda}'^n}{c!\;\; c^{n-c}\tilde{\mu}'^n} \tilde{P}_0' & \text{for all } n \geq c \end{cases}$$

$$\tilde{P}_n' = \begin{cases} \frac{1}{n!}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n \tilde{P}_0' & \text{for all } n < c \\ \frac{1}{c!\;\; c^{n-c}}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n \tilde{P}_0' & \text{for all } n \geq c \end{cases}$$

From the property of discrete random variable, we have

$$\sum_{n=0}^{\infty} \tilde{P}_n' = 1$$

$$\Rightarrow \sum_{n=0}^{c-1} \tilde{P}_n' + \sum_{n=c}^{\infty} \tilde{P}_n' = 1$$

$$\Rightarrow \sum_{n=0}^{c-1} \frac{1}{n!}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n \tilde{P}_0' + \sum_{n=c}^{\infty} \frac{1}{c!c^{n-c}}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n \tilde{P}_0' = 1$$

$$\Rightarrow \tilde{P}_0'\left(\sum_{n=0}^{c-1} \frac{1}{n!}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n + \sum_{n=c}^{\infty} \frac{1}{c!c^{n-c}}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n\right) = 1 \tag{12}$$

Consider

$$\sum_{n=c}^{\infty} \frac{1}{c!\,c^{n-c}}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n = \frac{1}{c!\,c^{-c}} \sum_{n=c}^{\infty} \left(\frac{\tilde{\lambda}'}{\tilde{\mu}'c}\right)^n$$

$$\sum_{n=c}^{\infty} \frac{1}{c!\,c^{n-c}}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n = \frac{1}{c!\,c^{-c}}\left(\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'c}\right)^c + \left(\frac{\tilde{\lambda}'}{\tilde{\mu}'c}\right)^{c+1} + \left(\frac{\tilde{\lambda}'}{\tilde{\mu}'c}\right)^{c+2} + \cdots\right)$$

$$\sum_{n=c}^{\infty} \frac{1}{c!\,c^{n-c}}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n = \frac{\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'c}\right)^c}{c!\,c^{-c}}\left(1 + \frac{\tilde{\lambda}'}{\tilde{\mu}'c} + \left(\frac{\tilde{\lambda}'}{\tilde{\mu}'c}\right)^2 + \left(\frac{\tilde{\lambda}'}{\tilde{\mu}'c}\right)^3 + \cdots\right)$$

$$\sum_{n=c}^{\infty} \frac{1}{c!\,c^{n-c}}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n = \frac{1}{c!}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^c \left(1 - \frac{\tilde{\lambda}'}{\tilde{\mu}'c}\right)^{-1}$$

$$\sum_{n=c}^{\infty} \frac{1}{c!\,c^{n-c}}\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^n = \frac{\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^c}{c!\left(1 - \frac{\tilde{\lambda}'}{\tilde{\mu}'c}\right)} \tag{13}$$

Substituting Equation (13) in Equation (12), we get

$$\tilde{P}_0' \left[\sum_{n=0}^{c-1} \frac{1}{n!} \left(\frac{\tilde{\lambda}'}{\tilde{\mu}'} \right)^n + \frac{\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'} \right)^c}{c! \left(1 - \frac{\tilde{\lambda}'}{\tilde{\mu}' c} \right)} \right] = 1$$

$$\tilde{P}_0' = \left[\sum_{n=0}^{c-1} \frac{1}{n!} \left(\frac{\tilde{\lambda}'}{\tilde{\mu}'} \right)^n + \frac{\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'} \right)^c}{c! \left(1 - \frac{\tilde{\lambda}'}{\tilde{\mu}' c} \right)} \right]^{-1}$$

Or, we can write

$$\tilde{P}_0'^{-1} = \sum_{n=0}^{c-1} \frac{1}{n!} \left(\frac{\tilde{\lambda}'}{\tilde{\mu}'} \right)^n + \frac{\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'} \right)^c}{c! \left(1 - \frac{\tilde{\lambda}'}{\tilde{\mu}' c} \right)}$$

Theorem 5.1. In the $(FM/FM/c)$: $(\infty/FIFO)$ queuing model, that there will be n patrons, where $n = 0$ (n is an integer) in the system at any time $(\tilde{t}' + \Delta\tilde{t}')$, then prove that the rate is $\tilde{\lambda}'\tilde{P}_0' = \tilde{\mu}'\tilde{P}_1'$.

Proof. For $n = 0$,
The possibility that somehow there won't be any patrons in the system at $(\tilde{t}' + \Delta\tilde{t}')$ will be equal to the sum of the following two separate probabilities:

i) $\tilde{P}_0'(\tilde{t}').\left(1 - \tilde{\lambda}'\Delta\tilde{t}'\right)$ is the statistical likelihood that there would be no patrons in the system at $\tilde{t}'$ and no arrival at $\Delta\tilde{t}'$

ii) $\tilde{P}_1'(\tilde{t}').\tilde{\mu}'\Delta\tilde{t}'.\left(1 - \tilde{\lambda}'\Delta\tilde{t}'\right) \cong \tilde{P}_1'(\tilde{t}')\tilde{\mu}'\Delta\tilde{t}' + o(\Delta\tilde{t}')$ is the statistical likelihood that one patron is present in the system at $\tilde{t}'$, one patron is serviced at $\Delta\tilde{t}'$ and no arrival at $\Delta\tilde{t}'$.

As a result, the probability of $n = 0$ is

$$\tilde{P}_0'(\tilde{t}' + \Delta\tilde{t}') = \tilde{P}_0'(\tilde{t}')\left(1 - \tilde{\lambda}'\Delta\tilde{t}'\right) + \tilde{P}_1'(\tilde{t}')\tilde{\mu}'\Delta\tilde{t}' + o(\Delta\tilde{t}') \quad (14)$$

$$\tilde{P}_0'(\tilde{t}' + \Delta\tilde{t}') - \tilde{P}_0'(\tilde{t}') = -\tilde{\lambda}'\Delta\tilde{t}'\tilde{P}_0'(\tilde{t}') + \tilde{P}_1'(\tilde{t}')\tilde{\mu}'\Delta\tilde{t}' + o(\Delta\tilde{t}')$$

Now divide the aforementioned Equation by $\Delta\tilde{t}'$ and assuming the limit as $\Delta\tilde{t}' \to 0$, hence it transforms into:

$$\Delta\tilde{t}' \to 0 \left[\frac{\tilde{P}_0'(\tilde{t}' + \Delta\tilde{t}') - \tilde{P}_0'(\tilde{t}')}{\Delta\tilde{t}'} \right] = -\tilde{\lambda}'\tilde{P}_0'(\tilde{t}') + \tilde{\mu}'\tilde{P}_1'(\tilde{t}'); n = 0$$

$$\tilde{P}_0'(\tilde{t}') = -\tilde{\lambda}'\tilde{P}_0'(\tilde{t}') + \tilde{\mu}'\tilde{P}_1'(\tilde{t}') \text{ for } n = 0 \quad (15)$$

In the scenario of a steady state where $\tilde{t}' \to \infty$, $\tilde{P}_n'(\tilde{t}') \to \tilde{P}_n'$ (independent of $\tilde{t}'$) and consequently $\tilde{P}_n'(\tilde{t}') \to 0$. Eventually, the steady state difference equations of the system are provided by

$$0 = -\tilde{\lambda}'\tilde{P}_0' + \tilde{\mu}'\tilde{P}_1' \quad \text{(From Equation (15))}$$

Furthermore, $\tilde{\lambda}'\tilde{P}_0' = \tilde{\mu}'\tilde{P}_1'; n = 0$ (16)

Hence the proof.

Theorem 5.2. In the $(FM/FM/c)$: $(\infty/FIFO)$ queuing model, at the state $1 \le n \le c - 1$, where $0 \le n \le c$ prove that the rate is $\left(\tilde{\lambda}' + n\tilde{\mu}'\right)\tilde{P}_n' = \tilde{\lambda}'\tilde{P}_{n-1}' + (n+1)\tilde{\mu}'\tilde{P}_{n+1}'$.

Proof. Let,

$$\tilde{\lambda}_n' = \tilde{\lambda}' \quad \text{for all } n$$

$$\tilde{\mu}_n' = \begin{cases} n\tilde{\mu}' & \text{for all } n < c \\ c\tilde{\mu}' & \text{for all } n \ge c \end{cases}$$

For $n = 1,2,3,\ldots c - 1$,

When independent conditions are incorporated together, the probability will be shown as follows:

i) At any time $\tilde{t}'$, there are n patrons in the system are $\tilde{P}_n'(\tilde{t}')$. However, at $\Delta\tilde{t}'$, there are no arrivals and no services respectively, $(1-\tilde{\lambda}'\Delta\tilde{t}')$ and $(1-n\tilde{\mu}'\Delta\tilde{t}')$. As a consequence, the probability is provided as

$$\Rightarrow [\tilde{P}_n'(\tilde{t}')]_{\tilde{\lambda}'=0;\tilde{\mu}'=0} = \tilde{P}_n'(\tilde{t}').(1-\tilde{\lambda}'\Delta\tilde{t}').(1-n\tilde{\mu}'\Delta\tilde{t}')$$

$$\Rightarrow [\tilde{P}_n'(\tilde{t}')]_{\tilde{\lambda}'=0;\tilde{\mu}'=0} = \tilde{P}_n'(\tilde{t}').(1-\tilde{\lambda}'\Delta\tilde{t}'-n\tilde{\mu}'\Delta\tilde{t}') + o_1(\Delta\tilde{t}')$$

$$\Rightarrow [\tilde{P}_n'(\tilde{t}')]_{\tilde{\lambda}'=0;\tilde{\mu}'=0} = \tilde{P}_n'(\tilde{t}')\left(1-\Delta\tilde{t}'(\tilde{\lambda}'+n\tilde{\mu}')\right) + o_1(\Delta\tilde{t}') \quad (17)$$

ii) At any time $\tilde{t}'$, there are $(n-1)$ patrons in the system are $\tilde{P}_{n-1}'(\tilde{t}')$; there is one entrance at $\Delta\tilde{t}'$ is $\tilde{\lambda}'\Delta\tilde{t}'$ and no assistance at $\Delta\tilde{t}'$ is $(1-n\tilde{\mu}'\Delta\tilde{t}')$. As a result, the probability is given as

$$\Rightarrow [\tilde{P}_{n-1}'(\tilde{t}')]_{\tilde{\lambda}'=1;\tilde{\mu}'=0} = \tilde{P}_{n-1}'(\tilde{t}').\tilde{\lambda}'\Delta\tilde{t}'.(1-n\tilde{\mu}'\Delta\tilde{t}')$$

$$\Rightarrow [\tilde{P}_{n-1}'(\tilde{t}')]_{\tilde{\lambda}'=1;\tilde{\mu}'=0} = \tilde{\lambda}'\Delta\tilde{t}'\tilde{P}_{n-1}'(\tilde{t}') + o_2(\Delta\tilde{t}') \quad (18)$$

iii) At any time $\tilde{t}'$, there are $(n+1)$ patrons in the system are $\tilde{P}_{n+1}'(\tilde{t}')$; there is no inflow at $\Delta\tilde{t}'$ is $\left(1-\tilde{\lambda}'\Delta\tilde{t}'\right)$ and one service at $\Delta\tilde{t}'$ is $(n+1)\tilde{\mu}'\Delta\tilde{t}'$. As an outcome, the probability is given as

$$\Rightarrow [\tilde{P}_{n+1}'(\tilde{t}')]_{\tilde{\lambda}'=0;\tilde{\mu}'=1} = \tilde{P}_{n+1}'(\tilde{t}').\left(1-\tilde{\lambda}'\Delta\tilde{t}'\right).(n+1)\tilde{\mu}'\Delta\tilde{t}'$$

$$\Rightarrow [\tilde{P}_{n+1}'(\tilde{t}')]_{\tilde{\lambda}'=0;\tilde{\mu}'=1} = (n+1)\tilde{\mu}'\Delta\tilde{t}'\tilde{P}_{n+1}'(\tilde{t}') + o_3(\Delta\tilde{t}') \quad (19)$$

Now adding the above three Equations (17), (18) and (19) independent compound probabilities, we obtain the probability of n patrons in the system at the time $(\tilde{t}'+\Delta\tilde{t}')$. We obtain

$$\tilde{P}_n'(\tilde{t}'+\Delta\tilde{t}') = \tilde{P}_n'(\tilde{t}')\left[1-\left(\tilde{\lambda}'+n\tilde{\mu}'\right)\Delta\tilde{t}'\right] + \tilde{\lambda}'\Delta\tilde{t}'\tilde{P}_{n-1}'(\tilde{t}') + (n+1)\tilde{\mu}'\Delta\tilde{t}'\tilde{P}_{n+1}'(\tilde{t}') + o(\Delta\tilde{t}'); 1 \le n \le c-1 \quad (20)$$

The Equation (20) now becomes

$$\tilde{P}_n'(\tilde{t}'+\Delta\tilde{t}') - \tilde{P}_n'(\tilde{t}') = -\left(\tilde{\lambda}'+n\tilde{\mu}'\right)\tilde{P}_n'(\tilde{t}')\Delta\tilde{t}' + \tilde{\lambda}'\tilde{P}_{n-1}'(\tilde{t}')\Delta\tilde{t}' + (n+1)\tilde{\mu}'\tilde{P}_{n+1}'(\tilde{t}')\Delta\tilde{t}' + o(\Delta\tilde{t}'); 1 \le n \le c-1 \quad (21)$$

Now, presuming that $\Delta\tilde{t}' \to 0$ is the limit and dividing it by $\Delta\tilde{t}'$ on both sides of the above equation, we obtain

$$\Delta\tilde{t}' \to 0\left[\frac{\tilde{P}_n'(\tilde{t}'+\Delta\tilde{t}') - \tilde{P}_n'(\tilde{t}')}{\Delta\tilde{t}'}\right] = -\left(\tilde{\lambda}'+n\tilde{\mu}'\right)\tilde{P}_n'(\tilde{t}') + \tilde{\lambda}'\tilde{P}_{n-1}'(\tilde{t}') + (n+1)\tilde{\mu}'\tilde{P}_{n+1}'(\tilde{t}'); 1 \le n \le c-1$$

By the definition of the first derivative, we get

$$\tilde{P}_n'(\tilde{t}') = -\left(\tilde{\lambda}'+n\tilde{\mu}'\right)\tilde{P}_n'(\tilde{t}') + \tilde{\lambda}'\tilde{P}_{n-1}'(\tilde{t}') + (n+1)\tilde{\mu}'\tilde{P}_{n+1}'(\tilde{t}'); 1 \le n \le c-1 \quad (22)$$

In the circumstance of a steady state where $\tilde{t}' \to \infty$, $\tilde{P}_n'(\tilde{t}') \to \tilde{P}_n'$ (independent of $\tilde{t}'$) and hence $\tilde{P}_n'(\tilde{t}') \to 0$. So, the system of steady-state difference equations is given by

$$0 = -\left(\tilde{\lambda}'+n\tilde{\mu}'\right)\tilde{P}_n' + \tilde{\lambda}'\tilde{P}_{n-1}' + (n+1)\tilde{\mu}'\tilde{P}_{n+1}'; 1 \le n \le c-1 \quad \text{(From Equation (22))}$$

Hence,

$$\left(\tilde{\lambda}'+\tilde{\mu}'\right)\tilde{P}_n' = \tilde{\lambda}'\tilde{P}_{n-1}' + (n+1)\tilde{\mu}'\tilde{P}_{n+1}'; 1 \le n \le c-1 \quad (23)$$

Hence the proof.

Theorem 5.3. In the $(FM/FM/c):(\infty/FIFO)$ queuing model, at the state $n \ge c$, (where n is an integer), proves that the rate is $\left(\tilde{\lambda}'+c\tilde{\mu}'\right)\tilde{P}_n' = \tilde{\lambda}'\tilde{P}_{n-1}' + c\tilde{\mu}'\tilde{P}_{n+1}'$.

Proof. Let,

$$\tilde{\lambda}'_n = \tilde{\lambda}' \qquad \text{for all } n$$

$$\tilde{\mu}'_n = \begin{cases} n\tilde{\mu}' & \text{for all } n < c \\ c\tilde{\mu}' & \text{for all } n \geq c \end{cases}$$

For $n \geq c$,

When independent conditions are incorporated together, the probability will be shown as follows:

i) At any time $\tilde{t}'$, there are n patrons in the system are $\tilde{P}'_n(\tilde{t}')$. However, at $\Delta\tilde{t}'$, there are no arrivals and no services respectively, $(1-\tilde{\lambda}'\Delta\tilde{t}')$ and $(1-c\tilde{\mu}'\Delta\tilde{t}')$. As a consequence, the probability is provided as

$$\Rightarrow [\tilde{P}'_n(\tilde{t}')]_{\tilde{\lambda}'=0;\tilde{\mu}'=0} = \tilde{P}'_n(\tilde{t}').(1-\tilde{\lambda}'\Delta\tilde{t}').(1-c\tilde{\mu}'\Delta\tilde{t}')$$

$$\Rightarrow [\tilde{P}'_n(\tilde{t}')]_{\tilde{\lambda}'=0;\tilde{\mu}'=0} = \tilde{P}'_n(\tilde{t}').(1-\tilde{\lambda}'\Delta\tilde{t}'-c\tilde{\mu}'\Delta\tilde{t}') + o_1(\Delta\tilde{t}')$$

$$\Rightarrow [\tilde{P}'_n(\tilde{t}')]_{\tilde{\lambda}'=0;\tilde{\mu}'=0} = \tilde{P}'_n(\tilde{t}')\left(1-\Delta\tilde{t}'(\tilde{\lambda}'+c\tilde{\mu}')\right) + o_1(\Delta\tilde{t}') \tag{24}$$

ii) At any time $\tilde{t}'$, there are $(n-1)$ patrons in the system are $\tilde{P}'_{n-1}(\tilde{t}')$; there is one entrance at $\Delta\tilde{t}'$ is $\tilde{\lambda}'\Delta\tilde{t}'$ and no assistance at $\Delta\tilde{t}'$ is $(1-c\tilde{\mu}'\Delta\tilde{t}')$. As a result, the probability is given as

$$\Rightarrow [\tilde{P}'_{n-1}(\tilde{t}')]_{\tilde{\lambda}'=1;\tilde{\mu}'=0} = \tilde{P}'_{n-1}(\tilde{t}').\tilde{\lambda}'\Delta\tilde{t}'.(1-c\tilde{\mu}'\Delta\tilde{t}')$$

$$\Rightarrow [\tilde{P}'_{n-1}(\tilde{t}')]_{\tilde{\lambda}'=1;\tilde{\mu}'=0} = \tilde{\lambda}'\Delta\tilde{t}'\tilde{P}'_{n-1}(\tilde{t}') + o_2(\Delta\tilde{t}') \tag{25}$$

iii) At any time $\tilde{t}'$, there are $(n+1)$ patrons in the system are $\tilde{P}'_{n+1}(\tilde{t}')$; there is no inflow at $\Delta\tilde{t}'$ is $(1-\tilde{\lambda}'\Delta\tilde{t}')$ and one service at $\Delta\tilde{t}'$ is $c\tilde{\mu}'\Delta\tilde{t}'$. As an outcome, the probability is given as

$$\Rightarrow [\tilde{P}'_{n+1}(\tilde{t}')]_{\tilde{\lambda}'=0;\tilde{\mu}'=1} = \tilde{P}'_{n+1}(\tilde{t}').(1-\tilde{\lambda}'\Delta\tilde{t}').c\tilde{\mu}'\Delta\tilde{t}'$$

$$\Rightarrow [\tilde{P}'_{n+1}(\tilde{t}')]_{\tilde{\lambda}'=0;\tilde{\mu}'=1} = c\tilde{\mu}'\Delta\tilde{t}'\tilde{P}'_{n+1}(\tilde{t}') + o_3(\Delta\tilde{t}') \tag{26}$$

Now adding the above three Equations (24), (25) and (26) independent compound probabilities, we obtain the probability of n patrons in the system at the time $(\tilde{t}'+\Delta\tilde{t}')$. We obtain

$$\tilde{P}'_n(\tilde{t}'+\Delta\tilde{t}') = \tilde{P}'_n(\tilde{t}')\left[1-\left(\tilde{\lambda}'+c\tilde{\mu}'\right)\Delta\tilde{t}'\right] + \tilde{\lambda}'\Delta\tilde{t}'\tilde{P}'_{n-1}(\tilde{t}') + c\tilde{\mu}'\Delta\tilde{t}'\tilde{P}'_{n+1}(\tilde{t}') + o(\Delta\tilde{t}'); n \geq c \tag{27}$$

The Equation (27) now becomes

$$\tilde{P}'_n(\tilde{t}'+\Delta\tilde{t}') - \tilde{P}'_n(\tilde{t}') = -\left(\tilde{\lambda}'+c\tilde{\mu}'\right)\tilde{P}'_n(\tilde{t}')\Delta\tilde{t}' + \tilde{\lambda}'\tilde{P}'_{n-1}(\tilde{t}')\Delta\tilde{t}' + c\tilde{\mu}'\tilde{P}'_{n+1}(\tilde{t}')\Delta\tilde{t}' + o(\Delta\tilde{t}'); n \geq c \tag{28}$$

Now, presuming that $\Delta\tilde{t}' \to 0$ is the limit and dividing it by $\Delta\tilde{t}'$ on both sides of the above equation, we obtain

$$\Delta\tilde{t}' \to 0\left[\frac{\tilde{P}'_n(\tilde{t}'+\Delta\tilde{t}')-\tilde{P}'_n(\tilde{t}')}{\Delta\tilde{t}'}\right] = -\left(\tilde{\lambda}'+c\tilde{\mu}'\right)\tilde{P}'_n(\tilde{t}') + \tilde{\lambda}'\tilde{P}'_{n-1}(\tilde{t}') + c\tilde{\mu}'\tilde{P}'_{n+1}(\tilde{t}'); n \geq c$$

By the definition of the first derivative, we get

$$\tilde{P}'_n(\tilde{t}') = -\left(\tilde{\lambda}'+c\tilde{\mu}'\right)\tilde{P}'_n(\tilde{t}') + \tilde{\lambda}'\tilde{P}'_{n-1}(\tilde{t}') + c\tilde{\mu}'\tilde{P}'_{n+1}(\tilde{t}'); n \geq c \tag{29}$$

In the circumstance of a steady state where $\tilde{t}' \to \infty$, $\tilde{P}'_n(\tilde{t}') \to \tilde{P}'_n$ (independent of $\tilde{t}'$) and hence $\tilde{P}'_n(\tilde{t}') \to 0$. So, the system of steady-state difference equations is given by

$$0 = -\left(\tilde{\lambda}'+c\tilde{\mu}'\right)\tilde{P}'_n + \tilde{\lambda}'\tilde{P}'_{n-1} + c\tilde{\mu}'\tilde{P}'_{n+1}; n \geq c \qquad \text{(From Equation (29))}$$

Hence,

$$\left(\tilde{\lambda}'+c\tilde{\mu}'\right)\tilde{P}'_n = \tilde{\lambda}'\tilde{P}'_{n-1} + c\tilde{\mu}'\tilde{P}'_{n+1}; n \geq c \tag{30}$$

Hence the proof.

VI. Suggested Queuing Framework and Its Execution

Let $\tilde{\lambda}$ and $\tilde{\lambda}'$ be the fuzzy and intuitionistic fuzzy arrival rates respectively. Let $\tilde{\mu}$ and $\tilde{\mu}'$ be the fuzzy and intuitionistic fuzzy service rates, respectively. At the steady-state, the FCFS discipline is upheld, and the capacity is unlimited.

The operational characteristics of the multi-server model are provided below:

i) Anticipated number of patrons in the system

$$\tilde{N}_s' = \frac{\tilde{\lambda}'\tilde{\mu}'\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^c}{(c-1)!\left(c\tilde{\mu}'-\tilde{\lambda}'\right)^2}\tilde{P}_o + \frac{\tilde{\lambda}'}{\tilde{\mu}'} \tag{31}$$

ii) Anticipated number of patrons standing up in the queue

$$\tilde{N}_q' = \frac{\tilde{\lambda}'\tilde{\mu}'\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^c}{(c-1)!\left(c\tilde{\mu}'-\tilde{\lambda}'\right)^2}\tilde{P}_o \tag{32}$$

iii) Mean waiting time a patron spends in the system

$$\tilde{T}_s' = \frac{\tilde{\mu}'\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^c}{(c-1)!\left(c\tilde{\mu}'-\tilde{\lambda}'\right)^2}\tilde{P}_o + \frac{1}{\tilde{\mu}'} \tag{33}$$

iv) Mean standing up time of a patron in the queue

$$\tilde{T}_q' = \frac{\tilde{\mu}'\left(\frac{\tilde{\lambda}'}{\tilde{\mu}'}\right)^c}{(c-1)!\left(c\tilde{\mu}'-\tilde{\lambda}'\right)^2}\tilde{P}_o \tag{34}$$

Where, $\tilde{P}_o = \frac{1}{\sum_{n=0}^{c-1}\frac{\left(\frac{\tilde{\lambda}}{\tilde{\mu}}\right)^n}{n!} + \frac{\left(\frac{\tilde{\lambda}}{\tilde{\mu}}\right)^c}{c!}\cdot\frac{c\tilde{\mu}}{c\tilde{\mu}-\tilde{\lambda}}}$

VII. Mathematical description

Imagine a bustling supermarket where customers approach the checkout counters at unpredictable intervals. In this model, there are four available checkout counters (servers), each operating at a unique service rate. Interpret the entry rate and the departure rate as both TFNs and TIFNs are symbolized by $\tilde{\lambda}, \tilde{\lambda}'$ and $\tilde{\mu}, \tilde{\mu}'$ respectively.

I. Single server fuzzy queuing model with infinite capacity

Let $\tilde{\lambda} = (12,13,14)$ is the arrival rate and $\tilde{\mu} = (22,23,24)$ is the service rate of the queuing model. Determine the TFN in the form of $(\tilde{m}, \tilde{\alpha}, \tilde{\beta})$ as $\tilde{\lambda} = (13,1,1)$ and $\tilde{\mu} = (23,1,1)$.

To determine the values of the number of patrons and their sojourn time in the queue as well as a system using suitable formulas among (31), (32), (33), & (34). It is necessary to use the appropriate arithmetic operations described in (1), (2), (3), (4), and (5) for add, sub, multiply, and divide, respectively.

The metrics of performance are calculated and tabulated in Table 1.

Table 1: *Performance Measures Using Triangular Fuzzy Numbers*

S. No	Parameters	Quantifiable Metrics Using TFN
1	$\widetilde{N}_s$	$(-0.4345, 0.5655, 1.5655)$
2	$\widetilde{N}_q$	$(-0.9997, 0.0003, 1.0003)$
3	$\tilde{T}_s$	$(-0.9566, 0.0434, 1.0434)$
4	$\tilde{T}_q$	$(-0.99997, 0.00003, 1.00003)$

The following Fig. 4 – 7 depict the visualizations of Table 1.

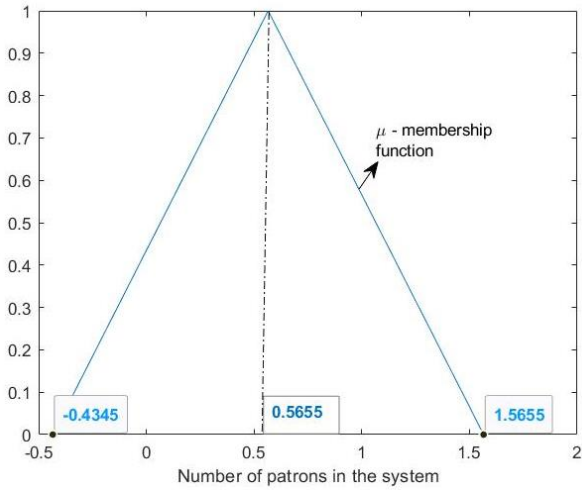


Figure 4: *Number of patrons in the system* $\widetilde{N}_s$

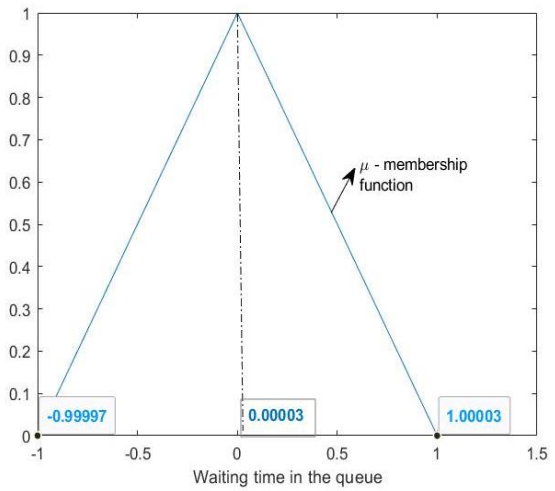


Figure 6: *Waiting time of patrons in the queue* $\tilde{T}_q$

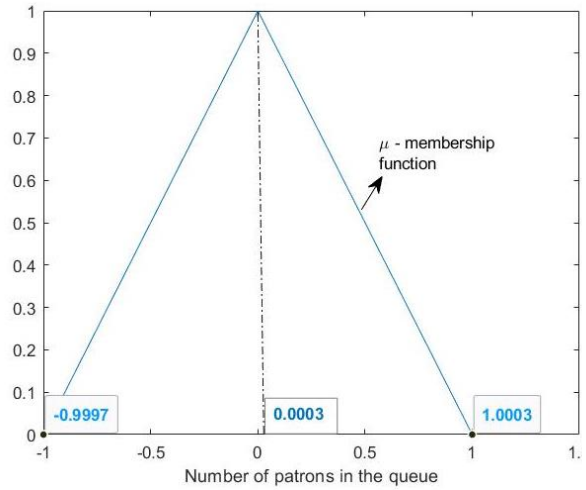


Figure 5: *Number of patrons in the queue* $\widetilde{N}_q$

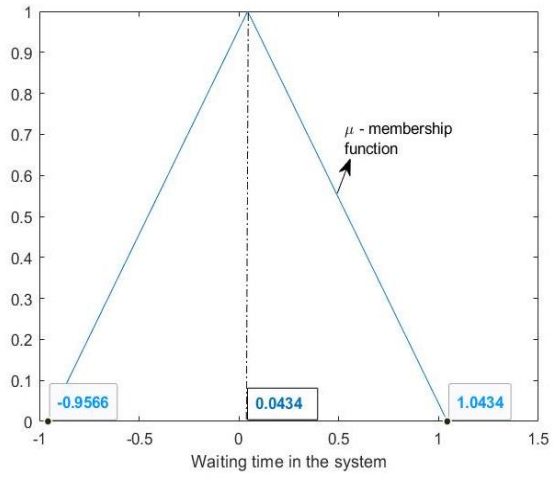


Figure 7: *Waiting time of patrons in the system* $\tilde{T}_s$

II. Single server intuitionistic fuzzy queuing model with infinite capacity

Let $\tilde{\lambda}' = (12,13,14; 11,13,15)$ is the arrival rate and $\tilde{\mu}' = (22,23,24; 21,23,25)$ is the service rate of the queuing model. Determine the TIFN in the guise of $(\tilde{m}, \tilde{\alpha}, \tilde{\beta}; \tilde{m}, \tilde{\alpha}', \tilde{\beta}')$ as $\tilde{\lambda}' = (13,1,1; 13,2,2)$ and $\tilde{\mu}' = (23,1,1; 23,2,2)$.

To determine the values of the number of patrons and their sojourn time in the queue as well as a system using suitable formulas among (31), (32), (33), & (34). It is necessary to use the appropriate arithmetic operations described in (6), (7), (8), (9), and (10) for addition, subtraction, multiplication, and division, respectively.

The metrics of performance are gauged and tabulated in Table 2.

Table 2: *Performance Measures using triangular intuitionistic fuzzy numbers*

S. No	Parameters	Quantifiable Metrics Using TIFN
1	$\tilde{N}_s'$	$(-0.4345,0.5655,1.5655;-1.4345,0.5655,2.5655)$
2	$\tilde{N}_q'$	$(-0.9997,0.0003,1.0003;-1.999,0.0003,2.0003)$
3	$\tilde{T}_s'$	$(-0.9566,0.0434,1.0434;-1.9566,0.0434,2.0434)$
4	$\tilde{T}_q'$	$(-0.9999,0.00003,1.00003;-1.99997,0.00003,2.00003)$

The following Fig. 8 – 11 depict the visualizations of Table 2.

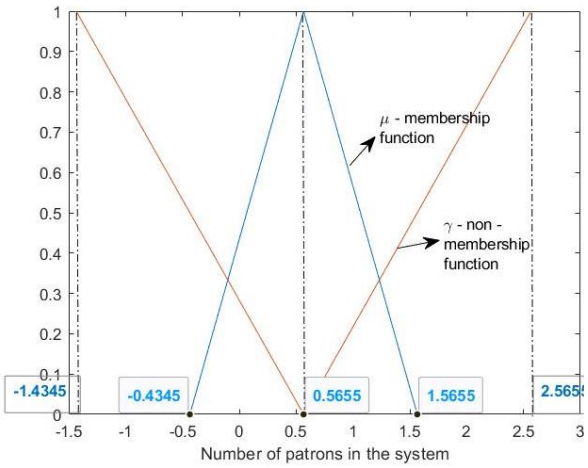


Figure 8: *The membership$(\tilde{\mu})$ and the non-membership$(\tilde{\gamma})$ functions of the number of patrons in the system $\tilde{N}_s'$*

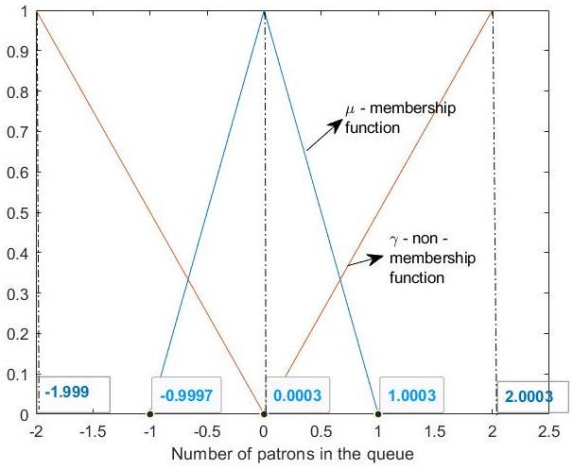


Figure 10: *The membership$(\tilde{\mu})$ and the non-membership$(\tilde{\gamma})$ functions of the number of patrons in the queue $\tilde{N}_q'$*

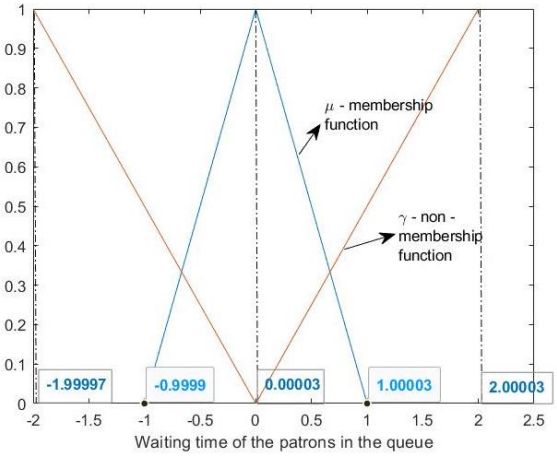


Figure 9: *The membership$(\tilde{\mu})$ and the non-membership$(\tilde{\gamma})$ functions of the waiting time of patrons in the queue $\tilde{T}_q'$*

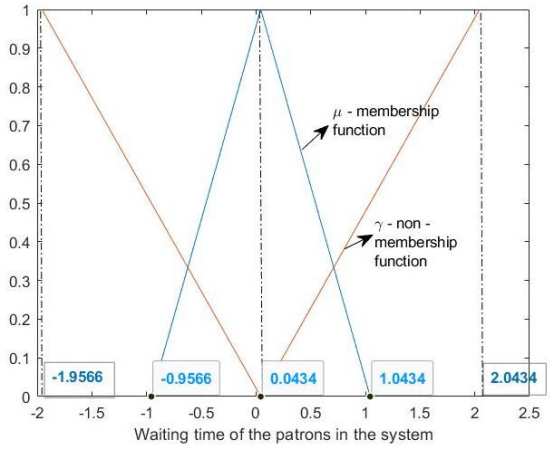


Figure 11: *The membership$(\tilde{\mu})$ and the non-membership$(\tilde{\gamma})$ functions of the waiting time of patrons in the system $\tilde{T}_s'$*

VIII. Conclusion

IFS show a predisposition toward addressing decision-making challenges. Oftentimes, decision-makers struggle to articulate their viewpoints due to insufficient, dependable, or precise data or their hesitancy to explicitly delineate the extent of preference for one option over others. While decision-makers may lean toward certain alternatives, they may not be entirely convinced. Many researchers have focused on IFS theory to tackle optimization problems characterized by ambiguity and inconsistency.

In this study, we delve into the investigation of the $(FM/FM/C)$: $(\infty/FCFS)$ queuing model using algebraic operations techniques. We presume that the entry and processing times are indeterminate. The outcomes of this procedure appear somewhat unclear. However, numerical demonstrations prove the efficacy of this technique. It has been observed that augmenting the convergence of variables can bolster the success of the queuing model. The proposed model is anticipated to assist businesses, distributors, and merchants in effectively identifying optimal performance metrics for queuing systems. There are numerous avenues for expanding the scope of this report, one of which involves interpreting arrival and service rates as fluctuating random variables or fuzzy random variables. Moreover, it is widely acknowledged in the realm of fuzzy logic literature that analytical findings derived within an intuitionistic fuzzy context offer more practical and enlightening insights for software developers and programmers compared to those obtained from a fuzzy model. The fuzzy and intuitionistic fuzzy queues under multiple servers are delineated with greater precision, employing prediction models to yield scientific outcomes. TFN and TIFN numerical descriptors are utilized to evaluate the robustness of the proposed queuing system.

Looking ahead, the proposed queuing model may eventually be expanded to encompass multiple objectives in Erlang. Furthermore, it can serve as a platform for exploring novel facets of intuitionistic set extensions, such as neutrosophic sets.

References

[1] Kumar, C. A. (2023). Analysis of a multiserver system of queue-dependent channel using genetic algorithm. *Mathematics and Computer Science*, 2(23):337–348. Doi:10.1002/9781119896715.

[2] Atanassov, K. T. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, 20:87–96. Doi: https://doi.org/10.1016/S0165-0114(86)80034-3.

[3] Sun, B., Lee, M. H., Dudin, S. A. and Dudin, N. A. (2014). Analysis of multiserver queuing system with opportunistic occupation and reservation of servers. *Mathematical Problems in Engineering*, 178108. Doi: https://doi.org/10.1155/2014/178108.

[4] Buckley, J. J. (1990). Elementary queuing theory based on possibility theory. *Fuzzy Sets and Systems*, 35:43–52.

[5] Chen, S. P. (2005). Parametric nonlinear programming approach to fuzzy queues with bulk service. *European Journal of Operations Research*, 163:434–444.

[6] Chen, S. P. (2006). A mathematics programming approach to the machine interference problem with fuzzy parameters. *Applied Mathematics and Computation*, 174:374–387.

[7] Olliaro, D., Marsan, M. A., Balsamo, S. and Marin, A. (2023). The saturated multiserver job queuing model with two classes of jobs: exact and approximate results. *Performance Evaluation*, 162:102370. Doi: https://doi.org/10.1016/j.peva.2023.102370.

[8] Dudin, A., Dudina, O., Dudin, S. and Samouylov, K. (2021). Analysis of multi-server queue with self-sustained servers. *Mathematics*, 9(17): Article ID 2134. Doi: https://doi.org/10.3390/math9172134.

[9] Kuaban, G. S., Kumar, R., Soodan, B. S. and Czekalski, P. (2020). A multi-server queuing model with balking and correlated reneging with application in health care management. *IEEE Access*, 8:169623–169639. Doi: 10.1109/ACCESS.2020.3024259.

[10] Li, R. J. and Lee, E. S. (1989). Analysis of fuzzy queues. *Computers and Mathematics with Applications*, 17(7):1143–1147. Doi: https://doi.org/10.1016/0898-1221(89)90044-8.

[11] Ma, M., Friedman, M. and Kandel, A. (1999). A new fuzzy arithmetic. *Fuzzy Sets and Systems*, 108(1):83–90.

[12] Negi, D. S. and Lee, E. S. (1992). Analysis and simulation of fuzzy queue. *Fuzzy Sets and Systems*, 46:321–330.

[13] Samouylov, K., Dudina, O. and Dudin, A. (2023). Analysis of multi-server queuing system with flexible priorities. *Mathematics*, 11(4): Article ID 1040. Doi: https://doi.org/10.3390/math11041040.

[14] Selvam, P., Rajkumar, A. and Easwari, J. S. (2017). Ranking of pentagonal fuzzy numbers applying incentre of centroids. *International Journal of Pure and Applied Mathematics*, 117(13):165–174.

[15] Shanmugasundaram, S. and Venkatesh, B. B. (2015). Fuzzy multi-server queuing model through DSW algorithm. *International Journal of Latest Trends in Engineering and Technology*, 5(3):452–457.

[16] Shaw, A. K. and Roy, T. K. (2012). Some arithmetic operations on triangular intuitionistic fuzzy number and its application on reliability evaluation. *International Journal of Fuzzy Mathematics and Systems*, 2(4):363–382.

[17] Srinivasan, V. (2014). Fuzzy queuing model using DSW algorithm. *International Journal of Advanced Research in Mathematics and Applications*, 1(1):57–62.

[18] Thamotharan, S. (2016). A study on multi-server fuzzy queuing model in triangular and trapezoidal fuzzy numbers using α cuts. *International Journal of Science and Research*, 5(1):226–230.

[19] Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, 8:338–353. Doi: https://doi.org/10.1016/S0019-9958(65)90241-X.

[20] Anggraito, A., Olliaro, D., Ajmone Marsan, M. and Marin, A. (2024). Stability Condition for the Multi-server Job Queuing Model: Sensitivity Analysis. In *International Conference on Analytical and Stochastic Modeling Techniques and Applications*, Cham: Springer Nature Switzerland, 148–163. Doi: https://doi.org/10.1007/978-3-031-70753-7_10.

[21] Golovin, A. and Rumyantsev, A. (2023). On Multiresource Queues with Multiserver Customers. In *2023 5th International Conference on Problems of Cybernetics and Informatics (PCI)*, IEEE, 1–3. Doi: https://doi.org/10.1109/PCI60110.2023.10325941.

[22] Anupama and Kumar, C. (2023). Analysis of a Multiserver System of Queue-Dependent Channel Using Genetic Algorithm. *Mathematics and Computer Science*, 2:337–348. Doi: https://doi.org/10.1002/9781119896715.ch23.

RT&A, No 2 (84)
Volume 20, June 2025

FORECASTING VOLATILITY IN INDIAN NATIONAL STOCK EXCHANGE USING A MARKOV SWITCHING GARCH MODEL

D. Pachiyappan[1,*], S. Jayabharathi[2], Y. Geeyha Evanjalin[3], and S. Kavitha[4]

•

[1] Assistant Professor, Department of Mathematics, Vel Tech Rangarajan Dr. Sagunthala R and D Institute of Science and Technology, Avadi, Tamil Nadu, India
Correspondences Mail: drpachiyappand@veltech.edu.in
[2]Associate Professor, Department of Mathematics, Sona college of technology, Salem-5, Tamil Nadu, India. jayaajay2005@yahoo.com
[3]Assistant Professor, Department of Mathematics, Sri Venkateswara college of Engineering and technology Thirupachur, Thiruvallur, Tamil Nadu, India.
[4]Assistant Professor, Department of Statistics, Periyar University, Salem-11, Tamil Nadu, India. pustatkavitha@gmail.com

Abstract

This paper analyzes various GARCH models in terms of their effectiveness in predicting volatility in the national Stock Exchange (NSE). This paper examines GARCH models that feature both Gaussian and fat-tailed residual conditional distributions, and evaluates their ability to characterize and predict volatility over time horizons ranging from 1 to 22 days. The AR(2)-MR-SGARCH-GED model demonstrates superior performance compared to other models at a one-day horizon. The AR(2)-MR-SGARCH-GED and AR(2)-MRSGARCH-t models surpass other models at the 5-day horizon. Within a ten-day timeframe, three AR(2)-MR-SGARCH models surpass the performance of alternative models. Regarding the 22-day forecast horizon, the results demonstrate no distinctions between MR-SGARCH models and standard GARCH models. When looking at risk management outside of samples (95% VaR), a number of models seem to give accurate and reasonable VaR estimates for a one-day period, with a coverage rate close to the nominal level. According to the risk management loss functions, no model is universally the most accurate. This variability suggests that model performance is context-dependent, influenced by the specific characteristics of the data and the underlying assumptions of each approach. Therefore, practitioners should consider employing a combination of models to enhance robustness in their risk assessments and decision-making processes.

Keywords: MRSGARCH, MLE, VaR-estimator, MSE, Forecasting, Risk assessment.

1. Introduction

Volatility forecasting is a crucial endeavor in financial markets, garnering the interest of scholars and professionals for the past twenty years [1]. Volatility is frequently employed as an indicator of risk related to financial returns, making it significant for portfolio managers, option traders, and market makers, among others. Financial market volatility can significantly impact the overall

economy. The predominant category of models utilized in econometric methods for volatility forecasting is GARCH models (refer to Engle [2], Bollerslev [3], Engle and Lee [4], Bollerslev, Engle, and Nelson [5] for a comprehensive overview). In empirical research, the parameters of GARCH models are typically presumed to be constant over time. The conditional distribution of financial returns varies between recessionary and expansionary periods [6]. Furthermore, GARCH models frequently indicate significant volatility persistence following individual shocks. Lamoureux and Lastrapes [7] contended that elevated persistence in volatility may result from structural alterations in the variance process. Cai [8] and Hamilton and Susmel [9] have independently developed the Markov Regime Switching ARCH model (MRSARCH), which integrates Hamilton's Markov Switching model [10, 11] with ARCH specifications.

The MRSARCH model was developed to detect regime shifts in volatility using an unobservable state variable that follows a first-order Markov Chain process. In the ARCH process, parameters may vary across different states. Cai [8] and Hamilton and Susmel [9] employed ARCH specifications to address the issue of infinite path dependence in the Markov Regime Switching GARCH model (MRSGARCH). Additionally, Dueker [12] employed a similar methodology as Gray [13] to address the issue of infinite path dependence and proposed several alternative MRSGARCH models. Klaassen [14] altered Gray's MRSGARCH model, asserting that his specification enhances the forecasting efficacy of MRSGARCH models. Haas, Mittnik, and Paolella [15] introduced a novel method distinct from Gray's [13] approach, asserting that the analytical tractability of their model facilitates the derivation of stationary conditions and dynamic properties.

Leon Li and William Lin [16] and Fong [17] utilized the MRSARCH model developed by Hamilton and Susmel [9] to investigate regime shifts and volatility persistence in the weekly Taiwan Stock Index (TAIEX) and the weekly Japanese Stock Index (TOPIX), respectively. The Fréchet distribution is a special case of the generalized extreme value distribution. This type-II extreme value distribution (Fréchet) case is equivalent to taking the reciprocal of values from a standard Weibull distribution. This distribution has been shown to be useful for modeling and analysis of several extreme events ranging from accelerated life testing to earthquakes, wind speeds, sea currents, floods, and rain fall.

The authors assert that the MRSARCH model offers a superior representation of the data and exhibits significantly reduced volatility persistence compared to uni-regime GARCH models. Kaufmann and Scheicher [18] utilized the MRSARCH model, implemented within a Bayesian framework, to characterize the daily German Stock Index (DAX). Gray [13] utilized weekly one-month US Treasury bill rates from 1970 to 1994. He asserts that the MRSARCH model surpasses basic uni-fregime models in forecasting efficacy and diminishes volatility persistence more effectively than the MRSARCH model proposed by Cai [8] and that of Hamilton and Susmel [9]. Marcucci [6] evaluated a series of GARCH models against a collection of MRSGARCH models regarding their efficacy in forecasting S&P100 volatility over a period ranging from one day to one month. Marcucci [6] concludes that the forecasting efficacy of MRSGARCH models surpasses that of uni-regime GARCH models at shorter time horizons. Standard asymmetric GARCH demonstrates superior performance over extended horizons. Furthermore, MRSARCH models have been utilized in international stock markets by Fornari and Mele [19], Norden and Schaller [20], Susmel [21], Bautista [22], and Leon Li [23].

Erb et al. [24] and Bekaert et al. [25] observe that emerging markets share a significant characteristic: a high degree of country risk, encompassing political, economic, and financial risks. Over the past decades, these markets have experienced currency devaluations, unsuccessful economic strategies, coups, financial shocks, regulatory alterations, and capital market reforms at varying intensities. This work integrates GARCH models within a regime-switching framework, which efficiently accommodates the presence of two distinct volatility regimes, each defined by varying levels of volatility. In both regimes, volatility exhibits a GARCH-like pattern, thereby preventing the actual variance from relying on the complete information set, as noted in Klaassen [14]. The scarcity of research on volatility forecasting in emerging markets, particularly the national

Stock Exchange (NSE), motivated the composition of this paper. The primary objective of this paper is to compare various GARCH models regarding their efficacy in characterizing and predicting the volatility of financial time series over horizons ranging from one day to 22 days.

This paper contributes by addressing the following critical questions: which GARCH models in emerging markets, such as the MSE, exhibit superior performance in volatility forecasting? Do empirical results in MSE differ from those in developed markets? This paper compares standard GARCH models and Markov Regime-Switching GARCH models utilizing normal, t-student, and GED distributions regarding their efficacy in forecasting volatility in the National stock exchange. Empirical findings indicate that there is no disparity in the forecasting accuracy of GARCH models across all time horizons in the National Stock Exchange.

This document is structured as follows. Section 2 delineates various models of stock return volatility, including standard GARCH models and Markov Regime-Switching GARCH models. Section 3 delineates the data. Section 4 delineates empirical findings and engages in a discourse regarding various volatility models, evaluating their efficacy in forecasting volatility. Section 5 offers a concise summary and conclusion.

2. Econometric Methodology

2. 1. GARCH Models

The rate of return R_t is defined as following:

$$R_t = 100[\log(y_t) - \log(y_{t-1})] \tag{1}$$

where y_t are denote stock market index, t denotes the daily closing observations and $t = -S+1, \dots, n$ the sample period consists of an estimation (or in-sample) period with S observations ($t = -S+1, \dots, 0$), and an evaluation (or out-of-sample) period with n observations ($t = 1, \dots, n$). The $\mathrm{GARCH}(1,1), \mathrm{AR}(1)-\mathrm{GARCH}(1,1)$ and $\mathrm{AR}(2)-\mathrm{GARCH}(1,1)$ models for the series of returns r_t are used that they can be written as following

$$\begin{cases} R_t = \mu + \varepsilon_t = \mu + \xi_t\sqrt{h_t} \\ R_t = \mu + \varphi_1 r_{t-1} + \varepsilon_t = \mu + \varphi_1 r_{t-1} + \xi_t\sqrt{h_t} \\ R_t = \mu + \varphi_1 R_{t-1} + \varphi_2 R_{t-2} + \varepsilon_t = \mu + \varphi_1 R_{t-1} + \varphi_2 R_{t-2} + \xi_t\sqrt{h_t} \end{cases} \tag{2}$$

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1 h_{t-1} \tag{3}$$

where $\alpha_0 > 0, \alpha_1 \geq 0$ and $\beta_1 \geq 0$ to guarantee a positive conditional variance, and the innovation is effectively represented as the product of an i.i.d. process with zero mean and unit variance (ξ_t) times the square root of the conditional variance (h_t) [6]. Glosten, Jagannathan and Runkle [26] put forward a modified GARCH model (GJR) to account for the 'leverage effect'. This is an asymmetric GARCH model that allows the conditional variance to respond differently to shocks of either sign and is defined as follows

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2\left[1 - I_{\{\varepsilon_{t-1}>0\}}\right] + \lambda\varepsilon_{t-1}^2 I_{\{\varepsilon_{t-1}>0\}} + \beta_1 h_{t-1} \tag{4}$$

where $I_{\{\omega\}}$ is the indicator function which is equal to one when ω is true and zero otherwise. Another common finding in the GARCH literature is the leptokurtosis of the empirical distribution of financial returns [6]. To model such fat-tailed distributions researchers have adopted the Student's t or the Generalized Error Distribution (GED). Therefore, in addition to the classic Gaussian assumption, in what follows the errors ε_t are also assumed to be distributed according to a Student's t or a GED distribution.

If a Student's t conditional distribution with v degrees of freedom is assumed, the probability density function (pdf) of ε_t, takes the form

$$f(\varepsilon_t) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{v}{2}\right)}(v-2)^{-\frac{1}{2}}(h_t)^{-\frac{1}{2}}\left[1+\frac{\varepsilon_t^2}{h_t(v-2)}\right]^{-\frac{(v+1)}{2}} \tag{5}$$

where $\Gamma(\cdot)$ is the Gamma function and v is the degree-of-freedom (or shape) parameter, constrained to be greater than two so that the second moments exist. With a GED conditional distribution instead, the pdf of the innovations becomes [6]

$$f(\varepsilon_t) = \frac{\text{vexp}\left[-\frac{1}{2}\left|\frac{\varepsilon_t}{\lambda h_t^{\frac{1}{2}}}\right|^v\right]}{h_t^{\frac{1}{2}}\lambda 2^{\left(1+\frac{1}{v}\right)}\Gamma\left(\frac{1}{v}\right)} \tag{6}$$

With $\lambda \equiv \left[\frac{2^{-\frac{2}{v}}\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)}\right]^{\frac{1}{2}}$, where $\Gamma(\cdot)$ is the Gamma function, v is the thickness-of-tail (or shape) parameter, satisfying the condition $0 < v \leq \infty$ and indicating how thick the tails of the distribution are, compared to the normal. When the shape parameter $v = 2$, the GED becomes a standard normal distribution, while for $v < 2$ and $v > 2$ the distribution has thicker and thinner tails than the normal respectively [6].

2. 2. Markov Switching GARCH Model

The main feature of regime-switching models is the possibility for some or all the parameters of the model to switch across different regimes according to a Markov process, which is governed by a state variable, denoted s_t. The state variable is assumed to evolve according to a first-order Markov chain, with transition probability [13,14,6]

$$\Pr(s_t = j \mid s_{t-1} = i) = p_{ij} \tag{7}$$

That indicates the probability of switching from state i at time $t-1$ into state j at t. Usually, these probabilities are grouped together into the transition matrix:

$$P = \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix} = \begin{bmatrix} p & 1-q \\ 1-p & q \end{bmatrix} \tag{8}$$

where for simplicity the existence of only two regimes has been considered. The ergodic probability (that is the unconditional probability) of being in state $s_t = 1$ is given by $\pi_1 = \frac{1-q}{2-p-q}$ [13,14,6]. The MRS-GARCH model in its most general form can be written as

$$r_t \,\middle|\, \Omega_{t-1} \sim \begin{cases} f\left(\theta_t^{(1)}\right) & w.P \cdot p_{1,t} \\ f\left(\theta_t^{(2)}\right) & w.P \cdot \left(1-p_{1,t}\right) \end{cases} \tag{9}$$

where $f(\cdot)$ represents one of the possible conditional distributions that can be assumed, that is Normal (N), student's t or GED, $\theta_t^{(i)}$ denotes the vector of parameters in the i th regime that characterize the distribution, $p_{1,t} = \Pr[s_t = 1 \mid \Omega_{t-1}]$ is the ex ante probability and Ω_{t-1} denotes the information set at time $t-1$, that is the σ-algebra induced by all the variables that are observed at $t-1$. More specifically, the vector of time-varying parameters can be decomposed into three

components [13,14,6].

$$\theta_t^{(i)} = \left(\mu_t^{(i)}, h_t^{(i)}, v_t^{(i)}\right) \tag{10}$$

where $\mu_t^{(i)} \equiv E(r_t \mid \Omega_{t-1}, s_t = i)$ is the conditional mean (or location parameter), $h_t^{(i)} \equiv \mathrm{Var}(r_t \mid \Omega_{t-1})$ is the conditional variance (or scale parameter), and $v_t^{(i)}$ is the shape parameter of the conditional distribution. Hence, the family of density functions of r_t is a location-scale family with time-varying shape parameters in the most general setting [13,14,6]. Therefore, the MRS-GARCH consists of four elements: the conditional mean, the conditional variance, the regime process and the conditional distribution. The conditional mean equation is modeled as a simple model with constant mean and AR (1) and AR (2) models:

$$\begin{cases} R_t^{(i)} = \mu^{(i)} + \varepsilon_t^{(i)} \\ R_t^{(i)} = \mu^{(i)} + \varphi_1^{(i)} R_{t-1} + \varepsilon_t^{(i)} \\ R_t^{(i)} = \mu^{(i)} + \varphi_1^{(i)} R_{t-1} + \varphi_2^{(i)} R_{t-2} + \varepsilon_t^{(i)} \end{cases} \tag{11}$$

where (i) is the symbol of each regime ($i = 1$ or 2), $\varepsilon_t^{(i)} = \eta_t^{(i)} \sqrt{h_t^{(i)}}$ and η_t is a zero mean, unit variance process. The conditional variance of r_t, given the whole regime path (not observed by the econometrician) $\tilde{s}_t = (s_t, s_{t-1}, \ldots)$, is $h_t^{(i)} = V[\varepsilon_t \mid \tilde{s}_t, \Omega_{t-1}]$. For this conditional variance the following GARCH (1,1)-like expression is assumed

$$h_t^{(i)} = \alpha_0^{(i)} + \alpha_1^{(i)} \varepsilon_{t-1}^2 + \beta_1^{(i)} h_{t-1} \tag{12}$$

In which h_{t-1} is a state-independent average of past conditional variances. Actually, in a regime switching context a GARCH model with a state-dependent past conditional variance would be infeasible. The conditional variance would in fact depend not only on the observable information Ω_{t-1} and on the current regime s_t which determines all the parameters, but also on all past states $\tilde{s}_{t-1}$. This would require the integration over a number of (unobserved) regime paths that would grow exponentially with the sample size rendering the model essentially intractable and impossible to estimate. Therefore, a simplification is needed to avoid the conditional variance be a function of all past states. To integrate out the past regimes by also taking into account the current one, Klaassen [14] adopts the following expression for the conditional variance [14]

$$h_t^{(i)} = \alpha_0^{(i)} + \alpha_1^{(i)} \varepsilon_{t-1}^2 + \beta_1^{(i)} E_{t-1}\left\{h_{t-1}^{(i)} \mid s_t\right\} \tag{13}$$

where the expectation is computed as [14] and the probabilities are calculated as [14]

$$E_{t-1}\left\{h_{t-1}^{(i)} \mid s_t\right\} = \tilde{p}_{ii,t-1}\left[\left(\mu_{t-1}^{(i)}\right)^2 + h_{t-1}^{(i)}\right] + \tilde{p}_{ji,t-1}\left[\left(\mu_{t-1}^{(j)}\right)^2 + h_{t-1}^{(j)}\right] - \left[\tilde{p}_{ii,t-1}\mu_{t-1}^{(i)} + \tilde{p}_{ji,t-1}\mu_{t-1}^{(j)}\right]^2 \tag{14}$$

$$\tilde{p}_{ji,t} = \Pr(s_t = j \mid s_{t+1} = i, \zeta_{t-1}) = \frac{p_{ji}\Pr(s_t = j \mid \zeta_{t-1})}{\Pr(s_{t+1} = i \mid \zeta_{t-1})} = \frac{p_{ji}p_{j,t}}{p_{i,t+1}} \tag{15}$$

where $i, j = 1,2$[14]. Klaassen's [14] regime-switching GARCH has two main advantages over the other models. Within the model, it allows higher flexibility in capturing the persistence of shocks to volatility. Furthermore, it allows having straightforward expression for the multi-step ahead volatility forecasts that can be calculated recursively as in standard GARCH models [14]. Since there is no serial correlation in the returns, the h-step ahead volatility forecast at time $T-1$ can be calculated as follows [14,6]

$$\hat{h}_{T,T+h} = \sum_{\tau=1}^{h} \hat{h}_{T,T+\tau} = \sum_{\tau=1}^{h} \sum_{i=1}^{2} \Pr(s_\tau = i \mid \Omega_{T-1}) \hat{h}_{T,T+\tau}^{(i)} \tag{16}$$

where $\hat{h}_{T,T+h}$ denotes the time T aggregated volatility forecast for the next h steps, and $\hat{h}_{T,T+\tau}^{(i)}$ indicates the τ-step-ahead volatility forecast in regime i made at time T that can be calculated recursively [14,6]

$$\hat{h}_{T,T+\tau}^{(i)} = \alpha_0^{(i)} + \left(\alpha_1^{(i)} + \beta_1^{(i)}\right) E_T\left\{\hat{h}_{T,T+\tau-1}^{(i)} \mid s_{T+\tau}\right\} \tag{17}$$

Therefore, the multi-step-ahead volatility forecasts are computed as a weighted average of the multi-step-ahead volatility forecasts in each regime, where the weights are the prediction probabilities. Each regime volatility forecast is obtained with a GARCH-like formula where the expectation of the previous period volatility is determined by weighting the previous regime volatility with the probabilities in (8). In general, to compute the volatility forecasts the filter probability at τ periods ahead $\Pr(s_{t+\tau} = i \mid \Omega_t) = P_{i,t+\tau} = P^\tau p_{i,t}$ is needed [14,6].

Typically, in the Markov regime-switching literature maximum likelihood estimation is adopted to estimate the numerous parameters. An essential ingredient is the ex-ante probability $p_{1,t} = \Pr[s_t = 1 \mid \Omega_{t-1}]$, i.e. the probability of being in the first regime at time t given the information at time $t-1$, whose specification is

$$\begin{aligned} p_{1,t} &= \Pr[s_t = 1 \mid \Omega_{t-1}] \\ &= (1-q)\left[\frac{f(r_{t-1} \mid s_{t-1} = 2)\left(1 - p_{1,t-1}\right)}{f(r_{t-1} \mid s_{t-1} = 1)p_{1,t-1} + f(r_{t-1} \mid s_{t-1} = 2)\left(1 - p_{1,t-1}\right)}\right] \end{aligned}$$

where p and q are the transition probabilities in (10) and $f(\cdot)$ is the likelihood given in (9) [14,6]. Thus, the log-likelihood function can be written as

$$l = \sum_{t=-S+w+1}^{T+w} \log\left[p_{1,t} f(r_t \mid s_t = 1) + \left(1 - p_{1,t}\right) f(r_t \mid s_t = 2)\right] \tag{19}$$

where $w = 0,1,\ldots,n, f(\cdot \mid s_t = i)$ is the conditional distribution given that regime i occurs at time t [6].

2. 3. Estimation Procedure

The conditional mean and conditional variance are estimated simultaneously by maximizing the log-likelihood function, calculated as the logarithm of the product of the conditional densities of the prediction errors, as demonstrated in (14). The maximum likelihood estimates are derived by maximizing the log-likelihood using the Broyden, Fletcher, Goldfarb, and Shanno (BFGS) quasi-Newton optimization algorithm within R's numerical optimization routines.

2. 4. Estimation of Volatility Forecasting

Forecast analysis is a crucial component of any forecasting endeavor. A widely used metric for assessing various forecasting models involves the minimization of a specific statistical loss function. Nevertheless, assessing the quality of competing volatility models can be quite challenging, as noted by Bollerslev, Engle, and Nelson [5] and Lopez [27], due to the absence of a singular criterion for identifying the optimal model. Consequently, despite being somewhat contentious, the majority of

existing literature has predominantly concentrated on a specific statistical loss function, the Mean Squared Error (MSE) [6]. This paper employs the MSE criterion to evaluate competing volatility models. This loss function is defined as follows:

$$MSE = n^{-1}\sum_{t=1}^{n}\left(\hat{\sigma}_{t+1}^2 - \hat{h}_{t+1|t}\right)^2 \tag{20}$$

where $\hat{\sigma}_{t+1}^2$ is the actual volatility that we have used the squared return for the measure of daily volatility and $\hat{h}_{t+1|t}$ is the forecasted volatility based on period of forecasting. We have used 1-day, 5-day, 10-day and 22-day step ahead for forecasting horizon.

2. 5. Value at Risk

VaR is one of the most important measures of market risk portfolio, which quantifies in monetary terms the likely losses that could arise from fluctuations in the market [6]. Brooks and Persand [30] and Sarma, Thomas and Shah [31] suggested to use VaR based loss functions [30,6,31]. The VaR at time t of model i at α% significance level is calculated as following:

$$\mathrm{VaR}_t^i[n,\alpha] = \mu_{t+n}^i + \vartheta(\alpha)\sqrt{h_{t+n}^i} \tag{21}$$

where $\vartheta(\alpha)$ is a cumulative distribution function, n is the investment horizon that we assumed $n = 1,5,10,20$ days, $\alpha = 1\%$ or $5\%, \mu_{t+n}^i$ is the volatility forecasts used as an input for the VaR [30,6]. In this paper, we have employed three methods to determine the adequacy of the volatility forecasts used as an input for the VaR. The TUFF test is based on the number of observation before the first exception. The relevant null is: $\alpha = \alpha_0$ and the corresponding LR test is [30,6,31].

$$LR_{\mathrm{TUFF}}\left(\tilde{T},\hat{\alpha}\right) = -2\log\left\{\hat{\alpha}(1-\hat{\alpha})^{\tilde{T}-1}\right\} + 2\log\left\{\frac{1}{\tilde{T}}\left(1-\frac{1}{\tilde{T}}\right)^{\tilde{T}-1}\right\} \tag{22}$$

where $\tilde{T}$ denotes the number of observations before the first exception. LR_{TUFF} is also asymptotically distributed as $\chi^2(1)$ [30,6,31]. A correctly specified VaR model should generate the pre-specified failure rate conditionally at every points in time [6]. Christoffersen [32] developed a framework for interval forecast evaluation [32,6]. The VaR is interpreted as a forecast that the portfolio return will lie in $(-\infty, VaR)$ with a pre-specified probability p [32,6]. Consider a sequence of one-period-ahead VaR forecasts $\{\vartheta_{t|t-1}\}_{t=1}^{T}$, estimated at a significance level $1-p$[32,6]. These forecasts are intended to be one-sided interval forecasts $(-\infty, \vartheta_{t|t-1}]$ with coverage probability p [32,6]. Given the realizations of the return series r_t and the ex-ante VaR forecasts, the following indicator variable can be calculated [32,6]

$$I_t = \begin{cases} 1, & r_t < \vartheta_t \\ 0, & \text{otherwise} \end{cases}$$

The stochastic process $\{I_t\}$ is called "failure process". The forecasts of VaR are explained to be efficient if they demonstrate "correct conditional coverage" i.e. if $E[I_{t|t-1}] = p, \forall t$ or, equivalently, if $\{I_t\}$ is i.i.d. with mean p[32,6]. Christoffersen [32] developed a three-step testing method to test for correct conditional coverage consisting of a test for correct unconditional coverage, a test for independence and a test for correct conditional coverage.

In the test for correct unconditional coverage the null hypothesis of the failure probability p is tested against the alternative hypothesis that the failure probability is different from p, under the assumption and independently distributed failure process [32,6]. In the test for independence, the

hypothesis of an independently distributed failure process is tested against the alternative hypothesis of a first order Markov failure process [32,6]. Finally, the test of correct conditional coverage is done by testing the null hypothesis of an independent failure process with failure probability p against the alternative of a first order Markov failure process [32,6]. All the three tests are carried out in the likelihood ratio (LR) framework. The likelihood ratio for each is as follows [32,6]. The LR statistic for the test of unconditional coverage [32,6]:

$$LR_{UC} = LR_{PF} = -2\log\left[\frac{p^{n_1}(1-p)^{n_0}}{\hat{\pi}^{n_1}(1-\hat{\pi})^{n_0}}\right] \sim \chi^2_{(1)}$$

where p is the tolerance level at which VaR measures are estimated, n_1 is the number of 1 's in the indicator series, n_0 is the number of 0 's in the indicator series, and $\hat{\pi} = \frac{n_1}{(n_0+n_1)}$ is the MLE estimate of p[32,6]. The LR statistic for the test of independence [32,6]:

$$LR_{\text{ind}} = -2\log\left[\frac{(1-\hat{\pi}_2)^{(n_{00}+n_{10})}(1-\hat{\pi}_2)^{(n_{01}+n_{11})}}{(1-\hat{\pi}_{01})^{n_{00}}\hat{\pi}_{01}^{n_{01}}(1-\hat{\pi}_{11})^{n_{10}}\hat{\pi}_{11}^{n_{11}}}\right] \sim \chi^2_{(1)} \quad (23)$$

where n_{ij} is the number of i value followed by a j value in the I_t series $(i,j = 0,1), \pi_{ij} = \Pr\{I_t = i \mid I_{t-1} = j\}(i,j = 0,1)$, $\hat{\pi}_{01} = \frac{n_{01}}{(n_{01}+n_{01})}, \hat{\pi}_{11} = \frac{n_{11}}{(n_{10}+n_{11})}, \hat{\pi}_2 = \frac{(n_{01}+n_{11})}{(n_{00}+n_{01}+n_{10}+n_{11})}$ [32,6]. The LR statistic for the test of correct conditional coverage [32,6]:

$$LR_{cc} = -2\log\left[\frac{(1-p)^{n_0}p^{n_1}}{(1-\hat{\pi}_{01})^{n_{00}}\hat{\pi}_{01}^{n_{01}}(1-\hat{\pi}_{11})^{n_{10}}\hat{\pi}_{11}^{n_{11}}}\right] \sim \chi^2_{(2)} \quad (24)$$

If we condition on the first observation, then these LR test statistics are related by the identity $LR_{cc} = LR_{\text{ind}} + LR_{UC}$[32,6]. All these tests will be utilized in the evaluation of the VaR-based forecast. Models are ranked in this manner for the interpretation of the results. Consistent with Brooks and Persand [30], we have posited that any model exhibiting percentage failures in the rolling hold-out sample exceeding the nominal threshold should be deemed inadequate and consequently rejected. Consequently, the models with the lowest rankings (the least effective) are those exhibiting the highest percentage of failures exceeding the nominal value [30,6]. Once these models have been depleted, we posit that any model exhibiting a failure rate significantly lower than the anticipated figure is less preferable than those models demonstrating a failure rate nearer to the nominal level [30,6]. Consequently, the optimal models under this loss function are those that produce a coverage rate lower than the nominal rate, yet as proximate to it as feasible [30,6]. Subsequently, we assess the accuracy of both unconditional and conditional coverage; a model that successfully meets these criteria is deemed suitable for risk management objectives [30].

3. Data and Statistical Analysis of National Stock Exchange

The NSE was officially inaugurated in June 1988 with merely six listed companies, in contrast to the 110 companies currently traded by individual and institutional investors. The initial decade of the NSE was characterized by vigorous activity, with capitalization increasing from INR 7.2 billion to INR 240 billion and the number of listed companies expanding to 105. Following 1988, the Indian revolution and Indian incursion into substantially diminished exchange activities, resulting in a decline in capitalization to INR 12.9 billion in 198. Following the conclusion of the India-China war, the NSE was regarded as a pivotal mechanism for promoting economic development by directing savings into investment [33].

This objective rapidly increased the count of listed companies from 56 in 1988 to 422 in 2006 [33]. From 2000 to 2004, the market capitalization of the NSE increased from INR 60 billion to INR 411 billion. The Financial index attained a record peak of 13,482 on August 4, 2004, but within two

years, a significant market correction reduced the index by 35% to 9,069 on July 26, 2006. The stock market decline was not entirely detrimental, as it was viewed as a necessary correction in a market that had outpaced its fundamentals and required enhanced regulation and increased transparency.

Table 1: *Results of Descriptive statistics*

Mean	Std	Min	Max	Skew	Kur	BJ	$Q^2(12)$	LM-(12)
0.3248	0.4920	-4.25	4.23	-0.58	61.75	643732.1	172.89	71.05
P-value:				0.000	0.000	0.000	0.000	0.000

Table 2: *Summary of Unit root tests for rate of return series*

Test type	Augmented Dickey-Fuller	Phillips-Perron
Statistic:	-9.15	-63.17
P-value:	0.00	0.00

The market correction resulted in significant reforms and facilitated the consolidation and merger of numerous smaller companies [33]. By 2007, the market capitalization had surpassed its 2004 level; however, the number of listed companies remained low due to sustained merger and acquisition activity. All NSE transactions must be conducted on the Trading Floor utilizing the Exchange's Automated Trading System (ATS) [34]. Trading occurs from 10:00 am to 12:00 pm for the cash market and from 9:00 am to 12:30 pm for the futures market, from Saturday to Wednesday, excluding public holidays, with a pre-opening period of 30 minutes commencing at 8:30 am. The NSE comprises two markets: the cash market and the futures market. The cash market is divided into two segments: the first market and the second market [39]. An independent market exists for Islamic bonds. The futures market was established in mid-2010 with the introduction of single stock futures.

The NSM has 30 characteristics, one of which is that stock prices are regulated and cannot exceed a certain range due to intervention. The history of NSC is relatively brief in comparison to other stock markets. The information dissemination in this market is exceedingly sluggish. Transactions on the National Stock Exchange are conducted through orders submitted by brokers. The trading days of the week are Saturday, Sunday, Monday, Tuesday, and Wednesday, excluding national holidays. The dataset comprises 3,068 daily observations of the closing values of the National Stock Exchange from August 29, 2014, to August 9, 2023.

The sample is partitioned into two segments. The initial 2567 observations are used as the in-sample for estimation purposes, while the remaining 500 observations are taken as the out-of-sample for forecast evaluation purposes.

The return is calculated as $R_t = 100\left[\log\left(\frac{y_t}{y_{t-1}}\right)\right]$ where y_t is the index value at time t. Table 1 shows some descriptive statistics of the MSE rate of return. The mean is quite small and the standard deviation is around 0.4. The Kurtosis (Ku) is significantly higher than the normal value of 3 indicating that fat-tailed distribution are necessary to correctly describe conditional distribution of R_t. The Skewness (Sk) is significant, small and negative, showing that the lower tail of empirical distribution of the return is longer than the upper tail, that means negative returns are more likely to be far below the mean than their counterparts.

LM (12) is the Lagrange Multiplier test for ARCH effects in the OLS residuals from the regression of the returns on a constant, while $Q^2(12)$ is the corresponding Ljung-Box [35,38] statistic on the squared standardized residuals. Both these statistics are highly significant suggesting the presence of ARCH effects in the MSE returns up to the twelfth order.

4. Empirical Results and Discussion

In the present section we present the empirical estimates of ARMA, single-regime GARCH and MRS-GARCH models, together with the in-sample and the out-of-sample forecast evaluation.

4.1. Single-Regime GARCH and ARMA Models

Initially, we have evaluated various ARMA models of return rates in terms of mean squared error (MSE). Table 3 presents the estimation outcomes of ARMA models. Subsequently, we have evaluated various GARCH models, including GARCH (1,1), AR(1)-GARCH(1,1), and AR(2)-GARCH(1,1) (refer to Table 4). The parameter estimates for various state GARCH models utilizing normal, t-student, and GED distributions are displayed in Tables 5-7. The initial 2567 observations (from 08/29/2014 to 11/27/2021) are designated as the in-sample for estimation, whereas the subsequent 500 observations (from 11/28/2021 to 09/09/2023) are utilized as the out-of-sample for forecasting assessment. All parameters for the different GARCH models are statistically significant concerning the conditional mean.

The estimates of conditional variance indicate that all parameters are highly significant. Furthermore, stationary conditions are present in various GARCH models. Furthermore, for the Student's t distribution, the degrees of freedom consistently exceed 3, indicating that all conditional moments up to the third order are present. The models incorporating GED innovations indicate that the conditional distribution exhibits fatter tails compared to the Gaussian, as all shape parameters significantly approximate one. The utilization of a t-distribution rather than a normal distribution is prevalent in the standard, single-regime GARCH literature.

Table 4: *Estimation of different ARMA* (p, q) *models of rate of return*

	AR (1)	MA (1)	AR (2)	ARMA (1,1)	ARMA (1,2)	ARMA (2,1)	ARMA (2,2)
M	0.023*	0.013*	0.032*	0.031*	0.041*	0.021*	0.041*
φ_1	0.389*	-	0.321*	0.714**	0.5677**	-0.015**	-0.175**
$\boldsymbol{\varphi_2}$	-	-	0.175	-	-	0.30**	0.527**
ϑ_1	-	0.275*	-	-0.391**	-0.237**	0.35**	0.500**
$\boldsymbol{\vartheta_2}$	-	-	-	-	0.068**	-	-0.189**
F-Statistic	459	302	277	271	181	187	142
AIC	-0.552	-0.495	-0.522	-0.571	-0.591	-0.583	-0.5855
SBC	-0.547	-0.450	-0.554	-0.572	-0.570	-0.576	-0.541
HQC	-0.550	-0.479	-0.597	-0.5766	-0.5758	-0.505	-0.514
Log(L)	-710.5	-643.2	749.7718	-746.02	-747.05	-752.7	-755.9

Table 5: *MLE estimates of standard GARCH* (1,1) *models with constant mean of return*

	GARCH-N	GARCH-T	GARCH-GED	GJR-N	GJR-T	GJR-GED
$\boldsymbol{\alpha}$	0.0045**	0.043**	0.0077**	0.0074**	0.043**	0.0388**
$\boldsymbol{\varphi 1}$	0.0358**	0.0267**	0.0316**	0.0343**	0.0232**	0.0207**
$\boldsymbol{\varphi 2}$	0.4753**	0.6612**	0.6055**	0.4509**	0.6156**	0.0814**
$\boldsymbol{\alpha 0}$	0.4575**	0.2739**	0.3073**	0.4852**	0.7687**	0.1945**
$\boldsymbol{\alpha 1}$	-	-	-	0.4079**	0.2812**	0.3079**
$\boldsymbol{\beta 1}$	-	3.8477**	0.9029**	-	3.83**	0.9488
Log(L)	-1289.1196	-838.53	-887.354	-1283.922	-938.1431	-980.0731
P-Value LR Test for Leverage Effect				0.21	0.00	0.00

In regime switching models, a t-distribution can be particularly advantageous. Ultimately, under normal circumstances, a significant innovation during a low volatility phase will precipitate a transition to a high-volatility regime sooner, even if it constitutes a solitary outlier amidst an otherwise stable period. Incorporating a t-distribution will consequently improve regimes stability.

Table 6: *MLE estimates of AR (1)-GARCH (1,1) models*

	GARCH-N	GARCH-T	GARCH-GED	GJR-N	GJR-T	GJR-GED
$\boldsymbol{\alpha}$	0.0176**	0.0119**	0.0457*	0.0905**	0.0096**	0.0280**
$\boldsymbol{\varphi 1}$	0.5188**	0.5349**	0.5633**	0.5163**	0.5368**	0.5638**
$\boldsymbol{\varphi 2}$	0.0061**	0.0027**	0.0027**	0.0069**	0.0025**	0.0036**
$\boldsymbol{\alpha 0}$	0.4796**	0.6635**	0.6203**	0.5078**	0.5337**	0.6039**
$\boldsymbol{\alpha 1}$	0.3882**	0.2963**	0.3575**	0.3896**	0.2922**	0.2532**
$\boldsymbol{\beta 1}$	-	-	-	0.6186**	0.5323**	0.4102**
$\boldsymbol{\gamma}$	-	3.5073**	0.9493**	-	3.5613**	0.9507**
Log(L)	-1299.945	-1814.682	-1774.192	-1300.345	-1814.808	-1774.270
P-Value LR Test for Leverage Effect				0.08	0.00	0.00

Table 7: *MLE estimates of* $AR(2) - GARCH(1,1)$ *models*

	GARCH-N	GARCH-T	GARCH-GED	GJR-N	GJR-T	GJR-GED
$\boldsymbol{\alpha}$	0.0163**	0.0865*	0.0104**	0.0194**	0.0205**	0.0190**
$\boldsymbol{\varphi 1}$	0.4630**	0.4757**	0.5125**	0.4600**	0.4766**	0.5132**
$\boldsymbol{\varphi 2}$	0.0921**	0.0986**	0.0740**	0.0932**	0.0983**	0.0748**
$\boldsymbol{\alpha 0}$	0.0062**	0.0026**	0.0024**	0.0063**	0.0029**	0.0020**
$\boldsymbol{\alpha 1}$	0.4664*	0.6696**	0.5085**	0.5009**	0.6108**	0.5758**
$\boldsymbol{\beta 1}$	0.4068**	0.3025**	0.4113**	0.4070**	0.3009**	0.2967**
$\boldsymbol{\gamma}$	-	-	-	0.0745**	0.0253**	0.4821**
V	-	3.5731**	0.9498**	-	3.5726**	0.9515**
Log (L)	-1305.098	-1824.771	-1781.028	-1305.692	-1824.785	-1781.136
P-Value LR Test for Leverage Effect				0.09	0.00	0.00

The t-distribution encompasses the normal distribution as its limiting case. According to "asymmetric" or "leverage" volatility models, positive and negative news exhibit differing predictability for future volatility. In the majority of these studies, researchers have provided compelling evidence that volatility is asymmetric in equity markets: negative returns typically correlate with significant upward adjustments of conditional volatility, whereas positive returns are linked to minimal upward or even downward adjustments of conditional volatility.

Researchers hypothesize that the asymmetry may result from alterations in leverage corresponding to fluctuations in equity value. Some have contended that the asymmetry may stem from the feedback mechanism between volatility and stock price, wherein fluctuations in volatility lead to alterations in risk premiums (refer to Kavitha et al. [37] and Duraisamy et al. [33]). Asymmetric volatility is most evident during a market crisis, characterized by substantial declines in stock prices accompanied by a notable rise in market volatility.

Asymmetric volatility may elucidate the negative skewness observed in stock return data, as articulated by Harvey and Siddique [49]. This paper demonstrates that a leverage effect is present in the National stock market, as evidenced by GJR models utilizing t-student and GED distributions, where the impact of adverse news on volatility (α_1) surpasses that of favorable news (λ).

The P-Value LR Test for Leverage Effect in Tables 5-7 demonstrates that the disparity between the α_1 and λ coefficients is not significant for the GJR-N model; however, it is significant at the 5% confidence level for GJR models utilizing t-student and GED distributions.

4.2. Markov Regime-Switching GARCH

Tables 8-10 present the parameter estimates for MRS-GARCH models. Both models, one with constant degrees of freedom and the other allowing degrees-of-freedom parameters to transition between two regimes, exhibit highly significant in-sample estimates. All estimates of the conditional mean and parameters of conditional variance are statistically significant.

The estimates indicate the presence of two distinct states: the first regime is marked by low volatility and minimal persistence of shocks in conditional volatility, while the second regime exhibits high volatility and greater persistence. The transition probabilities are significant, indicating that nearly all regimes exhibit considerable persistence in MRS-GARCH models. The results present the unconditional probabilities for each MRS-GARCH model. The unconditional probability π_1 of being in the first regime varies from 20% to 86% across different MRSGARCH models. The unconditional probability π_2 of being in the second regime varies from 14% to 80% across different MR-SGARCH models.

Table 8: *MLE estimates of MRS-GARCH models*

	MRS-GARCH-N	MRS-GARCH-t	MRS-GARCH-GED
$\boldsymbol{\delta^{(1)}}$	0.03302**	0.2379**	0.2917**
$\boldsymbol{\delta^{(2)}}$	-0.0296**	-0.0803**	0.0
$\boldsymbol{\alpha_0^{(1)}}$	0	0.0024**	0.0768**
$\boldsymbol{\alpha_0^{(2)}}$	0.4599**	0.0205**	0.0003**
$\boldsymbol{\alpha_1^{(1)}}$	0.4988**	0.1150**	0.8363**
$\boldsymbol{\alpha_1^{(2)}}$	0.6358**	0.7954**	0.0352**
$\boldsymbol{\beta_1^{(1)}}$	0.2744**	0.8403**	0.0051**
$\boldsymbol{\beta_1^{(2)}}$	0.3881**	0.1614**	0.7962**
P	0.8321**	0.9649**	0.9347**
Q	0.0	0.9745**	0.9559**
$\boldsymbol{v^{(1)}}$	-	3.3174**	0.7470**
Log (L)	-809.79925	-580.6226	-568.81478
$\boldsymbol{\pi_1}$	0.86	0.39	0.40
$\boldsymbol{\pi_2}$	0.14	0.61	0.60

The student's t variant of the MRS-GARCH models, characterized by constant degrees of freedom across regimes, exhibits a shape parameter of less than four, signifying the presence of conditional moments up to the third order. Allowing state-dependent parameters enables the modeling of the majority of leptokurtosis present in the data. In the GED case, the v parameter is below the threshold value of 2, indicating that the distribution exhibits thicker tails compared to the normal distribution.

The MRS-GARCH models incorporating Student's t innovations are introduced in a variant that permits the degrees of freedom to vary, indicating a time-dependent kurtosis, as discussed in Hansen [10] and Dueker [12].

Hansen proposes a model where the Student's t degrees-of-freedom parameter varies over time based on a logistic function of variables present in the information set up to time t-1. In contrast, Dueker [12] restricts this parameter to being state-dependent. This paper, following Marcucci [6], permits the degrees-of-freedom parameter to switch across regimes alongside all other parameters.

4.3 In Sample Statistics

Given the primary emphasis on predictive capability, we present only the mean squared error (MSE) for forecasting the rate of return and volatility in Table 11, omitting any formal tests. The optimal model for forecasting the rate of return of the National stock exchange, according to MSE criteria, is AR(2)-MRSGARCH with GED distribution. The second model is AR(2)-MRSGARCH with t distribution. Standard GARCH models, specifically AR(2)-GJR and AR(2)-GARCH with t distribution, demonstrate superior performance in forecasting rate of return compared to alternative models. The MSE criteria for forecasting rate of return demonstrate that GARCH models outperform ARMA models. The optimal model for forecasting volatility of National Stock Exchange, according to MSE criteria, is AR(2)-MRSGARCH with GED distribution, while the second-best model is AR(2)-MRSGARCH with t distribution. The AR(2)-GJR model with GED distribution outperforms other standard GARCH models.

Table 9: *MLE estimates of AR-(1)-MRS-GARCH models*

	MRS-GARCH-N	MRS-GARCH-t	MRS-GARCH-GED
$\boldsymbol{\delta^{(1)}}$	0.0215**	0.7012**	0.2917**
$\boldsymbol{\delta^{(2)}}$	-0.0332**	-0.06721**	-0.0093**
$\boldsymbol{\varphi_1^{(1)}}$	0.3503*	0.4843**	0.5213**
$\boldsymbol{\varphi_1^{(2)}}$	0.4320**	0.47054**	0.908**
$\boldsymbol{\alpha_0^{(1)}}$	0.0151**	0.0216**	0.0162**
$\boldsymbol{\alpha_0^{(2)}}$	0.1591**	0.0328**	0.0242**
$\boldsymbol{\alpha_1^{(1)}}$	0.3948**	0.13234**	0.6415*
$\boldsymbol{\alpha_1^{(2)}}$	0.6521**	0.710**	0.165**
$\boldsymbol{\beta_1^{(1)}}$	0.2433*	0.8316**	0.0033**
Q	0.473*	0.734**	0.568**
$\boldsymbol{v^{(1)}}$	-	3.392**	0.684**
Log (L)	-729.68410	-562.7126	-547.36478
$\boldsymbol{\pi_1}$	0.77	0.35	0.30
$\boldsymbol{\pi_2}$	0.23	0.65	0.70

4.4 Out-of-Sample Forecast Evaluation

A potential solution to the issues identified in the preceding section is to evaluate the models based on their out-of-sample forecasting performance. An out-of-sample test effectively addresses potential issues of overfitting or overparameterization, providing a robust framework for assessing the performance of competing models. Most models serve as simple approximations of the true data-generating process; thus, a good in-sample fit does not necessarily guarantee accurate or reliable forecasts. Researchers and practitioners prioritize accurate volatility forecasts over strong in-sample fits, which are often more achievable with highly parameterized models like MRS-GARCH. Table 12 presents the DM test results for the equal predictive ability of the GJR model with Student's t innovations compared to other GARCH models.

The results demonstrate that the DM test, which assesses the equal predictive ability of the AR-(2)-MRSGARCH model with GED innovations compared to other GARCH and MRSGARCH models, allows for the rejection of the null hypothesis of no difference in accuracy between the two competing forecasts at 1-, 5-, and 10-day horizon periods. The accuracy of forecasting does not differ between the AR-(2)-MRSGARCH-t and AR(2)-MRSGARCH-GED models at a 5-day period horizon. Furthermore, the accuracy of forecasting does not differ between the AR-(2)-MRSGARCH-t and AR-

(2)-MRSGARCH-N models when compared to the AR-(2)-MRSGARCH-GED model at a 10-day horizon. Consequently, according to MSE criteria, the AR-(2)-MRSGARCH-GED model demonstrates superior performance compared to alternative models at a one-day horizon.

The AR(2)-MRSGARCH-GED and AR(2)-MRSGARCH-t models demonstrate superior performance compared to other models at the 5-day horizon. Over a 10-day horizon, three AR(2)-MRSGARCH models demonstrate superior performance compared to alternative models.

Table 10: *MLE estimates of AR(2)-MRS-GARCH models*

	MRS-GARCH-N	MRS-GARCH-t	MRS-GARCH-GED
$\boldsymbol{\delta^{(1)}}$	0.0354**	0.1442**	0.3110**
$\boldsymbol{\delta^{(2)}}$	-0.0112**	-0.0031*	-0.004**
$\boldsymbol{\varphi_1^{(1)}}$	0.31274**	0.45671**	0.3743**
$\boldsymbol{\varphi_1^{(2)}}$	0.2343*	0.3334**	0.4690**
$\boldsymbol{\varphi_2^{(1)}}$	0.251**	0.3490**	0.2640**
$\boldsymbol{\varphi_2^{(2)}}$	0.1653**	0.234**	0.2784**
$\boldsymbol{\alpha_0^{(1)}}$	0.0226**	0.0267**	0.0356*
$\boldsymbol{\alpha_0^{(2)}}$	0.0371**	0.0312**	0.0056**
$\boldsymbol{\alpha_1^{(1)}}$	0.3211**	0.1715**	0.4531**
$\boldsymbol{\alpha_1^{(2)}}$	0.4212**	0.7305**	0.1872**
$\boldsymbol{\beta_1^{(1)}}$	0.2214**	0.1207**	0.0067**
$\boldsymbol{\beta_1^{(2)}}$	0.3301**	0.1564*	0.4698**
P	0.8626**	0.9547*	0.9271**
Q	0.6735**	0.8894**	0.9743**
$\boldsymbol{\nu^{(1)}}$	-	3.9142**	0.5731**
Log (L)	-668.3529	-551.6271	-534.37908
$\boldsymbol{\pi_1}$	0.70	0.20	0.26
$\boldsymbol{\pi_2}$	0.30	0.80	0.74

Table 11: *In sample forecasting criteria*

Model	N. of Par.	MSE of forecasting return	Rank	MSE of forecasting volatility	Rank
AR(1)	2	0.1834	29	-	-
MA(1)	2	0.1883	30	-	-
AR(2)	3	0.1806	26	-	-
ARMA(1,1)	3	0.1809	28	-	-
ARMA(1,2)	4	0.1808	27	-	-
GARCH-N	4	0.1797	23	0.051	15
GARCH-t	5	0.1797	23	0.058	19
GARCH-GED	5	0.1794	20	0.054	16
AR(1)-GARCH-N	5	0.1654	16	0.050	14
AR(1)-GARCH-t	6	0.1658	17	0.049	13
GJR-t	6	0.1795	21	0.056	18
GJR-GED	6	0.1794	20	0.055	17
AR(1)-GJR-N	6	0.1653	15	0.049	13
AR(2)-MRSGARCH-N	14	0.0788	5	0.034	4

The DM test conducted over a 22-day forecast horizon reveals no significant difference in

forecasting accuracy between MRSGARCH models and standard GARCH models. Table 12 presents the out-of-sample evaluation of risk management for our competing GARCH models across 1-day, 1-week, 2-week, and 1-month horizons. For each forecast horizon, five statistics are provided: the TUFF, the proportion of failures (PF), the test of correct unconditional coverage (LR_PF), which assesses whether PF is significantly greater than the nominal rate; the "LR_{ind} " which evaluates independence; and the "LR_{cc}" which examines the correct conditional coverage.

The ranking for each forecast horizon reflects the order based on the percentage PF. Models exhibiting a PF exceeding the coverage probability of 5% are deemed inadequate. The theoretical TUFF at 5% is expected to be 20. We can thus see that only for 95%VaR at one-day ahead, there are values on the theoretical ones except for AR(2)-MRSGARCH-t and AR(2)-MRSGARCHGED. At all other forecast horizons, the TUFF exceeds 20. Furthermore, if the aim is to account for 95% of future losses, numerous models appear insufficient, particularly at both the shortest and longest forecast horizons.

Table 12: *Risk management out-of-sample evaluation* 95% *VaR*

Step Model	Rank	LRPF	LRind	LRcc
GARCH-N	26	6.53 *	87.53 *	143.86 *
GARCH-t	18	2.45	98.45	65.45
GARCH-GED	21	7.53*	79.53*	147.86*
AR(1)-GARCH-N	23	1.45	83.45*	156.35*
AR(1)-GARCH-t	22	8.86 *	88.86*	163.66 *
AR(1)-GARCH-GED	17	2.10	98.10*	171.57*
AR(2)-GARCH-N	15	9.94	110.94*	34.85 *
AR(2)-GARCH-t	20	0.86	123.86*	67.64*
AR(2)-GARCH-GED	11	1.45	78.90	89.76*
GJR-N	27	8.86*	45.86*	87.45*
GJR-t	19	2.97	33.97 *	98.86*
GJR-GED	25	9.65 *	87.65*	187.97*
AR(1)-GJR-N	14	2.34	45.34	169.65
AR(1)-GJR-t	9	11.10*	120.10*	143.34
AR(1)-GJR-GED	24	2.94	93.94 *	165.67*
AR(2)-GJR-N	10	7.86*	91.86*	176.94 *
AR(2)-GJR-t	16	1.45	111.45 *	181.86*
AR(2)-GJR-GED	13	9.86*	121.86*	196.45*
MRSGARCH-N	12	12.53*	211.45	232.67 *
MRSGARCH-t	7	17.25*	231.86*	224.94*
MRSGARCH-GED	8	21.53*	243.97*	276.86 *
AR(1)-MRSGARCH-N	6	12.45 *	257.65*	287.86*
AR(1)-MRSGARCH-t	5	14.86*	262.34	301.45*
AR(1)-MRSGARCH-GED	3	22.10	261.67*	365.86*
AR(2)-MRSGARCH-N	4	15.94*	234.97	341.35*
AR(2)-MRSGARCH-t	2	13.86 *	218.84*	356.66*
AR(2)-MRSGARCH-GED	1	11.45 *	217.43*	378.57*

The final three LR tests reject nearly all models at extended horizons. At all horizons, the AR(2)-MRSGARCH-t and AR(2)-MRSGARCH-GED models are consistently rejected based on the statistical forecast evaluation criteria due to an excessively high PF. The performance of these models remains ambiguous. No model can consistently pass all tests over extended time horizons. Consequently, several models appear to yield reasonable and precise Value at Risk (VaR) estimates for a 1-day horizon, exhibiting a coverage rate that closely aligns with the nominal level. There is no universally most accurate model based on the risk management loss functions. The results

corroborate the findings of Brooks and Persand [30], who reported an absence of definitive answers for the majority of the series they investigated.

5. Conclusion

This paper compares standard GARCH models with Markov Regime-Switching GARCH models regarding their effectiveness in forecasting volatility in the National Stock Exchange (NSE). This includes AR-GARCH and ARMA-GARCH models with various residual distributions (Normal, t-student, and GED) on one hand, and AR-MSGARCH and ARMA-MSGARCH models with different residual distributions (Normal, t-student, and GED) on the other. The forecasts encompass volatility predictions for 1-day, 5-day, 10-day, and 22-day intervals. We have acquired the out-of-sample MSE statistic for the various models. We employed the DM statistic to compare the significant differences in MSE across various models. The findings demonstrate that the AR(2)-MRSGARCH-GED model surpasses all alternative models at the one-day horizon.

The AR(2)-MRSGARCH-GED and AR(2)-MRSGARCH-t models demonstrate superior performance compared to other models at the 5-day horizon. Over a 10-day horizon, three models of AR(2)-MRSGARCH with varying residual distributions demonstrate superior performance compared to alternative models. The DM test conducted over a 22-day forecast horizon reveals no significant difference in forecasting accuracy between MRSGARCH models and standard GARCH models. The out-of-sample evaluation of risk management using 95% VaR reveals ambiguity regarding which model performs optimally. Over extended timeframes, no model can consistently meet all evaluation criteria. Consequently, several models appear to yield reasonable and accurate Value at Risk (VaR) estimates for a 1-day horizon, exhibiting a coverage rate that closely aligns with the nominal level. There is no universally most accurate model based on risk management loss functions.

References

[1] Kristjanpoller, W. and Michell, K. (2018). A stock market risk forecasting model through integration of switching regime, ANFIS and GARCH techniques. Applied Soft Computing, 67:106-116.

[1] Granger, C.W.J. and Poon, S.H. (2003). Forecasting financial market volatility: A review, J. Econom. Lit. 41(2):478–539.

[2] Engle, R. (1982). Autoregression conditional heteroskedasticity with estimates of the variance of the UK inflation. Econometrica, 50:987–1007.

[3] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. J. Econometrics, 31:307–327.

[4] Engle, R. and G. Lee, A. (1993). permanent and transitory component model of stock return volatility, Discussion Paper, University of San Diego, 92–44R.

[5] Bollerslev, T., Engle, R., and Nelson, D. (1994). ARCH models, in: R. Engle, D. McFadden (Eds.), Handbook of Econometrics, North Holland Press, Amsterdam, 2959–3038.

[6] J. Marcucci, (2005). Forecasting stock market volatility with regime-switching GARCH models, Stud. Nonlinear Dyn. Econom. 9:1–53.

[7] Lamoureux, C. and Lastrapes, W. (1990). Heteroskedasticity in stock return data: Volume versus GARCH effects, J. Finance 45: 221–229.

[8] Cai, J. (1994) Markov model of unconditional variance in ARCH, J. Bus. Econom. Statist. 12: 309–316.

[9] Hamilton, J. and Susmel, R. (1994). Autoregressive conditional Heteroskedasticity and changes in regime, J. Econometrics 64:307–333.

[10] Hamilton, J.D. (1989). A new approach to the economic analysis of non-stationary time series and the business cycle, Econometrica 57:357–384.

[11] Hamilton, J.D. (1990). Analysis of time series subject to changes in regimes, J. Econometrics 45: 39–70.

[12] Dueker, M. (1997). Markov switching in GARCH processes in mean reverting stock market volatility, J. Bus. Econom. Statist. 15:26–34.

[13] Gray, S. (1996). Modeling the conditional distribution of interest rates as a regime-switching process, J. Financ. Econ. 42:27–62.

[14] Klaassen, F. (2002). Improving GARCH volatility forecasts with regime-switching GARCH, Empir. Econom. 27:363–394.

[15] Haas, M., Mittnik, and Paolella, S. (2004) A new approach to Markov-switching GARCH models, J. Financ. Econ. 2:493–530.

[16] Leon Li, M. Y. and William Lin, H.W. (2003). Examining the volatility of Taiwan stock index returns via a three-volatility-regime Markov-switching ARCH model, Rev. Quant. Financ. Account. 21 (1):123–139.

[17] Fong, W.M. (1997). Volatility persistence and switching ARCH in Japanese stock returns, Financ. Eng. Japan. Mark. 4:37–57.

[18] Kaufmann, S. and Scheicher, M. (2006). A switching ARCH model for the German DAX index, Stud. Nonlinear Dyn. Econom. 10(4).

[19] Fornari, F. and Mele, A. (1997). Sign and volatility-switching ARCH models: Theory and application to international stock markets, J. Appl. Econometrics 12:49–67.

[20] Norden, S. V. and Schaller, H. (1997). Regime switching in stock market returns, Appl. Financ. Econ. 7:177–191.

[21] Susmel, R. (2000). Switching volatility in international equity markets, Int. J. Finance Econ. 5:265–283.

[22] Bautista, C.C. (2003). Stock market volatility in the Philippines, Appl. Econ. Lett. 10(5):315–318.

[23] Leon Li, M.Y. (2007). Volatility states and international diversification of international stock markets, Appl. Econ. 39(14):1867–1876.

[24] Erb, C.B., Harvey, C.R., and Viskanta, T.E. (1997). The making of an emerging market, Emerg. Mark. Quart. 1(1):14–19.

[25] Bekaert, G., Erb, C.B., Harvey, C.R., and Viskanta, T.E. (1997). What matters for emerging equity market investments, Emerg. Mark. Quart. 1(1): 17–46.

[26] Glosten, L.R., Jagannathan, R., and Runkle, D.E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks, J. Finance 48:1779–1801.

[27] Lopez, J.A. (2001). Evaluating the predictive accuracy of volatility models, J. Forecast. 20(1):87–109.

[28] Mohamed Imath., Jayant Giri., Faruq Mohammad., Chinnasamy Ragavendran. (2025). Eco-friendly synthesis of ZnO nanoparticles fabricated using Fioria vitifolia L. and their biomedical potentials. Microbial Pathogenesis. 199:107-139.

[29] Harvey, D., Leybourne, S., and Newbold, P. (1997). Testing the equality of prediction mean squared errors, Int. J. Forecast. 13:281–291.

[30] Brooks, C. and Persand, G. (2003) The effect of asymmetries on stock index return value-at-risk estimates, J. Risk Finance 4(2):29–42.

[31] Sarma, M., Thomas, S., Shah, A. (2003). Selection of value-at-risk models, J. Forecast. 22(4): 337–358.

[32] Christoffersen, P.F. (1998). Evaluating interval forecasts, Internat. Econom. Rev. 841–862.

[33] Duraisamy, P. Alam, MS. Uddin, MA. Kuppusamy, A. Palanisamy, M. (2022). Modelling volatility of petroleum productions by using markov-switching GARCH models. International Journal of Agricultural Statistical Science, 18(1):1657-67.

[34] Poshakwale, S. and Murinde, V. (2001). Modelling volatility in East European emerging stock markets: evidence on Hungary and Poland. Applied Financial Economics, 11(4):445-456.

[35] Xiao, L. and Aydemir, A. Volatility modelling and forecasting in finance, Forecasting

volatility in the financial markets, Oxford: Elsevier Butterworth-Heinemann, (2007).

[36] Kanasro, H. A. Rohra, C. L. and Junejo, M. A. (2009). Measurement of stock market volatility through ARCH and GARCH models: a case study of Karachi stock exchange, Australian Journal of Basic and Applied Sciences, *3*(4):3123-3127.

[37] Kavitha, S., Rayalu, G. M., Pachiyappan, D., and Manigandan, P. (2023). FORECASTING OF EXTREME RISK USING MARKOVSWITCHING GARCH MODELS: EVIDENCE FROM GLOBAL ENERGY MARKETS. *Reliability: Theory & Applications*, *18*(4 (76)): 325-337.

[38] Alam, M. S., Murshed, M., Manigandan, P., Pachiyappan, D., and Abduvaxitovna, S. Z. (2023). Forecasting oil, coal, and natural gas prices in the pre-and post-COVID scenarios: Contextual evidence from India using time series forecasting tools. Resources Policy, 81, 103342.

[39] Uddin, M. A., Geetha, C., Khan, M. A., Srividhya, K., and Duraisamy, P. Exploring the impact of economic growth and sources of electricity production in India: New insights from sectoral analysis using NARDL approach. Journal of Infrastructure, Policy and Development, 8(10).

MULTI REFERENCE SKIP-LOT SAMPLING PLAN

Meby Joseph Manoj [1], Azarudheen S [2]

•

[1,2] Department of Statistics and Data Science
CHRIST (Deemed to be University),
Bengaluru, India
mebyjosephmanoj0328@gmail.com, sazaru1990@gmail.com

Abstract

Skip-lot sampling plans have become significant in modern quality control due to rising production volumes and the demand for cost-effective inspection methods that will yield high-quality outputs. When inspecting a submitted lot, a skip-lot plan is economically favourable and guarantees high quality. Thus, this approach benefits both producers and consumers. The skip-lot sampling plan generally utilizes the same sampling plan as the reference plans for both skipping and normal inspection. However, using the same plan in both phase favours either the producer or the consumer in the most essential situations. This article introduces a novel approach, the Multi Reference Skip-Lot Sampling Plan with the provision of having two different reference plans in the normal and skipping phases of the skip-lot plan. The paper explores the efficacy of this approach by deriving performance measures using a power series approach. To evaluate the proposed plan, a comparison is made with existing skip-lot sampling plans that use single sampling plans or double sampling plans as reference plans. This comparison is based on operational characteristics and average sample number values, accompanied by graphical representations. The comparative analysis demonstrates that the new plan effectively balances the satisfaction of both producers and consumers. Additionally, the study offers a strategy for selecting the plan parameters using the unity value approach, supported by a table providing unity values.

Keywords: Quality control, acceptance sampling plan, skip-lot sampling plan, operational characteristic function, average sample number

1. Introduction

Quality plays a crucial role in determining a positive brand image and reputation. Enhancing the quality is directly equal to the improvement of the level of satisfaction among customers. As industries navigate the complex production process, the significance of Statistical Quality Control (SQC) becomes concrete. SQC serves as a guiding compass to ensure consistency of quality through various statistical analyses. An important part of SQC is acceptance sampling, where quality is assessed on the basis of sample quality or statistical characteristics. Accordingly, it can be categorized into two types: attribute sampling plans, which are based on the number of defectives, and variable sampling plans, which are based on sample averages and deviations. A detailed study has been conducted by [15] about the acceptance sampling. These acceptance sampling plans are extensively used in businesses to maintain high product quality with minimal inspection costs. Various types of plans have emerged for this purpose, such as single sampling plans (SSP), double sampling plans (DSP), multiple sampling plans, chain sampling plans, etc. Numerous modifications have been developed in plans to achieve new heights in quality control. However, the quest for a plan which can maintain consistent top quality with low inspection costs in all kind of situations remains.

The basic methodology of the lot-by-lot sampling plan involves inspecting samples from individual lots and deciding whether to accept or reject the lot based on these samples. In contrast, cumulative results plans, which include continuous sampling plans, skip lot sampling plans, chain sampling plans, etc. rely on the acceptance results from the previous lots to make decisions about the current lot being inspected. The skip-lot sampling plan (SkSP-1) was designed by [8], incorporating the concepts of a continuous sampling plan of type CSP-1 into a continuous series of product lots from the same trusted source. Later, [14] introduced the SkSP-2, utilizing reference plans for lot inspection and derived performance measures for the plan through the power series approach and the Markov chain method. The introduction of skip-lot sampling plans significantly reduces inspection costs, making them economically advantageous. Then, [13] demonstrated that single sampling, double sampling, and multiple sampling plans could serve as reference plans for SkSP-2. Also, [9] highlighted that the most efficient plan for situations requiring expensive or detrimental inspections is SkSP-2 with a single sampling plan (SSP) as the reference plan, specifically one with an acceptance level of zero. Further development of the SkSP-2 using SSP with a zero-acceptance limit as the reference plan was designed by [20]. Later, [22] developed SkSP-2 with reference plan as double sampling plan (DSP) to reduce the sample size required for inspection. Introduction of an SkSP-2 with reference plan as the two-stage group sampling and constructed tables for parameter selection was done by [1] . Later on, [2] considered the Intervened Poisson distribution (IPD) in SSP as the reference plan and presented a corresponding design for SkSP-2.

A modification on the skip-lot plan using the principles of CSP-2, introducing the SkSP-3 was done by [16]. They demonstrated that SkSP-3 with a single sampling plan (SSP) as the reference plan is much more efficient than the conventional SSP. More studies about the SkSP-3 can be referred in [21], [3], [17], [11], etc. Skip-lot sampling plans have been modified over time to meet specific requirements and situations because these plans can significantly reduce the expense of sample inspection while maintaining the necessary level of quality. Such a modification is done by [12] where a modified skip-lot sampling plan known as MSkSP-1 is developed, which will inspect at least one unit is from the lot, ensuring safety against "out of control" lots. Over the time, more such modifications were made in skip-lot sampling plans and one can refer to [4], [5], [18], [19], [6], [10], etc.

Historically, skip-lot sampling plans and their modifications have utilized the same reference plan for both normal and skipping inspections. When SkSP-2 uses a single sampling plan (SSP) as the reference for both inspections, it tends to favor consumer satisfaction. The acceptance probability for low proportion defective is minimum in the case of SkSP with SSP, which leads to rejection of some lots with acceptable quality level which in turn causes loss for the producers. Conversely, using a double sampling plan (DSP) as the reference for both inspections tends to favor the producer. The acceptance rate for the lots with high proportion is also high, this will result in the acceptance of lots with rejectable quality level making a loss for the customers who receive the defective products. This paper introduces a Multi Reference Skip-Lot Sampling Plan of type MR-SkSP-2, with two distinct reference plans for normal and skipping inspections. The introduction of separate reference plans aims to reduce the risk of missing defects while optimizing the risk for both producers and consumers. The content of the paper is structured as follows: Section 2 details the design and operation of the MR-SkSP plan, including its performance measures. Section 3 outlines the method for selecting the parameters of the plan, followed by a numerical example in Section 4. A comparative study has been performed in Section 5, between the MR-SkSP plan with conventional skip-lot plans using graphs and tables to illustrate performance metrics. In conclusion, Section 6 provides key takeaways and closing remarks. Additionally, a unity value table for selecting the parameters is included.

2. Designing of the Skip Lot Sampling Plan with the Provision of Having Two Different Reference Plans

The skip-lot sampling plan of type SkSP-2 was derived from SkSP-1 by introducing a reference plan for both normal and skipping inspections. This modification allowed the plan to accommodate the preferences of both producers and consumers during lot inspections. The proposed MR-SkSP-2 plan employs different reference plans in the normal and skipping inspection. This multi-reference approach ensures a more balanced consideration of both producer and consumer interests. The operating steps for the MR-SkSP-2 are as given;

1. Inspect i consecutive lots submitted continuously for inspection without interruption (normal inspection), employing the SSP as the reference plan
2. If i consecutive lots are accepted, proceed to the skipping inspection. There, only a fraction f of the lots is inspected at random, using the DSP as the reference plan.
3. If no lot is rejected during the skipping inspection, continue with the skipping inspection. If any lot is rejected, switch back to normal inspection.
4. Examine the rejected lots and replace any defective products with intact ones.

The MR-SkSP-2 has the following parameters;

- The normal inspection reference plan SSP parameters c_1 & n_1 .
- The skipping inspection reference plan DSP parameters c_1, c_2, n_1 & n_2.
- The sampling frequency, f.
- The clearance interval, i.

Performance measures are essential for evaluating the effectiveness and efficiency of the sampling plan in ensuring product quality and demonstrating superiority over conventional plans. These measures help to assess how effectively the sampling plan accomplishes its goals and maintains the expected standards of quality. Additionally, they guarantee a balance between quality assurance and operational efficiency. A vital performance measure for evaluating the effectiveness of a sampling plan is the Operating Characteristic (OC) function and the OC curve plotted from it.

The derivation of OC function for the SkSP-2 was done by [13], using the power series approach by finding the number of lots inspected during normal inspection and skipping inspection. Accordingly, for an MR-SkSP-2, the expected number of lots during the normal inspection is:

$$U = \frac{1 - P_{\text{SSP}}^i}{P_{\text{SSP}}^i(1 - P_{\text{SSP}})} \tag{1}$$

Similarly, during the normal inspection, the expected number of failure sequences before having i successive accepted lots is given as;

$$G = \frac{1 - P_{\text{SSP}}^i}{P_{\text{SSP}}^i} \tag{2}$$

Furthermore, the expected number of lots during skipping inspection is;

$$V = \frac{1}{f(1 - P_{\text{DSP}})} \tag{3}$$

Using these values, the probability of rejection, P_r, is defined as,

$$P_r = \frac{\text{expected number of lots rejected}}{\text{expected number of lots submitted}} = \frac{G + 1}{U + V} \tag{4}$$

Therefore, the probability of acceptance, P_a, is given by,

$$P_a = 1 - P_r = 1 - \frac{G + 1}{U + V} \tag{5}$$

where the P_{SSP} is the probability of acceptance of SSP and P_{DSP} is that of DSP.

Also, the average fraction of lots inspected (AFI) during the whole plan comprises of the lots from the normal phase and a fraction f of lots from the skipping inspection, and is given as;

$$F = AFI = \frac{U + fV}{U + V} \tag{6}$$

The average sample number for the proposed plan is given by;

$$ASN = \frac{\text{ASN}_{SSP} \cdot U + f \cdot \text{ASN}_{DSP} \cdot V}{U + V} \tag{7}$$

3. Determination of Optimal Parameters for MR-SkSP-2

The selection of optimal parameters of the MR-SkSP-2 plan can be determined based on two defined points on the OC curve viz., acceptable quality level (AQL) and limiting quality level (LQL) and are denoted as p_1 and p_2 respectively. The risks associated with these points are known as producers risk (α) and consumer's risk (β). The producer wants to avoid the rejection of acceptable lots and the consumers want to halt the acceptance of lots having poor quality. The optimal designing of the sampling plan should then address both the producer's and consumer's concern. This can be done by choosing the parameters for the plan while the OC curve pass through pre-determined points, say $(p_1, 1-\alpha)$ & (p_2, β). Thus, the parameters of the MR-SkSP-2 plan can be determined when the following inequalities are satisfied;

$$P_a(p_1) \geq 1 - \alpha \tag{8}$$

$$P_a(p_2) \leq \beta \tag{9}$$

The plan parameters can be found using the unity value approach introduced by [7]. The resulting np values that satisfy the equations (8) & (9) are known as the unity values. The unity values for various combinations of f, i, c_1 & c_2 along with the $P_a(p)$ are provided in the Table 1. The operating ratio ($R = np_2/np_1$) for the values, $\alpha = 0.05$ and $\beta = 0.10$ which is used to distinguish the sampling plans are also given in the same table. An example illustrating the procedure for selecting the parameters is presented in detail in the next section.

Table 1: *Unity values & operating ratio for the MR-SkSP-2.*

f	i	c_1	c_2	P_a							$R = np_2/np_1$ $(\alpha = 0.05, \beta = 0.10)$
				0.99	0.95	0.75	0.50	0.25	0.10	0.05	
0.25	5	1	5	0.9206	1.1836	1.5691	1.9425	2.7148	3.8792	4.7250	3.2774
0.33				0.8903	1.1381	1.5209	1.9046	2.7087	3.8786	4.7249	3.4080
0.50				0.8362	1.0764	1.4547	1.8551	2.7022	3.8780	4.7249	3.7517
0.67				0.8091	1.0336	1.4092	1.8229	2.6989	3.8777	4.7248	3.6027
0.75				0.7801	1.0149	1.3908	1.8104	2.6977	3.8776	4.7248	3.8207
0.25	4	1	5	0.9793	1.2535	1.6850	2.0713	2.7690	3.8837	4.7255	3.0983
	6			0.8987	1.1250	1.4805	1.8520	2.7690	3.8822	4.7253	3.4507
	8			0.8418	1.0384	1.3473	1.7449	2.6927	3.8893	4.7261	3.7453
	10			0.7942	0.9714	1.2551	1.6979	2.6922	3.8893	4.7261	4.0037
	12			0.7547	0.9182	1.1866	1.6819	2.6922	3.8893	4.7261	4.2359
0.25	5	0	1	0.1247	0.2470	0.4933	0.7886	1.3944	2.3004	2.9942	9.3142
			2	0.2406	0.3729	0.5974	0.8433	1.3990	2.3005	2.9942	6.1685
			3	0.3653	0.5019	0.7131	0.9234	1.4094	2.3007	2.9942	4.5843
			4	0.4930	0.6334	0.8357	1.0197	1.4296	2.3011	2.9942	3.6328
			5	0.6207	0.7648	0.9605	1.1264	1.4662	2.3019	2.9942	3.0099
0.25	5	0	6	0.7551	0.8955	1.0868	1.2405	1.5232	2.3031	2.9943	2.5718
		1		1.0930	1.3464	1.7028	2.0307	2.7257	3.8801	4.7251	2.8817
		2		1.3764	1.7500	2.3338	2.9008	3.9346	5.3155	6.2826	3.0375
		3		1.6359	2.1810	3.0475	3.8518	5.1232	6.6752	7.7447	3.0606
		4		1.9571	2.7057	3.8711	4.8538	6.2898	7.9890	9.1479	2.9527

4. Numerical Example and Interpretation

Consider implementing an MR-SkSP-2 plan in an inspection process, with parameter thresholds set as $p_1 = 0.004$ and $p_2 = 0.02$. These thresholds correspond to a producer's risk (α) of 0.05 and a consumer's risk (β) of 0.10. Thus, the parameters of the plan are determined using the given steps,

Step 1 Using the given values of p_1 and p_2, calculate the operating ratio, R, as;

$$R = \frac{p_2}{p_1} = \frac{0.02}{0.004} = 5 \tag{10}$$

Step 2 From the Table 1 of unity values and operating ratio, find the R-value closest to the calculated R-value (10) and it is found as 4.5843.

Step 3 The parameter values from the table corresponding to the R-value for the plan are: f = 1/4, i = 5, $c_1 = 0$ & $c_2 = 3$.

Step 4 The np values corresponding to the obtained R-value is found as 0.5019 (np_1) and 2.3007 (np_2). From this, the sample size is calculated as largest of $(np_1)/(p_1)$ or $(np_2)/(p_2)$. Thus,

$$n = \frac{np_1}{p_1} \quad \text{or} \quad \frac{np_2}{p_2} = \frac{0.5019}{0.004} \quad \text{or} \quad \frac{2.3007}{0.02} = 125.475 \quad \text{or} \quad 115.035 \tag{11}$$

Among the n_1 and n_2, we choose the largest one, thus the required sample size is approximately 126 products. Here, it's assumed that $n_1 = n_2$, thus both SSP and DSP is done using the same sample size. Hence the required sampling plan for the given specifications is f = 1/4, i = 5, c_1 = 0, c_2 = 3, $n_1 = n_2$ = 126.

5. Comparative Analysis of the Plans

The advantages of the proposed plan from the conventional SkSP-2 plans can be determined by comparing the performance measures of the plans such as P_a, ASN, and ATI. To analyze the characteristic of the proposed plan, OC curves are constructed with parameters $(n = 100, c_1 = 2, c_2 = 5, i = 5, f = 0.25)$ and it is evaluated against the conventional plans such as SkSP with SSP as the reference plan (SkSP-SSP) with parameters $(n = 100, c_1 = 1, i = 5, f = 0.25)$ and SkSP with DSP as the reference plan (SkSP-DSP) with parameters $(n_1 = n_2 = 100, c_1 = 2, c_2 = 5, i = 5, f = 0.25)$ and presented in Figure 1. It is noted that the proposed plan depicts the general behavior of the acceptance sampling plans, i.e., the probability of acceptance P_a decreases as the proportion defective increases. It is also noted that the proposed MR-SkSP plan lies between the conventional plans, indicating that it has the combined behavior of the conventional plans together. It is clear from the curves that when the proportion defective is high, the MR-SkSP-2 plan closely resembles the behavior of the SkSP-SSP plan. This arises because both plans highlight stricter screening and inspection, resulting in similar OC curve behavior under such conditions. Conversely, at lower levels of proportion defective, the MR-SkSP-2 plan mirrors the SkSP-DSP plan's behavior. This occurs because both plans allow for more tolerant inspection strategies or reduced sampling frequency, yielding comparable OC curves for lower defective proportions. Hence, the advantage of using the proposed plan over the conventional ones is its flexibility and adaptability in balancing inspection strictness and effectiveness.

Figure 2 gives an evaluation of the Average Sample Number (ASN) curves of the above-mentioned plans. It is seen that when the proportion defective is high, the ASN curve of the MR-SkSP plan matches closely with that of the SkSP-SSP plan. This indicates that the MR-SkSP procedure results in intensive analysis, requiring large sample sizes to detect and control higher defect rates effectively. This behavior reduces the risk of accepting defective items. It is also noted that for lower levels of p, the ASN curve of the MR-SkSP plan imitates the performance of the SkSP-DSP plan, needing fewer samples. This shows how the plan is able to relax inspection effort when defect levels are low, thereby saving resources and reducing inspection costs.

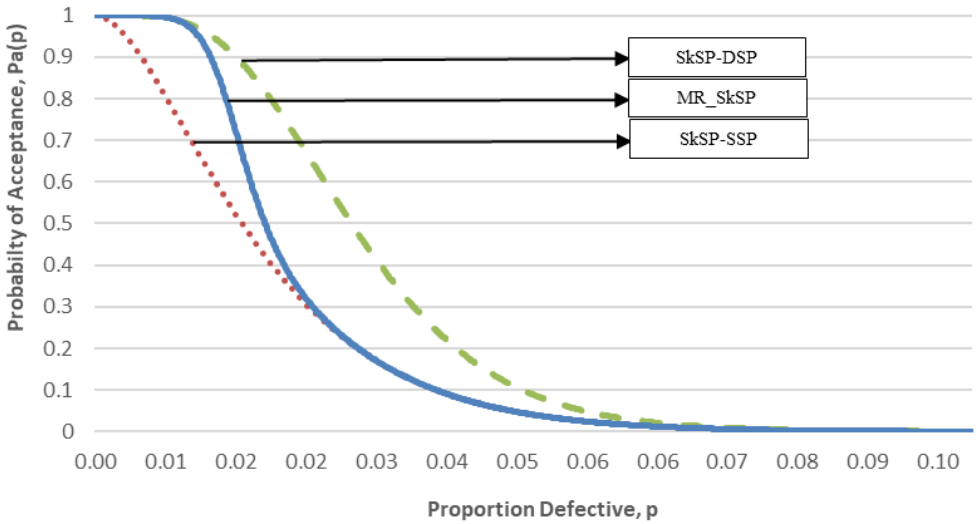


Figure 1: *Operating Characteristic (OC) Curves*

For further comparison, tables are provided with values of P_a, the ASN, and ATI for the aforementioned plans. Table 2 presents the performance measure values for varying proportions of defective, p. From the table, consider the case where the proportion defective is $p = 0.01$. The P_a value for the SkSP-SSP is 0.8395, i.e., 83.95%. This means that for the SkSP-SSP, the producer's risk for getting a good lot rejected is 16.05%. Again, when $p = 0.01$, the SkSP-DSP has an acceptance value of 99.6% ($P_a = 0.9960$). Thus, the producer's risk is reduced to 0.4%. For the

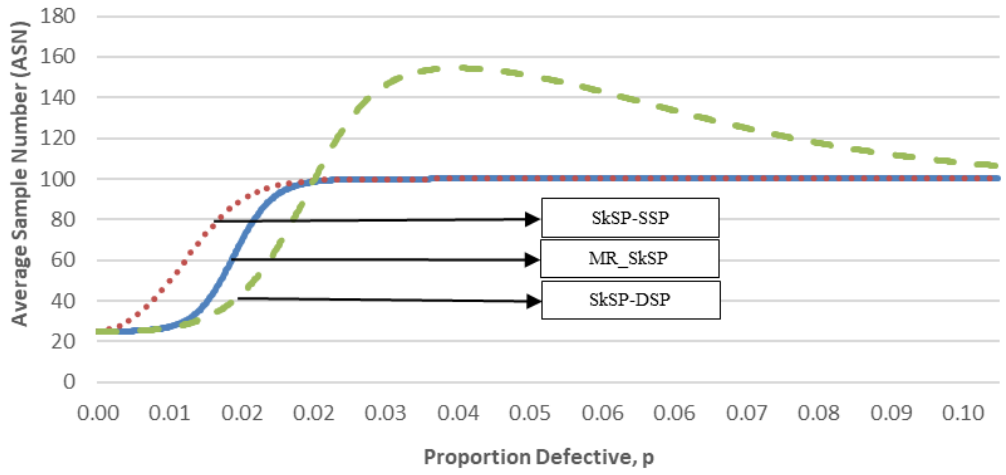


Figure 2: *Average Sample Number (ASN) Curves*

proposed plan, MR-SkSP-2, the acceptance value at $p = 0.01$ is 98.35% ($P_a = 0.983$). Similar to the SkSP-DSP plan, the MR-SkSP-2 also has a low producer's risk of 1.65% for a low proportion defective case. Meanwhile, when the proportion defective is elevated at $p = 0.05$, the SkSP-SSP has an acceptance of 4.04% ($P_a = 0.0404$). Thus, the consumer's risk is only 4.04% for accepting a lot with a high fraction defective. However, the consumer's risk for the SkSP-DSP is at 8.85% ($P_a = 0.0885$) for the same $p = 0.05$, leading to a higher acceptance of lots with defective products. Here, the proposed MR-SkSP-2 plan has a low acceptance of 4.04% ($P_a = 0.0404$), similar to the SkSP-SSP. Thus, the consumer's risk is kept at a minimum level for the plan even when the defective proportion is high. This numerical illustration states that by integrating the strengths of the two different reference plans used in normal and skipping inspections, the MR-SkSP plan ensures a moderate approach that safeguards the interests of both parties involved. A similar behavior is observed in the case of ASN and ATI as well. Table 3 provides the same performance measure values for different parameter values such as sampling frequency (f) and clearance interval (i).

Table 2: *Measures tabulated against p for SkSP-SSP, SkSP-DSP, MR-SkSP-2*

p	P_a			ASN			ATI		
	SkSP-SSP	SkSP-DSP	MR-SkSP-2	SkSP-SSP	SkSP-DSP	MR-SkSP-2	SkSP-SSP	SkSP-DSP	MR-Sk0SP-2
0.01	0.8395	0.9960	0.9835	60.7222	28.5511	30.5765	0.0084	0.0100	0.0098
0.015	0.6195	0.9745	0.7982	86.0558	39.3873	60.0287	0.0093	0.0146	0.0120
0.02	0.4250	0.9005	0.4702	96.7964	65.5656	92.0643	0.0085	0.0180	0.0094
0.025	0.2915	0.7299	0.2969	99.4162	106.6065	99.0067	0.0073	0.0182	0.0074
0.03	0.1999	0.5137	0.2004	99.9061	138.9470	99.8798	0.0060	0.0154	0.0060
0.035	0.1360	0.3422	0.1361	99.9861	151.7832	99.9852	0.0048	0.0120	0.0048
0.04	0.0916	0.2236	0.0916	99.9981	154.4508	99.9982	0.0037	0.0089	0.0037
0.045	0.0611	0.1425	0.0611	99.9997	152.9084	99.9998	0.0027	0.0064	0.0027
0.05	0.0404	0.0885	0.0404	100.0000	149.1284	100.0000	0.0020	0.0044	0.0020
0.055	0.0266	0.0539	0.0266	100.0000	144.0540	100.0000	0.0015	0.0030	0.0015
0.06	0.0174	0.0323	0.0174	100.0000	138.3711	100.0000	0.0010	0.0019	0.0010
0.07	0.0073	0.0114	0.0073	100.0000	127.1072	100.0000	0.0005	0.0008	0.0005
0.08	0.0030	0.0041	0.0030	100.0000	117.7482	100.0000	0.0002	0.0003	0.0002
0.09	0.0012	0.0015	0.0012	100.0000	110.9458	100.0000	0.0001	0.0001	0.0001
0.10	0.0005	0.0006	0.0005	100.0000	106.4317	100.0000	0.0000	0.0001	0.0000

These tables offer a comprehensive insight into the performance of each plan under different conditions, aiding in the evaluation and comparison of their effectiveness and efficiency. In scenarios with a low proportion of defects, the MR-SkSP plan behaves similarly to the SkSP-DSP,

thereby reducing the rejection of quality lots and minimizing the cost for inspection by providing a lower sample size. Conversely, in conditions with a high proportion of defects, the MR-SkSP plan matches the performance of the conventional SkSP-SSP plan, thereby mitigating the risk of accepting low-quality lots while still using a moderately small sample size. Thus, within the same plan, the MR-SkSP balances the safety of manufacturers and customers under desirable conditions by carefully balancing lot acceptance and average sample number efficiently.

Table 3: *Measures tabulated against parameter values for SkSP-SSP, SkSP-DSP, MR-SkSP-2*

f	i	P_a			ASN			ATI		
		SkSP-SSP	SkSP-DSP	MR-SkSP-2	SkSP-SSP	SkSP-DSP	MR-SkSP-2	SkSP-SSP	SkSP-DSP	MR-SkSP-2
0.25	2	0.89930	0.99616	0.99316	38.10942	27.60917	27.86001	0.00899	0.00996	0.00993
	5	0.83955	0.99603	0.98346	60.72225	28.55108	30.57653	0.00840	0.00996	0.00983
	7	0.80429	0.99595	0.97099	74.06515	29.19033	34.07228	0.00804	0.00996	0.00971
0.25	5	0.83955	0.99603	0.98346	60.72225	28.55108	30.57653	0.00840	0.00996	0.00983
0.33		0.81537	0.99475	0.97831	69.87004	37.82546	40.11235	0.00815	0.00995	0.00978
0.50		0.78263	0.99222	0.96847	82.26293	56.02399	58.29159	0.00783	0.00992	0.00968
0.67		0.76147	0.98975	0.95924	90.26842	73.77007	75.37095	0.00761	0.00990	0.00959

6. Conclusion

The present research explores, a multi-reference skip-lot sampling plan, MR-SkSP, incorporating two different reference plans for normal and skipping inspections, has been designed and evaluated. The operational procedure and the unity value approach for parameter selection for the plan have been presented. The effectiveness of the MR-SkSP plan has been judged through comparative analyses with conventional skip-lot plans employing same reference plans in both phases. The Performance measures have been scrutinized using various graphs and tables. Findings reveal that the MR-SkSP plan demonstrates remarkable adaptability and effectiveness across varying proportions of defectiveness. In circumstances with a low proportion of defects, the MR-SkSP plan offers enhanced protection to producers by providing a high acceptance probability and lower ASN. Conversely, in situations with a higher proportion of defects, the plan prioritizes consumer satisfaction by offering a lower acceptance probability and higher ASN value. Notably, the MR-SkSP plan cleverly synthesizes the properties of conventional plans at the critical phases of inspection, ensuring optimal outcomes irrespective of the proportion of defects. Thus, the study underscores the significance of the MR-SkSP plan in achieving balanced and effective quality control, catering to the diverse needs of both producers and consumers alike. As the importance for the SkSP plans to maintain high quality along with low inspection time and cost is growing day by day, the proposed MR-SkSP has a significant role by enabling the suitable adaptation of the traditional SkSPs in the serious situations. While it is clear that using two different reference plan allows the adaptation of the traditional plans with a single reference plan, the method can be applied in the other skip-lot sampling plans like, SkSP-3. SkSP-V, SkSP-T etc.

References

[1] Muhammad Aslam, Saminathan Balamurali, Muhammad Azam, and Chi-Hyuck Jun. Skip-lot sampling plan of type sksp-2 with two-stage group acceptance sampling plan as reference plan. *Communications in Statistics-Simulation and Computation*, 43(4):777–789, 2014.

[2] S Azarudheen and K Pradeepaveerakumari. Determination of skip lot sampling plan with single sampling plan as reference plan under the conditions of intervened poisson distribution. *Journal of Statistics and Management Systems*, 19(6):771–780, 2016.

[3] S Azarudheen and K Pradeepa Veerakumari. Designing of skip lot sampling plan (sksp–3) with single sampling plan as reference plan under the conditions of intervened poisson distribution. *Asia Mathematika*, 1(1):23–29, 2017.

[4] S Balamurali and Chi-Hyuck Jun. A new system of skip-lot sampling plans having a provision for reducing normal inspection. *Applied Stochastic Models in Business and Industry*, 27(3):348–363, 2011.

[5] Saminathan Balamurali, Muhammad Aslam, and Chi-Hyuck Jun. A new system of skip-lot sampling plans including resampling. *The Scientific World Journal*, 2014(1):192412, 2014.

[6] Saminathan Balamurali and Usha Mahalingam. Developing of a new system of skip-lot reinspection schemes. *Communications in Statistics-Simulation and Computation*, 46(2):1481–1491, 2017.

[7] Joseph M Cameron. Tables for constructing and for computing the operating characteristics of single sampling plans. *Industrial Quality Control*, 9(1):37–39, 1952.

[8] Harold F Dodge. Skip-lot sampling plan. *Journal of Quality Technology*, 9(3):143–145, 1977.

[9] K Govindaraju. Contributions to the study of certain special purpose plans. 1984.

[10] N Murugeswari, P Jeyadurga, and S Balamurali. Evaluation and optimal designing of a two-level skip-lot sampling plan for resubmitted lots. *Communications in Statistics-Theory and Methods*, 53(6):2043–2066, 2022.

[11] N Murugeswari, P Jeyadurga, S Sridevi, and S Balamurali. Evaluation and optimal designing of a new system of skip-lot reinspection schemes. *Journal of Statistical Computation and Simulation*, 93(9):1420–1440, 2023.

[12] Rodger D Parker and Larry Kessler. A modified skip-lot sampling plan. *Journal of Quality Technology*, 13(1):31–35, 1981.

[13] Robert L Perry. Skip-lot sampling plans. *Journal of Quality Technology*, 5(3):123–130, 1973.

[14] Robert Leonard Perry. *A system of skip-lot sampling plans for lot inspection*. Rutgers The State University of New Jersey, School of Graduate Studies, 1970.

[15] Edward G Schilling and Dean V Neubauer. *Acceptance sampling in quality control*. Chapman and Hall/CRC, 2009.

[16] V Soundararajan and R Vijayaraghavan. A new system of skip-lot inspection plans of type sksp-3. *Quality for Process and Development*, 1989.

[17] S Sridevi and S Balamurali. Evaluation and design of a new system of tightened single-level skip-lot sampling plan. *Communications in Statistics-Simulation and Computation*, pages 1–18, 2023.

[18] K Pradeepa Veerakumari and S Suganya. A new system of sksp-t plan with single sampling plan as reference plan. *Research Journal of Mathematical and Statistical Sciences*, 4(4):1–6, 2016.

[19] K Pradeepa Veerakumari and S Suganya. Characteristic evaluation of sksp-t with special type double sampling plan as reference plan. *International Journal of Productivity and Quality Management*, 28(3):360–371, 2019.

[20] R Vijayaraghavan. Construction and selection of skip-lot sampling inspection plans of type sksp-2 indexed by indifference quality level and maximum allowable per cent defective. *Journal of Applied Statistics*, 21(5):417–423, 1994.

[21] R Vijayaraghavan. Design and evaluation of skip-lot sampling plans of type sksp-3. *Journal of Applied Statistics*, 27(7):901–908, 2000.

[22] R Vijayaraghavan and V Soundararajan. Design and evaluation of skip-lot sampling inspection plans with double-sampling plan as the reference plan. *Journal of Applied Statistics*, 25(3):341–348, 1998.

RT&A, No 2 (84)
Volume 20, June 2025

MARSHALL – OLKIN TWO PARAMETER SUJATHA DISTRIBUTION AND ITS APPLICATIONS

V. Sreejith [1], Ajitha Sasi [2], B. Vineshkumar [3]

•

1. Department of Statistics, Government Arts College, Thiruvananthapuram, Kerala, India.
2. School of Mathematics & Statistics, Mahatma Gandhi University, Kottayam, Kerala, India.
3. Department of Statistics, Government Arts College, Thiruvananthapuram, Kerala, India.

Email: [1]sreejithv@gactvm.ac.in; [2] ajitha_karthi@yahoo.co.in [3]vinesh@gactvm.ac.in

Abstract

Sujatha distribution, often discussed in the context of reliability engineering and survival analysis, is a probability distribution that is useful for modeling lifetimes of objects and systems. In this paper, we introduce a three parameter distribution called the Marshall – Olkin two parameter Sujatha distribution. Marshall – Olkin generalization is a novel method of enlarging a known family of distributions. It leads to a rich class of distributions that can capture different shapes and behaviours and thus allows for better fitting to empirical data, thereby establishing its importance in reliability engineering. We derive many useful statistical properties such as the hazard rate function, reverse hazard rate function, order statistics, moments, measure of skewness and measure of kurtosis of the Marshall – Olkin two parameter Sujatha distribution. The stress – strength reliability of this distribution has also been derived. The application of this new distribution is established using real life data sets. We also compare the performance of this distribution with that of some existing distributions.

Keywords: Lifetime Distributions, Sujatha Distribution, Reliability, Marshall – Olkin Distribution.

1. Introduction

Statistical analysis and modeling of lifetime data has now become a topic of relevance because of its widespread applicability in the fields of engineering, insurance and medical science. Numerous statistical methods which includes both parametric and non-parametric models have been evolved so far. Still the scope and importance of the topic are being expanded beyond limits. Exponential, gamma, lognormal, Weibull, Lindley, Akash, Shanker, Sujatha are some probability distributions which are very widely used in modeling lifetime data. Each of these distributions has its own merits and demerits while applying to a real life situation. For example, the survival functions of lognormal and gamma distributions cannot be written in closed forms. The exponential distribution has a constant hazard rate. At the same time, the hazard rate of Lindley, Shanker, Akash and Sujatha distributions are monotonically increasing.

Sujatha distribution has been introduced by [20]. It has been originally presented as a one parameter continuous distribution with probability density function

$$f(x,\theta) = \frac{\theta^3}{\theta^2+\theta+2}(1+x+x^2)e^{-\theta x};\ x>0,\ \theta>0 \tag{1}$$

and cumulative distribution function

$$F(x,\theta) = 1-\left[1+\frac{\theta x(\theta x+\theta+2)}{\theta^2+\theta+2}\right]e^{-\theta x};\ x>0,\ \theta>0. \tag{2}$$

In fact, this distribution is a three-component mixture of an exponential distribution with scale parameter θ, a gamma distribution with shape parameter 2 and scale parameter θ, and a gamma distribution with shape parameter 3 and scale parameter θ. A two parameter Sujatha distribution (TPSD) having parameters θ and α was proposed by [21]. The probability density function $f(x)$ and cumulative distribution function $F(x)$ of the two parameter Sujatha distribution are given by

$$f(x,\theta,\alpha) = \frac{\theta^3}{\alpha\theta^2+\theta+2}(\alpha+x+x^2)e^{-\theta x};\ x>0,\ \theta>0,\ \alpha\geq 0 \tag{3}$$

and

$$F(x,\theta) = 1-\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x};\ x>0,\ \theta>0,\ \alpha\geq 0. \tag{4}$$

Here θ is a scale parameter and α is a shape parameter. It is easy to verify that the two parameter Sujatha distribution reduces to the Sujatha distribution for $\alpha = 1$.

The hazard rate function, mean residual life function, stochastic ordering, Bonferroni and Lorenz curves, stress-strength reliability and many other important properties of Sujatha distribution as well as two parameter Sujatha distribution has been discussed in detail by [20] and [21] respectively.

In this manuscript, we present the Marshall – Olkin generalization of the two parameter Sujatha distribution. A new method of expanding a known family of distributions was put forward by [17]. This method consists of considering the survival function, say $\bar{F}(x)$ of an existing distribution and constructing a new function $\bar{G}(x)$ defined as

$$\bar{G}(x) = \frac{\bar{F}(x)}{\gamma+(1-\gamma)\bar{F}(x)}, \tag{5}$$

$\bar{G}(x)$ is a proper survival function and forms a new family of survival functions. This method of extending an existing distribution usually gives rise to distributions having useful hazard functions. As an immediate consequence, they can be employed for modeling real life situations in a better way than the already known distributions. Because of this wide applicability, the Marshall - Olkin extended family of distributions had been studied in detail by many researchers. Numerous examples can be seen in [12], [13], [1], [2], [3], [4], [5], [8], [6], [11], [7], [14], [8], [15],[16],[18], [19], [9]. The Marshall – Olkin Sujatha (MOS) distribution introduced by [10] has many useful statistical properties with respect to survival rate function, hazard rate function, reversed hazard rate function, cumulative hazard rate function, stochastic ordering, Shannon and Renyi entropies, order statistics etc.

2. Marshall-Olkin two parameter Sujatha distribution

The survival function $\bar{G}(x,\ \theta,\ \alpha,\ \gamma)$ of the Marshall-Olkin two parameter Sujatha distribution (MOTPSD) computed using (5) is given by

$$\bar{G}(x,\ \theta,\ \alpha,\ \gamma) = \frac{\bar{F}(x,\ \theta,\ \alpha)}{\gamma+(1-\gamma)\bar{F}(x,\ \theta,\ \alpha)} = \frac{\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}}{\gamma+(1-\gamma)\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}}.$$

The cumulative distribution function of MOTPSD is given by

$$\begin{aligned} G(x,\ \theta,\ \alpha,\ \gamma) &= 1-\bar{G}(x,\ \theta,\ \alpha,\ \gamma) \\ &= \frac{\gamma\left\{1-\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}\right\}}{\gamma+(1-\gamma)\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}}. \end{aligned} \tag{6}$$

The probability density function of $MOTPSD$ $g(x,\ \theta,\ \alpha,\ \gamma)$ is

$$g(x,\ \theta,\ \alpha,\ \gamma) = \frac{\gamma\theta^3(\alpha + x^2 + x)e^{-\theta x}}{(\alpha\theta^2 + \theta + 2)\left\{\gamma + (1-\gamma)\left[1 + \frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}\right\}^2}; x > 0,\ \alpha \geq 0,\ 0 < \gamma < 1. \tag{7}$$

The plots of probability density function for different values of the parameters are given in figure 1.

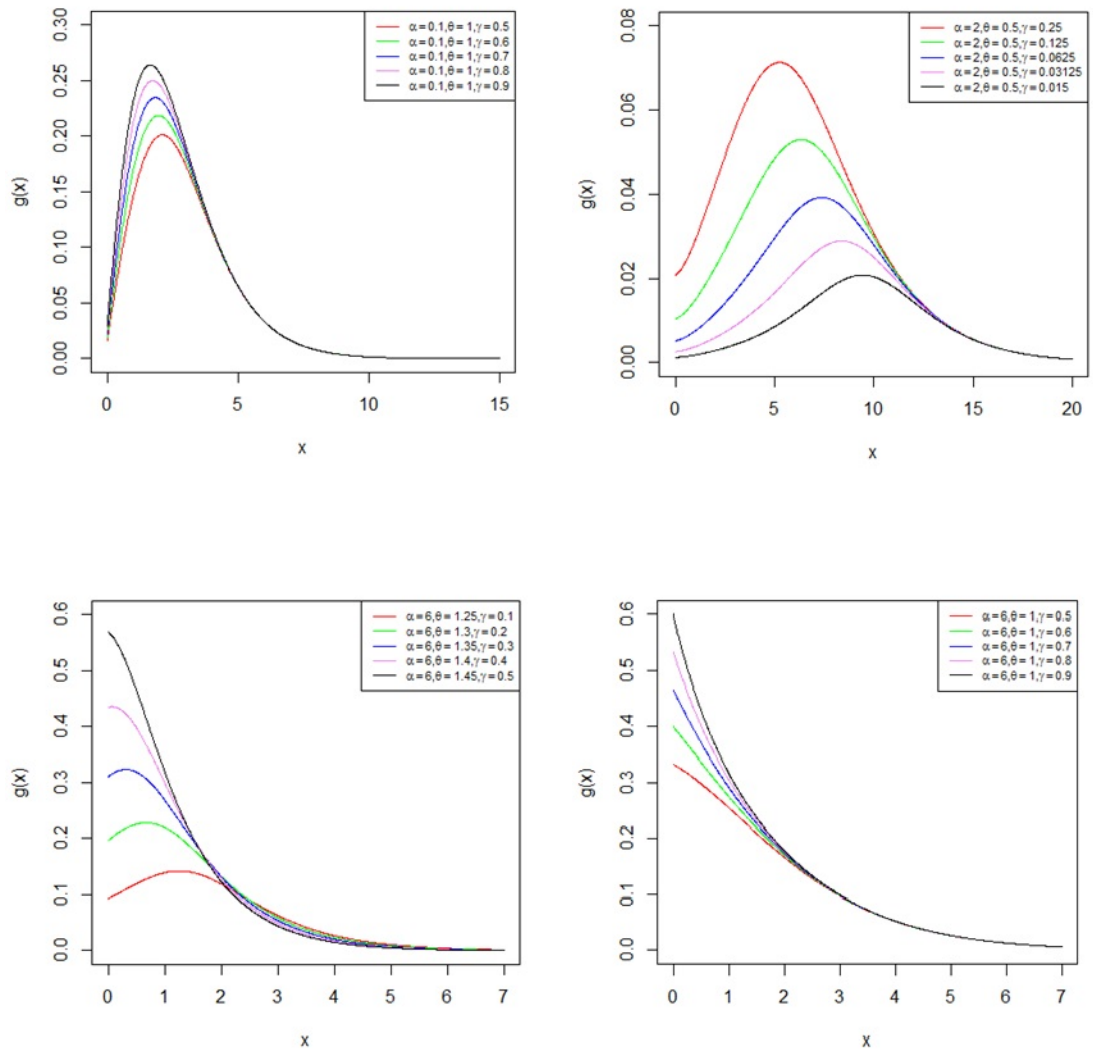


Figure 1: *Plots of probability density functions of MOTPSD for various parameter values.*

The hazard rate function is

$$h(x,\ \theta,\ \alpha,\ \gamma) = \frac{g(x,\ \theta,\ \alpha,\ \gamma)}{\bar{G}(x,\ \theta,\ \alpha,\ \gamma)} = \frac{\gamma\theta^3(\alpha + x^2 + x)e^{-\theta x}}{(\alpha\theta^2 + \theta + 2)\left[1 + \frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]} \times \frac{1}{\left\{\gamma + (1-\gamma)\left[1 + \frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}\right\}}. \tag{8}$$

The plot of the hazard rate function is given in figure 2.

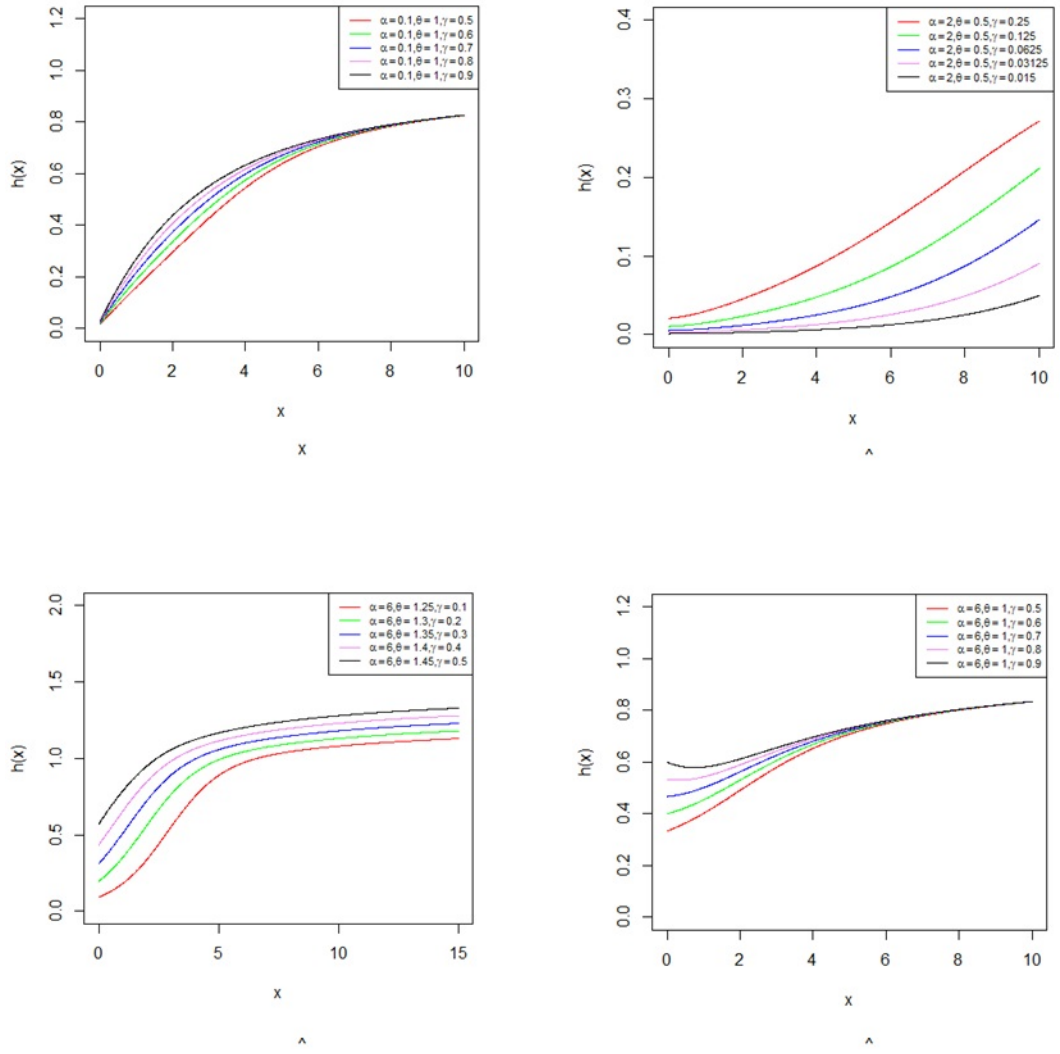


Figure 2: *Plots of hazard rate function of MOTPSD for various parameter values.*

The reverse hazard function is

$$h_r(x,\ \theta,\ \alpha,\ \gamma) = \frac{g(x,\ \theta,\ \alpha,\ \gamma)}{G(x,\ \theta,\ \alpha,\ \gamma)}$$

$$= \frac{\theta^3(\alpha + x^2 + x)e^{-\theta x}}{(\alpha\theta^2 + \theta + 2)\left\{1 - \left[1 + \frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}\right\}\left\{\gamma + (1-\gamma)\left[1 + \frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}\right\}}. \tag{9}$$

3. Order Statistics

Let $X_{(1)}$, $X_{(2)}$, ..., $X_{(n)}$ be the order statistics based on a random sample X_1, X_2, ..., X_n of size n from the population having probability density function given by (7). Then the probability density function of the r^{th} order statistic is given by

$$\begin{aligned} g_{x_{(r)}}(x,\,\theta,\,\alpha,\,\gamma) &= \frac{n!}{(r-1)!(n-r)!} g(x,\,\theta,\,\alpha,\,\gamma)[G(x,\,\theta,\,\alpha,\,\gamma)]^{r-1}[1-G(x,\,\theta,\,\alpha,\,\gamma)]^{n-r} \\ &= \frac{n!\gamma^r\theta^3(\alpha+x^2+x)e^{-\theta x}\left\{\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}\right\}^{n-r}}{(r-1)!(n-r)!(\alpha\theta^2+\theta+2)} \\ &\times \frac{\left\{1-\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}\right\}^{r-1}}{\left\{\gamma+(1-\gamma)\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}\right\}^{n+1}};\ x>0,\ \alpha\geq 0,\ 0<\gamma<1. \quad (10) \end{aligned}$$

Now, the probability density function of the first order statistic $X_{(1)}$ and the n^{th} order statistic $X_{(n)}$ are

$$\begin{aligned} g_{x_{(1)}}(x,\,\theta,\,\alpha,\,\gamma) &= \frac{n\gamma\theta^3(\alpha+x^2+x)e^{-n\theta x}\left\{\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]\right\}^{n-1}}{(\alpha\theta^2+\theta+2)\left\{\gamma+(1-\gamma)\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}\right\}^{n+1}}; \\ &\quad x>0,\ \alpha\geq 0,\ 0<\gamma<1. \quad (11) \end{aligned}$$

and

$$\begin{aligned} g_{x_{(n)}}(x,\,\theta,\,\alpha,\,\gamma) &= \frac{n\gamma^n\theta^3(\alpha+x^2+x)e^{-\theta x}\left\{1-\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]\right\}^{n-1}}{(\alpha\theta^2+\theta+2)\left\{\gamma+(1-\gamma)\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}\right\}^{n+1}}; \\ &\quad x>0,\ \alpha\geq 0,\ 0<\gamma<1. \quad (12) \end{aligned}$$

4. Moments and related measures

We have

$$\begin{aligned} \{\gamma+(1-\gamma)\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}\Big\}^{-2} &= \gamma^{-2}\left\{1+\left(\frac{1-\gamma}{\gamma}\right)\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x}\right\} \\ &= \gamma^{-2}\sum_{k=0}^{\infty}(-1)^k\binom{k+1}{k}\left(\frac{1-\gamma}{\gamma}\right)^k\left[1+\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]^k e^{-\theta kx} \\ &= \gamma^{-2}\sum_{k=0}^{\infty}\sum_{i=0}^{k}(-1)^k\left(\frac{1-\gamma}{\gamma}\right)^k\binom{k+1}{k}\binom{k}{i}\left[\frac{\theta x(\theta x+\theta+2)}{\alpha\theta^2+\theta+2}\right]^i e^{-\theta kx} \end{aligned}$$

$$\begin{aligned} &= \gamma^{-2}\sum_{k=0}^{\infty}\sum_{i=0}^{k}(-1)^k\left(\frac{1-\gamma}{\gamma}\right)^k\binom{k+1}{k}\binom{k}{i}\left[\frac{\theta x}{\alpha\theta^2+\theta+2}\right]^i \\ &\times\ e^{-\theta kx}\sum_{j=0}^{i}\binom{i}{j}(\theta+2)^{i-j}(\theta x)^j e^{-\theta kx} \\ &= \gamma^{-2}\sum_{k=0}^{\infty}\sum_{i=0}^{k}\sum_{j=0}^{i}(-1)^k\left(\frac{1-\gamma}{\gamma}\right)^k\binom{k+1}{k}\binom{k}{i}\binom{i}{j}(\theta)^{i+j} \\ &\times\ (\alpha\theta^2+\theta+2)^{-i}(\theta+2)^{i-j}x^{i+j}e^{-\theta kx}. \end{aligned}$$

Therefore

$$g(x,\ \theta,\ \alpha\ ,\gamma) = \frac{\theta^3}{\gamma(\alpha\theta^2+\theta+2)}\sum_{k=0}^{\infty}\sum_{i=0}^{k}\sum_{j=0}^{i}P_{kij}x^{i+j}e^{-\theta(k+1)x}(\alpha+x^2+x);$$
$$x>0,\ \alpha\geq 0;\ \theta>0,\ 0<\gamma<1, \tag{13}$$

where

$$P_{kij}=(-1)^k\left(\frac{1-\gamma}{\gamma}\right)^k\binom{k+1}{k}\binom{k}{i}\binom{i}{j}(\theta)^{i+j}(\alpha\theta^2+\theta+2)^{-i}(\theta+2)^{i-j}.$$

Thus, the rth moment about the origin is given by

$$\begin{aligned}\mu_r &= \frac{\theta^3}{\gamma(\alpha\theta^2+\theta+2)}\int_0^{\infty}\sum_{k=0}^{\infty}\sum_{i=0}^{k}\sum_{j=0}^{i}P_{kij}x^{i+j+r}e^{-\theta(k+1)x}(\alpha+x+x^2)\\ &= \sum_{k=0}^{\infty}\sum_{i=0}^{k}\sum_{j=0}^{i}P_{kij}\frac{\theta^3}{\gamma(\alpha\theta^2+\theta+2)}\\ &\times \left\{\frac{\alpha\Gamma(i+j+r+1)}{[\theta(k+1)]^{i+j+r+1}}+\frac{\Gamma(i+j+r+2)}{[\theta(k+1)]^{i+j+r+2}}+\frac{\Gamma(i+j+r+3)}{[\theta(k+1)]^{i+j+r+3}}\right\}.\end{aligned} \tag{14}$$

To study the behaviour of the distribution regarding dispersion, skewness and kurtosis, the coefficient of variation (CV), the moment measure of skewness ($\sqrt{\beta_1}$) and the moment measure of kurtosis (β_2) have been computed for different values of the parameters and are shown in Table 1 below.

Table 1: *CV, $\sqrt{\beta_1}$ and β_2 of MOTPSD for different values of α, θ, γ.*

Parameters	μ_1'	μ_2'	μ_3'	μ_4'	CV	$\sqrt{\beta_1}$	β_2
$\alpha=1$, $\theta=2$, $\gamma=0.5$	1.188	2.1881	5.2428	15.2408	1.1653	4.7739	1.3580
$\alpha=1.5$, $\theta=1$, $\gamma=0.95$	2.1574	7.5865	35.4228	203.672	1.2755	5.2179	1.6270
$\alpha=2$, $\theta=1.8$, $\gamma=0.90$	0.9143	1.5475	3.6860	11.1401	0.922605	1.6160	6.5671
$\alpha=6$, $\theta=1$, $\gamma=0.75$	1.7864	5.7403	25.5035	142.184	0.893743	1.509087	6.049135

5. Estimation of Parameters

In this section, we estimate the parameters of Marshall – Olkin two parameter Sujatha distribution using the method of maximum likelihood estimation. Let $X_1,\ X_2,\ \ldots,\ X_n$ be a random sample of size n from the population having probability density function given by (7)

Then the likelihood function can be written as

$$L(\underline{x};\ \theta,\ \alpha,\ \gamma)=\left[\frac{\gamma\theta^3}{(\alpha\theta^2+\theta+2)}\right]^n e^{-\theta\sum_{1=1}^{n}x_i}\prod_{i=1}^{n}\left\{\frac{\alpha+x_i^2+x_i}{\left\{\gamma+(1-\gamma)\left[1+\frac{\theta x_i(\theta x_i+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x_i}\right\}^2}\right\}. \tag{15}$$

The log likelihood function is

$$\ln L(\underline{x};\ \theta,\ \alpha,\ \gamma) = n\ln(\gamma\theta^3) - n\ln(\alpha\theta^2+\theta+2) - \theta\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\ln(\alpha+x_i^2+x_i)$$
$$-2\sum_{1=1}^{n}\ln\left\{\gamma+(1-\gamma)\left[1+\frac{\theta x_i(\theta x_i+\theta+2)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x_i}\right\}. \qquad (16)$$

The log likelihood function does not have an explicit solution. Hence the likelihood equations are to be solved numerically with the help of computer software.

6. Stress Strength Reliability

Let X and Y be two independent random variables denoting the strength and stress of a component. If we assume that X and Y have the MOTPS distribution with parameters $(\theta_1, \alpha_1, \gamma_1)$ and $(\theta_2,\ \alpha_2,\ \gamma_2)$ respectively, then the stress – strength reliability (R) can be obtained as

$$\begin{aligned} R = P(Y<X) &= \int_0^\infty g(x,\ \theta_1,\alpha_1,\gamma_1)G(x,\theta_2,\ \alpha_2,\ \gamma_2)dx \\ &= \int_0^\infty \left[\frac{\gamma_2\left[1-\left(1+\frac{\theta_2 x(\theta_2 x+\theta_2+2)}{\alpha_2\theta_2^2+\theta_2+2}\right)e^{-\theta_2 x}\right]}{\gamma_2+(1-\gamma_2)\left[1+\frac{\theta_2 x(\theta_2 x+\theta_2+2)}{\alpha_2\theta_2^2+\theta_2+2}\right]e^{-\theta_2 x}}\right] \\ &\times \left[\frac{\gamma_1\theta_2^2(\alpha_1+x^2+x)e^{-\theta_1 x}}{(\alpha_1\theta_1^2+\theta_1+2)\left[\gamma_1+(1-\gamma_1)\left[1+\frac{\theta_1 x(\theta_1 x+\theta_1+2)}{\alpha_1\theta_1^2+\theta_1+2}\right]e^{-\theta_1 x}\right]}\right]dx. \end{aligned}$$

Since this integral is difficult to compute, we employ some numerical methods for computation for some set of values of the parameters and the values of R so obtained is shown in table 2.

Table 2: *Values of Stress – Strength Reliability(R).*

Sl.No	θ_1	α_1	γ_1	θ_2	α_2	γ_2	**R**
1.	0.5	2	0.25	0.5	2	0.125	0.386294
2.	1	5	0.5	1.25	5	0.6	0.625242
3.	1	6	0.5	1.25	5	0.5	0.583306
4.	1	6	0.5	1.25	5	0.4	0.549194
5.	1.25	6	0.1	1.3	6	0.2	0.635063

7. Application

In this section, we appl a real lifetime data set in the Marshall – Olkin two parameter Sujatha distribution to determine its goodness of fit and then comparison has been developed in order to show that the Marshall – Olkin two parameter Sujatha distribution provides a better fit over Sujatha distribution, two parameter Sujatha distribution and another two parameter Sujatha distribution. The real data set given in table 3 represents the remission time (in months) of 50 breast cancer women subjected to treatment using trastzuzumab as medication reported by cancer registry department, university of Benin teaching hospital, Benin, Edo state.

Table 3: *Data regarding the remission time (in months) of 50 breast cancer women reported by university cancer registry department of Benin teaching hospital.*

50	74	35	39	21	37	27	35	30	35
26	28	34	34	26	41	61	33	33	26
25	41	35	34	34	33	60	61	42	30
80	31	24	49	26	31	28	41	37	41
61	33	26	34	50	73	45	80	39	21

To estimate the unknown parameters and for model comparison, R software is used. In order to compare the performance of MOTPSD with Length Biased Another two parameter Sujatha distribution, Another two parameter Sujatha distribution, two parameter Sujatha distribution and Sujatha distributions, we consider the criteria such as Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC) and Akaike Information Criterion Corrected (AICC). The distribution which shows lesser values of AIC, BIC, AICC is better. For computing the criteria values such as AIC, BIC, AICC and $-2\log L$ the following formulae are used:

$$AIC = 2k - 2\ln L,\ BIC = k\ln(n) - 2\ln L \text{ and } AICC = AIC + \frac{2k(k+1)}{n-k-1},$$

where n is the sample size, k is the number of parameters in the statistical model.

Table 4: *Model Comparison.*

Distribution	**MLE**	**AIC**	**BIC**	**AICC**
Sujatha Distribution	$\hat{\theta} = 0.074719546$	422.4818	424.3938	422.5651
Two parameter	$\hat{\alpha} = 0.001$			
Sujatha Distribution	$\hat{\theta} = 0.074847071$	424.2968	428.1208	424.5521
Another Two parameter	$\hat{\alpha} = 2.189802$			
Sujatha Distribution	$\hat{\theta} = 7.7485049$	424.2974	428.1215	424.5527
Marshall – Olkin	$\hat{\alpha} = 3.165$			
Two parameter	$\hat{\theta} = 0.154855$	416.0321	421.7682	416.5538
Sujatha Distribution	$\hat{\gamma} = 0.1$			

8. Simulation

Next we assess the performance of the maximum likelihood estimators of the parameters of $MOTPSD(\alpha,\ \theta\ \gamma)$. We simulated a sample of size $n = 20, 50, 100, 500, 1000$ and $10,000$ based on the following sets of parameters.

- $\alpha = 2,\ \theta = 1,\ \gamma = 0.5$
- $\alpha = 3,\ \theta = 0.15,\ \gamma = 0.1$
- $\alpha = 5,\ \theta = 3,\ \gamma = 0.75$

Then we computed the maximum likelihood estimates of these parameters and a comparison of the average bias and mean square of all the parameters have been carried out. The results obtained are summarized in table 5 shown below.

9. Conclusion

A new lifetime distribution known as Marshall – Olkin two parameter Sujatha distribution has been introduced. Its statistical properties such as hazard rate function, reverse hazard rate function, distribution of order statistics, skewness and kurtosis have been investigated. The method of maximum likelihood for estimation of parameters has been discussed. A comparison of this distribution with Sujatha, Two Parameter Sujatha and Another two parameter Sujatha

Table 5: *Average Bias and Mean Squared Errors of MLEs of* $MOTPSD(\alpha,\ \theta,\ ,\gamma)$.

Parameter		$n=20$	$n=50$	$n=100$	$n=500$	$n=1000$	$n=10000$
$\alpha=2$	Avg.	1.970704	1.43478	1.15029	0.45921	0.36780	0.15733
	Bias						
	MSE	9.43113	4.89108	2.40008	1.79432	0.93313	0.39167
$\theta=1$	Avg.	0.06633	0.04608	0.00139	0.00134	-0.00106	-0.00101
	Bias						
	MSE	0.06016	0.03186	0.10022	0.00213	0.00204	0.00022
$\gamma=0.5$	Avg.	-0.06676	-0.06401	-0.04781	0.00558	-0.00029	-0.00025
	Bias						
	MSE	0.17423	0.06407	0.02701	0.02857	0.01788	0.00500
$\alpha=3$	Avg.	113.28788	111.71993	72.66700	41.30937	7.46501	0.00709
	Bias						
	MSE	17444.57814	14692.92071	8499.1245	4953.06184	750.05078	1.95245
$\theta=0.15$	Avg.	0.00866	0.00464	0.00342	0.00218	0.00021	-0.00020
	Bias						
	MSE	0.00071	0.00020	0.00015	4.39864×10^{-5}	1.75124×10^{-5}	1.563×10^{-5}
$\gamma=0.1$	Avg.	-0.04827	0.39261	-0.03555	-0.002109	-0.00339	0.00102
	Bias						
	MSE	0.00377	3.97591	0.00272	0.00166	0.000039	1.524×10^{-5}
$\alpha=5$	Avg.	2.45105	1.27315	0.87924	0.14098	0.04309	0.03736
	Bias						
	MSE	35.76077	15.92763	9.074422	8.41508	8.14613	3.27945
$\theta=3$	Avg.	-0.10186	0.07054	0.03867	-0.01533	0.00552	-0.00030
	Bias						
	MSE	1.89987	0.40468	0.05986	0.05369	0.05259	0.00678
$\gamma=0.75$	Avg.	0.48633	0.06119	0.01544	0.00262	5.02×10^{-5}	4.255×10^{-5}
	Bias						
	MSE	1.11601	0.17848	0.03189	0.02028	0.00584	0.00056

distributions is also given with respect to a real lifetime data set. A simulation study has been done to check the consistency of estimators for different sets of values of parameters.

References

[1] Alice, T. and Jose, K. K. (2002). Multivariate Minification Processes. *STARS International Journal*, 3: 1-9.

[2] Alice, T. and Jose, K. K. (2003). Marshall-Olkin Pareto Processes. *Far East Journal of Theoretical Statistics*, 9(2): 117-132.

[3] Alice, T. and Jose, K. K. (2004 a). Bivariate semi-Pareto Minification Processes. *Metrica*, 59: 305-313.

[4] Alice, T. and Jose, K. K. (2004 b). Marshall-Olkin Pareto distributions and its Reliability Applications. *IAPQR Trans.*, 29(1): 1-9.

[5] Alice, T. and Jose, K. K. (2004 c). Marshall-Olkin Exponential Time series modelling. *STARS International Journal*, 5(1): 12-22.

[6] Ghitany, M. E, Al-Hussaini, E. K. and Al-Jarallah (2005). Marshall-Olkin Extended Weibull Distribution and its Application to Censored Data. *Journal of Applied Statistics*, 32(10): 1025-1034.

[7] Ghitany, M. E., Kotz, S. (2007). Reliability properties of Extended Linear Failure-rate Distributions. *Probability in the Engineering and informational Sciences*, 21: 441-450.

[8] Gupta, R. C., Ghitany, M. E., Al-Mutairi (2010). Estimation of reliability from Marshall-Olkin Extended Lomax Distribution. *Journal of Statistical Computation and Simulation*, 80(8): 937-947.

[9] Haq, M. A., Usman, R. M., Hashmi, S., Al-Omeri, A. I. (2019). The Marshall-Olkin length-biased exponential distribution and its applications. *Journal of King Saud University – Science*, 31(2): 246-251.

[10] Ikechukwu, A. F. and Eghwerido, J. T. (2022). Marshall – Olkin Sujatha Distribution and its Applications. *Thailand Statistician*, 20(1): 36-52.

[11] Jayakumar, K. and Thomas Mathew (2006). On a generalisation to Marshall-Olkin Scheme and its Application to Burr type XII Distribution. *Statistical Papers*, preprint.

[12] Jose, K. K., Alice, T. (2001). Marshall-Olkin Generalised Weibull distributions and applications. *STARS International Journal*, 2(1): 1-8.

[13] Jose, K. K., Alice, T. (2005). Marshall-Olkin Family of Distrbutions, Applications in Time series Modelling and Reliability. *J.C Publications*, Palakkad.

[14] Jose, K. K., Uma, P. (2009). On Marshall-Olkin Mittag-Leffler Distributions and Processes. *Far East Journal of Theoretical Statistics*, 28: 189-199.

[15] Jose, K. K., Naik, S. R., Ristic, M. M. (2010). Marshall – Olkin - Weibull Distribution and Max-Min Process. *Statistical Papers*, 51: 837-851.

[16] Jose, K. K. (2011). Marshall-Olkin Family of Distributions and Their Applications in Reliability Theory, Time Series Modeling and Stress Strength Analysis. *Proc. ISI* 58^{th} *World Statistical Congress, Int Stat Inst*, 21st-26th August 2011, Dublin, 3918-3923.

[17] Marshall, A. W. and Olkin, I. (1997). A New Method of Adding a Parameter to a Family of Distributions with Application to the Exponential and Weibull Families. *Biometrika*, 84: 641-652.

[18] Pogàny, T. K., Saboor, A., Provost, S. (2015). The Marshall-Olkin exponential Weibull distribution. *Hacettepe Journal of Mathematics and Statistics*, 44(2015): 1579-1594.

[19] Ristiċ, M. M., Kundu, D. (2015). Marshall-Olkin Generalized Exponential Distribution. *Metron*, 73(3): 317-333.

[20] Shanker, R. (2016). Sujatha Distribution and its Applications. *Statistics in Transition New Series*, 17(3): 391-410.

[21] Tesfay, M. and Shanker, R. (2018a). A two parameter Sujatha distribution. *Biometrics & Biostatistics International Journal*, 7(3): 188-197.

A NEW GENERALIZATION OF PARETO DISTRIBUTION:PROPERTIES AND APPLICATIONS

TABASUM AHAD[1], S.P.AHMAD[2]

•

[1,2] Department of Statistics,University of Kashmir, Srinagar, India
tabasumjan1234@gmail.com, sprvz@yahoo.com

Abstract

This manuscript introduces a new extension of the Pareto distribution using the SMP transformation technique, known as SMP Pareto distribution (SMPP). The SMP technique is named after (Shamshad, Murtiza, Parvaiz) who pioneered this approach to enhance the flexibility and applicability of statistical models. This new distribution provides a better fit for data than many existing models. The shapes of the density function exhibit great flexibility. It can support different hazard shapes, such as increasing, decreasing and constant shapes. Various statistical properties of the proposed distribution such as moments, quantile function, entropy and order statistics were presented. The maximum likelihood technique is used to estimate the model parameters. An extensive simulation study is carried out to illustrate the behavior of MLEs. In addition, two real-world datasets are analyzed to showcase the applicability of the proposed approach. The results indicate that the SMP Pareto distribution (SMPP) is more flexible and offers a superior fit for describing data compared to several existing forms of the Pareto distribution.

Keywords: SMP technique, Pareto distribution, Order statistics, Maximum likelihhood estimation.

1. Introduction

The field of probability distributions is fundamental in the area of statistics. A probability distribution is a mathematical function that tells us the likelihood of different outcomes in an experiment. Many distributions have been studied extensively in the literature for analyzing real-life data. These models play a crucial role across various fields, including physics, medicine, business management, engineering, and food science. In addressing complex real-world problems where data doesn't adhere to classical or standard models, developing new generalized probability models that offer greater flexibility for accurately modeling such data becomes essential. Over the years, statisticians have focused extensively on improving classical distributions by introducing additional parameters through generators or by combining existing distributions, making them more adaptable for analyzing empirical data. A methodology for introducing a new parameter to existing distributions was developed by [16], [5] introduced the T-X family of continuous distributions, where the PDF of the Beta distribution was replaced with the PDF of any continuous random variable. Alpha power transformation technique was proposed by [15]. The Rayleigh distribution was extended using the sine-G family of distributions by [17].

More recently,[20] proposed a novel transformation for generating probability distributions by introducing an extra parameter to continuous distributions, known as the SMP transformation technique. The idea was introduced to enhance the flexibility of the baseline distribution. The inverted exponential distribution was extended using the SMP transformation technique by [2], while the Kumaraswamy distribution was extended by [12] using the same transformation.

The Cumulative Distribution Function (CDF) of the SMP method is given as:

$$G_{SMP}(x) = \begin{cases} \frac{e^{\log(\alpha)\overline{F}(x)}-\alpha}{1-\alpha}, & \alpha \neq 1, \alpha > 0 \\ F(x), & \alpha = 1 \end{cases} \tag{1}$$

where $\overline{F}(x) = 1 - F(x)$

and the corresponding Probability Density Function (PDF) is obtained as follows:

$$g_{SMP}(x) = \begin{cases} \frac{e^{\log(\alpha)\overline{F}(x)}\log(\alpha)f(x)}{\alpha-1}, & \alpha \neq 1, \alpha > 0 \\ f(x) \quad , & \alpha = 1 \end{cases} \tag{2}$$

The Pareto distribution is named after the Italian economist Vilferdo Pareto (1848-1923) a well-known probability model for modeling and forecasting many socioeconomic elements. The Pareto distribution is often used to fit heavy-tailed phenomena. It has many applications in actuarial science, survival analysis, economics, reliability analysis and engineering.

The Probability Density Fuction (PDF) given by [11] is defined as

$$f(x) = \frac{\beta}{x^{\beta+1}}; \quad x \geq 1, \beta > 0, \tag{3}$$

and corresponding Cumulative Distribution Function (CDF) as follows:

$$F(x) = 1 - \frac{1}{x^{\beta}}; \quad x \geq 1, \beta > 0. \tag{4}$$

Where β is a scale parameter and is greater than 0.

In the literature, various extensions of Pareto distributions have been proposed. Alpha power pareto distribution was studied by [11]. Cubic transmuted pareto distribution was obtained by [19]. The beta pareto distribution was proposed by [4]. On a new generalization of pareto distribution was studied by [3]. The Alpha power transformation generalized pareto distribution were discussed by [6]. Sine exponential pareto distribution was studied by [1]. Topp leone exponentiated pareto distribution was obtained by [7].

In this study, we create and investigate a novel extension of Pareto distribution using the SMP transformation technique. The proposed distribution is named as SMP Pareto (SMPP) distribution. The primary motivation of developing SMPP distribution are described via:

- The proposed model is highly efficient and adaptable for incorporating a new parameter to generalize the existing distribution.
- The proposed model offers better fits compared to other competitive models.
- The proposed model can be used to model diverse nature of datasets.
- The proposed model offers more flexible shapes in terms of hazard and density plots.

The rest of this paper is organised as follows: Section 2 defines the PDF and CDF of the proposed model, i.e., SMPP Distribution. Section 3 discusses the reliability measures of the model. In Section 4, we derive various mathematical properties. Section 5 investigates order statistics. The maximum likelihood estimation of the model parameters is addressed in Section 6. Section 7 presents the simulation study, while Section 8 discusses the applicability of the model. Finally, concluding remarks are provided in Section 9.

2. SMPP Distribution

A random variable X is said to follow SMPP distribution with scale parameter $\beta > 0$ and shape parameter $\alpha > 0$, if its CDF is given as

$$G_{SMPP}(x;\alpha,\beta) = \begin{cases} \frac{e^{\frac{\log\alpha}{x^\beta}}-\alpha}{1-\alpha}, & \alpha \neq 1, \alpha > 0 \\ 1-\frac{1}{x^\beta} & , \quad \alpha = 1 \end{cases} \tag{5}$$

The corresponding PDF of SMPP distribution is given as

$$g_{SMPP}(x;\alpha,\beta) = \begin{cases} \frac{\log(\alpha)}{\alpha-1}\cdot\frac{\beta}{x^{\beta+1}}\cdot e^{\frac{\log(\alpha)}{x^\beta}}, & \alpha \neq 1, \alpha > 0 \\ \frac{\beta}{x^{\beta+1}} & , \quad \alpha = 1 \end{cases} \tag{6}$$

Figure (1) displays the PDF plot of the SMPP distribution for different parameter values. It demonstrates that the proposed distribution is unimodal, decreasing, and right-skewed in nature.

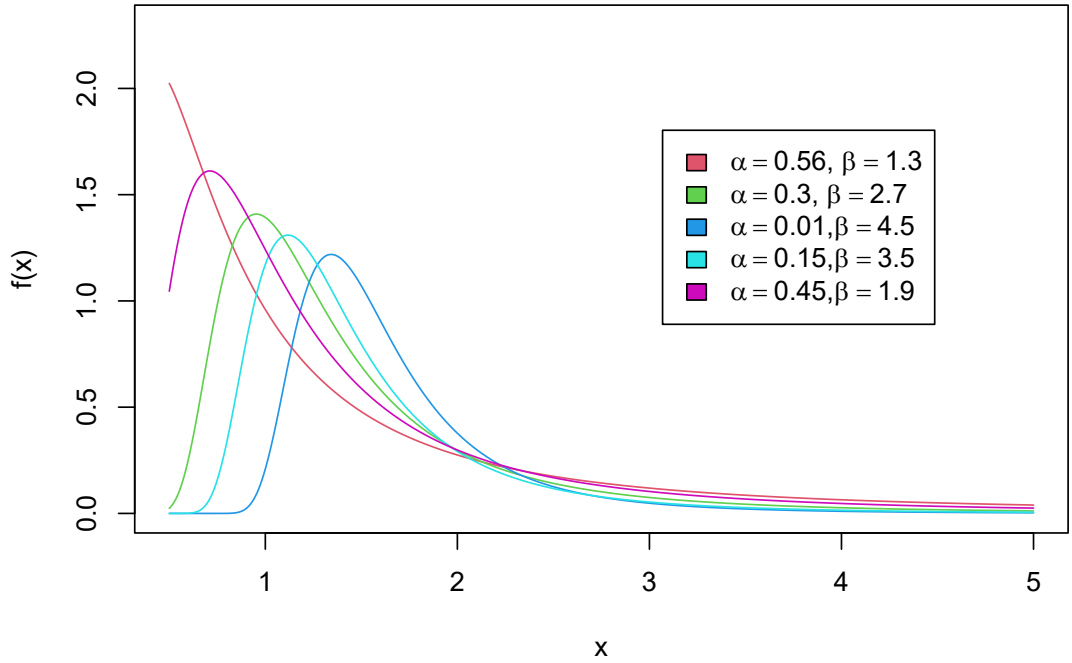


Figure 1: *PDF plot for SMPP distribution*

3. Reliability Analysis Of SMPP Distribution

In this section, we derive the expressions for the reliability function, hazard rate, reverse hazard function and Mills ratio for the SMPP distribution.

3.1. Reliability function

The reliability function represents the probability that an item does not fail before a given time, say x. For the SMPP distribution, it is expressed as

$$R(x;\alpha,\beta) = 1 - G_{SMPP}(x;\alpha,\beta) = \frac{1-e^{\frac{\log\alpha}{x^\beta}}}{1-\alpha}, \quad \alpha \neq 1 \tag{7}$$

3.2. Hazard Rate

The hazard rate function is an important measure used to characterize a life phenomenon. It is given by

$$h(x;\alpha,\beta) = \frac{\frac{\log\alpha}{\alpha-1}\cdot\frac{\beta}{x^{\beta+1}}\cdot e^{\frac{\log\alpha}{x^{\beta}}}}{e^{\frac{\log\alpha}{x^{\beta}}}-1} \tag{8}$$

Figure (2) illustrates the hazard rate graphs of the SMPP distribution for various parameter combinations. It reveals that the hazard function can exhibit unimodal, increasing, decreasing, or constant shapes depending on the parameter values.

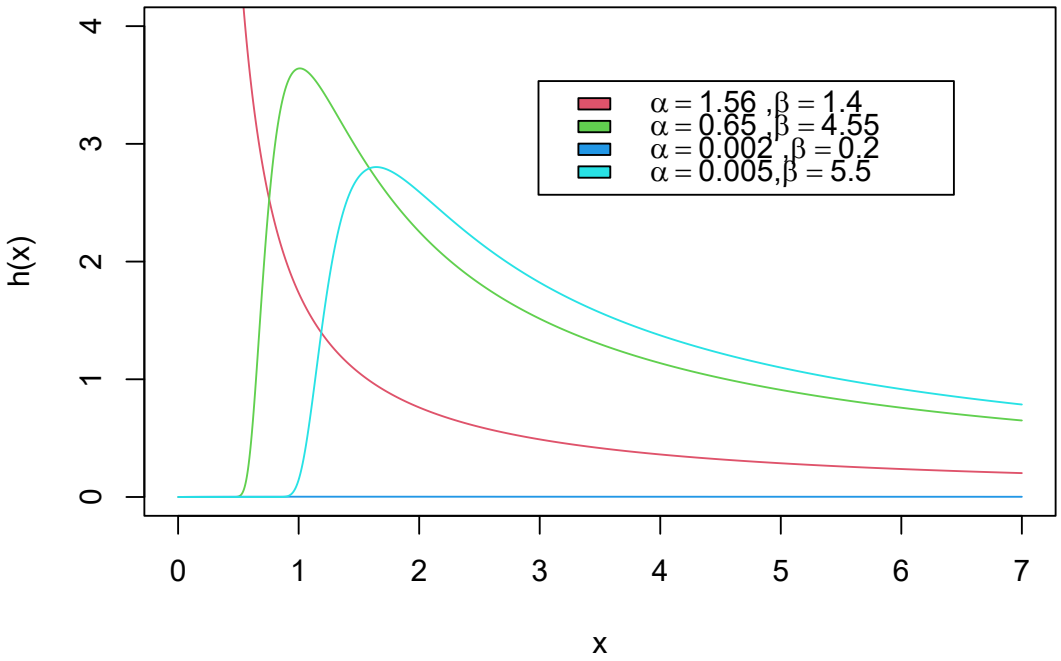


Figure 2: *Hazard plot for SMPP distribution*

3.3. Reverse Hazard Rate

The expression for the reverse hazard rate of the SMPP distribution is given as

$$h_r(x;\alpha,\beta) = \frac{(\log\alpha)e^{\frac{(\log\alpha)}{x^{\beta}}}\frac{\beta}{x^{\beta+1}}}{\alpha - e^{\frac{\log\alpha}{x^{\beta}}}} \tag{9}$$

3.4. Mills Ratio

The mills ratio for the SMPP is defined as

$$M.R = \frac{e^{\frac{\log\alpha}{x^{\beta}}}-\alpha}{1-e^{\frac{\log\alpha}{\beta}}} \tag{10}$$

3.5. Quantile function

Theorem 1. If $X \sim SMPP(\alpha, \beta)$ distribution, then the quantile function of X is given as

$$x = \left[\frac{\log \alpha}{\log\{u(1-\alpha)+\alpha\}}\right]^{\frac{1}{\beta}} \tag{11}$$

where u is a uniform random variable, $0 < u < 1$

Proof. Let $G_{SMPP}(x; \alpha, \beta) = u$. The quantile function of SMPP distribution can be obtained as follows.

$$\Rightarrow \frac{e^{\frac{log\alpha}{x^\beta}} - \alpha}{1-\alpha} = u$$

$$\Rightarrow e^{\frac{\log \alpha}{x^\beta}} = u(1-\alpha)+\alpha$$

Taking logarithm on both sides and simplifying further, we obtain the required quantile function as

$$x = \left[\frac{\log \alpha}{\log\{u(1-\alpha)+\alpha\}}\right]^{\frac{1}{\beta}} \tag{12}$$

Where u follows a uniform (0,1) distribution. The q^{th} quantile function of SMPP distribution is expressed as:

$$x_q = \left[\frac{\log \alpha}{\log\{u(1-\alpha)+\alpha\}}\right]^{\frac{1}{\beta}}$$

The median can be obtained as

$$x_{0.5} = \left[\frac{\log \alpha}{\log\{0.5(1-\alpha)+\alpha\}}\right]^{\frac{1}{\beta}}$$

■

4. Statistical properties of SMPP distribution

Some of the statistical properties of SMPP will be discussed in this section.

4.1. Moments

The r^{th} moment for SMPP distribution can be obtained as

$$\mu_r{}' = E(x^r) = \int_1^\infty x^r g(x; \alpha, \beta)\, dx$$

$$E(x^r) = \beta \frac{\log \alpha}{\alpha - 1} \int_1^\infty x^{-\beta-1+r} e^{\frac{\log \alpha}{x^\beta}}$$

using expansion $e^x = \sum_{j=0}^{\infty} \frac{x^j}{j!}$

$$\Rightarrow \mu_r{}' = \frac{\beta}{\alpha - 1} \sum_{j=0}^{\infty} \frac{(log\alpha)^{j+1}}{j!} \left[\frac{1}{\beta + \beta_j - r}\right] \tag{13}$$

4.2. Mode

The mode of the distribution is determined by solving the following equation

$$\log g(x) = \log\left(\frac{\log \alpha}{\alpha - 1} \cdot \frac{\beta}{x^{\beta+1}} \cdot e^{\frac{\log \alpha}{x^\beta}}\right) \tag{14}$$

differentating above equation w.r.t x and equate to zero we get,

$$x = \left(\frac{\beta \log \alpha}{\beta + 1}\right)^{\frac{1}{\beta}}$$

4.3. Moment Generating Function of SMPP

Theorem 2. Let $X \sim \text{SMPP}(\alpha, \beta)$, then the moment generating function, $M_X(t)$ is,

$$M_X(t) = \frac{\beta}{\alpha - 1} \sum_{r=0}^{\infty} \sum_{j=0}^{\infty} \frac{t^r}{r!} \frac{(\log \alpha)^{j+1}}{j!} \frac{1}{(\beta + \beta_j - r)} \tag{15}$$

Proof. The moment-generating function of SMPP distribution is given as

$$M_X(t) = \int_1^{\infty} e^{tx} g(x; \alpha, \beta)\, dx$$

$$M_X(t) = \int_1^{\infty} \left(1 + tx + \frac{(tx)^2}{2!} + \dots\right) g(x; \alpha, \beta)\, dx$$

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \int_1^{\infty} x^r g(x; \alpha, \beta)\, dx$$

$$\Rightarrow M_X(t) = \frac{\beta}{\alpha - 1} \sum_{r=0}^{\infty} \sum_{j=0}^{\infty} \frac{t^r}{r!} \frac{(\log \alpha)^{j+1}}{j!} \frac{1}{(\beta + \beta_j - r)} \tag{16}$$

■

4.4. Characteristic Function of SMPP distribution

Theorem 3. Let $X \sim \text{SMPP}(\alpha, \beta)$, then the characteristic function, $\phi_X(t)$ is given as

$$\phi_X(t) = \frac{\beta}{\alpha - 1} \sum_{r=0}^{\infty} \sum_{j=0}^{\infty} \frac{(it)^r}{r!} \frac{(\log \alpha)^{j+1}}{j!} \frac{1}{(\beta + \beta_j - r)}$$

Proof. The characteristic function of SMPP distribution is given as

$$\phi_X(t) = \int_1^{\infty} e^{itx} g(x; \alpha, \beta)\, dx$$

$$\phi_X(t) = \int_1^{\infty} \left(1 + itx + \frac{(itx)^2}{2!} + \dots\right) (x; \alpha, \beta)\, dx$$

$$\phi_X(t) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \int_1^{\infty} x^r g(x; \alpha, \beta)\, dx$$

$$\Rightarrow \phi_X(t) = \frac{\beta}{\alpha - 1} \sum_{r=0}^{\infty} \sum_{j=0}^{\infty} \frac{(it)^r}{r!} \frac{(\log \alpha)^{j+1}}{j!} \frac{1}{(\beta + \beta_j - r)} \tag{17}$$

■

Lemma 1. Let us suppose that a random variable X follows SMPP (α, β) with PDF given in equation (6) and let $I_r(t) = \int_1^t x^r g(x; \alpha, \beta) dx$ denote the r^{th} incomplete moment, then we have

$$I_r(t) = \frac{\beta}{\alpha - 1} \sum_{j=0}^{\infty} \frac{(log\alpha)^{j+1}}{j!} \left[\frac{t^{r-\beta-\beta_j} - 1}{r - \beta - \beta_j}\right] \tag{18}$$

Proof. Using the PDF of SMPP given in equation (6), we have

$$I_r(t) = \frac{\beta}{\alpha - 1} \sum_{j=0}^{\infty} \frac{(log\alpha)^{j+1}}{j!} \int_1^t x^{r-\beta-1-\beta_j}\, dx \tag{19}$$

$$I_r(t) = \frac{\beta}{\alpha - 1} \sum_{j=0}^{\infty} \frac{(log\alpha)^{j+1}}{j!} \left[\frac{t^{(r-\beta-\beta_j)} - 1}{r - \beta - \beta_j}\right] \tag{20}$$

4.5. Mean Residual Life (MRL)

The MRL, say $\mu(t)$, of SMPP distribution can be determined as follows:

$$\mu(t) = \frac{1}{R(t)}\left[E(x) - \int_1^t xf(x;\alpha,\beta)\,dx\right] - t$$

$$\mu(t) = \frac{1-\alpha}{1-e^{\frac{\log\alpha}{x^\beta}}}\frac{\beta}{\alpha-1}\sum_{j=0}^{\infty}\frac{(\log\alpha)^{j+1}}{j!}\left[\frac{1}{\beta+\beta_j-1} - \frac{t^{1-\beta-\beta_j}}{1-\beta-\beta_j}\right] - t$$

4.6. Mean Waiting Time (MWT)

The MWT of x, say $\bar{\mu}(t)$,is determined by

$$\overline{\mu}(t) = t - \frac{1}{F(t)}\int_1^t xf(x;\alpha,\beta)\,dx$$

$$\overline{\mu}(t) = t - \frac{1-\alpha}{e^{\frac{log\alpha}{x^\beta}} - \alpha}\cdot\frac{\beta}{\alpha-1}\cdot\sum_{j=0}^{\infty}\frac{(log\alpha)^{j+1}}{j!}\left[\frac{t^{1-\beta-\beta_j}-1}{1-\beta-\beta_j}\right]$$

4.7. Renyi Entropy

The entropy of a random variable is defined as the measure of uncertainty. The Renyi entropy is defined as,

$$I_v = \frac{1}{1-v}\log\int_1^{\infty} g^v(x)\,dx$$

Using PDF given in Equation (6), we have

$$I_v = \frac{1}{1-v}\log\left(\frac{\beta\log\alpha}{\alpha-1}\right)^v\int_1^{\infty}\sum_{j=0}^{\infty}\frac{v^j(\log\alpha)^j}{j!x^{\beta_j}}\frac{1}{x^{(\beta+1)v}}\,dx$$

$$I_v = \frac{1}{1-v}\log\left(\frac{\beta\log\alpha}{\alpha-1}\right)^v\sum_{j=0}^{\infty}\frac{v^j(\log\alpha)^j}{j!}\frac{1}{(\beta_v+\beta_j+v-1)}$$

which is required expression of Renyi entropy for SMPP distribution.

4.8. Stress Strength Reliability

Suppose X_1 and X_2 be two continous random variables, where $X_1 \sim$ SMPP(α_1,β) and $X_2 \sim$ SMPP(α_2,β), then the stress-strength parameter,say R, is defined as

$$SSR = \int_{-\infty}^{\infty} f_1(x)F_2(x)\,dx. \tag{21}$$

Using Equations (5) and (6) in Equation (21), the stress-strength parameter R, can be obtained as

$$SSR = \frac{\beta\log\alpha_1}{(\alpha_1-1)(1-\alpha_2)}\int_1^{\infty} e^{\frac{(\log\alpha_1)}{x^\beta}}\left(e^{\frac{(\log\alpha_2)}{x^\beta}} - \alpha_2\right)\,dx$$

$$SSR = \frac{\beta\log\alpha_1}{(\alpha_1-1)(1-\alpha_2)}\sum_{j=0}^{\infty}\frac{(\log\alpha_1)^j}{j!}\left[\sum_{k=0}^{\infty}\frac{(\log\alpha_2)^k}{k!}\frac{1}{\beta_j+\beta_k-1} - \frac{\alpha_2}{\beta_j-1}\right]$$

5. Order Statistics

Let $x_{(1)},x_{(2)},x_{(3)}, \ldots ,x_{(n)}$ be the random sample of size n and let $X_{r:n}$ denote the r^{th} order statistics, then the PDF of $X_{r:n}$ is given by

$$f_{r:n}(x) = \frac{n!}{(r-1)!(n-r)!} F(x)^{r-1}(1-F(x))^{n-r} g(x) \tag{22}$$

Using equation (5) we have,

$$f_{r:n}(x) = \frac{n!}{(r-1)!(n-r)!} \left(\frac{e^{\frac{(log\alpha)}{x^\beta}} - \alpha}{1-\alpha} \right)^{r-1} \left(\frac{1-e^{\frac{(log\alpha)}{x^\beta}}}{1-\alpha} \right)^{n-r} g(x)$$

$$f_{r:n}(x) = \frac{n!}{(r-1)!(n-r)!(1-\alpha)^{n-1}} \left(e^{\frac{(log\alpha)}{x^\beta}} - \alpha \right)^{r-1} \left(1 - e^{\frac{(log\alpha)}{x^\beta}} \right)^{n-r} \times \left(\frac{\log(\alpha)}{\alpha - 1} \frac{\beta}{x^{\beta+1}} e^{\frac{log\alpha}{x^\beta}} \right) \tag{23}$$

The expressions for the PDF of the smallest (minimum) and largest (maximum) order statistics of the SMPP distribution are obtained by setting $r = 1$ and $r = n$, respectively, in the above equation.

6. Estimation of Parameters

6.1. Maximum Likelihood Estimation

Let $x_1, x_2, x_3, \ldots, x_n$ be a random sample of size n having PDF in equation (6). Then the likelihood function is given by

$$L = \prod_{i=1}^{n} \frac{\log \alpha}{\alpha - 1} \cdot \frac{\beta}{x_i^{\beta+1}} \cdot e^{\frac{\log(\alpha)}{x_i^\beta}}$$

The log likelihood function is given by;

$$\log L(x : \alpha, \beta) = -n \log(\alpha - 1) + n \log(\log \alpha) + n \log \beta - \sum_{i=1}^{n} (\beta + 1) log x_i + (\log \alpha) \sum_{i=1}^{n} x_i^{-\beta}$$

The MLEs of α, β are obtained by partially differentiating above equation with respect to model parameters and equating to zero, we have

$$\frac{\partial \ell}{\partial \alpha} = -\frac{n}{\alpha - 1} + \frac{n}{\alpha \log \alpha} + \frac{1}{\alpha} \sum_{i=1}^{n} x_i^{-\beta} = 0 \tag{24}$$

$$\frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} - \sum_{i=1}^{n} \log (x_i) - (\log \alpha) \sum_{i=1}^{n} (\log x_i \cdot x_i^{-\beta}) = 0 \tag{25}$$

As the above equations are not in closed form, we will use the Newton-Raphson method and employ R software to solve the equations and estimate the parameters.

7. Simulation Study

This section presents a simulation study conducted using R software to examine the behaviour of the MLEs. Random samples of sizes n=25, 75, 150, 300, 500 are generated using the quantile function, with the process repeated 1000 times. Two different parameter combinations, (α, β)=(0.25, 0.94) and (0.50, 0.35) are considered. The average values of the MLEs, along with their biases

and mean square errors (MSEs), are summarised in Tables 1 and 2. The results indicate that the estimates are consistent and converge closely to the true parameter values across the given sample sizes. Additionally, the MSE and bias consistently decrease as the sample size increases in all cases.

Table 1: *MLE, Bias, and MSE for the parameters α and β*

Sample size	Parameters		MLE		Bias		MSE	
n	α	β	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
25	0.25	0.94	0.42332	1.00330	0.33034	0.22074	0.63895	0.07852
75			0.29609	0.96187	0.15788	0.12391	0.06768	0.02465
150			0.27572	0.94916	0.10828	0.08893	0.02251	0.01197
300			0.26238	0.94300	0.07090	0.05965	0.00896	0.00565
500			0.25658	0.94241	0.05176	0.04546	0.00449	0.00319

Table 2: *MLE, Bias, and MSE for the parameters α and β*

Sample size	Parameters		MLE		Bias		MSE	
n	α	β	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
25	0.50	0.35	2.16491	0.38053	1.99298	0.09651	0.57886	0.01555
75			0.61870	0.35945	0.34101	0.05291	0.40309	0.00453
150			0.55311	0.35700	0.22113	0.03808	0.10438	0.00232
300			0.52724	0.35187	0.14309	0.02502	0.03608	0.00098
500			0.51296	0.35105	0.11196	0.02005	0.02309	0.00066

8. Application

The application of the SMPP distribution is demonstrated using two datasets to highlight its importance, usefulness, and flexibility compared to other existing distributions.The performance of the SMPP distribution was compared to the Generalized pareto distribution given by [18] , Exponentiated Generalized pareto distribution by [14] ,Inverse Pareto distribution by [9] and basic Pareto distribution by [13] with the following pdfs.

- Generalized Pareto Distribution

$$\frac{1}{\beta}\left(1+\frac{\alpha x}{\beta}\right)^{-\left(\frac{1}{\alpha}+1\right)}$$

- Exponentiated Generalized Pareto Distribution(ExGPD)

$$\frac{e^x}{\beta}\left(1+\frac{ae^x}{\beta}\right)^{-\left(\frac{1}{\alpha}+1\right)}$$

- Inverse Pareto Distribution

$$\frac{\alpha\beta x^{\alpha-1}}{(\beta+x)^{\alpha+1}}$$

- Basic Pareto Distribution(BP)

$$f(x)=\frac{\beta}{x^{\beta+1}}$$

The goodness-of-fit test is used to assess the performance of the SMPP distribution. The evaluation criteria include the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Corrected Akaike Information Criterion (AICC), and -2 log-likelihood (-2LL), along with the results of the Kolmogorov-Smirnov (KS) test and the p- value. In general, the model is to be considered the best for which these goodness-of-fit statistics have the least and the p-value is greater.

Data Set 1. Patients receiving an analgesic. The data set is taken [8] which consists of 20 observations of patients receiving an analgesic. The values are as follows.

$$1.1, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7,$$
$$4.1, 1.8, 1.5, 1.2, 1.4, 3.0, 1.7, 2.3, 1.6, 2.0$$

Data set 2. Precipitation data (in inches) The second data set consists of 25 observations for the precipitation (in inches) from Jug Bridge, Maryl and previosly used by [10] and is given as follows.

$$1.01, 1.11, 1.13, 1.15, 1.16, 1.17, 1.17, 1.20, 1.52, 1.54, 1.54, 1.57, 1.64$$
$$1.73, 1.79, 2.09, 2.09, 2.57, 2.752.93, 3.19, 3.54, 3.57, 5.11, 5.62$$

Table 3: *Comparison of SMPP and competitive models for Data set 1*

Model	$\hat{\alpha}$	$\hat{\beta}$	$-2ll$	AIC	BIC	AICC	K-S	p-value
SMPP	0.00326	3.94075	30.71710	34.71076	36.70222	35.41666	0.10530	0.97940 IPP
57.43931	0.01096	65.45431	69.45431	71.44578	70.1602	0.38689	0.00502	
GPD	0.00100	1.89999	65.67433	69.64433	71.66579	70.3801	0.43951	0.00088
EGP	0.14408	8.02475	53.06459	57.06459	59.05606	57.07704	0.30554	0.04779
BP		1.69706	42.41429	44.41429	45.41000	45.12018	0.28505	0.07754

Table 4: *Comparison of SMPP and competitive models for Data set 2*

Model	$\hat{\alpha}$	$\hat{\beta}$	$-2ll$	AIC	BIC	AICC	K-S	p-value
SMPP	0.29724	2.02042	57.07541	61.07541	63.43152	61.64684	0.14161	0.72170
IPP	53.11791	0.01120	84.42617	88.42617	90.78228	88.9976	0.29780	0.02833
GPP	0.00100	2.18124	85.43527	89.43527	91.79138	90.0067	0.37063	0.00273
EPD	0.73112	8.50710	81.15739	85.15739	87.51350	85.72882	0.25156	0.09590
BP		1.54771	58.04799	65.04799	66.22605	65.61942	0.14360	0.60540

The results presented in Tables 3 and 4 demonstrate that the AIC, BIC, AICC, and -2LL values are lower for the SMPP distribution compared to the other fitted distributions. Consequently, the SMPP outperforms both the competing models and the base model. The promising performance of the proposed distribution is further illustrated in Figures 3 and 4.

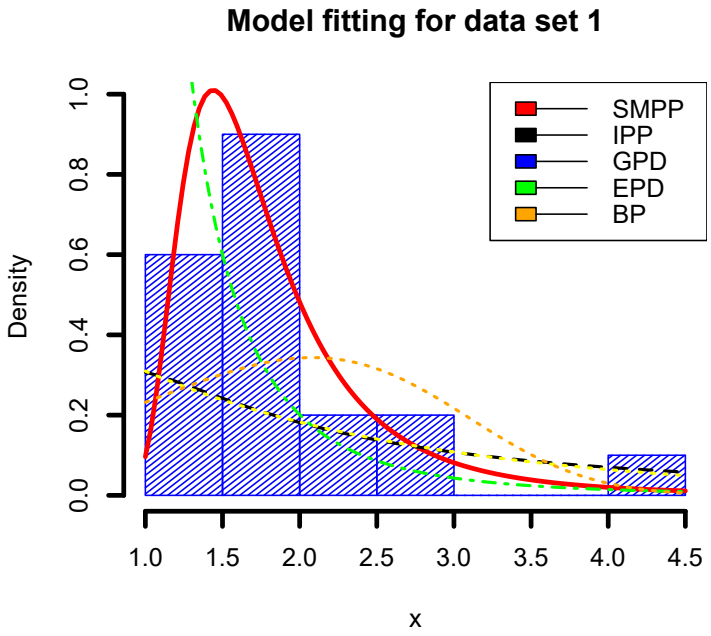


Figure 3: *Fitted density plots for data set 1*

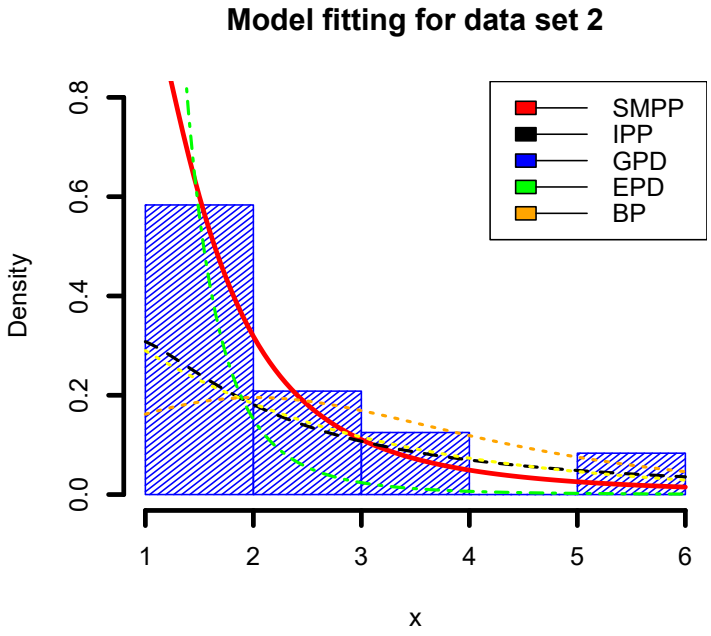


Figure 4: *Fitted density plots for data set 2*

9. Conclusion

In this paper, we propose a new model called the SMPP distribution, which extends the basic Pareto distribution. The applied transformation generalizes the standard distribution to enhance its flexibility for modeling real-world data. The key statistical properties of the proposed model have been derived. Its hazard rate function demonstrates greater flexibility and accommodates complex forms. The parameters are estimated using the maximum likelihood estimation (MLE) method, which is shown to be accurate and consistent through simulation analysis. Compared to other competing distributions, the proposed model has the potential to provide a superior fit for real-world applications. We believe this approach will find extensive use in the field of statistics.

References

[1] D. abd Khalil ELhertaniy. Sine exponential pareto distribution: properties, estimation, and applications. *Jordan Journal of Mathematics and Statistics*, 17(3), 2024.

[2] T. Ahad and S. Ahmad. New generalization of inverted exponential distribution: Properties and its applications. *Reliability: Theory & Applications*, 19(4 (80)):541–551, 2024.

[3] H. H. Ahmad and E. Almetwally. Generalizations of pareto distribution with applications to lifetime data. In *Journal of Physics: Conference Series*, volume 1943, page 012141. IOP Publishing, 2021.

[4] A. Akinsete, F. Famoye, and C. Lee. The beta-pareto distribution. *Statistics*, 42(6):547–563, 2008.

[5] A. Alzaatreh, C. Lee, and F. Famoye. A new method for generating families of continuous distributions. *Metron*, 71(1):63–79, 2013.

[6] S. O. Bleed, R. A.-E. Attwa, R. F. M. Ali, and T. Radwan. On alpha power transformation generalized pareto distribution and some properties. *Journal of Applied Mathematics*, 2024(1):6270350, 2024.

[7] F. M. Correa, B. J. Odunayo, I. Sule, and O. A. Bello. Topp-leone exponentiated pareto distribution: Properties and application to covid-19 data. *Journal of Statistical Theory and Applications*, pages 1–19, 2024.

[8] A. J. Gross and V. Clark. Survival distributions: reliability applications in the biomedical sciences. *(No Title)*, 1975.

[9] L. Guo and W. Gui. Bayesian and classical estimation of the inverse pareto distribution and its application to strength-stress models. *American Journal of Mathematical and Management Sciences*, 37(1):80–92, 2018.

[10] N. Henze and B. Klar. Goodness-of-fit tests for the inverse gaussian distribution based on the empirical laplace transform. *Annals of the Institute of Statistical Mathematics*, 54:425–444, 2002.

[11] S. Ihtisham, A. Khalil, S. Manzoor, S. A. Khan, and A. Ali. Alpha-power pareto distribution: Its properties and applications. *PloS one*, 14(6):e0218027, 2019.

[12] M. Jan and S. Ahmad. A new extension of kumaraswamy distribution for improved data modeling: Properties and applications. *Reliability: Theory & Applications*, 19(4 (80)):654–665, 2024.

[13] N. L. Johnson and S. Kotz. Continuous univariate distributions-2 boston. *MA: Houghton Mifflin Company*, 1970.

[14] S. Lee and J. H. Kim. Exponentiated generalized pareto distribution: Properties and applications towards extreme value theory. *Communications in Statistics-Theory and Methods*, 48(8):2014–2038, 2019.

[15] A. Mahdavi and D. Kundu. A new method for generating distributions with an application to exponential distribution. *Communications in Statistics-Theory and Methods*, 46(13):6543–6557, 2017.

[16] A. W. Marshall and I. Olkin. A new method for adding a parameter to a family of distributions with application to the exponential and weibull families. *Biometrika*, 84(3):641–652, 1997.

[17] A. A. Mir and S. Ahmad. Modeling and analysis of sine power rayleigh distribution: Properties and applications. *Reliability: Theory & Applications*, 19(1 (77)):703–716, 2024.

[18] J. Pickands III. Statistical inference using extreme order statistics. *the Annals of Statistics*, pages 119–131, 1975.

[19] M. M. Rahman, B. Al-Zahrani, and M. Q. Shahbaz. Cubic transmuted pareto distribution. *Annals of Data Science*, 7:91–108, 2020.

[20] S. U. Rasool, M. Lone, and S. Ahmad. An innovative technique for generating probability distributions: A study on lomax distribution with applications in medical and engineering fields. *Annals of Data Science*, pages 1–17, 2024.

FUZZY DATA-DRIVEN ESTIMATION FOR DUAL WEIBULL POPULATIONS: EM AND BAYESIAN METHODS WITH GIBBS SAMPLING

Abbarapu Ashok[1] and Nadiminti Nagamani[2,*]

•

[1,2]Department of Mathematics, School of Advanced Sciences, VIT-AP University, Amaravati, India.
[1]ashokabbarapu@gmail.com,[2]*maniinadiminti999@gmail.com

Abstract

In statistical analysis, accurately estimating parameters in the presence of uncertain and imprecise data is critical, particularly when dealing with complex, dual population models. Fuzzy data, which effectively represents real-world ambiguity, provides a framework for handling such uncertainties. While parameter estimation for single-population models with fuzzy data has been explored extensively, extending these methods to dual populations remains challenging. This study addresses this gap by developing estimation techniques for two Weibull populations that share a common shape parameter but differ in scale parameters, under fuzzy data conditions. We employ the Expectation-Maximization (EM) algorithm for Maximum Likelihood estimation and a Bayesian framework with TK approximation for parameter estimation. To refine our Bayesian estimates, we use Gibbs sampling to compute posterior densities. Through Monte Carlo simulations, for generated data as well as for a real dataset, we evaluate the accuracy and robustness of the proposed estimators, demonstrating their practical utility in applications where data imprecision is a significant factor. This research highlights the importance of robust methodologies for dual-population Weibull models, contributing to enhanced reliability in statistical analysis across diverse fields.

Keywords: Gibbs sampling; EM algorithm; TK approximation; Fuzzy data; Weibull distribution.

1. Introduction

Statistical estimation of parameters, under uncertain and imprecise conditions, becomes even more crucial in cases with complex multi-population models. Most real-world scenarios are characterized by uncertainty and vagueness, which cannot be adequately captured by classical data representations. For management of such uncertainties, fuzzy data provides a richer framework that allows researchers to make better inferences. At the same time, modeling through fuzzy sets enables one to manage data variability in a systematic and mathematically constructive manner so that statistical analysis could attain better precision.

Due to its suitability in modeling with limited precise data, the Weibull distribution has vital usages in reliability and survival analysis. For example, survival times of patients or the materials' failure times depend quite heavily on various external factors or incomplete information thereby making an imprecision such as these. Apart from classical literature, the model incorporates fuzzy data in a Weibull model so that such uncertainties are realistically represented. In this study, two Weibull populations have been considered with similar shape parameters hence suggesting similar survival or failure rates and distinct from each other with respect to scale parameters that provide different averages for these populations.

The Weibull distribution is commonly used in failure modeling, reliability determination, and life data analysis. Its introduction in various situations can be traced back to Waloddi Weibull, who presented it in 1951. The variety of applications for this distribution in biology, engineering, and weather forecasting is due to its adaptability and capability to model a variety of data types. The p.d.f of weibull distribuiton with scale parameter σ and shape parameter k is,

$$f(x;\sigma,k) = k\sigma x^{k-1}e^{-\sigma x^k}, \quad x > 0, \quad k > 0, \quad \sigma > 0.$$

By looking into some conclusions drawn by different scholars on the Weibull distribution. In 1977 [1], the maximum likelihood method was used to estimate parameters of Weibull distribution, and they also provided a method of calculating confidence intervals. Lai in 2011 [2] has given extensive explanations of various extensions, parameter estimation techniques, and basic Weibull distribution properties useful in modeling complex data sets. N. Balakrishnan and Kateri proposed a new method based on a simple and easy-to-apply graphical method that easily shows the existence and uniqueness of maximum likelihood estimators [14]. To the best of our knowledge, the only work in this direction that discuses the estimation of a common shape parameter of several Weibull populations is considered by Yang and Lin [15]. The estimation of parameters of various lifetime distributions including weibull distribution is considered by [16]. In 2009 [17] came to the conclusion that there is no analytical solution for the restricted MLE of the scale parameter for a given shape parameter, and he finally developed a new approach that is thought to be more effective and efficient at handling interval data than regular MLE methods, directly developed EM algorithms, as well as genetic algorithms.

Recently, the application of fuzzy sets in estimation theory has attracted considerable attention from researchers because it can simultaneously take care of uncertainties inherent in lifespan data. Such uncertainties are often of human error nature or can originate from technical failures or unforeseen circumstances, which could be easily modeled using fuzzy sets. The idea was formed in 1991 by Coppi et al. [4], who studied different applications of fuzzy techniques in statistical analysis. Building on this work, Denoeux [5] later investigated the use of the EM approach for maximum likelihood estimations based on fuzzy data. This was followed by significant advances, including Huang's contribution in 2006 [3], where a new method was devised to compute membership functions and reliability for multiparameter lifetime distributions. Pak et al. [6], [7] extended this work by conducting several experiments for the development of inferential methods for lifespan distributions based on fuzzy data, setting a deep base for forthcoming research. For one a recent study of, Ashok et al. [9] studied the parameter estimation of gamma distribution when the available information as imprecise data, while Khoolenjani et al. [8] offered an estimate of the mean parameter of the exponential distribution based on Type-II censoring. The use of fuzzy clustering methods for estimating parameters has a much longer history. Like, researchers such as Bezdek and Dunn (1975), Gath and Geva (1989), and Yang (1994) [10] already tested these methods for normal distributions, which just emphasizes the flexibility of fuzzy logical application in statistical modeling. More recently, Basharat et al. [12] showed how fuzzy data may estimate parameters of a linear combination of two exponentially distributed random variables.

Given the complex task of estimating under fuzzy data conditions, the study employs several sophisticated statistical techniques, chief among them being the Expectation-Maximization (EM) algorithm for Maximum Likelihood Estimation and the Bayesian approach, which is further conditioned by the TK approximation. Further Bayesian estimates are achieved using Gibbs sampling to derive posterior densities so that robust and more reliable parameter estimates are achieved. Performance of the proposed methodologies is evaluated through Monte Carlo simulation demonstrating convincingly that they are capable of handling data imprecision. These methodologies are practical in very applicable areas, such as medical research, industrial reliability testing, and the modeling of the environmental conditions of uncertainty.

2. Maximum Likelihood Estimation

Suppose two independent random samples are taken from two Weibull populations with a common shape parameter k and possibly different scale parameters σ_1 and σ_2, respectively. Specifically, let $\mathbf{X} = (X_1, X_2, \ldots, X_m)$ and $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ be m and n random samples taken from two Weibull populations Weibull(σ_1, k) and Weibull(σ_2, k), respectively. Here, Weibull(σ_1, k) denotes the Weibull population with the probability density function,

$$f(x; \sigma_1, k) = k\sigma_1 x^{k-1} e^{-\sigma_1 x^k}, \quad x > 0, \quad k > 0, \quad \sigma_1 > 0$$

likewise, Weibull(σ_2, k) denotes the population having the probability density function

$$f(y; \sigma_2, k) = k\sigma_2 y^{k-1} e^{-\sigma_2 y^k}, \quad y > 0, \quad k > 0, \quad \sigma_2 > 0$$

the joint probability density function of $\mathbf{X}$ and $\mathbf{Y}$ is obtained as follows,

$$f(x, y; ,\sigma_1, \sigma_2) = k^{m+n} \sigma_1^m \sigma_2^n \prod_{i=1}^{m} x_i^{k-1} \prod_{j=1}^{n} y_j^{k-1} e^{-\left(\sigma_1 \sum_{i=1}^{m} x_i^k + \sigma_2 \sum_{j=1}^{n} y_j^k\right)}$$

taking into account the scenario where the information that is currently accessible regarding $x$and y are fuzzy data rather than crisp data, that is, $\tilde{X} = \tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_m$, $\tilde{Y} = \tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_n$ denotes its observations, where $\tilde{x}_i$, $\tilde{y}_j$ are the fuzzy observed number for the random variable x_i,y_j the likelihood function based on fuzzy observations $\tilde{X}$ and $\tilde{Y}$ is defined as, with the membership function $\mu_{\tilde{x}_i}(x)$ and $\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)$ as a function of θ. The observed likelihood function L_0 and observed log likelihood function L_* are as follows,

$$L_0 = f(\tilde{x}, \tilde{y}; k, \sigma_1, \sigma_2) = k^{m+n} \sigma_1^m \sigma_2^n \prod_{i=1}^{m} \int x_i^{k-1} e^{-\sigma_1 x_i^k} \mu_{\tilde{x}_i}(x)\, dx \prod_{j=1}^{n} \int y_j^{k-1} e^{-\sigma_2 y_j^k} \mu_{\tilde{y}_j}(y)\, dy$$

$$\begin{aligned} L_* &= (m+n)\log k + m\log\sigma_1 + n\log\sigma_2 + \sum_{i=1}^{m} \log \int x_i^{k-1} e^{-\sigma_1 x_i^k} \mu_{\tilde{x}_i}(x)\, dx \\ &\quad + \sum_{j=1}^{n} \log \int y_j^{k-1} e^{-\sigma_2 y_j^k} \mu_{\tilde{y}_j}(y)\, dy \end{aligned}$$

by setting the partial derivatives of the log-likelihood with respect to the parameters σ_1 , σ_2 and k to zero, two resulting equations can be obtained for solving the parameters.

$$\frac{\partial L_*}{\partial \sigma_1} = \frac{m}{\sigma_1} - \Sigma \frac{\int x^{2k-1} e^{-\sigma_1 x^k} \mu_{\tilde{x}_i}(x)\, dx}{\int x^{k-1} e^{-\sigma_1 x^k} \mu_{\tilde{x}_i}(x)\, dx}$$

$$\frac{\partial L_*}{\partial \sigma_2} = \frac{n}{\sigma_2} - \Sigma \frac{\int y^{2k-1} e^{-\sigma_2 y^k} \mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\, dy}{\int y^{k-1} e^{-\sigma_2 y^k} \mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\, dy}$$

$$\begin{aligned} \frac{\partial L_*}{\partial k} &= \frac{m+n}{k} + \sum_{i=1}^{m} \frac{\int \left(x^{k-1} - \sigma_1 x^{2k-1}\right) \log x\, e^{-\sigma_1 x^k} \mu_{\tilde{x}_i}(x)\, dx}{\int x^{k-1} e^{-\sigma_1 x^k} \mu_{\tilde{x}_i}(x)\, dx} + \\ &\quad \sum_{j=1}^{n} \frac{\int \left(y^{k-1} - \sigma_2 y^{2k-1}\right) \log y\, e^{-\sigma_2 y^k} \mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\, dy}{\int y^{k-1} e^{-\sigma_2 y^k} \mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\, dy} \end{aligned}$$

to obtain the estimation using the ML-method, it is tough to figure out the above nonlinear equations. In this case, an iterative numerical method called an EM-algorithm and the Newton-Raphson method is used to obtain the MLEs.

2.1. EM Algorithm

By employing the whole data log-likelihood function, the EM method continually tackles the problem of maximizing the observed data log-likelihood function. The Expectation phase and the Maximization step are the two steps that make up each iteration of the procedure. Iteratively computing the expectation of the observed data log-likelihood function using the full data log-likelihood function is part of the E-step. The following is the expression for the whole data log-likelihood function:

$$\log L = (m+n)\log k + m\log\sigma_1 + n\log\sigma_2 + (k-1)\sum\log x_i + (k-1)\sum\log y_j - \sigma_1\sum x_i^k - \sigma_2\sum y_j^k$$

by differentiating the above equation with respect to k, σ_1 and $\sigma_2$simplifying we get,

$$\frac{m+n}{k} = \sigma_1\sum x_i^k\log x + \sigma_2\sum y_j^k\log y - \sum\log x - \sum\log y$$

$$\frac{m}{\sigma_1} = \sum x_i^k \quad \text{and} \quad \frac{n}{\sigma_2} = \sum y_j^k$$

the iterative process of the EM algorithm is as follows,

- Give initial values (calculated by sample mean and variance) of $\rho = \rho^{(0)}$ and $\eta = \eta^{(0)}$, set $h = 0$.
- In the $(h+1)$th iterative process, the EM approach aims to maximize the log-likelihood function $L_*(\theta)$ of the observed data by repeatedly processing the log-likelihood function of the entire data.
- **E-step:** The E-step requires the computation of conditional expectations utilizing the expression.

$$E_{1i} = E\left(\log x_i \mid \tilde{x}_i\right) = \frac{\int x_i^{k-1}\log x_i\, e^{-\sigma_1 x_i^k}\mu(x)\,dx}{\int x_i^{k-1}e^{-\sigma_1 x_i^k}\mu(x)\,dx}$$

$$E_{2i} = E\left(\log y_j \mid \tilde{y}_j\right) = \frac{\int y_j^{k-1}\log y_j\, e^{-\sigma_2 y_j^k}\mu(y)\,dy}{\int y_j^{k-1}e^{-\sigma_2 y_j^k}\mu(y)\,dy}$$

$$E_{3i} = E\left(x_i^k\log x_i \mid \tilde{x}_i\right) = \frac{\int x^{2k-1}\log x\, e^{-\sigma_1 x^k}\mu(x)\,dx}{\int x^{k-1}e^{-\sigma_1 x^k}\mu(x)\,dx}$$

$$E_{4i} = E\left(y_j^k\log y_j \mid \tilde{y}_j\right) = \frac{\int y^{2k-1}\log y\, e^{-\sigma_2 y^k}\mu(y)\,dy}{\int y^{k-1}e^{-\sigma_2 y^k}\mu(y)\,dy}$$

$$E_{5i} = E\left(x_i^k \mid \tilde{x}_i\right) = \frac{\int x_j^k e^{-\sigma_2\ x^k}\,\mu(x)\,dx}{\int x^{k-1}e^{-\sigma_2\ x^k}\,\mu(x)\,dx}$$

$$E_{6i} = E\left(y_j^k \mid \tilde{y}_j\right) = \frac{\int y_j^k e^{-\sigma_2\ y^k}\,\mu(y)\,dy}{\int y^{k-1}e^{-\sigma_2\ y^k}\,\mu(y)\,dy}$$

now, the non-linear equations will be replaced by,

$$\frac{m+n}{k} = \sigma_1\sum E_{3i} + \sigma_2\sum E_{4i} - \sum(E_{1i} - E_{2i})$$

$$\frac{m}{\sigma_1} = \sum E_{5i} \quad \text{and} \quad \frac{n}{\sigma_2} = \sum E_{6i}$$

- **M-step:**In the M-step after solving the above Non-linear eq's,

$$k^{(h+1)} = \frac{m+n}{\sigma_1^{(h+1)} \sum E_{3i} + \sigma_2^{(h+1)} \sum E_{4i} - \sum (E_{1i} - E_{2i})}$$

$$\sigma_1^{(h+1)} = \frac{m}{\sum E_{5i}} \quad \text{and} \quad \sigma_2^{(h+1)} = \frac{n}{\sum E_{6i}}$$

- **Convergence:** If convergence is achieved, the current values of $\sigma_1^{(h+1)}, \sigma_2^{(h+1)} and \eta^{(h+1)}$ are the maximum likelihood estimators (MLEs) of $\sigma_1, sigma_2$ and k. Otherwise, repeat steps 2 and 3.

Once the parameters' MLEs have been established, the 95% asymptotic confidence intervals for the parameters may be computed as follows.

$$\hat{\theta} \pm Z_{\alpha/2} \sqrt{I(\theta)^{-1}}$$

the inverse of the fisher information matrix will be,

$$I(\sigma_1, \sigma_2, k)^{-1} = \begin{bmatrix} \frac{T_{22}T_{33} - T_{23}^2}{D} & \frac{-T_{31}T_{23}}{D} & \frac{T_{22}T_{31}}{D} \\ \frac{-T_{31}T_{23}}{D} & \frac{T_{11}T_{33} - T_{31}^2}{D} & \frac{-T_{11}T_{32}}{D} \\ \frac{T_{22}T_{31}}{D} & \frac{-T_{11}T_{32}}{D} & \frac{T_{11}T_{22}}{D} \end{bmatrix}$$

here,

$$D = T_{11}(T_{22}T_{33} - T_{32}^2) + T_{13}^2 T_{22}$$

now, we construct 95% asymptotic confidence intervals using the information matrix for the model parameters σ_1, σ_2 and k are as follows:

$$\hat{\sigma_1}_{\text{ML}} \pm 1.96 \sqrt{\frac{T_{22}T_{33} - T_{23}^2}{D}},$$

$$\hat{\sigma_2}_{\text{ML}} \pm 1.96 \sqrt{\frac{T_{11}T_{33} - T_{31}^2}{D}},$$

$$\hat{k}_{\text{2ML}} \pm 1.96 \sqrt{\frac{T_{11}T_{22}}{D}}.$$

The Appendix contains all of the notations and calculations. For varying sample sizes, all of these findings are calculated quantitatively and displayed in Section 5.

3. Bayes Estimation

Being an even powerful and viable alternative to the traditional statistical approaches, Bayesian perspective has received considerable attention for decades on statistical inference. In this section, the Bayesian estimate of the model-associated parameters is discussed. When the prior distribution of parameters is known beforehand, the Bayes estimator plays a crucial role. Assume that $\rho_1(k)$ is the shape parameter k prior density function. For the scale parameters σ_1 and σ_2, let $\rho_2(\sigma_1)$ and $\rho_3(\sigma_2)$ be the gamma prior densities, respectively.

$$\rho_1(k) = \frac{b_1^{a_1}}{\Gamma(a_1)} k^{a_1 - 1} \exp(-kb_1), \quad k > 0$$

$$\rho_2(\sigma_1) = \frac{b_2^{a_2}}{\Gamma(a_2)}\sigma_1^{a_2-1}\exp(-\sigma_1 b_2), \quad \sigma_1 > 0$$

$$\rho_3(\sigma_2) = \frac{b_3^{a_3}}{\Gamma(a_3)}\sigma_2^{a_3-1}\exp(-\sigma_2 b_3), \quad \sigma_2 > 0$$

the joint probability density function of (k, σ_1, σ_2) based on the above priors is,

$$f(k,\sigma_1,\sigma_2) = (m+n)\log k + m\log\sigma_1 + n\log\sigma_2 +$$

$$\sum_{i=1}^{m}\log\int x_i^{k-1}e^{-\sigma_1 x_i^k}\mu_{\tilde{x}_i}(x)\,dx + \sum_{j=1}^{n}\log\int y_j^{k-1}e^{-\sigma_2 y_j^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\,dy\rho_1(k).\rho_2(\sigma_1).\rho_3(\sigma_2)$$

the joint posterior density function of $\theta = (k, \sigma_1, \sigma_2)$ given the data (X, Y) is,

$$f(\sigma_1,\sigma_2,k \mid x,y) = \frac{f(\sigma_1,\sigma_2,k,x,y)}{\int_0^\infty\int_0^\infty\int_0^\infty f(\sigma_1,\sigma_2,k,x,y)\,d\sigma_1\,d\sigma_2\,dk}$$

the Bayes estimator of $g(\sigma_1, \sigma_2, k)$ with respect to the squared error loss function is the mean of the posterior distribution and is given by the following.

$$\hat{g}_B(\sigma_1,\sigma_2,k) = \mathbb{E}[g(\sigma_1,\sigma_2,k) \mid x,y]$$

$$= \frac{\int_0^\infty\int_0^\infty\int_0^\infty g(\sigma_1,\sigma_2,k)f(\sigma_1,\sigma_2,k,x,y)\,d\sigma_1\,d\sigma_2\,dk}{\int_0^\infty\int_0^\infty\int_0^\infty f(\sigma_1,\sigma_2,k,x,y)\,d\sigma_1\,d\sigma_2\,dk}$$

analytical methods for finding the ratio of integrals in above equation are difficult to use. However, one might obtain a numerical quantity with certain approximations. In this paper, we employ Tierney and Kadane's TK approximation to estimate the Bayes estimators under squared error loss function.

3.1. TK approximation

In this part, we use Tierney and Kadane's [13] approximation approach to find the Bayes estimators for the common Shape parameter k and the other scale parameters σ_1, σ_2 . Tierney and Kadane's approximation simplifies the computation of the posterior expectation $E(u(\theta)|\text{data})$. The approximation is given by,

$$E(u(\theta) \mid \tilde{X},\tilde{Y}) = \left[\frac{\det\Sigma^*}{\det\Sigma}\right]^{\frac{1}{2}}\exp\{k\big(L_*(\theta_*) - L_0(\theta_0)\big)\},$$

where:

- $L_0(\theta_0)$: The log-likelihood function maximized over the parameter space θ,
- $L_*(\theta_*) = L_0(\theta_0) + \frac{1}{k}\log u(\theta)$,
- Σ^*: Minus the inverse Hessian of $L_*(\theta)$ at the mode θ^*,
- Σ: Minus the inverse Hessian of $L_0(\theta_0)$ at the mode θ^*.

$$L_0(\theta_0) = \frac{1}{k}[L(\theta) + logv(\theta] = \frac{1}{k}[L(k,\sigma_1,\sigma_2|\tilde{x}) + logv(\theta)]$$

$$L_*(\theta) = L_0(\theta_0) + \frac{1}{k}\ logv(\theta)$$

In our case the approximation will be,

$$E(u(\theta) \mid \tilde{X}, \tilde{Y})) = \left[\frac{\det \Sigma^*}{\det \Sigma}\right]^{\frac{1}{2}} \exp\{L_*(\hat{k}_*, \hat{\sigma}_{1_*}, \hat{\sigma}_{2_*}) - L_0(\hat{k}_0, \hat{\sigma}_{1_0}, \hat{\sigma}_{2_0})\}.$$

The Hessians are calculated as the negative second derivatives of $L_*(k, \sigma_1, \sigma_2)$ and $L_0(k, \sigma_1, \sigma_2)$. Let the parameter vector be $\boldsymbol{\theta} = (k, \sigma_1, \sigma_2)$.

- For $L_*(\hat{k}_*, \hat{\sigma}_{1_*}, \hat{\sigma}_{2_*})$:

$$\Sigma^* = -\begin{bmatrix} \frac{\partial^2 L_*}{\partial k^2} & \frac{\partial^2 L_*}{\partial k \partial \sigma_1} & \frac{\partial^2 L_*}{\partial k \partial \sigma_2} \\ \frac{\partial^2 L_*}{\partial \sigma_1 \partial k} & \frac{\partial^2 L_*}{\partial \sigma_1^2} & \frac{\partial^2 L_*}{\partial \sigma_1 \partial \sigma_2} \\ \frac{\partial^2 L_*}{\partial \sigma_2 \partial k} & \frac{\partial^2 L_*}{\partial \sigma_2 \partial \sigma_1} & \frac{\partial^2 L_*}{\partial \sigma_2^2} \end{bmatrix}^{-1} = (\Sigma^*_{ij})$$

- For$L_0(\hat{k}_0, \hat{\sigma}_{1_0}, \hat{\sigma}_{2_0})$:

$$\Sigma = -\begin{bmatrix} \frac{\partial^2 L_0}{\partial k^2} & \frac{\partial^2 L_0}{\partial k \partial \sigma_1} & \frac{\partial^2 L_0}{\partial k \partial \sigma_2} \\ \frac{\partial^2 L_0}{\partial \sigma_1 \partial k} & \frac{\partial^2 L_0}{\partial \sigma_1^2} & \frac{\partial^2 L_0}{\partial \sigma_1 \partial \sigma_2} \\ \frac{\partial^2 L_0}{\partial \sigma_2 \partial k} & \frac{\partial^2 L_0}{\partial \sigma_2 \partial \sigma_1} & \frac{\partial^2 L_0}{\partial \sigma_2^2} \end{bmatrix}^{-1} = (\Sigma_{ij}).$$

It should be noted that the elements of L_0 and L_* are being evaluated at $(\hat{\sigma}_{1_0}, \hat{\sigma}_{2_0}, \hat{k}_0)$ and $(\hat{\sigma}_{1_*}, \hat{\sigma}_{2_*}, \hat{k}_*)$, respectively. Now it is simple to evaluate for the Bayes estimate of k. The same procedure can be used for obtaining the Bayes estimates for $u(\sigma_1, \sigma_2, k) = \sigma_1$ and $u(\sigma_1, \sigma_2, k) = \sigma_2$.

3.2. Gibbs sampling

Gibbs sampling is one of the most popular MCMC methods in Bayesian statistics, simplifying sampling from complex distributions, especially in high dimensions or with intractable forms. It constructs a Markov chain that converges to the desired posterior by iteratively sampling from conditional distributions.

This paper applies Gibbs sampling to estimate the parameters of the Weibull distribution, which is widely used in survival and reliability research. Gibbs sampling handles fuzzy data effectively by iteratively refining parameter estimates and demonstrates its flexibility in managing uncertainty within probabilistic models such as the Weibull distribution. The Gibbs sampling procedure can be outlined as follows:

- Assume that $X \sim \text{Weibull}(k, \sigma_1)$ and $Y \sim \text{Weibull}(k, \sigma_2)$. The joint density function is given by:

$$L_* = (m+n)\log k + m\log\sigma_1 + n\log\sigma_2 + \sum_{i=1}^{m} \log \int x_i^{k-1} e^{-\sigma_1 x_i^k} \mu_{\tilde{x}_i}(x)\, dx +$$

$$+ \sum_{j=1}^{n} \log \int y_j^{k-1} e^{-\sigma_2 y_j^k} \mu_{\tilde{x}_i}(x) \mu_{\tilde{y}_j}(y)\, dy$$

- Assume the following prior distributions for the parameters:

$$\begin{aligned} k &\sim \text{Gamma}(b_1, a_1) \\ \sigma_1 &\sim \text{Gamma}(b_2, a_2) \\ \sigma_2 &\sim \text{Gamma}(b_3, a_3) \end{aligned}$$

- Derive the full conditional posterior distributions for k, σ_1, and σ_2:

$$\begin{aligned} \pi(k \mid \sigma_1, \sigma_2, \tilde{X}, \tilde{Y}) &\propto L_*(k, \sigma_1, \sigma_2 \mid \tilde{X}, \tilde{Y})\ \pi(k) \\ \pi(\sigma_1 \mid k, \tilde{X}) &\propto L_*(k, \sigma_1 \mid \tilde{X})\ \pi(\sigma_1) \\ \pi(\sigma_2 \mid k, \tilde{Y}) &\propto L_*(k, \sigma_2 \mid \tilde{Y})\ \pi(\sigma_2) \end{aligned}$$

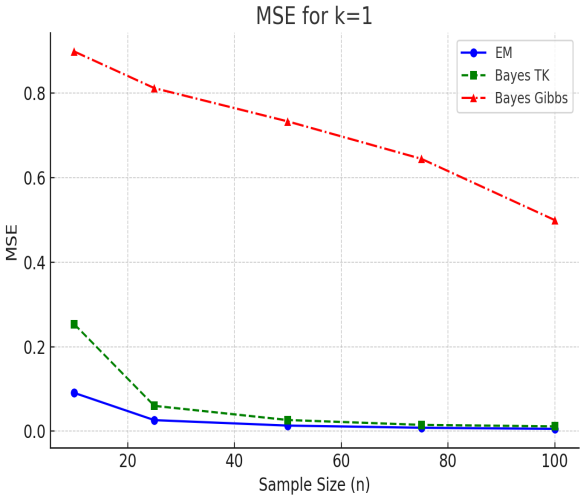
- Estimate the parameters iteratively using the following steps:
  1. Start with initial values for the parameters $k^{(0)}$, $\sigma_1^{(0)}$, and $\sigma_2^{(0)}$.
  2. For each iteration $t = 1, 2, \ldots, T$.
  3. Sample $k^{(t+1)}$ from its conditional posterior distribution $\pi(k \mid \sigma_1^{(t)}, \sigma_2^{(t)}, \tilde{x}, \tilde{y})$.
  4. Sample $\sigma_1^{(t+1)}$ from its conditional posterior distribution $\pi(\sigma_1 \mid k^{(t+1)}, \tilde{x})$.
  5. Sample $\sigma_2^{(t+1)}$ from its conditional posterior distribution $\pi(\sigma_2 \mid k^{(t+1)}, \tilde{y})$.

The samples obtained can be used for posterior inference once the Gibbs sampler runs for sufficiently many iterations to guarantee convergence. This implies estimating posterior means, credible intervals, and other relevant statistics with respect to the parameters under consideration.

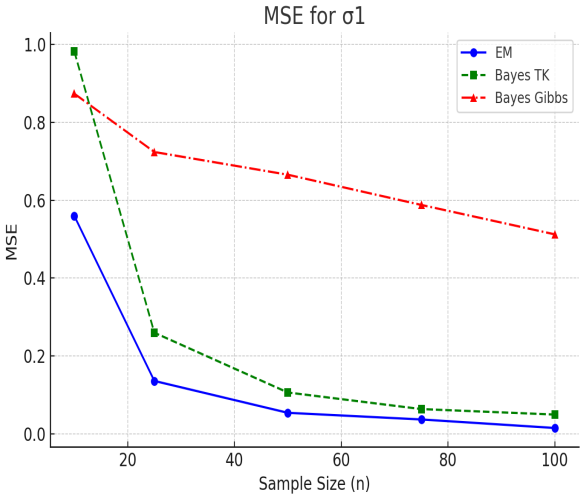
4. Simulation study

This section presents experimental results regarding the evaluation of the ranged performance of different estimation methods on the basis of sample sizes. The estimation is aimed at the common shape parameter k and possibly distinct scale parameters σ_1 and σ_2. These estimations were conducted using maximum likelihood (ML), implementing it via EM algorithm, and Bayesian estimation approaches with TK approximation and Gibbs sampling method. All the computations were carried out under R software version 4.4.1 along with the curve plots, supporting the analysis. Ten thousand random samples are drawn from two Weibull populations for each combination of parameter values and sample sizes for comparisons across the estimators. These samples form the fuzzy values using the triangular membership function. Three sets of parameter values are analyzed for a different sample size. A summary of the results is given as in Table 1 compares estimates and mean square errors (MSE) for various sample sizes using the EM method, Bayes (TK), and Bayes (Gibbs) methods, assuming parameter values $k = 1$, $\sigma_1 = 1.5$, and $\sigma_2 = 2$. Table 2 presents similar comparisons for parameter values $k = 1$, $\sigma_1 = 2$, and $\sigma_2 = 2.5$. Table 3 evaluates results for parameter values $k = 2$, $\sigma_1 = 3$, and $\sigma_2 = 4.5$. Moreover, Table 4 and Table 5 give asymptotic confidence intervals and the highest posterior density (HPD) intervals, arising from the Fisher information matrix and using Gibbs sampling, respectively. The results are summarized in several tables as follows:

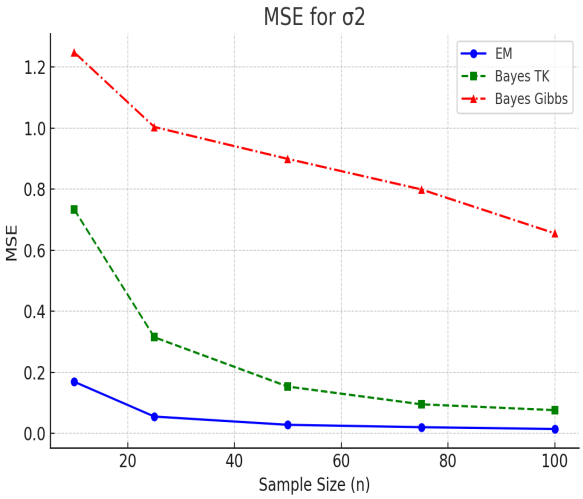
- The analytical results showed that increasing the sample size gave all estimators more accuracy as it led to considerable reductions in bias and mean squared error (MSE).
- As the sample numbers increased, the widths of the intervals decreased, indicating a higher degree of precisions, with all parameter estimates coming under their respective confidence intervals.
- The sample size, along with the selected confidence level, significantly shaped the confidence intervals. All other things being equal, a larger sample infinitesimally shrinks the confidence intervals at any fixed confidence level.
- Estimation accuracy was demonstrated by narrowing estimations of the confidence interval with increased sample size.
- The figures 1(a), 1(b) and 1(c) in 1 clearly indicates that, EM algorithm is the better one when compared to the remaining estimation methods.
- The Bayes (Gibbs) method exhibited relatively higher MSE values, particularly for small sample sizes, indicating that it requires larger datasets to achieve accuracy comparable to the EM algorithm.
- The confidence intervals derived from Gibbs sampling were generally wider than those obtained through asymptotic methods, suggesting greater uncertainty in Bayesian estimates, especially when dealing with smaller sample sizes.



(a)



(b)



(c)

Figure 1: *MSE values of different parameters at various sample sizes*

Table 1: *Comparison of estimates and mean square errors (MSE) for various sample sizes using EM, Bayes(TK), and Bayes(Gibbs) methods for the parameters $k = 1$, $\sigma_1 = 1.5$, and $\sigma_2 = 2$.*

(m,n)	Parameter	EM		Bayes(TK)		Bayes(Gibbs)	
		Estimate	MSE	Estimate	MSE	Estimate	MSE
(10,10)	$k = 1$	1.0887	0.0908	1.1988	0.2534	1.3187	0.8988
	$\sigma_1 = 2$	2.2599	0.5599	2.2681	0.9824	2.4235	0.8739
	$\sigma_2 = 2.5$	2.4054	0.1689	2.2734	0.7331	3.0273	1.2482
(25,25)	$k = 1$	1.0359	0.0263	1.0741	0.0601	1.2944	0.8122
	$\sigma_1 = 2$	2.1085	0.1357	2.1485	0.2596	2.3804	0.7239
	$\sigma_2 = 2.5$	2.5110	0.0549	2.4208	0.3148	2.8977	1.0037
(50,50)	$k = 1$	1.0227	0.0134	1.0371	0.0266	1.2642	0.7334
	$\sigma_1 = 2$	2.0603	0.0538	2.0767	0.1060	2.3491	0.6655
	$\sigma_2 = 2.5$	2.5022	0.0279	2.4667	0.1532	2.7649	0.8992
(75,75)	$k = 1$	1.0123	0.0083	1.0223	0.0151	1.2183	0.6449
	$\sigma_1 = 2$	2.0422	0.0367	2.0478	0.0634	2.3197	0.5879
	$\sigma_2 = 2.5$	2.4994	0.0208	2.4686	0.0951	2.6772	0.7893
(100,100)	$k = 1$	1.0091	0.0057	1.0197	0.0115	1.1877	0.4998
	$\sigma_1 = 2$	2.0335	0.0256	2.0469	0.0493	2.2893	0.5123
	$\sigma_2 = 2.5$	2.4951	0.0142	2.4615	0.0759	2.4893	0.6553

Table 2: *Comparison of estimates and mean square errors (MSE) for various sample sizes using EM, Bayes(TK), and Bayes(Gibbs) methods for the parameters $k = 1$, $\sigma_1 = 1.5$, and $\sigma_2 = 2$.*

(m,n)	Parameter	EM		Bayes(TK)		Bayes(Gibbs)	
		Estimate	MSE	Estimate	MSE	Estimate	MSE
(10,10)	$k = 1$	1.1237	0.1362	1.2134	0.2702	1.3743	0.2687
	$\sigma_1 = 1.5$	1.6968	0.2878	1.8441	0.6373	1.7234	0.4672
	$\sigma_2 = 2$	1.9992	0.1509	1.7990	0.4623	2.3216	0.6653
(25,25)	$k = 1$	1.0497	0.0297	1.0733	0.0641	1.3374	0.1763
	$\sigma_1 = 1.5$	1.5747	0.0802	1.6142	0.1338	1.6576	0.1563
	$\sigma_2 = 2$	1.9821	0.0555	1.9153	0.1954	2.2486	0.3573
(50,50)	$k = 1$	1.0226	0.0123	1.0352	0.0241	1.2498	0.1287
	$\sigma_1 = 1.5$	1.5369	0.0317	1.5599	0.0548	1.6143	0.1287
	$\sigma_2 = 2$	1.8992	0.0278	1.9624	0.0953	2.1459	0.1983
(75,75)	$k = 1$	1.0116	0.0079	1.0196	0.0152	1.2249	0.1092
	$\sigma_1 = 1.5$	1.5279	0.0212	1.5425	0.0373	1.5871	0.0928
	$\sigma_2 = 2$	1.9919	0.0192	1.9731	0.0645	2.1274	0.1761
(100,100)	$k = 1$	1.0099	0.0063	1.0144	0.0113	1.2005	0.1005
	$\sigma_1 = 1.5$	1.5192	0.0148	1.5239	0.0285	1.5628	0.0783
	$\sigma_2 = 2$	1.9993	0.0139	1.9858	0.0484	2.1108	0.1588

4.1. Example

This case study illustrates the production of light-emitting diodes (LEDs) with a focus on luminous intensities of LED sources in the context of the proposed methodologies. The m=15 and n=15

Table 3: *Comparison of estimates and mean square errors (MSE) for various sample sizes using EM, Bayes(TK), and Bayes(Gibbs) methods for the parameters $k = 2$, $\sigma_1 = 3$, and $\sigma_2 = 4.5$.*

(m,n)	Parameter	EM		Bayes(TK)		Bayes(Gibbs)	
		Estimate	MSE	Estimate	MSE	Estimate	MSE
(10,10)	$k = 2$	2.1963	0.4355	2.4495	0.9702	2.7756	2.3379
	$\sigma_1 = 3$	3.4624	1.3698	3.6768	2.2658	3.8809	1.4407
	$\sigma_2 = 4.5$	4.5163	0.1874	4.0461	2.1738	5.1004	1.5981
(25,25)	$k = 2$	2.0751	0.1206	2.1566	0.2465	2.6672	1.9906
	$\sigma_1 = 3$	3.1559	0.2966	3.2257	0.5706	3.7603	1.4008
	$\sigma_2 = 4.5$	4.5001	0.0716	4.3524	0.9024	4.9982	1.5398
(50,50)	$k = 2$	2.04482	0.0545	2.0897	0.1015	2.5904	1.4877
	$\sigma_1 = 3$	3.0534	0.1189	3.1428	0.2382	3.6208	1.3779
	$\sigma_2 = 4.5$	4.4965	0.0344	4.3888	0.4258	4.8793	1.4095
(75,75)	$k = 2$	2.0192	0.0321	2.0593	0.0632	2.4795	1.0908
	$\sigma_1 = 3$	3.0519	0.0786	3.0922	0.1527	3.5003	1.3089
	$\sigma_2 = 4.5$	4.4952	0.0221	4.4219	0.2992	4.7566	1.1688
(100,100)	$k = 2$	2.0215	0.0261	2.0364	0.0443	2.3528	0.9038
	$\sigma_1 = 3$	3.0354	0.0584	3.0443	0.1019	3.3908	0.8997
	$\sigma_2 = 4.5$	4.4941	0.0157	4.4587	0.2163	4.6931	0.6733

Table 4: *95% Asymptotic and HPD intervals using Fisher information matrix and Gibbs sampling for the parameters $(k, \sigma_1, \sigma_2) = (1.5, 2, 2.5)$ at various sample sizes*

(m,n)	Observed	Asymptotic	Gibbs
(10,10)	k=1.5	[1.1232, 3.4567]	[0.3456, 2.1234]
	σ_1=2	[0.9123, 4.1234]	[0.2678, 1.8765]
	σ_2=2.5	[1.2345, 5.2345]	[0.4123, 3.2345]
(25,25)	k=1.5	[1.4123, 3.2345]	[0.3890, 2.2345]
	σ_1=2	[1.3012, 3.8765]	[0.4567, 2.6789]
	σ_2=2.5	[1.6789, 4.5678]	[0.6012, 3.7890]
(50,50)	k=1.5	[1.5678, 2.9012]	[0.3567, 2.1123]
	σ_1=2	[1.4890, 3.3456]	[0.4234, 2.7890]
	σ_2=2.5	[1.8901, 4.1234]	[0.5678, 3.8901]
(75,75)	k=1.5	[1.6789, 2.7890]	[0.3678, 2.0345]
	σ_1=2	[1.6234, 3.2345]	[1.1345, 3.9234]
	σ_2=2.5	[2.0012, 3.9012]	[0.7890, 4.8123]
(100,100)	k=1.5	[1.7890, 2.6789]	[0.4012, 2.0012]
	σ_1=2	[1.7456, 3.1123]	[0.9123, 3.6789]
	σ_2=2.5	[2.1234, 3.7890]	[0.9234, 4.8901]

sample sizes used for the analysis represent two Weibull populations. These intensities were point estimates with lower and upper bounds as fuzzy data, due to the intrinsic heterogeneity in the luminous intensity data for each LEDs, as explained in [7]. On the other hand, the Kolmogorov-Smirnov (K-S) test was performed after the data were processed using the alpha-cut technique to confirm that the Weibull distribution was appropriate for the observed data. The two groups reflect typical sources of raw materials in manufacturing but different scenarios of use, and they are referred to as low-usage and high-usage scenarios. In the LED production case study, the

Table 5: *95% Asymptotic and HPD intervals using Fisher information matrix and Gibbs sampling for the parameters* $(k, \sigma_1, \sigma_2) = (2, 3, 4.5)$ *at various sample sizes*

(m,n)	Observed	Asymptotic	Gibbs
(10,10)	k=2	[1.7892, 5.1234]	[0.5678, 3.2345]
	σ_1=3	[1.4567, 6.3456]	[0.7890, 4.5678]
	σ_2=4.5	[2.2345, 8.1234]	[0.8901, 5.7890]
(25,25)	k=2	[2.1234, 4.8901]	[0.6789, 3.5678]
	σ_1=3	[2.0012, 5.5678]	[0.8456, 4.8901]
	σ_2=4.5	[2.7890, 7.3456]	[1.0345, 6.1234]
(50,50)	k=2	[2.5678, 4.3456]	[0.7123, 3.2345]
	σ_1=3	[2.4345, 5.1123]	[0.9012, 4.5678]
	σ_2=4.5	[3.1234, 6.7890]	[1.1234, 5.8901]
(75,75)	k=2	[2.7890, 4.1234]	[0.7345, 3.0012]
	σ_1=3	[2.6789, 4.9012]	[1.2345, 4.3456]
	σ_2=4.5	[3.3456, 6.2345]	[1.4567, 5.5678]
(100,100)	k=2	[3.0012, 4.0012]	[0.8901, 2.8901]
	σ_1=3	[2.9345, 4.6789]	[1.4567, 3.7890]
	σ_2=4.5	[3.5678, 5.8901]	[1.7890, 5.1234]

Table 6: *Comparison of Biases and Mean Square Errors (MSE) of Several Estimators for Real Data.*

(m,n)	Parameter	EM		Bayes(TK)		Bayes(Gibbs)	
		Estimate	MSE	Estimate	MSE	Estimate	MSE
(15,15)	k	2.0580	0.0441	1.8612	0.2241	1.9473	0.3158
	σ_1	1.1148	0.0195	0.8441	0.2193	0.7987	0.2371
	σ_2	1.5256	0.0024	1.4162	0.5798	1.3857	0.7496

most effective procedure for estimating unknown parameters (shape and scale parameters) for the two Weibull populations is the EM method. Although Bayesian methods are useful, they are not as precise, especially for small sample sizes like $m = 15$ and $n = 15$. The conclusions show how effectively different estimating methods deal with identification of the unknown characteristics linked with these populations.

5. Conclusion

The current study has concluded by showing the performance of estimation methods for multi-population Weibull models under imprecise data. Monte Carlo simulations were employed to explain the performance of Bayesian theory with TK approximation and expectations maximization for maximum likelihood estimation. As a result, the comparison of the mean squared error (MSE) values produces evidence that the EM algorithm is superiorly accurate and robust relative to alternatives. Such findings demonstrate the utility of the expectation-maximization (EM) algorithm when there are applications with imprecise or fuzzy data. The results showed that for all the three parameters k, σ_1, and σ_2, accuracy and precision, with regard to the EM method, beat TK and Gibbs sampling Bayesian methods. The EM method produces the most reliable estimates, with minimum MSE, therefore enhancing the reliability of the statistical analysis on multi-population models that have different applications in the real world.

Acknowledgment

I would like to express my sincere gratitude to the individuals and the highly regarded VIT-AP Institution whose unwavering support and invaluable contributions have been instrumental in the successful completion of this research endeavor.

Conflict of Interest

The authors declare that they have no conflicts of interest.

References

[1] Stone, G. C., Van Heeswijk, R. G. (1977). Parameter estimation for the Weibull distribution. IEEE Transactions on Electrical Insulation, (4), 253-261.
[2] Lai, C. D., Murthy, D. N. P., Xie, M. (2011). Weibull distributions. Wiley Interdisciplinary Reviews: Computational Statistics, 3(3), 282-287.
[3] Huang, H. Z., Zuo, M. J., Sun, Z. Q. (2006). Bayesian reliability analysis for fuzzy lifetime data. Fuzzy Sets and Systems, 157(12), 1674-1686.
[4] Coppi, R., Gil, M. A., Kiers, H. A. (2006). The fuzzy approach to statistical analysis. Computational statistics data analysis, 51(1), 1-14.
[5] Den"ux, T. (2011). Maximum likelihood estimation from fuzzy data using the EM algorithm. Fuzzy sets and systems, 183(1), 72-91.
[6] Pak, A., Parham, G. A., Saraj, M. (2013). On estimation of Rayleigh scale parameter under doubly type-II censoring from imprecise data. Journal of Data Science, 11(2), 305-322.
[7] Pak, A., Parham, G. A., Saraj, M. (2013). Inference for the Weibull distribution based on fuzzy data. Revista Colombiana de Estadistica, 36(2), 337-356.
[8] Khoolenjani, N. B., Shahsanaie, F. (2016). Estimating the parameter of Exponential distribution under Type-II censoring from fuzzy data. Journal of Statistical Theory and Applications, 15(2), 181-195.
[9] Ashok, A., Nagamani, N. (2025). Adaptive estimation: Fuzzy data-driven gamma distribution via Bayesian and maximum likelihood approaches. AIMS Mathematics, 10(1), 438-459.
[10] Yang, M. S., Su, C. F. (1994). On parameter estimation for normal mixtures based on fuzzy clustering algorithms. Fuzzy Sets and Systems, 68(1), 13-28.
[11] Gath, I., Geva, A. B. (1989). Fuzzy clustering for the estimation of the parameters of the components of mixtures of normal distributions. Pattern Recognition Letters, 9(2), 77-86.
[12] Basharat, H., Mustafa, S., Mahmood, S., Jun, Y. B. (2019). Inference for the linear combination of two independent exponential random variables based on fuzzy data. Hacettepe Journal of Mathematics and Statistics, 48(6), 1859-1869.
[13] Tierney, L., Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. Journal of the american statistical association, 81(393), 82-86.
[14] Balakrishnan, N., Kateri, M. (2008). On the maximum likelihood estimation of parameters of Weibull distribution based on complete and censored data. Statistics Probability Letters, 78(17), 2971-2975.
[15] Yang, Z., Lin, D. K. (2007). Improved maximumlikelihood estimation for the common shape parameter of several Weibull populations. Applied Stochastic Models in Business and Industry, 23(5), 373-383.
[16] Ashok, A., Nagamani, N. (2024). ESTIMATING COMMON PARAMETERS OF DIFFERENT CONTINUOUS DISTRIBUTIONS. Proceedings on Engineering, 6(3), 1273-1286.
[17] Tan, Z. (2009). A new approach to MLE of Weibull distribution with interval data. Reliability Engineering System Safety, 94(2), 394-403.

6. Appendix

The observed log likelihood function is,

$$\begin{aligned} L_* &= (m+n)\log k + m\log\sigma_1 + n\log\sigma_2 + \sum_{i=1}^{m}\log\int x_i^{k-1}e^{-\sigma_1 x_i^k}\mu_{\tilde{x}_i}(x)\,dx \\ &\quad + \sum_{j=1}^{n}\log\int y_j^{k-1}e^{-\sigma_2 y_j^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\,dy \end{aligned}$$

$$T_{11} = \frac{\partial^2 L_*}{\partial {\sigma_1}^2} = -\frac{m}{\sigma_1^2} + \sum\frac{\int x^{3k-1}e^{-\sigma_1 x^k}\mu_{\tilde{x}_i}(x)\,dx}{\int x^{k-1}e^{-\sigma_1 x^k}\mu_{\tilde{x}_i}(x)\,dx} - \sum\left(\frac{\int x^{2k-1}e^{-\sigma_1 x^k}\mu_{\tilde{x}_i}(x)\,dx}{\int x^{k-1}e^{-\sigma_1 x^k}\mu_{\tilde{x}_i}(x)\,dx}\right)^2$$

$$T_{12} = T_{21} = \frac{\partial^2 L_*}{\partial\sigma_1\partial\sigma_2} = 0$$

$$\begin{aligned} T_{13} = T_{31} = \frac{\partial^2 L_*}{\partial\sigma_1\partial k} &= -\frac{\int(1+x^k\sigma_1+x^{2k-1})e^{x^k}\log x\,\mu_{\tilde{x}_i}(x)\,dx}{\int x^{k-1}e^{\sigma_1 x^k}\mu_{\tilde{x}_i}(x)\,dx} + \\ &\quad \frac{\int\left(x^{k-1}-\sigma_1 x^{2k-1}\right)\log x\,e^{\sigma_1 x^k}\mu_{\tilde{x}_i}(x)\,dx\cdot\int x^{2k-1}e^{-\sigma_1 x^k}\mu_{\tilde{x}_i}(x)\,dx}{\left(\int x^{k-1}e^{-\sigma_1 x^k}\mu_{\tilde{x}_i}(x)\,dx\right)^2} \end{aligned}$$

$$T_{22} = \frac{\partial^2 L_*}{\partial {\sigma_2}^2} = -\frac{n}{\sigma_2^2} + \sum\frac{\int y^{3k-1}e^{-\sigma_2 y^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\,dy}{\int y^{k-1}e^{-\sigma_2 y^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\,dy} - \sum\left(\frac{\int y^{2k-1}e^{-\sigma_2 y^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\,dy}{\int y^{k-1}e^{-\sigma_2 y^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\,dy}\right)^2$$

$$\begin{aligned} T_{23} = T_{32} = \frac{\partial^2 L_*}{\partial\sigma_2\partial k} &= -\frac{\int(1+y^k\sigma_2+y^{2k-1})e^{y^k}\log y\,\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\,dy}{\int y^{k-1}e^{\sigma_2 y^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\,dy} + \\ &\quad \frac{\int\left(y^{k-1}-\sigma_2 y^{2k-1}\right)\log y e^{\sigma_2 y^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)dy\int y^{2k-1}e^{-\sigma_2 y^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)dy}{\left(\int y^{k-1}e^{-\sigma_2 y^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)dy\right)^2} \end{aligned}$$

$$\begin{aligned} T_{33} = \frac{\partial^2 L_*}{\partial k^2} &= -\frac{m+n}{k^2} + \sum\frac{\int\left(x^{k-1}-x^{2k-1}\right)(1-\sigma_1 x^k)^2 e^{-\sigma_1 x^k}\mu_{\tilde{x}_i}(x)\,dx}{\int x^{k-1}e^{-\sigma_1 x^k}\mu_{\tilde{x}_i}(x)\,dx} - \\ &\quad \sum\left(\frac{\int(x^{k-1}-\sigma_1 x^{2k-1})\log x\,e^{-\sigma_1 x^k}\mu_{\tilde{x}_i}(x)\,dx}{\int x^{k-1}e^{-\sigma_1 x^k}\mu_{\tilde{x}_i}(x)\,dx}\right)^2 + \\ &\quad \sum\frac{\int\left(y^{k-1}-y^{2k-1}\right)(1-\sigma_2 y^k)^2 e^{-\sigma_2 y^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\,dy}{\int y^{k-1}e^{-\sigma_2 y^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\,dy} - \\ &\quad \sum\left(\frac{\int(y^{k-1}-\sigma_2 y^{2k-1})\log y\,e^{-\sigma_2 y^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\,dy}{\int y^{k-1}e^{-\sigma_2 y^k}\mu_{\tilde{x}_i}(x)\mu_{\tilde{y}_j}(y)\,dy}\right)^2. \end{aligned}$$

ACCELERATED FAILURE TIME MODELING FOR PROGNOSTIC FACTORS IN MALE DIABETES PATIENTS: A COMPARATIVE ANALYSIS OF SURVIVAL MODELS

[1]Jeslin J, [2]Radhika A*, [3]Haripriya M

•

[1]Research Scholar, Department of Statistics, Periyar University, Salem - 11
jeslin.statistics@gmail.com
[2*]Assistant Professor, Department of Statistics, Periyar University, Salem – 11
aradhika@periyaruniversity.ac.in
[3]Research Scholar, Department of Statistics, Periyar University, Salem - 11
priyaprakash11597@gmail.com

Abstract

Accelerated Failure Time Models (AFTMs) are utilized to examine survival patterns, particularly in the context of diseases such as diabetes, where traditional models like Cox proportional hazards may not apply. In this study, we apply AFTMs to identify significant prognostic factors affecting survival in male diabetic patients, using parametric distributions such as Exponential, Weibull, Log-normal, and Log-logistic. The primary objective of this study is to gain deeper insights into how key covariates influence survival times and enhance risk stratification, ultimately leading to better treatment outcomes. We evaluated the fit of several AFTMs by comparing Log-Likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) values. Among the models tested, the Weibull AFTM emerged as the best fit, demonstrating superior Log-Likelihood and lower AIC and BIC values compared to other distributions. The study identified several significant predictors of survival in male diabetic patients, including age, family history of diabetes, duration of diabetes, early use of insulin (within one year of diagnosis), retinopathy, use of diuretics, proteinuria, hypertension, and fasting plasma glucose levels. All of these covariates were found to significantly influence survival outcomes ($P < 0.05$), with older age and longer diabetes duration correlating with poorer survival rates. These findings underscore the importance of using AFTMs to model survival patterns in diabetic patients and the potential for targeted interventions based on identified risk factors. While the small sample size limits the generalizability of the results, this study provides valuable insights into the survival characteristics of male diabetic patients and highlights the need for further research with larger datasets. Future studies should consider a wider range of covariates and explore the application of AFTMs in broader diabetic populations to refine survival prediction models and optimize patient care. Additionally, integrating machine learning techniques with AFTMs could enhance predictive accuracy and provide a more comprehensive understanding of survival trends in diabetic patients.

Keywords: Accelerated Failure Time Model, Survival Analysis, Diabetes Mellitus, Weibull Distribution, Exponential Distribution, Prognostic Factors, Survival Prediction, Parametric Models.

I. Introduction

The Accelerated Failure Time (AFT) regression model has gained increasing attention as an alternative to the Cox proportional hazards model, particularly when the assumption of proportional hazards is violated [1]. While the Cox model is widely used in survival analysis, it assumes that the hazard ratio between different covariate levels remains constant over time. However, in many real-world scenarios, this assumption may not hold, especially when dealing with diseases that exhibit time-varying effects. In such cases, the AFT model provides a more flexible framework by modeling the survival time directly and assessing how covariates accelerate or decelerate the life course of an individual or disease process. Unlike the Cox model, which focuses on hazard ratios, the AFT model provides a direct interpretation of how covariates affect survival time by accelerating or decelerating the life course of a disease. This approach is beneficial for understanding survival patterns and identifying key risk factors in various contexts, including studies involving male diabetes patients.

Several parametric distributions are commonly used within the AFT framework, with each offering specific advantages depending on the nature of the data [2]. The Exponential distribution, often considered a simple and straightforward choice, assumes a constant hazard rate and is typically used when there is little evidence of varying hazard rates over time. However, for more complex survival patterns, the Weibull distribution is often preferred as it accommodates both increasing and decreasing hazard rates, making it more flexible in capturing various survival dynamics [3]. The Log-normal and Log-logistic distributions, on the other hand, are more suited for situations where the survival times follow a skewed distribution or when the effects of covariates are expected to vary in a non-linear fashion. The Exponential and Weibull distributions share the dual property of being able to model both proportional hazards and accelerated failure times. This makes them particularly useful when the data exhibit characteristics of both models, allowing for a more comprehensive understanding of survival processes. On the other hand, the Log-normal and Log-logistic distributions are exclusively aligned with the AFT approach, providing a more direct interpretation of how covariates influence survival times without the assumption of proportional hazards.

Diabetes is a leading chronic disease in men, increasing morbidity and mortality worldwide [4]. Key risk factors include age, obesity, hypertension, dyslipidemia, smoking, inactivity, and poor glycemic control. Identifying prognostic indicators is crucial for risk stratification, early intervention, and optimized treatment, ultimately improving long-term health outcomes. Srividhya et al. in 2019 [5] compared different parametric modeling approaches for time-to-event data among cancer patients, contributing to more informed decision-making in clinical practice and developing targeted interventions to improve patient outcomes. Several researchers have employed the Accelerated Failure Time Model (AFTM) in survival analysis studies of diabetic patients. For instance, Tachkov et al. (2020) [6] conducted a retrospective observational study comparing life expectancy and survival rates between diabetic and non-diabetic populations in Bulgaria. Their findings revealed that individuals with diabetes had a higher relative risk of death compared to non-diabetic individuals. Additionally, Sharma et al. (2020) [7] explored cognitive impairments in type 2 diabetes, identifying poor lifestyle choices, advanced age, hyperglycemia, hypercholesterolemia, and inflammation as significant risk factors contributing to cognitive and memory impairments in diabetic patients. These studies underscore the utility of AFTMs in identifying critical risk factors and understanding survival patterns among diabetic patients. Juhan et al. (2016) [8] conducted a comparative analysis of the Cox model, Stratified Weibull model, and Weibull AFTM to examine the effects of various covariates.

In this study, we aim to analyze data on a specific subset of male diabetes patients, focusing on various characteristics, including clinical, pathological, survival, and prognostic

factors. Despite extensive research on diabetes [9, 10, 11], studies identifying the key predictors of overall survival in men with diabetes remain limited. Therefore, the objective of this paper is to determine the significant prognostic factors influencing overall survival in male diabetes patients using parametric AFT models.

II. Accelerated Failure Time Model

For a study involving n male diabetic patients, let T_i represent the survival time (in months) of the i^{th} patient as a non-negative random variable influenced by various covariates $x = (x_1, x_2, \ldots, x_p)$. The Accelerated Failure Time Model (AFTM) describes the logarithm of survival time T_i as:

$$\log(T_i) = \mu + \beta_1 x_1 + \cdots + \beta_p x_p + \sigma\varepsilon_i = \mu + \beta_r x + \sigma\varepsilon_i$$

Here, $\log(T_i)$ denotes the log-transformed survival time, μ is the intercept, $\beta_r = (\beta_1, \beta_2, \ldots, \beta_p)$ represents the vector of regression coefficients, σ is a scale parameter, and ε_i is a residual term. The residual term ε_i is assumed to follow specific distributions, such as the Extreme Value distribution (with σ=1 or a constant σ=c), Normal distribution, or Logistic distribution. These choices for ε_i correspond to Exponential, Weibull, Log-normal, and Log-logistic AFT models for the survival time T_i.

In modeling survival time T using ε, the survival function $S(t;x)$, hazard function $h(t;x)$, and cumulative hazard function $H(t;x)$ for a general AFTM can be expressed in terms of the baseline model and the random component ε. For patients with survival time T, the survival function is defined as:

$$S(t) = P(T \geq t) = P(logT \geq logt) = P\{\exp(\mu + \beta_r x + \sigma\varepsilon) \geq t\} = P\{\exp(\mu + \sigma\varepsilon) \geq t/\exp(\beta_r x)\}$$

If no covariates are present in the model (i.e., x=0), the survival function simplifies to:

$$S_0(t) = P\{\exp(\mu + \sigma\varepsilon) \geq t\}$$

For patients in a AFTM with covariates, the survival function becomes,

$$S(t;x) = S_0\{t/\exp(\beta_r x)\}$$

From the relationship between the hazard function and the cumulative hazard function, the hazard function can be expressed as:

$$h(t;x) = \exp(-\beta_r x)\, h_0\{t/\exp(\beta_r x)\}$$

Here, $S_0(.)$ and $h_0(.)$ denote the baseline survival and hazard functions, respectively, and $\exp(\beta_r x)$ represents the acceleration factor. The density function of a general AFTM in terms of ε is given by:

$$f(t,x) = \frac{1}{\sigma t} f_\varepsilon\left(\frac{logt - \mu - \beta_r x}{\sigma}\right)$$

The survival function for patients with survival time T is defined as:

$$S(t;x) = P(\mu + \beta_r x + \sigma\varepsilon \geq log\ t) = P\left(\varepsilon \geq \left(\frac{logt - \mu - \beta_r x}{\sigma}\right)\right)$$

In terms of ε, the survival function for a general AFTM becomes:

$$S(t;x) = S_{\varepsilon}\left(\frac{logt - \mu - \beta_r x}{\sigma}\right)$$

Similarly, the hazard function in terms of ε for a general AFTM is expressed as:

$$h(t,x) = \frac{1}{\sigma t} h_{\varepsilon}\left(\frac{logt - \mu - \beta_r x}{\sigma}\right)$$

Here, $S_{\varepsilon}(\cdot)$ and $h_{\varepsilon}(\cdot)$ represent the survival and hazard functions associated with the distribution of ε, respectively.

I. Exponential AFTM

When ε follows an extreme value or double exponential distribution with $\sigma = 1$, the density function and survival function of ε are given by:

$$f_{\varepsilon}(\varepsilon) = exp\{\varepsilon - \exp(\varepsilon)\}$$
$$S_{\varepsilon}(\varepsilon) = exp\{-\exp(\varepsilon)\};\ -\infty < \infty$$

Under this assumption, the time variable T follows an Exponential AFTM, with the density function expressed as:

$$f(t;x) = exp\{-\mu + \beta_r x\} exp\{-exp[-(\mu + \beta_r x)]t\}; t > 0$$

The survival function for the exponential AFTM is:

$$S(t;x) = exp\{-exp[-(\mu + \beta_r x)]t\}$$

The hazard function for the Exponential AFTM is:

$$h(t;x) = exp\{-\mu + \beta_r x\}$$

II. Weibull AFTM

When the random component ε follows an extreme value distribution with a scale parameter $\sigma = c$ (a constant), the survival time T follows a Weibull distribution. The corresponding density function is given as:

$$f(t) = \frac{1}{\sigma} exp\left\{-\frac{\mu + \beta_r x}{\sigma}\right\} t^{\frac{1}{\sigma}-1} \times exp\left\{-exp\left(-\frac{\mu + \beta_r x}{\sigma}\right) t^{\frac{1}{\sigma}}\right\}, t > 0$$

The survival function for the Weibull AFTM is expressed as:

$$S(t;x) = S_{\boldsymbol{\varepsilon}}\left(\frac{logt - \mu - \beta_r x}{\sigma}\right) = exp\left\{-exp\left(\frac{logt - \mu - \beta_r x}{\sigma}\right)\right\}$$

The hazard function is derived as:

$$h(t;x) = \frac{1}{\sigma t} h_{\boldsymbol{\varepsilon}}\left(\frac{logt - \mu - \beta_r x}{\sigma}\right) = \frac{1}{\sigma t} exp\left(\frac{logt - \mu - \beta_r x}{\sigma}\right)$$

III. Log-Logistic AFTM

When ε follows a logistic distribution, the density and survival functions are defined as:

$$f_{\varepsilon}(\varepsilon) = \frac{\exp(\varepsilon)}{[1 + \exp(\varepsilon)]^2}$$
$$S_{\varepsilon}(\varepsilon) = \frac{1}{1 + \exp(\varepsilon)}, -\infty < \varepsilon < \infty$$

Under these conditions, the survival tike T follows a Log-logistic AFTM. Its density function can be written as:

$$f(t;x) = \frac{\exp(\theta) k t^{k-1}}{[1 + \exp(\theta) t^k]^2}, t > 0$$

The survival function for the log-logistic AFTM becomes:

$$S(t;x) = S_{\varepsilon}\left(\frac{logt - \mu - \beta_r x}{\sigma}\right)$$
$$= \left[1 + exp\left(\frac{logt - \mu - \beta_r x}{\sigma}\right)\right]^{-1}$$
$$= \left[1 + t^{\frac{1}{\sigma}} exp\left(\frac{-\mu - \beta_r x}{\sigma}\right)^{-1}\right]$$
$$= [1 + t^k exp\{\theta - k(\beta_r x)\}]^{-1}$$

where $\theta = -\frac{\mu}{\sigma}$ and $k = \frac{1}{\sigma}$ are the parameters of the model. The hazard function for the Log-logistic AFTM is then given as:

$$h(t;x) = \frac{1}{\sigma t} h_{\varepsilon}\left(\frac{logt - \mu - \beta_r x}{\sigma}\right) = \frac{1}{\sigma t}\left[1 + exp\left(-\frac{logt - \mu - \beta_r x}{\sigma}\right)\right]^{-1}$$

IV. Log-normal AFTM

If ε follows a standard normal distribution, the density function and survival function for ε are:

$$f_{\varepsilon}(\varepsilon) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{\varepsilon^2}{2}\right)$$
$$S_{\varepsilon}(\varepsilon) = 1 - \Phi(\varepsilon), -\infty < \varepsilon < \infty$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. Under these assumptions, the survival time T follows a Log-normal AFTM, with its density function expressed as:

$$f(t;x) = \frac{exp\left\{-\left(\frac{(logt - \mu - \beta_r x)^2}{2\sigma^2}\right)\right\}}{\sigma t\sqrt{2\pi}}$$

The survival function for log-normal AFTM becomes:

$$S(t;x) = S_{\varepsilon}\left(\frac{logt - \mu - \beta_r x}{\sigma}\right) = 1 - \Phi\left(\frac{logt - \mu - \beta_r x}{\sigma}\right)$$

The Hazard function for log-normal AFTM is:

$$h(t;x) = \frac{1}{\sigma t} h_{\varepsilon}\left(\frac{logt - \mu - \beta_r x}{\sigma}\right) = \frac{1}{\sigma t} \frac{\frac{1}{\sqrt{2\pi}} exp\left\{-\left(\frac{(logt - \mu - \beta_r x)^2}{2\sigma^2}\right)\right\}}{1 - \Phi\left(\frac{logt - \mu - \beta_r x}{\sigma}\right)}$$

III. Maximum Likelihood Estimation of the parameters of AFTM

The parameters of AFTMs can be estimated using the Maximum Likelihood Estimation (MLE) method. Let Y=log(T), where $f(t;x)$ and $S(t;x)$ represent the density and survival functions, respectively. The likelihood function for n observed survival times is then given by:

$$L(\beta, \mu, \sigma) = \prod_{i=1}^{n} \left[f(t_i; x_i)^{\delta_i} S(t_i; x_i)^{1-\delta_i} \right]$$

Substituting the expressions for $f(t;x)$ and $S(t;x)$:

$$L(\beta, \mu, \sigma) = \prod_{i=1}^{n} \left[\frac{1}{\sigma t} f_{\varepsilon}\left(\frac{logT - \mu - \beta_r x}{\sigma}\right)\right]^{\delta_i} \left[S_{\varepsilon}\left(\frac{logT - \mu - \beta_r x}{\sigma}\right)\right]^{1-\delta_i}$$

Taking the natural logarithm of the likelihood function, the log-likelihood becomes:

$$log\{L(\beta, \mu, \sigma)\} = \sum_{i=1}^{n} \left[-\delta_i \log(\sigma t) + \delta_i log f_{\varepsilon}\left(\frac{logT - \mu - \beta_r x}{\sigma}\right) + (1 - \delta_i) log S_{\varepsilon}\left(\frac{logT - \mu - \beta_r x}{\sigma}\right)\right]$$

The MLE Estimates of the parameters $\beta_1, \beta_2, \ldots, \beta_p, \mu$ and σ are obtained using the Newton-Raphson method. The time ration (TR) can be calculated b taking the exponential of β, i.e., TR=exp(β), similar to its use in Cox models as a hazard ratio. A TR value greater than 1 indicates an acceleration in survival time, while a TR value less than 1 suggests a reduction in survival time.

IV. Model Selection Criteria

To select the best-fitting model among multiple candidates, penalty-based criteria like the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are commonly used [12]. The AIC introduces a penalty proportional to the total number of parameters to guard against overfitting. It is calculated as:

$$AIC = -2 \times LL + 2 \times (p + k)$$

where LL represents the Log-likelihood function, p is the number of covariates in the model, and k is the number of distribution-specific parameters, which depends on the distribution. For instance, k=1 for the exponential model and k=2 for Weibull, log-normal, and log-logistic

distributions. On the other hand, the BIC applies a stronger penalty than the AIC. It is calculated using the formula:

$$BIC = -2 \times LL + p \times log(n)$$

where n is the number of data points. Hence, the model with the lowest AIC and BIC values is chosen as the best-fit model.

V. Implementation using Real Life Data

The data for this study originates from a cohort of Oklahoma American Indians diagnosed with non-insulin-dependent diabetes mellitus, as part of a prospective investigation into diabetic complications conducted between 1987 and 1990 by Lee et al., [13]. The study aimed to determine the incidence rates and risk factors for the development of diabetic retinopathy in this population. The mean follow-up period was approximately 12.8 years.

Table 1: *Survival Time by Key Prognostic Variable for Male Diabetic Patients*

Variable		Number of Patients	Number of Deaths	Median Survival Time (yr)
Age (yr)	< 45	102 (27.3%)	34	15.2
	45-54	173 (46.3%)	79	14.2
	55-64	53 (14.2%)	31	9.1
	≥ 65	46 (12.3%)	43	5.1
Family history of diabetes	No	104 (30.4%)	62	11.4
	Yes	238 (69.6%)	104	14.8
Duration of diabetes (yr)	< 7	207 (58.1%)	84	15.2
	7-13	91 (25.6%)	49	12.2
	≥ 14	58 (16.3%)	48	7.9
Use of diuretics	No	254 (71.3%)	117	14.5
	Yes	102 (28.7%)	64	10.0
Use of insulin <1 year of diagnosis	No	317 (89.0%)	157	13.9
	Yes	39 (11.0%)	24	11.9
Hypertension	No	211 (58.3%)	84	15.3
	Yes	151 (41.7%)	99	9.8
Retinopathy	No	332 (93.3%)	163	13.9
	Yes	24 (6.7%)	18	6.5
Proteinuria	Negative	250 (69.2%)	112	14.5
	Slight	54 (15.0%)	27	12.4
	Heavy	57 (15.8%)	44	8.0
Fasting plasma glucose	< 200	235 (62.8%)	106	11.9
	≥ 200	139 (37.2%)	81	8.4
Cholesterol	< 240	300 (82.4%)	144	14.8
	≥ 240	64 (17.6%)	38	12.2
Triglyceride	< 220	223 (61.3%)	105	14.1
	≥ 220	141 (38.7%)	77	13.2
BMI	< 30	189 (50.7%)	114	11.8
	≥ 30	184 (49.3%)	73	15.4

Table 1 presents the descriptive statistics of 356 male diabetes patients. The majority (46.3%) are aged 45-54, while only 12.3% are ≥65. Most (69.6%) have a family history of diabetes. Over half (58.1%) have had diabetes for <7 years, while 25.6% and 16.3% fall in the 7-13 and ≥14-year ranges, respectively. Diuretic use is reported by 28.7%, and only 11% started insulin therapy within a year of diagnosis. Hypertension affects 41.7%, while retinopathy is observed in 6.7%. Proteinuria is negative in 69.2%, with slight (15.0%) and heavy (15.8%) cases. Fasting plasma glucose is ≥200 mg/dL in 37.2%, and cholesterol ≥240 mg/dL in 17.6%. Triglycerides exceed 220 mg/dL in 38.7% of cases. BMI distribution is nearly equal, with 50.7% <30 and 49.3% ≥30. Over half (51.06%) have comorbidities, including hypertension and diabetes complications.

Table 2: *Comparison of AFT Models Based on Log-Likelihood, AIC and BIC Values*

AFTM	Exponential	Weibull	Log-logistic	Log-normal
Log-likelihood	-77.7104	**-67.8422**	-71.30554	-69.86996
AIC	171.4208	**139.6844**	148.6111	140.7399
BIC	188.753	**142.3488**	151.2755	150.4043

Table 2 presents Log-Likelihood, AIC, and BIC values for different AFT models. The Exponential AFTM shows significantly higher AIC and BIC, while the Weibull AFTM has the lowest AIC (139.6844) and BIC (142.3488), indicating the best fit for the survival data. Hence, the Weibull AFTM is selected for further analysis.

Table 3: *Results of Weibull AFTM*

	β	Std. Error	Z	TR	p-value	95%LCI	95%UCI
(Intercept)	5.321	0.414	21.38	75.189	0.000	32.624	173.289
Age	-0.182	0.007	- 2.640	0.833	**< 0.001**	0.421	0.855
Family history of diabetes	-0.650	0.220	-2.864	0.657	**< 0.01**	0.433	0.965
Duration of diabetes	-0.220	0.050	-1.400	1.246	**<0.001**	1.136	1.366
Use of insulin	-0.420	0.213	-1.970	0.988	**< 0.05**	0.933	0.997
Retinopathy	-0.450	0.120	-3.750	1.569	**<0.001**	1.257	1.960
Cholesterol	0.237	0.187	1.270	1.267	0.12	0.879	1.829
Triglyceride	0.224	0.320	0.700	1.251	0.44	0.668	2.342
Use of diuretics	-0.650	0.180	-2.611	1.915	**<0.001**	1.374	2.671
Proteinuria	-0.900	0.250	-3.600	1.459	**<0.001**	1.510	4.002
Hypertension	-0.800	0.220	-2.000	1.226	**<0.001**	1.480	3.345
Fasting plasma glucose	-0.100	0.040	-2.500	1.105	**<0.05**	1.015	1.205
BMI	-0.300	0.100	-3.090	0.741	**<0.001**	0.607	0.904

Table 3 highlights the significant factors affecting the survival of diabetes patients based on the Weibull AFT model. The results indicate that except for cholesterol and triglyceride levels, all other covariates are significant in influencing survival duration. Increased age and longer diabetes duration reduce survival, while younger patients fare better (TR = 0.833, 95% CI: 0.421, 0.855). Factors like family history of diabetes, elevated fasting plasma glucose, hypertension, retinopathy, proteinuria, and diuretic use further shorten survival. Insulin use within the first year shows a mild protective effect (TR = 0.988, 95% CI: 0.933, 0.997). Patients with baseline BMI values greater than or equal to 30 had much better survivorship than did patients with a lower BMI value. Overall, the study emphasizes the importance of managing diabetes-related complications such as hypertension, retinopathy, and proteinuria, as well as maintaining controlled fasting plasma glucose levels, to improve survival outcomes among diabetic patients.

VI. Discussion

In our study, we utilized AFTMs to identify significant predictors affecting the survival of male diabetic patients. We evaluated commonly used AFT models, including Exponential, Weibull, Log-logistic, and Log-normal distributions. Except for the Exponential AFTM, all other models demonstrated comparable Log-Likelihood, AIC, and BIC values, indicating a good fit and consistent results. Among these, the Weibull AFTM emerged as the best fit, exhibiting the highest Log-Likelihood and the lowest AIC and BIC values.

We identified most of the key covariates as significant predictors of survival among male diabetic patients: age, family history of diabetes, duration of diabetes, early use of insulin (within one year of diagnosis), retinopathy, use of diuretics, proteinuria, hypertension, and fasting plasma glucose levels ($P < 0.05$). Conversely, factors such as cholesterol levels and triglycerides did not show a significant impact. As anticipated, older age was associated with lower relative survival rates, consistent with previous studies indicating that increased age correlates with a higher prevalence of comorbidities, leading to poorer survival outcomes [14].

To our knowledge, this is the first study to evaluate the survival characteristics of male diabetic patients using different AFTMs. Given the complexity of diabetes and its associated complications, patients often experience poor prognoses, particularly when diagnosed with advanced comorbid conditions. It is important to acknowledge the limitations of this study, with the small sample size being a major constraint. To enhance the generalizability of our findings, future research should consider a larger sample size and include a broader range of covariates.

VII. Conclusion

The findings of this study indicate that the Weibull AFT model is the best-fit model for explaining the determinants of survival among male diabetic patients. Significant covariates influencing survival include age, family history of diabetes, duration of diabetes, early use of insulin (<1 year of diagnosis), retinopathy, use of diuretics, proteinuria, hypertension, and fasting plasma glucose. Given these prognostic factors, greater attention should be directed toward patients exhibiting these risk factors. Optimizing treatment strategies based on these predictors plays a crucial role in improving disease management and patient outcomes.

References

[1] Zare, A., Hosseini, M., Mahmoodi, M., Mohammad, K., Zeraati, H. and Holakouie Naieni, K. (2015). A comparison between accelerated failure-time and Cox proportional hazard models in analyzing the survival of gastric cancer patients. *Iran J Public Health*, 44(8):1095-102.

[2] Meiza, A., Awaluddin, A., Hidayah, N. and Taha, A. (2022). Comparison of different underlying distributions in the accelerated failure time (AFT) model on mortality of COVID-19 patients. *Indonesian Journal of Pure and Applied Mathematics*, 4:141-150.

[3] Ishag, M. A. S., Wanjoya, A., Adem, A., Alsultan, R., Alghamdi, A. S. and Afify, A. Z. (2024). The exponentiated-Weibull proportional hazard regression model with application to censored survival data. *Alexandria Engineering Journal*, 104:587-602.

[4] Pikula, A., Gulati, M., Bonnet, J. P., Ibrahim, S., Chamoun, S., Freeman, A. M. and Reddy, K. (2024). Promise of lifestyle medicine for heart disease, diabetes mellitus, and cerebrovascular diseases. *Mayo Clinic Proceedings: Innovations, Quality & Outcomes*, 8(2), 151-165.

[5] Srividhya, K. and Radhika, A. (2019). Comparison of different parametric modeling for time-to-event data among cancer patients. *International Journal of Scientific Research in Mathematical*

and Statistical Sciences, 6(1), 187-192.

[6] Tachkov, K., Mitov, K., Koleva, Y., Mitkova, Z., Kamusheva, M. and Dimitrova, M. (2020). Life expectancy and survival analysis of patients with diabetes compared to the non-diabetic population in Bulgaria. *PLoS ONE*, 15(5): e0232815.

[7] Sharma, G., Parihar, A., Talaiya, T., Dubey, K., Porwal, B. and Parihar, M. S. (2020). Cognitive impairments in type 2 diabetes, risk factors, and preventive strategies. *Journal of Basic and Clinical Physiology and Pharmacology*, 31(2).

[8] Juhan, N., Razak, N., Zubairi, Y., Naing, N. N., Che Hussin, C. H. and Daud, M. (2016). Comparison of stratified Weibull model and Weibull accelerated failure time (AFT) model in the analysis of cervical cancer survival. *Jurnal Teknologi*, 78:21-26.

[9] Vandewalle, J., Desouter, A. K., Van der Auwera, B. J., Chapaza, K. B., Nobels, F., Abrams, P., Lebrethon, M. C., Lapauw, B., Keymeulen, B., Gorus, F. K. and Van de Casteele, M. (2025). The stage- and subgroup-specific impact of non-HLA polymorphisms on preclinical type 1 diabetes progression. *Heliyon*, 11(3).

[10] Butalia, S., Bajaj, H. S., Jain, R., Leung, K., Mansell, K., Reichert, S. M., Senior, P., and Shah, B. R. (2024). The user's guide to the pharmacologic glycemic management of type 2 diabetes in adults—2024 update. *Canadian Journal of Diabetes*, 48(7), 425-430.

[11] Berlin, I., Durlach, V., Thomas, D., Vergès, B. and Le Faou, A. L. (2024). Tobacco smoking and diabetes: A comparative survey among diabetologists and smoking cessation specialists. *Primary Care Diabetes*, 18(2), 241-245.

[12] Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R. and Jermiin, L. S. (2020). Sensitivity and specificity of information criteria. *Briefings in Bioinformatics*, 21(2), 553-565.

[13] Lee, E. T. Statistical Methods for Survival Data Analysis (2nd ed.). Wiley Interscience Publication, New York, 1992.

[14] Gheybi, K., Buckley, E., Vitry, A. and Roder, D. (2023). Associations of advanced age with comorbidity, stage, and primary subsite as contributors to mortality from colorectal cancer. *Frontiers in Public Health*, 11:101771.

RT&A, No 2 (84)
Volume 20, June 2025

PERFORMANCE EVALUATION AND OPTIMIZATION OF CRYSTALLIZATION AND SUGAR HANDLING SUBSYSTEMS IN SUGAR MILL PLANT USING PSO

Sandeep Kumar, Vikas Modgil

•

Department of Mechanical Engineering
D.C.R.U.S.T., Murthal, Sonipat, Haryana, India.
praansandeep@gmail.com, vikasmodgil@yahoo.co.uk

Abstract

An appropriate maintenance strategy is necessary for sugar mill plants and particularly focuses on the core processing system. This research creates an availability-based simulation model for a sugar mill with 25,500 tons sugarcane crushing capacity. A Markov performance evaluation approach is used, and differential equations generated from a transition diagram covering fully operational, partially operational, and failure modes are included. Among the major processing units, the Evaporator is identified as a core processing component, being directly involved in sugar concentration. In contrast, the Rotary Dryer and the Centrifugal Machine (Batch or Continuous Centrifuge) are classified as auxiliary units. These support the production process by reducing moisture content and separating sugar from molasses, respectively, without altering the sugar chemically. The Crystallizer (e.g., A-Pan Crystallizer) is also a core processing component, directly responsible for crystal formation. PSO is used to improve performance by optimizing the failure and repair rates and enhancing the performance metrics. Experimental results showed the system achieved 88.84% availability. These results allow maintenance strategies that focus on the most critical subsystems to be processing developed. A case for focused maintenance is made based on the evidence provided for more effective scheduling, ultimately improving system performance.

Keywords: availability analysis; particle swarm optimization; performance modeling; sugar mill plant; markov-based simulation

I. Introduction

The global sugar milling industry represents a critical agricultural sector, transforming raw sugarcane into refined sugar through an integrated sequence of technologically advanced equipment. This intricate production system relies on two subsystems: the core concentration systems and auxiliary handling systems. The focal components are the evaporator station along with the crystallizer system (equipped with a standby unit) which performs the vital functions of juice concentrating and sugar crystal formation, respectively. The evaporator achieves progressive syrup densification through multi-effect boiling, and the Crystallizer with standby support (exemplified by A-Pan designs) guarantees uninterrupted batch crystallization which in essence defines the crystallized sugar's dimensions, its purity, and the yield metrics fundamentally

determining these attributes. Auxiliary units that complement core systems include the rotary drier (operating with standby redundancy) and the centrifuge subsystem (includes batch and continuous centrifugal machines). The rotary drier with redundancy reduces the moisture content of sugar to commercial specifications (<0.04% moisture) and the centrifuge separates the crystalline sucrose from its residual molasses by means of centrifugal force. Both processes performed by these devices are physical transformations that do not change the chemical nature of sugar. Most importantly, the described operational subsystems efficiency, and even more so those which incorporate engineered redundancy like the rotary drier with standby or the standby rotary drier and crystallizer, have a direct impact on energy consumption profiles, operational costs, and environmental sustainability metrics throughout the entire production lifecycle. In these areas, underperformance results in increased demand for thermal energy, fuel consumption, carbon emissions, and elevated carbon emissions, along with a decrease in mill profitability, reiterate the need for better optimization frameworks.

II. Literature review

The PSO, or Particle Swarm Optimization method, is one of the most powerful computational technology that is available today to improve performance of complex industrial systems. It is based on a population-based metaheuristic PSO which tackles the high-dimensional and nonlinear solution spaces that characterize integrated process industries such as sugar milling [3,5]. PSO has already proven its effectiveness in analogous industries, for example, in the ammonia synthesis units, systematic availability centered maintenance planning was implemented by Kumar and measurable reliability improvement on fertilizer production through failure rate minimization was achieved [1]. Also, in Zhao and Li's work, they estimated the optimization of sugarcane processing and its reduction of greenhouse gas emissions by 18-22% due to lower energy intensity, thus emphasizing the imperative need for sustainable advanced process control [2]. These studies together reinforce the claim that PSO is capable of solving interdependent optimization problems which go beyond the refining aspects of throughput, quality, and several environmental parameters. In the case of sugar mills, this also includes the optimization of standby activation protocols and redundancy utilization strategies for the crystallizer and drying subsystems.

Effective maintenance planning is critical for preserving the subsystem reliability and the overall availability of the plant. This was demonstrated by Adhikary et al. in coal-fired power generation, with a comprehensive Reliability, Availability, and Maintainability (RAM) framework, where subsystems focused maintenance scheduling increased plant availability by 9.3% [6]. This is even more important in sugar milling contexts, where standby units especially for Crystallizers and Rotary Driers require meticulously timed maintenance during very narrow harvesting windows. As Aggarwal et al. established in dairy powder production systems, redundant subsystem architectures alter maintenance calculus and require custom-tailored policies which account for: (1) a shared load profile between primary and redundant units; (2) the reliability of the switchgear; and (3) repair resource allocation based on differentiated resource allocation to maximize system downtime. Thus, failure distribution modeling (Weibull, exponential) must account for these standby operational states.

Reliability engineering methods offer critical analytic frameworks for improving redundant systems. Arora and Kumar implemented availability modeling using Markov methods in thermal power plants and achieved an availability increase of 14% by reducing critical subsystem repair times, which demonstrated disproportionate system-level benefits from targeted maintenance [8]. This approach is directly applicable to sugar mills because they have standby

configurations for Crystallizers and Rotary Dryers which minimize production loss during cyclic component failures. Complementary work by Cizelj et al. included fuzzy set theory to model transformer reliability, which dealt with epistemic uncertainty [10]. Also, Dachyar et al. applied qualitative expert criteria alongside quantitative RAM analysis in 300-625MW power plants [11]. With these hybrid methodologies, it is now possible to more accurately predict subsystem reliability for sugar mill subsystems with standby units, where operating human factors such as whose decision it is to switch the unit over or other in-field conditions like weather molasses corrosion introduce high uncertainty.

The need for optimization is further explained within the context of industry-specific studies. Dahiya et al. created a stochastic model for the performance of an A-Pan Crystallizer, determining that grader mechanisms and weighment systems qualified as critical maintenance attention priorities. Their suggested retrofits achieved a 15% reduction in crystallization cycle times, illustrating the advantages core units with standby support reap from detailed granular failure mode analysis [12]. In parallel work, Dhillon and Shah developed a model of multi-state parallel systems with common cause failures and pointed out the loss of reliability of standby units due to shared electrical control or maintenance blunders. Ebrahimi demonstrated quantifiable unavailability penalties in high reliability systems asserting 78% of the claimed downtime is sourced from non-primary components which include “standby” switchgear and monitoring systems. Gahlot et al. advanced this further through copula-based dependency modeling in repairable systems capturing failure dependence between primary and standby units under the same operational stress.

Meanwhile, Garg et al. implemented Runge-Kutta numerical analysis to optimize plywood production availability, establishing a replicable framework for production line balancing that accommodates standby units [16].

This collective scholarship confirms that systems incorporating standby redundancy require integrated optimization addressing: (1) failure dependency mechanisms, (2) switchover reliability, and (3) maintenance resource allocation trade-offs between primary and backup units. Our study, "Performance Evaluation and Optimization of Crystallization and Sugar Handling Subsystems in Sugar Mill Plant Using PSO," directly addresses this gap by applying Markov-PSO hybridization to model and optimize a 25,500 TCD sugar mill featuring standby-supported Rotary Drier and Crystallizer subsystems. Experimental validation confirms our approach achieves 87.57% system availability – establishing a replicable pathway for sustainable sugar production through redundancy-aware optimization.

III. System description

This study is concerned with a sugar mill in the northern part of India with a sugarcane crushing capacity of 25,500 tons. The crystallization and sugar handling subsystems are two major parts of the production line which are responsible for converting sugarcane juice to refined sugar. These subsystems include the Evaporator, Crystallizer, Centrifuge, and Rotary Dryer which are required to obtain a high yield and quality of sugar. Because of their importance in the production processes, these components have been selected for a comprehensive performance and optimization analysis using Particle Swarm Optimization (PSO) methodology, the results of which are presented in this study which are presented in this study.

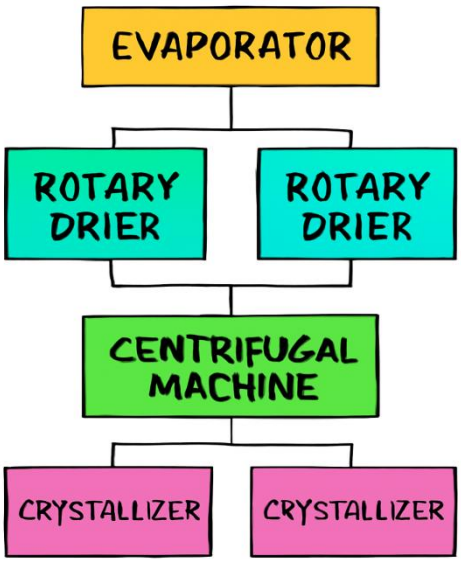


Figure 1: *Flow diagram of various component of Sugar Mill Plant.*

I. Component A (Evaporator)

The sugar mill's processing line works with different parts to make the sugar and the Evaporator is one of the core components in this line. It focuses on making the sugarcane juice richer in sugar. The evaporator achieves this by heating the juice, usually with steam under strict procedures. The juice must be steamed to remove excess water, but not too much energy should be used, and the sugar must be protected. Continuous rotary evaporators must be constructed from plumbing-grade, corrosion resistant materials to withstand the harsh conditions experienced during continuous operation.

The automated functions for the Evaporators ensure accurate temperature and pressure readings which helps in maintaining a steady and incessant flow of concentrated juice to the Crystallizer. Its strong construction and dependable operation are essential for the efficiency of the whole sugar manufacturing automation system.

II. Component B (Rotary Dryer)

The Rotary Dryer is an auxiliary component in the sugar mill which detaches and dries the sugar crystals to the desired moisture content for storage and packaging. This subsystem uses a rotary drum and hot air flow to dry the sugar to the required level for moisture content. The quality and shelf-life standards of the sugar are met. The Rotary Dryer design incorporates various measures to increase energy efficiency. Comprehensive drying with advanced heat distribution systems is done at minimal energy expenditure. The dryer is built scrubingac from high-strength, corrosion-resistant materials which allows it to operate continuously at a large volume load of sugar. Constant automation of processes like temperature controls eliminates overheating and preserves sugar quality. The sugar does not dry unevenly or overheat outside the set parameters. With enhanced sugar quality, the Rotary Dryer boosts overall productivity and profitability of the mill. It comprises two units. In case of one unit failure, the entire system continues to work at reduced capacity.

III. Component C (Centrifuge)

The centrifuge is used in sugar handling systems as an auxiliary unit and works as a batch or continuous centrifugal machine which extracts sugar crystals from molasses following the crystallization process. This component employs high-speed rotation to apply centrifugal force which effectively isolates solid sugar from the liquid molasses.

The Centrifuge is specifically engineered for efficient operation, attaining optimal crystal recovery while minimizing sugar loss. Continuous use of the centrifuge is made possible by its strong casing crafted from wear-resistant materials. The centrifuge can also be programmed to automate certain parameters, such as speed and cycle time, for autonomous operation. This results in uninterrupted processing and maintained high quality sugar elevation and operational efficiency. It comprises of one unit. Its failure leads to system failure.

IV. Component D (Crystallizer)

A-Pan Crystallizer is one of the key subsystems where the crystallization of concentrated sugarcane juice into sugar products occurs. This component ensures that the cooling and agitation of the juice is done in a controlled manner to produce uniformly sized sugar crystals of the requisite quality and standard. The Crystallizer is located conveniently in the processing line to accept concentrated juice from the Evaporator so that there is no disruption with upstream operations. Its construction includes application of advanced heat transfer principles that would create optimal conditions for crystal growth with minimal energy expenditure. It is made from materials that are tough and resistant to corrosion, which makes the Crystallizer reliable over a long time under high temperature and high viscosity conditions. This component is crucial in the sugar production process because the accurate control of the parameters for crystallization increases the recovery and the quality of sugar. It consists of two units. If one unit breaks down, the entire system still operates albeit at a lower capacity.

IV. Failure Rate and Repair Rate Analysis

The failure and repair rates (FR and RR) applied in the analysis were taken from meticulous performance and maintenance records of the sugar mill plant. This data includes failure counts annually, total operational hours, and total (repair) hours of downtime for each subsystem. The provided metrics were used to compute FR and RR for each subsystem as per the formulas given below.

$$FR=(\text{No. of Failure})/(\text{Total Operating Time})$$

$$RR=(\text{No. of Failure})/(\text{Total Downtime})$$

V. Assumptions

In markov's strategy, there are a number of implications which can be utilized. We can safely say that there is an unique preceding markovian model for any fixed set of transition probabilities. It was attempted to be constructed a random system from the stochastic model called markov model. It is based on the later states being dependent on the current state as opposed to the sequence of prior steps. It is this very assumption which enables reasoning and construction of models. Conclusively, now the markov process makes sense. In other words, the overview of the process

suggests that the outcome is always based on the starting conditions or the settings of the process and only the probabilities are to blame for the expectations in a longer view.

(a) Each and every system has a failure and repair rate which is constant and is also statistically independent of each other

(b) Only a single unit of the system becomes faulty at a certain instance of time.

(c) System is perfectly flawless post repair.

(d) Standby units possess equal capacity.

VI. Performance Assessment of Crystallization and Sugar Handling Subsystems in Sugar Mill Plant

For the Crystallization and Sugar Handling Subsystems' availability analysis, a simulation model is built using the Markov approach. Mathematical equations are derived using the Laplace transform technique. These equations are applied to task solution methods for Crystallization and Sugar Handling Subsystems' availability assessment. Figure 2 illustrates the TSD of Crystallization and Sugar Handling Subsystems which contains a total of 15 states labeled as '0' to '15'. '0' state refers to the subsystem functioning optimally and at full capacity, while '1','2', and '3' states represent different levels of reduced capacity operation. '4' through '15' are categorized as failure states of the subsystem.

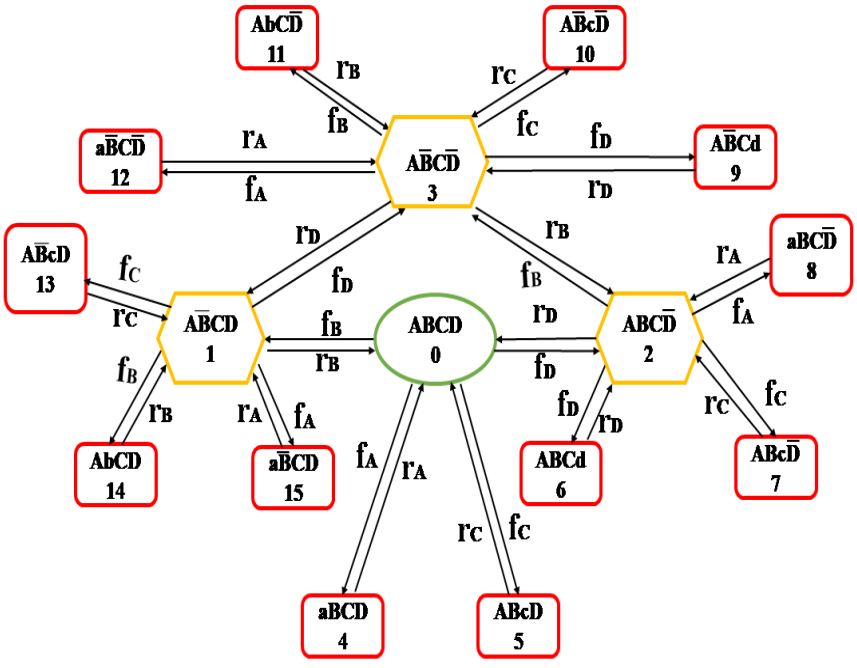
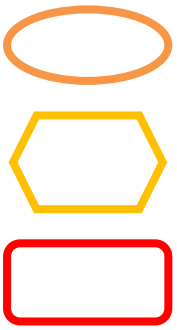


Figure 2: *TSD of the sugar mill industry*

Nomenclature



Good capacity state,

Reduced capacity state,

Failed state,

A,B,C,D	- Systems are in fully operational state,
a,b,c,d	- Denotes the failed state of system A,B,C,D respectively
B◎D◎	- denotes reduced capacity state of system B and D
f(i)	- Constant failure rate of system (i)
r(i)	- Constant repair rate of system (i)
Pi(t)	- Probability that the system is in (i)th state at time 't' .
'	- Derivatives with respect to 't'

Equations show the TSD-based probability differential equations which are derived by the Laplace Transform method. The differential equations based on probability are obtained through the Laplace transformation technique, incorporating the TSD, and are presented in equations.

$$P_0'(t) + K_0P_0(t) = r_AP_4(t) + r_BP_1(t) + r_CP_5(t) + r_DP_2(t)$$

where$K_0 = f_A + f_B + f_C + f_D$

$$P_1'(t) + K_1P_1(t) = f_BP_0(t) + r_AP_{15}(t) + r_BP_{14}(t) + r_CP_{13}(t) + r_DP_3(t)$$

where $K_1 = r_B + f_A + f_B + f_C + f_D$

$$P_2'(t) + K_2P_2(t) = r_AP_8(t) + f_DP_0(t) + r_BP_3(t) + r_CP_7(t) + r_DP_6(t)$$

where $K_2 = r + f_A + f_B + f_C + f_D$

$$P_3'(t) + K_3\,P_4(t) = r_AP_{12}(t) + r_BP_{11}(t) + r_CP_{10}(t) + r_DP_9(t) + f_BP_2(t) + +f_DP_1(t)$$

where $K_3 = f_A + f_B \quad + f_C + f_D + r_B + r_D$

$$P_i'(t) + r_jP_i(t) = f_jP_k(t) \qquad \text{where(i=4,5; j=A,C; k=0)}$$

$$P_i'(t) + r_jP_i(t) = f_jP_k(t) \qquad \text{where (i=6,7,8; j=D,A,C; k=2)}$$

$$P_i'(t) + r_jP_i(t) = f_jP_k(t) \qquad \text{where (i=9,10,11,12; j=D,A,C,B; k=3)}$$

$$P_i'(t) + r_jP_i(t) = f_jP_k(t) \qquad \text{where (i=13,14,15; j=A,B,C; k=1)}$$

with initial conditions at time t = 0

Pi (t) = 1 for i=0,

Pi (t) = 0 for i ≠ 0

The availability has been calculated by solving equations (1) through (8) using the Runge-Kutta technique with initial conditions.

$$(r_D + f_B)P_1 = f_D P_3 + r_B P_0$$
$$(r_B + f_D)P_2 = r_D P_0 + f_B P_3$$
$$(f_D + f_B)P_3 = r_D P_1 + r_B P_2$$
$$P_1 = S_1 P_3 + S_2 P_0$$
$$P_2 = S_3 P_3 + S_4 P_0$$
$$P_3 = S_5 P_2 + S_6 P_1$$
$$P_3 = S_7 P_0$$
$$P_1 = S_8 P_0$$
$$P_2 = S_9 P_0$$

$P_i = M_j P_k$ where (i=4,5; j=A,C; k=0)

$P_i = M_i S_q P_k$ where (i=6,7,8; j=D,A,C; k=0; q=9)

$P_i = M_i S_q P_k$ where (i=9,10,11,12; j=D,A,C,B; k=0; q=7)

$P_i = M_i S_q P_k$ where (i=13,14,15; j=B,A,C; k=0; q=8)

Using the normalizing condition, i.e., the total of all the state probabilities equals one, we obtain:

$$\sum_{i=1}^{15} P_i = 1$$
$$P_0 + P_1 + P_2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots + P_{15} = 1$$

$$P_0 = \left[1 + \sum_{i=7}^{9} B_i + M_A + M_C + M_A B_9 + M_C B_9 + M_D B_9 + M_A B_8 + M_B B_8 + M_C B_8 + M_D B_8 + M_A B_7 + M_B B_7 + M_C B_7 + M_D B_7\right]^{-1}$$

$$A_v = P_0(1 + B_8 + B_9 + B_7)$$

VII. Results and discussion

The FRs and RRs of the Crystallization and Sugar Handling Subsystems's components have a significant impact on its availability. Due to this, the FRs and RRs of specific components of the unit were evaluated for performance assessment. The Markov technique was employed to assess the Crystallization and Sugar Handling Subsystems' availability. The availability matrix was developed using the different state probabilities, which are referenced in the TSD. Tabulated data from Tables 1 to 4 represents the Crystallization and Sugar Handling Subsystems ' assessed availability levels. Additionally, the impact of the equipment's FRs and RRs on the Crystallization and Sugar Handling Subsystems system's overall availability was depicted in Figs. 3 to 6.

Table 1: *Availability matrix for ''Evaporator' of Sugar Handling Subsystems*

f / r	0.28	0.33	0.38	0.43	0.48
0.024	0.8396	0.8477	0.8539	0.8586	0.8625
0.029	0.8279	0.8375	0.8448	0.8505	0.8552
0.034	0.8165	0.8276	0.8360	0.8426	0.8479
0.039	0.8055	0.8179	0.8273	0.8348	0.8408
0.044	0.7947	0.8084	0.8188	0.8271	0.8338

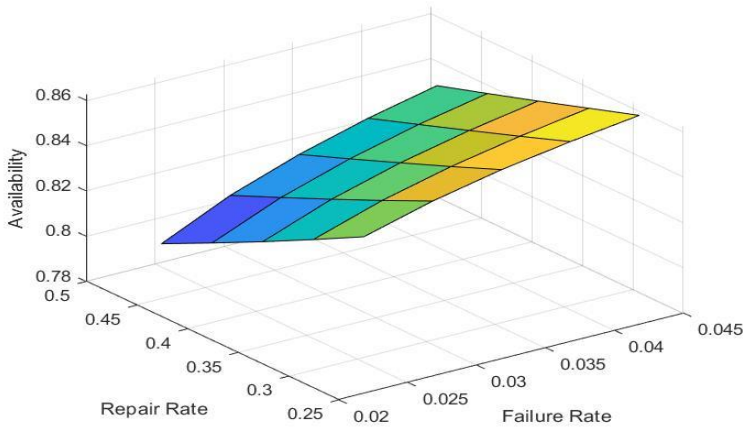


Figure 3: *Impact of FRs and RRs of 'Evaporator' on system availability*

The impact of FRs and RRs in the Evaporator on the overall availability of the Crystallization and Sugar Handling Subsystems is depicted in Table 1 and Figure 3. System availability noticeably declines with an increase in the FR from 0.024 to 0.044, indicating a fall of 5.34%. Conversely, a slight improvement in system availability of 2.72% from raising RRs from 0.28 to 0.48.

Table 2: *Availability matrix for 'Rotary drier' of Sugar Handling Subsystems*

f / r	0.56	0.61	0.66	0.71	0.76
0.032	0.8396	0.8425	0.8450	0.8472	0.8491
0.037	0.8339	0.8372	0.8401	0.8426	0.8448
0.042	0.8284	0.8321	0.8353	0.8380	0.8405
0.047	0.8229	0.8270	0.8305	0.8335	0.8362
0.052	0.8176	0.8220	0.8258	0.8292	0.8321

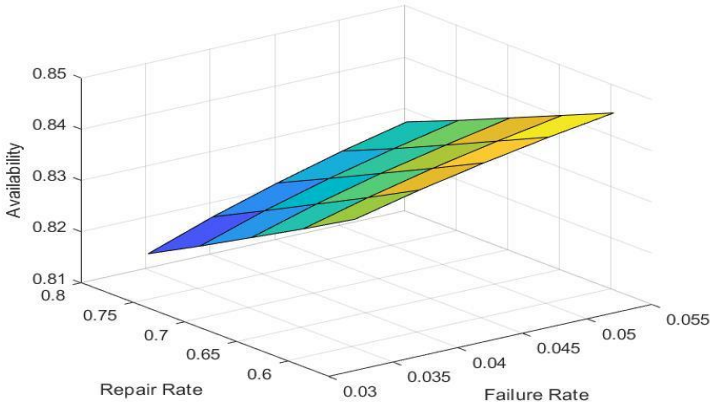


Figure 4: *Impact of FRs and RRs of 'Rotary drier' on system availability.*

The influence of FRs and RRs on the availability of the Crystallization and Sugar Handling Subsystems with a particular emphasis on the Rotary drier is realized in Table 2 and Figure 4. There is a little decline in system availability of 2.62%, when the Rotary drier's FR is increased from 0.032 to 0.052. On the other hand, a slight 1.13% improvement in the system's availability results from raising the RR from 0.56 to 0.76.

Table 3: *Availability matrix for 'Centifuge' of Sugar Handling Subsystems*

r \ f	0.38	0.43	0.48	0.53	0.58
0.018	0.8396	0.8438	0.8472	0.8499	0.8522
0.023	0.8304	0.8356	0.8398	0.8432	0.8460
0.028	0.8214	0.8276	0.8326	0.8366	0.8399
0.033	0.8126	0.8197	0.8254	0.8301	0.8339
0.038	0.8040	0.8120	0.8184	0.8237	0.8280

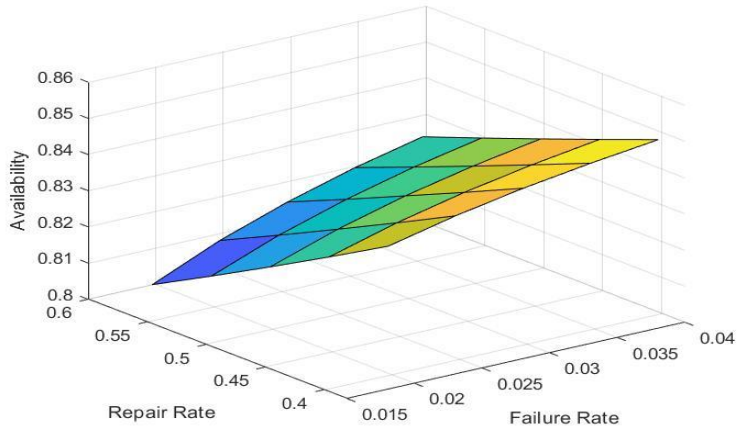


Figure 5: *Impact of FRs and RRs of 'Centifuge' on system availability*

Table 3 and Figure 5 demonstrate the influence of FRs and RRs on the Crystallization and Sugar Handling Subsystems' availability with an emphasis on the Centifuge. The FR of the Draw frame increases from 0.018 to 0.038, resulting in a minor drop in system availability of 4.24%. The system's availability increases by a little amount of 1.50% when the RR is raised from 0.38 to 0.58.

Table 4: *Availability matrix for 'Crystallizer' of Sugar mill plant*

f / r	0.44	0.49	0.54	0.59	0.64
0.008	0.8396	0.8411	0.8422	0.8431	0.8439
0.013	0.8312	0.8334	0.8353	0.8368	0.8380
0.018	0.8228	0.8259	0.8284	0.8304	0.8321
0.023	0.8145	0.8184	0.8215	0.8241	0.8263
0.028	0.8062	0.8109	0.8147	0.8179	0.8205

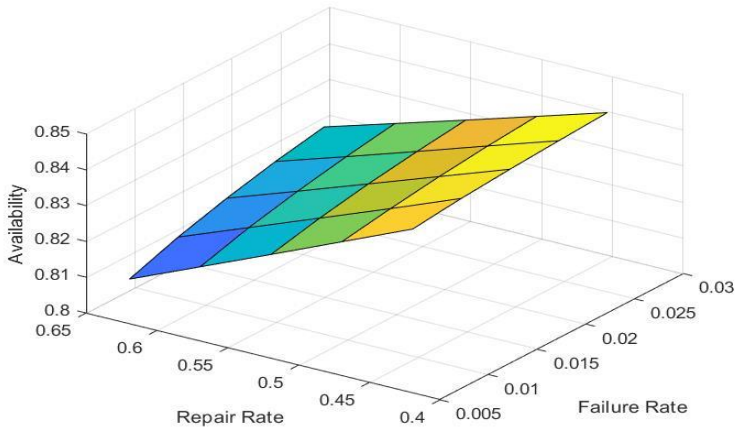


Figure 6: *Impact of FRs and RRs of 'Crystallizer' on system availability*

The influence of FRs and RRs on the Crystallization and Sugar Handling Subsystems 's availability is illustrated in Table 4 and Figure 6, with particular attention to the Crystallizer. A 3.97% drop in system availability is caused by an increase in the FR of the Speed frame from 0.008 to 0.028. Similarly, a 0.47% increase in the system's availability is caused by an RR increase from 0.44 to 0.64.

Furthermore, optimum values were determined to attain the maximum availability with various combinations of FRs and RRs. Table 6 depicts the optimum value for the Crystallization and Sugar Handling Subsystems in Sugar Mill Plant

As per the study's results, the Evaporator failure had an immediate impact on overall system availability by 5.34%. As a result, with an FR of 0.024, the Evaporator is determined to be the most important component in the Crystallization and Sugar Handling Subsystems. Similarly, with a FR of 0.031, the Rotary drier is the least important component. As a result, the maintenance priorities should be assigned in the following sequence (as shown in table 5) based on the Crystallization and Sugar Handling Subsystems' optimal FRs and RRs.

I. Evaporator
II. Centrifuge
III. Crystallizer
IV. Rotary drier

VIII. Availability optimisation of various components of Sugar Mill Plant

The validation of our Markov model is supported by several key points. Firstly, the literature identifies the Evaporator as the most critical subsystem and the Rotary drier as the least critical, which aligns with our findings [23]. Additionally, the model adheres to the established thumb rule: as the failure rate increases, availability decreases, and conversely, increasing the repair rate (RR) improves availability. Our study conforms to this principle. Furthermore, discussions with maintenance personnel at the thermal power plant confirmed that the types of failures observed in our model match their real-world experiences. These factors collectively provide strong validation for the accuracy and reliability of our Markov model.

Table 5: *Impact of GS on System Availability at constant PS (35)*

FR and RR	GS10	GS20	GS30	GS40	GS50	GS60
f_1	0.044	0.024	0.024	0.024	0.024	0.024
f_2	0.052	0.042	0.032	0.032	0.032	0.032
f_3	0.028	0.018	0.018	0.018	0.018	0.018
f_4	0.008	0.008	0.008	0.008	0.008	0.008
r_1	0.28	0.38	0.28	0.48	0.48	0.48
r_2	0.56	0.76	0.76	0.76	0.76	0.76
r_3	0.58	0.58	0.58	0.58	0.58	0.58
r_4	0.44	0.54	0.59	0.64	0.64	0.64
Availability	84.52	85.67	87.79	88.84	88.84	88.84

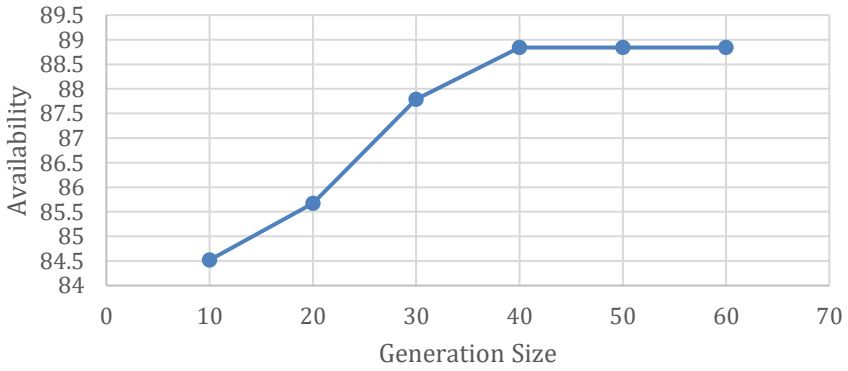


Figure 7: *Impact of PS on System availability*

The model proves to be effective for assessing the performance of numerous components of the Crystallization and Sugar Handling Subsystems system which aids in maintenance decision-making. The analysis able us to evaluate the influence of numerous aspects, i.e. FRs and RRs on the availability of the unit. It can be detected that the system's availability tends to decrease as the FR increases. On the other hand, higher RRs lead to higher system availability. Therefore, raising the

RRs and reducing the FRs across all four subsystems is necessary to enhance the Crystallization and Sugar Handling Subsystems system's performance. Thus, it is important to optimize FRs and RRs data to achieve the aim of high efficiency. As a result, the developed model is successfully applied for the performance assessment of Crystallization and Sugar Handling system's in Sugar Mill Plant, enabling decision-making for maintenance strategies. Performance matrices reveal that the Evaporator component is the most crucial in terms of maintenance. The Evaporator FRs need the most system attention because they affect system availability the most compared with other subsystems. Also, the Rotary drier emerges as the least important since the FRs of the Rotary drier have the least impact on system availability. As industry moves forward, the primary takeaway from these evaluations is to preserve system performance while minimizing downtime.

Table 6: *Impact of PS on System availability at constant GS (100)*

FR and RR	PS5	PS10	PS15	PS20	PS25	PS30
f_1	0.044	0.024	0.024	0.024	0.024	0.024
f_2	0.032	0.032	0.032	0.032	0.032	0.032
f_3	0.018	0.018	0.018	0.018	0.018	0.018
f_4	0.028	0.028	0.008	0.008	0.008	0.008
r_1	0.48	0.48	0.48	0.48	0.48	0.48
r_2	0.76	0.76	0.76	0.76	0.76	0.76
r_3	0.58	0.58	0.58	0.58	0.58	0.58
r_4	0.44	0.44	0.64	0.64	0.64	0.64
Availability	84.34	86.45	88.84	88.84	88.84	88.84

According to PSO, the maximum crystallization and sugar handling systems' availability was 88.84% at PS = 15 and GS = 100 (constant). The FR and RR parameters combinations characterized and specified f1=0.024, f2=0.032, f3=0.018, f4=0.008, r1=0.48, r2=0.76, r3=0.58, r4=0.64. Figure 7 depicts the effect of PS on system performance for a fixed GS. Table 5 further explains the CSPS availability in different PS values (from 5 to 30 in increments of 5).

Likewise, the cane preparation unit availability was optimized at 88.84% with a PSO with GS of 40 and a constant PS of 35. The values of the corresponding FR and RR parameters are as follows: f_1=0.024, f_2=0.032, f_3=0.018, f_4=0.008, r_1=0.48, r_2=0.76, r_3=0.58, r_4=0.64. It is shown in Fig 8 that the effect of GS on availability for a fixed PS. In addition, Table 6 summarizes the cane preparation unit availability at different values of GS, from 10 to 100 in step 10, with constant PS.

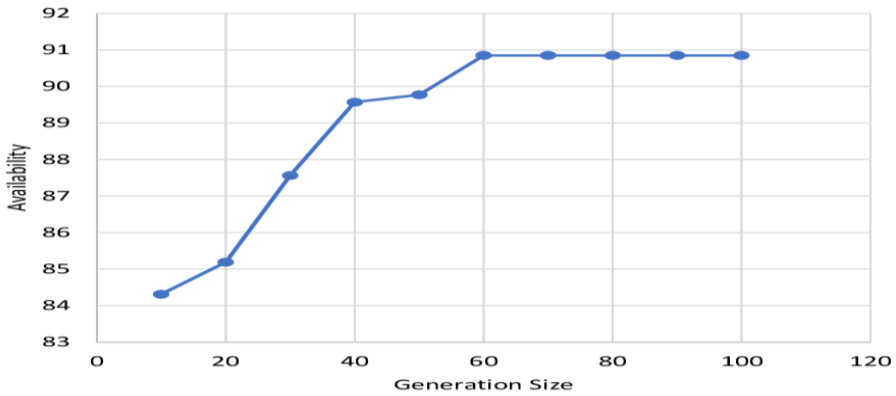


Figure 8: *Effect of GS on System availability*

IX. Conclusion

This study focused on providing a complete assessment and improvement of the crystallization and sugar handling processes of a sugar mill plant. In this case, a Markov process and Particle Swarm Optimization (PSO) were used to pinpoint the most crucial components for overall system availability. The findings indicate that maintenance attention is best concentrated on the Evaporator since it impacts the system's availability most. In contrast, the Rotary Dryer affects availability the least; therefore, maintenance scheduling should prioritize subsystems with higher failure rates and extensive repair durations.

From this perspective, the study also underscores the need for customized maintenance action plans guided by failure rate (FR) and repair rate (RR) metrics to enhance plant reliability and reduce operational downtimes. In this particular case, optimizing these parameters increased the plant's system availability to 88.84%. Under such conditions, the application of PSO method improves the operational efficiency of sugar mills, thus leading to higher productivity and lower energy consumption.

The findings also give some actionable recommendations for further work, like concentrating on improving the Evaporator and Centrifuge units, while managing some less important parts, like the Rotary Dryer, with more lenient timeframes. This is helping to improve maintenance strategies and sustainability in the operation of sugar mills by demonstrating that thorough Analysis and systematic optimization can improve plant performance and the environmental footprint of the processes substantially.

References

[1] Gahlot, M., Singh, V.V., Ayagi, H.I. and Goel, C.K. (2018). Performance assessment of repairable system in series configuration under different types of failure and repair policies using copula linguistics. *Int. J. Reliab. Saf.*, vol. 12, no. 4, pp. 348–363.

[2] Zhao, D. and Li, Y.R. (2015). Climate change and sugarcane production: potential impact and mitigation strategies. *Int. J. Agron.*, vol. 2015, no. 1, p. 547386.

[3] Dhillon, B.S. and Shah, A.S. (2007). Availability analysis of a generalized maintainable three-state device parallel system with human error and common-cause failures. *J. Qual. Maint. Eng.*, vol. 13, no. 4, pp. 411–432.

[4] Kumar, D. (1988). Availability of the feeding system in the sugar industry. vol. 28, no. 6, pp. 867–871.

[5] Kumar, A. (2018). Performance evaluation of multi-state degraded industrial production system and selection of performance measure using PSO: A case study. *Int. J. Product. Qual. Manag.*, vol. 25, no. 1, pp. 1–17.

[6] Adhikary, D.D., Bose, G.K., Mitra, S., and Bose, D. (2010). Reliability, Maintainability & Availability analysis of a coal fired power plant in eastern region of India. in *Proceedings of the 2nd International Conference on Production and Industrial Engineering (CPIE 2010)*, pp. 1505–1513.

[7] Aggarwal, A.K., Kumar, S. and Singh, V. (2013). Performance modeling and availability analysis of skim milk powder production system of a dairy plant. *Int. J. Agric. Food Sci. Technol.*, vol. 4, no. 9, pp. 851–858, 2013.

[8] Arora, N. and Kumar, D. (1997). Availability analysis of steam and power generation systems in the thermal power plant. *Microelectron Reliab.*, vol. 37, no. 5, pp. 795–799.

[9] Arulmozhi, G. (2002). Exact equation and an algorithm for reliability evaluation of k-out-of-n: G system. *Reliab. Eng. Syst. Saf.*, vol. 78, no. 2, pp. 87–91.

[10] Cizelj, R.J., Mavko, B. and Kljenak, I. (2001). Component reliability assessment using

quantitative and qualitative data. *Reliab. Eng. Syst. Saf.*, vol. 71, no. 1, pp. 81–95.

[11] Dachyar, M., Nurcahyo, R. and Tohir, Y. (2018). Maintenance strategy selection for steam power plant in range of capacity 300-625 MW in Indonesia. *ARPN J. Eng. Appl. Sci.*, vol. 13, no. 7, pp. 2571–2580.

[12] Dahiya, O., Kumar, A. and Saini, M. (2019). Mathematical modeling and performance evaluation of A-pan crystallization system in a sugar industry.*SN Appl. Sci.*, vol. 1, pp. 1–9.

[13] Sihmar, P. and Modgil, V. (2025). Optimization of combed-sliver production system availability for a Yarn Production Unit using PSO. *Life Cycle Reliab. Saf. Eng.*, vol. 14, pp. 225–235.

[14] Ebrahimi, N.B. (2003). Indirect assessment of system reliability. *IEEE Trans. Reliab.*, vol. 52, no. 1, pp. 58–62.

[15] Kumar, S., Tewari, P. C. and Kumar, S. (2009). Performance evaluation and availability analysis of ammonia synthesis unit in a fertilizer plant. *J. Ind. Eng. Int.*, vol. 5, no. 9, pp. 17–26.

[16] Garg, S., Singh, J. and Singh, D.V. (2010). Availability and maintenance scheduling of a repairable block-board manufacturing system. *Int. J. Reliab. Saf.*, vol. 4, no. 1, pp. 104–118.

RT&A, No 2 (84)
Volume 20, June 2025

ASSESSING THE PREDICTIVE RESILIENCE OF COMPOSITE EXPERIMENTAL DESIGNS UNDER MISSING OBSERVATIONS: THE INFLUENCE OF A ALPHA VALUE VARIATIONS

A.R. Gokul[1] and M. Pachamuthu[1*]

[1]Department of Statistics, Periyar University, India
gokularphd@periyaruniversity.ac.in
pachamuthu@periyaruniversity.ac.in

Withdrawal

Abstract

The field of robust missing observations has become a sig
Regarding Response Surface Methodology, a well
experimental design, missing observations in
and challenging. In Response Surface M
are primarily used. These are the
Minimum Run-Resolution V
the design and the pr
encompass rotatab
intended to
loss cri
r

alpha,
ne minimax
n Space plot. The
n types, even without

ction Variance, Optimality,
esign.

Introduction

(RSM) is a potent combination of statistical and mathematical building. It's specifically designed to evaluate the effects of several and determine their optimal values to achieve the most favourable results. is particularly beneficial when the goal is to optimize a product or process. It omprehensive understanding of the relationships between different variables and the re , facilitating efficient and effective optimization. The empirical model is constructed using data gathered directly from the system or process under study. RSM involves building empirical models using multiple regression and statistical techniques [1]. Drapery et al. in 1961 [2] examines the studies on sturdy methods for handling missing observations in response surface design and acknowledges the pioneer who formulated the parameter estimation equation. Indeed, Akhtar et al. in 1986 [3] introduced a minimax loss criterion to deal with missing observations. This approach has become the most used method in response surface designs. It's a significant contribution to the field as it provides a robust and effective way to handle missing data, thereby improving the accuracy and reliability of the designs. Indeed, Akhtar et al. in 2001 [4] conducted a study on five-factor

Central Composite Designs *(CCD)*, which had two missing observations. This study was carried out in three different settings, providing valuable insights into the impact of missing observations on the design's performance. Alrweili et.al in 2019 [5] employ the minimax loss criterion to develop more resilient models for missing data. They achieve this by integrating the most recent CCDs from GSA and AEK, both novel designs. Hayat et al. [6] delve into designs derived from both regular and irregular structure subsets. They evaluate how these designs handle missing points using the minimax loss criterion. Furthermore, they scrutinize their alphabetic optimality and predictive capabilities through Fraction of Design Space *(FDS)* plots depicting the variance in response difference. Alanazi et al. [7] introduce closed-form expressions that account for two missing observations. These expressions are based on α, the axial value utilized in CCDs that handle up to 10 factors. In their research, Rashid et al. [8] explore methods to handle a single missing data point in both Augmented Box-Behnken Design (ABBD) and Augmented Fractional Box-Behnken Design *(AFBBD)*. They use the minimax loss criterion and assess the relative D and G efficiency. Her thi et al. [9] examine the ability of sequential third-order rotatable design to manage m a without significant information loss. Additionally, the study quantifies the loss of i one or two absent experimental runs at varying distances from the design' Chigbu et al. [10] compared CCD, Small Composite Design *(SCD)*, and *(MinResV)* designs. They focus on spherical regions with k = 3 to criteria and the Variance Dispersion Graph *(VDG)* as bench reveal that none of these designs consistently outperform CCD, SCD, and MinRes*V* designs. These designs are different axial values. To analyze the prediction plots and box plots. Onwuamaeze et al. [12] They use these methods to assess the r designs within the hypercube re of each factorial, axial, and cente of factors k = 2 to 7. They assesse of the design space region, meas the Box-Behnken Design with th an the classical BBD, to investigat

T ssing observations in CCD with are scrutinized with sufficient center DIX 2. For full CCD, k ranges from 3 to 7, traction, k ranges from 8 to 9, and for CCD 1⁄8 , for MinRes V, k ranges from 6 to 10. We consider wly introduced alpha known as αPE. The minimax loss icitly stated alpha values, not sequentially. All these design iations are meticulously analyzed using the FDS plot, which of SPV for various scenarios and combinations under missing

divided into several sections. Section 2 outlines SOCD and alpha (α) values and ie minimax loss criterion methodology. The results and discussions are presented in . Section 4 interprets the FDS plot, and Section 5 concludes the study.

Withdrawal

II. Methodology

I. Description of Second-Order Composite Design and alpha (α) values

In many cases where we apply response surface methodology, the relationship between the predictor variables and the response might be unclear. A first-order model, while applicable in some cases, might not be able to accurately capture the curvature of the response function due to its linear

nature. This is where higher-degree polynomial models, such as second-order models, come into play. These models can capture more complex relationships and better evaluate curvature in optimization experiments. They provide a more nuanced understanding of the data, allowing for more accurate predictions and more effective optimization experiments. For k quantitative factors denoted by $x_1, x_2, \ldots, x_k$, a second-order model is

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j + \varepsilon$$

In this context, β_0, β_i, β_{ii} , and β_{ij}represent the intercept, linear, quadratic, and bilinear coefficients, respectively. ϵ_i is a random error term with a mean of zero, a variance of σ^2, and is independent for each pair of runs. The total number of parameters that need to be estimated, denoted as p, is calculated as $p = k + k + \binom{k}{2} + 1$. To ensure enough degrees of freedom to estimate these model coefficients, the number of runs, represented as n, must equal or exceed p.

II. Types of Design CCD, SCD and MinRes *V*

Central Composite Design is extensively utilized to estimate second-order
its introduction by [15], the CCD has been the subject of num
widespread use in various fields. Indeed, the design is based on
a full or fractional factorial design with a MinRes *V*. In
design as a more economical, cost-effective, and
framework of this design, known as the SCD, is li
the factorial component of the SCD is of res
main effects), unlike the CCD of resol
level irregular fractions of resoluti ıey
maintain the precise number ş with a
more significant numbe ıne distance
between each axial r ıonal freedom for
estimating the i

Withdrawal

III. T

$= F^{1/4}$, called rotatable alpha, where F is the
ɔsign is rotatable if the variance remains the same
ıne design's center. This characteristic allows for more
ıerical Alpha: Indeed, Myers et.al and Anderson et al. [18,
herical alpha. In this approach, the value of alpha is set to the
ıctors, represented as $\alpha_S = k^{1/2}$. Indeed, in αs, all the design points are
ɔometric sphere. Although these designs are not exactly rotatable which
ıability of the forecasted outcome is constant at all locations that are the same
ıe midpoint they are nearly rotatable. Practical Alpha: If the Variance Inflation
ıes) exceed a certain threshold (say, VIF > 10), it might lead to instability in the coefficient
esı ıtes, provided they can be estimated in the first place. Hence, it's suggested that equilibrium between the cube face's axial level and the spherical alpha be maintained. A feasible alpha value denoted as $\alpha_P$and calculated as the fourth root of $\alpha_P = k^{1/4}$, is put forward. This alpha value keeps the VIF within an acceptable range for designs incorporating more than five factors. This method guarantees the design's stability and dependability. Practical Efficient Alpha: Bharat et.al [20] noted that all previously suggested alpha values were associated with only one factorial design parameter: the factorial level, the number of factorial runs, or the number of factors. In this context, they

introduced a practical efficient alpha value $\alpha_{PE} = (nk)^u$ that considers the number of factorial runs and the number of factors. It is equivalent to α_P value at $u = 1/12$ and α_R or α_S alpha value at $u = 1/6$. The α_{PE} is calculated as the eighth root of the product of the number of factors.

IV. Scaled Prediction Variance and Relative G - Efficiency

Borkoski et al. [21] gives an analytical form for calculating scaled prediction variance values of central composite design and Box-Behnken design. SPV enables accurate prediction of response variables at different points of interest within the experimental area. The prediction variance at a point x is given by

$$v(x) = \frac{n \cdot var[\ y_{(x)}]}{\sigma_2} = n \cdot X^{(h)'}\left(X'X\right)^{-1}X^{(h)}$$

The vector $x^{(h)}$ represents the array of coordinates of a point in the design space that has been magnified to align with the model form, where n is quantity of experimental runs design the observation error.

G-optimality is a standard that seeks to reduce the highest possible varianc value across the entire experimental domain. Iwundu et al. [22] studied multiple missing observations on cuboidal designs' Relative A-, D efficiencies can be understood as the proportion between the dete for a specific design and the determinant of the information m

$$G_{\text{eff}} = \frac{p}{n \cdot M'}$$

In this context, p represents the e observations in the correspondin the variance of the predicted r RE_G, is determined by the ratio design.

esign, while n_r denotes the size of the that yields a higher value of RE_G would be ndicates a more efficient design relative to the

Withdrawal

ol for comparing multiple designs. It visually represents the design's sum ice distributions through a single curve. This allows for a comprehensive a evaluation of the efficiency of different designs. An FDS plot is generated by ubstantial number of samples, represented as n, from the entire design area and ca ating the associated SPV values [23]. The underlying concept is that the design quality improves if a greater portion of the design space is near the minimum SPV value.

Furthermore, a flatter line indicates a more stable design. The FDS plot effectively encapsulates the range and distribution of SPV values in the design space, facilitating the comparison of designs through a single curve. Furthermore, it equips the investigator with a unique graph for contrasting designs or examining the characteristics of a particular design [24]. As a result, the FDS method can be utilized in both regular and irregular design areas.

V. Minimax Loss Criterion

Akhtar et al. [3] conducted a study on the effects of missing observations, utilizing the minimax loss criterion. This criterion minimizes the most tremendous loss associated with a missing design point. Herzberg et al. [25] examined the link between the determinant of the information matrix (6) from the reduced design $|X\,X_r|$ and that of the full design $|X'X|$, considering a scenario where one observation is missing. The model matrix is divided into $X = \begin{pmatrix} X \\ X_r \end{pmatrix}^{'m}$, where X_m represents the m^{th} missing observation and X_r represents the reduced design. Obviously,

$$X'X = X\,X_m + X\,X_r'$$

The determinants of the reduced information matrix is

$$|X_r'X_r| = \left|1_p - (X\,X_m)(X'X)^{-1}\right||X'X|\ |X$$
$$\frac{X_r|}{X'X|} = \left|1_p - (X\,X_m)(X'X)^{-1}\right|\ |$$

Akhtar et al. [3] described this relationship as

$$\frac{|X_r'X_r|}{|X'X|} = E_m$$

where $E_m = \left|1_p - (X_m{}'X_m)(X'X)^{-1}\right|$ is the dia
missing design points. E_m represents the effic
points compared to the complete desig
symbolized by L_m, refers to the de
compared to the scenario where

$$\left|1_p - (X_m{}'X_m)(X'X)^{-1}\right|$$

d Discussion

ax loss criterion from Table 1, it is stated that when k = 3, neither the
sign points meet the average efficiency. However, the efficiency of axial
passes that of factorial observations. The center run shows good performance
or alpha values of both α_R and α_{PE}. When k = 4, the efficiency of handling missing
and axial observations remains consistent for both αR and α_S values, as they share the same alpha value. However, when the α_{PE} value is used, the axial design points outperform the factorial ones. In the case of the α_P value, the factorial design point undergoes a total information loss, with a value of 0.9973, while the axial design point experiences a lower-than-average information loss. Regarding factors (k = 5, 6,) and (7), a steady pattern is observed in the efficiency of absent factorial design points for all alpha values. Here, the efficiency percentages are roughly identical across different alpha values for each factor. The efficiency of missing axial design points varies from

Withdrawal

a minimum of 43.94% to a maximum of 52.60% for factors (k = 5, 6,) and (7). The only exception is factor (k = 5), where the α_P is the sole cause of total information loss. Lastly, the center run consistently exhibits high performance for all factors and alpha types, except for factor (k = 3), where only the α_S and α_{PE} values show below-average results.

Table 1: *Efficiency and Loss of information of CCD Factor K = 3 to 7*

		Efficiency (E_m)			Loss of information (L_m)		
Factors	Alpha(α)	Factorial(f)	Axial(a)	Centre(c)	Factorial(f)	Axial(a)	Centre(c)
K = 3	1.6818 (α_R)	0.3302	0.3927	0.8337	0.6698	0.6073	0.1663
	1.7321 (α_S)	0.3393	0.3810	0.3810	0.6607	0.6190	0.6190
	1.4877 (α_{PE})	0.2931	0.4347	0.8411	0.7069	0.5653	0.1589
	1.3161 (α_P)	0.2602	0.4656	0.4656	0.7398	0.5344	0.5344
K = 4	2.0000 (α_R)	0.4167	0.4167	0.8333	0.5833	0.5833	0.1667
	2.0000 (α_S)	0.4167	0.4167	0.8333	0.5833	0.5833	0.1667
	1.6818 (α_{PE})	0.3893	0.4610	0.8473	0.6107	0.5390	0.15
	1.4142 (α_P)	0.0027	0.4857	0.8714	0.9973	0.5143	
K = 5	2.3784 (α_R)	0.5496	0.4394	0.8367	0.4504	0.56	
	2.2361(α_S)	0.5446	0.4571	0.8333	0.4554		
	1.8859 (α_{PE})	0.5322	0.4862	0.8512	0.4678		
	1.4954 (α_P)	0.5204	0.0000	0.8873	0 4		
K = 6	2.8284 (α_R)	0.6785	0.4568	0.9091			
	2.4495 (α_S)	0.6735	0.4912	0.90			
	2.1040 (α_{PE})	0.6690	0.5096				
	1.5651 (α_P)	0.6633	0.52				
K = 7	3.3636(α_R)	0.7829					
	2.6458 (α_S)	0.7796					
	2.3390 (α_{PE})	0 7					+1
	1.6266 (α_P)						0.0652

Withdrawal

II. CCD ½ f

T- ugh efficiency, even with the loss of
ests that the central point's robustness is
rence in experimental designs. When *k*=5, the
e missing factorial design points for every alpha
tegic placement yields better performance under these

es more critical as the information loss surpasses 50% for all alpha
ctorial and axial design points, highlighting a significant impact of
sign's overall efficiency. For *k*=7, the factorial observation with the α_{PE} value
st efficiency at 0.1426, which is notably lower compared to the moderate efficiency
ved in the absence of other design points.

. the scenario where *k*=8, the efficiency of the missing factorial observations, except for the α_S value, remains above 70% across all different alpha values. This indicates that despite the absence of some design points, the efficiency retains a relatively high level, with the 70% efficiency representing the highest among all factors. This resilience in efficiency, even with missing factorial design points, underscores the robustness of the design under varying conditions.

Table 2: *Efficiency and Loss of information of CCD ½ fraction K = 5 to 8*

			Efficiency (E_m)			Loss of information (L_m)		
Factors	Alpha(α)		Factorial(f)	Axial(a)	Centre(c)	Factorial(f)	Axial(a)	Centre(c)
K = 5	2.0000	(α_R)	0.1212	0.4015	0.8409	0.8788	0.5985	0.1591
	2.2361	(α_S)	0.1442	0.3692	0.8333	0.8558	0.6308	0.1667
	1.7294	(α_{PE})	0.0944	0.4309	0.8634	0.9056	0.5691	0.1366
	1.4954	(α_P)	0.0004	0.4520	0.8853	0.9996	0.5480	0.1147
K = 6	2.3784	(α_R)	0.3693	0.4317	0.8755	0.6307	0.5683	0.1245
	2.4495	(α_S)	0.3722	0.4242	0.8750	0.6278	0.5758	0.1250
	1.9294	(α_{PE})	0.3512	0.4669	0.8949	0.6488	0.5331	0.1051
	1.5651	(α_P)	0.3386	0.4861	0.0000	0.6614	0.5139	1.0000
K = 7	2.8284	(α_R)	0.5721	0.4543	0.9022	0.4279	0.5457	0.0978
	2.6458	(α_S)	0.5693	0.4689	0.9000	0.4307	0.5311	0.1000
	2.1449	(α_{PE})	0.1426	0.4070	0.8817	0.8574	0.5930	0.1[illegible]
	1.6266	(α_P)	0.5555	0.5072	0.9344	0.4445	0.4928	
K = 8	3.3636	(α_R)	0.7216	0.4659	0.9180	0.2784	0.5[illegible]	
	2.8284	(α_S)	0.0000	0.5000	0.9000	1.0000		
	2.3784	(α_{PE})	0.7164	0.5115	0.9114	0.2836		
	1.6818	(α_P)	0.7137	0.5159	0.9397	0.2[illegible]		

III. CCD ¼ fraction for factors K = 8 to 9 and [illegible]

Table 3: *Efficiency and Loss of information CCD ¼ [illegible]*

			Efficiency ([illegible]				
Factors	Alpha(α)		Factor[illegible]				[illegible]e(c)
K = 8	2.8284	(α_R)	0				0.1000
	2.8284	(α[illegible]					0.1000
	2.1810					[illegible]5038	0.0607
	1					0.5038	0.0607
					[illegible]	0.5337	0.0911
					0.3497	0.5129	0.1000
					0.3532	0.4984	0.0826
				[illegible]9441	0.3562	0.4918	0.0559
				0.9028	0.4245	0.5346	0.0972
			[illegible]257	0.9000	0.4259	0.5243	0.1000
			0.4935	0.9234	0.4305	0.5065	0.0766
			0.5021	0.9479	0.4338	0.4979	0.0521

[illegible]raction of the axial design points for factors k = 8 and k = 9 maintain an [illegible]level. However, the absence of a factorial observation for factor k = 9 results in [illegible]ncy than factor k = 8. When a center run is missing in the CCD ¼ fraction for factor k = 8, [illegible]ghest efficiency is achieved exclusively for the α_{PE} value, while all other alpha values experience a moderate loss of information. In the CCD $^1/_8$ fraction of factor k = 10, the absence of a factorial observation significantly surpasses the axial efficiency, and the center design point attains the highest efficiency with only a slight loss of information.

Withdrawal

IV. SCD for Factors K = 3 to 7

Table 4 of SCD exhibits that all factors and alpha values from factorial observations show minimal efficiency or significant information loss. This applies to the axial design point, except when the factor k = 4. For this case, the only α_{PE} value of factor k = 4, among all alpha values, renders an efficiency of 39.30% for the axial design point. Notably, the α_P value of k = 5 results in a complete loss of information, and the α_S and α_{PE} values of k = 7 demonstrate subpar efficiency estimation.

Table 4: *Efficiency and Loss of information of SCD factors K = 3 to 7*

		Efficiency (E_m)			Loss of information (L_m)		
Factors	Alpha(α)	Factorial(f)	Axial(a)	Centre(c)	Factorial(f)	Axial(a)	Centre(c)
K = 3	1.4142 (α_R)	0.0938	0.0938	0.8437	0.9062	0.9062	0.1563
	1.7321 (α_S)	0.0000	0.0667	0.8333	1.0000	0.9333	0.1667
	1.3643 (α_{PE})	0.0839	0.0968	0.8473	0.9161	0.9032	0.1527
	1.3161 (α_P)	0.0745	0.0993	0.8511	0.9255	0.9007	
K = 4	1.6818 (α_R)	0.0905	0.0774	0.8451	0.9095	0.92	
	2.0000 (α_S)	0.0000	0.0625	0.8333	1.0000		
	1.5422 (α_{PE})	0.0747	0.3930	0.8566	0.9253		
	1.4142 (α_P)	0.0000	0.0789	0.8684	1.0		
K = 5	1.8212(α_R)	0.0189	0.0619	0.8528			
	2.2361 (α_S)	0.0352	0.0507	0.83			
	1.7294 (α_{PE})	0.0154	0.0618				
	1.4954 (α_P)	0.0326	0.059				
K = 6	2.0000 (α_R)	0.0130					
	2.4495 (α_S)	0.0268					
	1.9294 (α_{PE})	0.0114					.562
	1.5651 (α_P)						0.1145
K = 7	2.1657 (α					.956	0.9556
	2.64					1.0000	0.9575
						0.9591	0.1118
					.984	0.9984	0.9672

Withdrawal

observed that axial design points demonstrate greater across all factors and alpha values. This indicates that the aintaining the integrity of the design's efficiency under various , irrespective of the factors or alpha values, exhibit high efficiency, point's robustness and stability.

e are notable exceptions where this high efficiency does not hold true. k=7k = 7 for the αS value, the efficiency significantly decreases, suggesting a ity in the design at this particular combination of factor levels and alpha value. F ermore, at k=10k = 10, the design's efficiency is compromised for αR, αS, and αP values, indicating that the design struggles to maintain high efficiency under these conditions. Only the αPE value experiences a significant loss of information, which highlights a critical point of inefficiency in the design.

Addressing these inefficiencies could involve re-evaluating the placement of design points or considering alternative design strategies to bolster robustness under these conditions.

Table 5: *Efficiency and Loss of information of MinRes V factors K = 6 to 10*

		Efficiency (E_m)			Loss of information (L_m)		
Factors	Alpha(α)	Factorial(f)	Axial(a)	Alpha(α)	Factorial(f)	Axial(a)	Alpha(α)
K = 6	2.1657 (α_R)	0.0727	0.3330	0.8436	0.9273	0.6670	0.1564
	2.4495 (α_S)	0.0887	0.2980	0.8333	0.9113	0.7020	0.1667
	1.8411 (α_{PE})	0.0545	0.3710	0.8722	0.9455	0.6290	0.1278
	1.5651 (α_P)	0.0409	0.4020	0.8970	0.9591	0.5980	0.1030
K = 7	2.3404 (α_R)	0.1168	0.3739	0.8456	0.8832	0.6261	0.1544
	2.6458 (α_S)	0.1394	0.3418	0.0000	0.8606	0.6582	1.0000
	1.9511 (α_{PE})	0.1426	0.4070	0.8817	0.8574-	0.5930	0.1183
	1.6266 (α_P)	0.0651	0.4326	0.9088	0.9349	0.5674	0.0912
K = 8	2.4828 (α_R)	0.1116	0.3554	0.8497	0.8884	0.6446	0.15[illegible]
	2.8284 (α_S)	0.1213	0.3232	0.8333	0.8787	0.6768	
	2.0434 (α_{PE})	0.0994	0.3899	0.8917	0.9006	0.61[illegible]	
	1.6818 (α_P)	0.0904	0.4189	0.9187	0.9096	[illegible]	
K = 9	2.6043 (α_R)	0.0558	0.3580	0.8539	0.9442		
	3.0000 (α_S)	0.0649	0.3236	0.8333	0.93[illegible]		
	2.1239 (α_{PE})	0.0434	0.3936	0.8996			
	1.7321 (α_P)	0.0326	0.4235	0.926[illegible]			
K = 10	2.7356 (α_R)	0.0588	0.3770	[illegible]			
	3.1623 (α_S)	1.0000	0.3458				
	2.2056 (α_{PE})	0.0429	0.[illegible]				
	1.7783 (α_P)	0.0306					

Withdrawal

IV

I. Relative G [illegible] for Factors K = 3 to[illegible]

[illegible]ndicates that a consistent increase in G [illegible]t an αR value of design points, regardless of [illegible]nts with an αS value show an increasing trend in [illegible]. Beyond that, this upward trend is not observed for [illegible]now a decreasing G efficiency trend as the factor increases. [illegible]ency demonstrates superior G efficiency and scaled prediction [illegible]ral, the absence of an axial run outperforms the absence of a factorial [illegible]3, where the factorial run performs well. Design points with αP values [illegible]PV value for higher factors.

I. [illegible]ve G efficiency and max SPV for α_R, α_S, α_{PE} and α_P CCD ½ fraction for Factors K = 5 to 8

In the case of CCD ½ fraction from Appendix 1 of Table A2 results, specific alpha design points exhibit the same G efficiency and a similar maxSPV value regardless of whether an observation is missing. This occurs for factors k = 6 when α_P = 1.5651, k = 7 when α_S = 2.6458, and k = 8 when α_S = 2.8284 and α_P = 1.6818. This implies that the presence or absence of a missing observation does not significantly influence this situation. Among all factors, alpha values, and design points, only the

factorial points for factor k = 5 display the lowest G efficiency and maximum SPV across all alpha values. Factors k = 7 and 8 show a notable max SPV value for the α_P value. This analysis provides valuable insights into the behaviour of these factors under different conditions.

III. Relative G efficiency and max SPV for α_R, α_S, α_{PE} and α_P CCD ¼ fraction for Factors K = 8 to 9 and CCD $^1/_8$ fraction for Factor K = 10

Based on the CCD ¼ fraction data from Appendix 1 of Table A3, we deduced that there are 90 runs for factor k = 8 and 156 runs for factor k = 10, with each factor including 10 center points. If an observation is absent for all alpha values, the G efficiency escalates as the factor increases from k = 8 to 9, except for the α_P value design points. Regarding SPV performance, the α_R and α_S outperform the α_P and α_{PE} values. The factorial missing observation yields better results than the central observation, a scenario that fits perfectly with CCD $^1/_8$ fraction k = 10, which has [illegible] absence of a center run does not significantly impact the experimental conditions.

IV. Relative G efficiency and max SPV for α_R, α_S, α_{PE}, and α_P [illegible] to 7

Appendix 1 of Table A4 shows that the Small Comp[illegible] significantly impact the prediction performanc[illegible] estimating parameters across all factors an[illegible] only the α_S ones exhibit a marginally r[illegible] k = 3 and 4. Observations from or[illegible] [illegible]ncy α_S value.

V. Relative G e[illegible] [illegible]V for Factors K = 6 to 10

Accord[illegible] [illegible]es V design is merely average w[illegible] [illegible]rmance of the α_S value is superior to [illegible]ciency alone, α_R value outperforms α_{PE} [illegible] results in a slightly improved prediction [illegible] observation. It's also worth noting that the G [illegible]e roughly equivalent.

V. FDS Plot Interpretation

[illegible] to 7

[illegible]wn in Figure 1a, when k = 3, each design point regardless of alpha values possesses [illegible]s than 10, but this is only true for up to half of the Fraction of Design Space region. As w[illegible]ove towards the center run across all alpha values, there's a gradual rise in the SPV value starting from the midpoint of the design space region. Consequently, the center run has the highest SPV value compared to other design points.

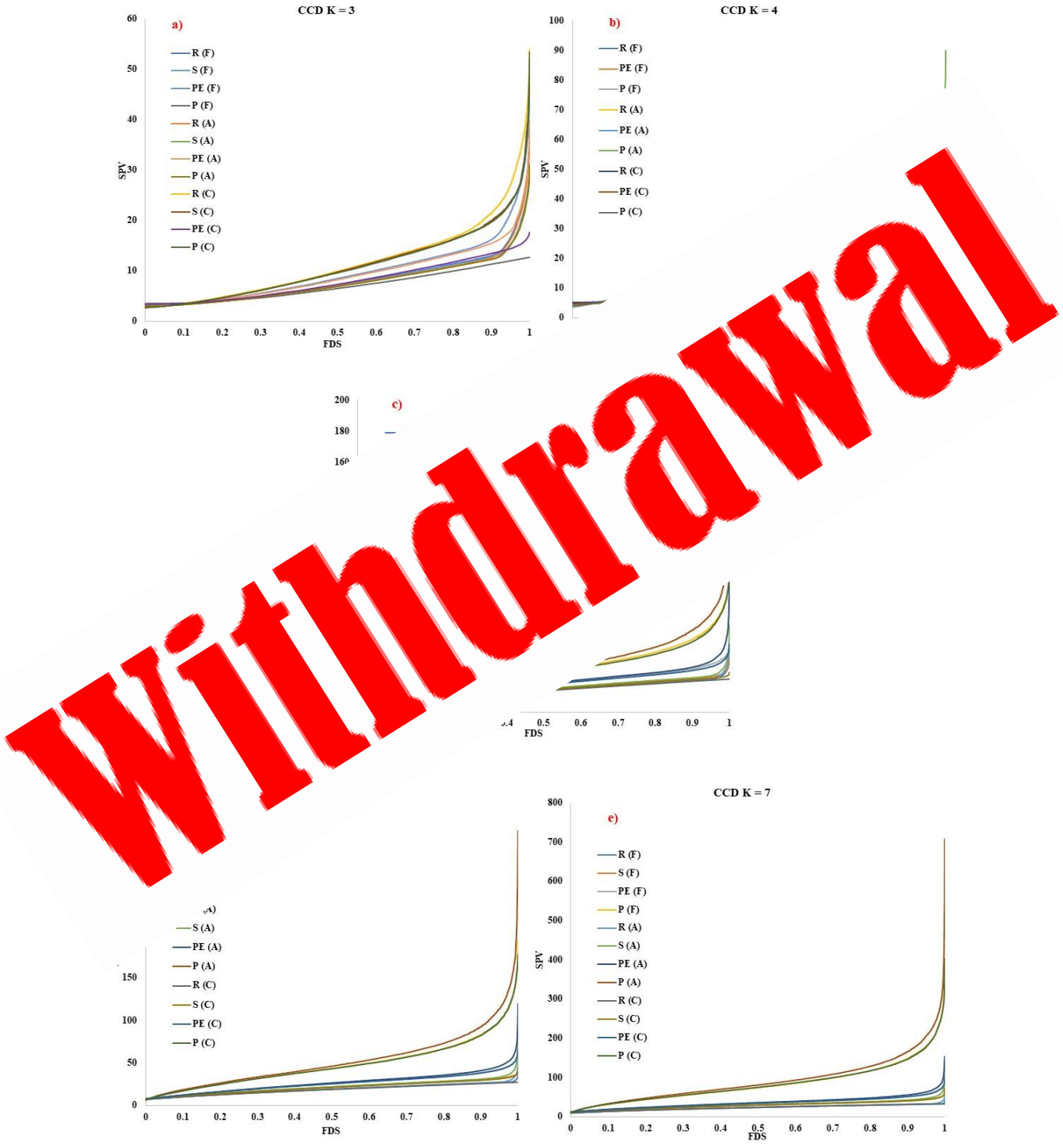
In CCD of k = 4 from Figure 1b, say, the maximum SPV is 90. A quarter of this maximum value gives an SPV of 22.5. Apart from the design points of α_P, all other design points of α_R, α_S, and α_{PE} approximately have an SPV value of 22.5 or less, but this is only applicable to up to 90% of the Fraction of Design Space (FDS) design region.

For CCD with k = 5, Figure 1c clearly illustrates that the alpha values of design points α_R and α_S have the lowest level of SPV compared to α_{PE}; this α_{PE} is superior compared to the alpha value α_P.

Withdrawal

We observe a certain degree of consistency in the trend line for SPV values of α_R and α_S in a specific region of FDS, that is, 60% to 95% of the FDS design region, showing a slightly steady increasing trend, which is appreciable when compared to the design points of the alpha value α_{PE}.

Figures 1d and 1e exhibit that the performance of factors k = 6 and k = 7 is nearly identical, with slight variations throughout the entire FDS region. Except for the α_P value, all other α_R, α_S, and α_{PE} values cover 95% of the FDS region with an SPV less than 50. We observe minor oscillations in SPV values ranging from 50% to 95% for α_R, α_S, and α_{PE} values.



Withdrawal

Figure 1: *(a) FDS for CCD K = 3. (b) FDS for CCD K = 4. (c) FDS for CCD K = 5. (d) FDS for CCD K = 6. (e) FDS for*

II. CCD ½ fraction of factors K = 5 to 8

Upon examining all the factors k = 5, 6, 7, 8, from Figures 2a,2b,2c, and 2d, it's noticeable that all alpha design points start to diverge with an increase in the factor. This divergence is attributed to the fluctuation in SPV values. This observation suggests a higher factor produces a more pronounced disparity among alpha values. In general, it's noted that both αR and αS consistently hold their minimum SPV value when compared to the αPE and αP values. When we compare factors k = 5,6,7,8 with an SPV below 50, for k = 5, all types of alpha design points, except for the αP (f) value, exhibit an SPV of 50 or lower for as much as 92% of the FDS design region. We see that for k = 6, the αP value of design points exceeds 50 SPV when it reaches 80% of the FDS. For k = 7, it surpasses 50 SPV at 45% of the FDS. For k = 8, αR and αS value's design points keep the SPV value under 50 up to 95% of the FDS region. However, for αPE and αP design points SPV value crosses above 100 from the midpoint of the design space region, resulting in max SPV. Only for k = 7 do αPE design points cross an SPV of 50 above for the but all other factors of αPE value maintain their SPV below 50 up to 95%

Withdrawal

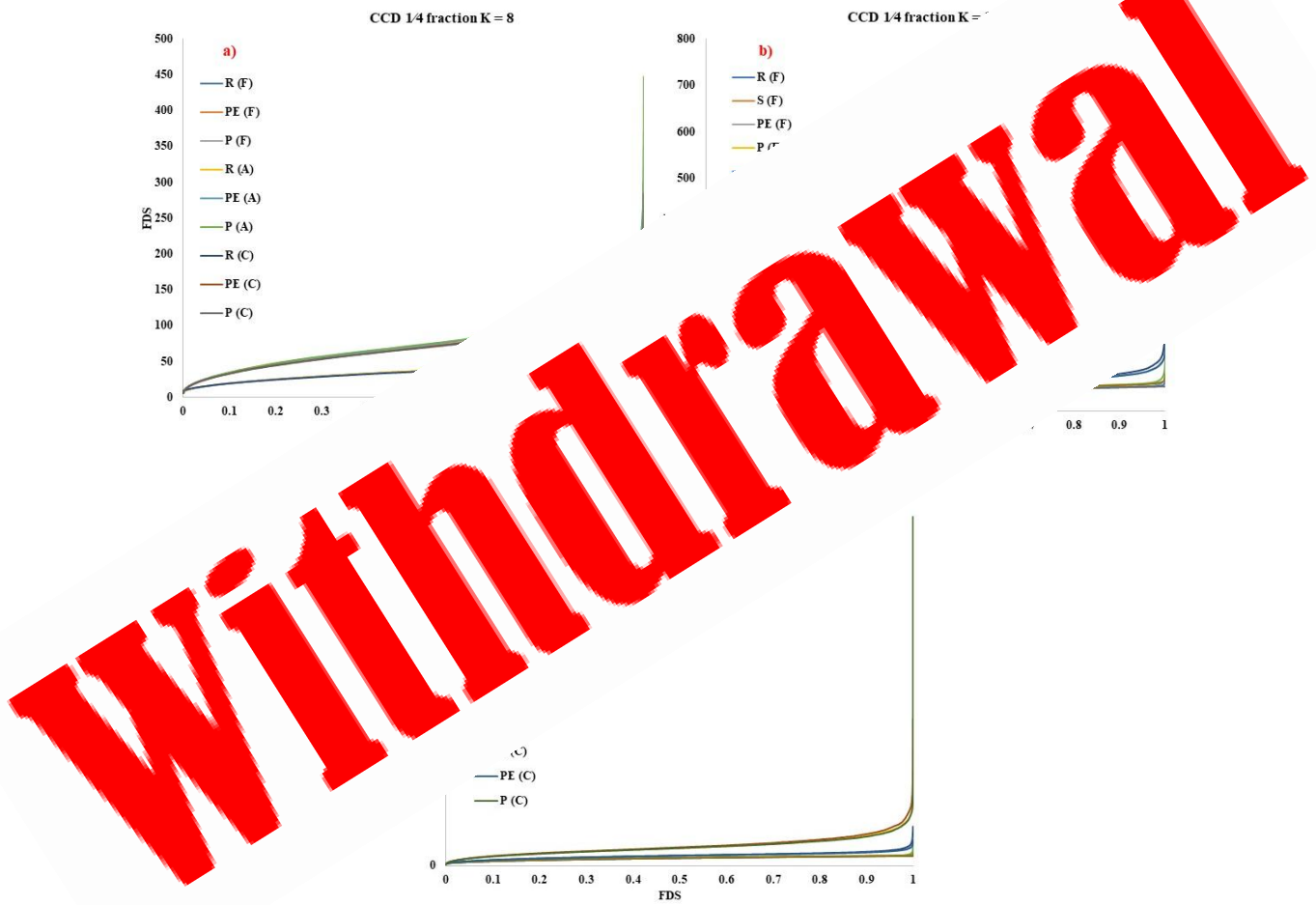


Figure 2: *(a) FDS for CCD ½ K = 5. (b) FDS for CCD ½ K = 6. (c) FDS for CCD ½ K = 7. (d) FDS for CCD ½ K = 8.*

III. CCD ½ fraction of factors K = 5 to 8

In the case of CCD ¼ fraction of k = 8 from Figure 3a, we can evaluate the FDS plot's performance by comparing α_{PE}, α_P Vs α_R values. A significant difference exists in SPV value among α_{PE}, α_P Vs α_R values. We use a 50 SPV value as a benchmark; the α_R value is at or below 50 SPV for up to 90% of the FDS region. In contrast, the performance of the other two alpha values is entirely unrelated to the α_R value, which surpasses the 50 SPV line in over 23% of the FDS region. Indeed, it's interesting to note that the α_R value exhibits a superior scaled prediction variance even in the absence of an observation. This suggests that it may have a more robust performance under these conditions.

In Figure 3b of CCD $^1/_8$ fraction of k = 10, which has a similar performance of CCD ¼ fraction of k =9 from Figure 4a, the α_R and α_S values follow the same trajectory up to an FDS of 1, the maximum region. The α_{PE} values, on the other hand, closely align with the paths of the α_R a[illegible] α_S values. In up to 90% of the design space, all these alpha values remain below 100 SPV. Ho[illegible]e α_P values exceed an SPV of 100 in less than 40% of the design space region.



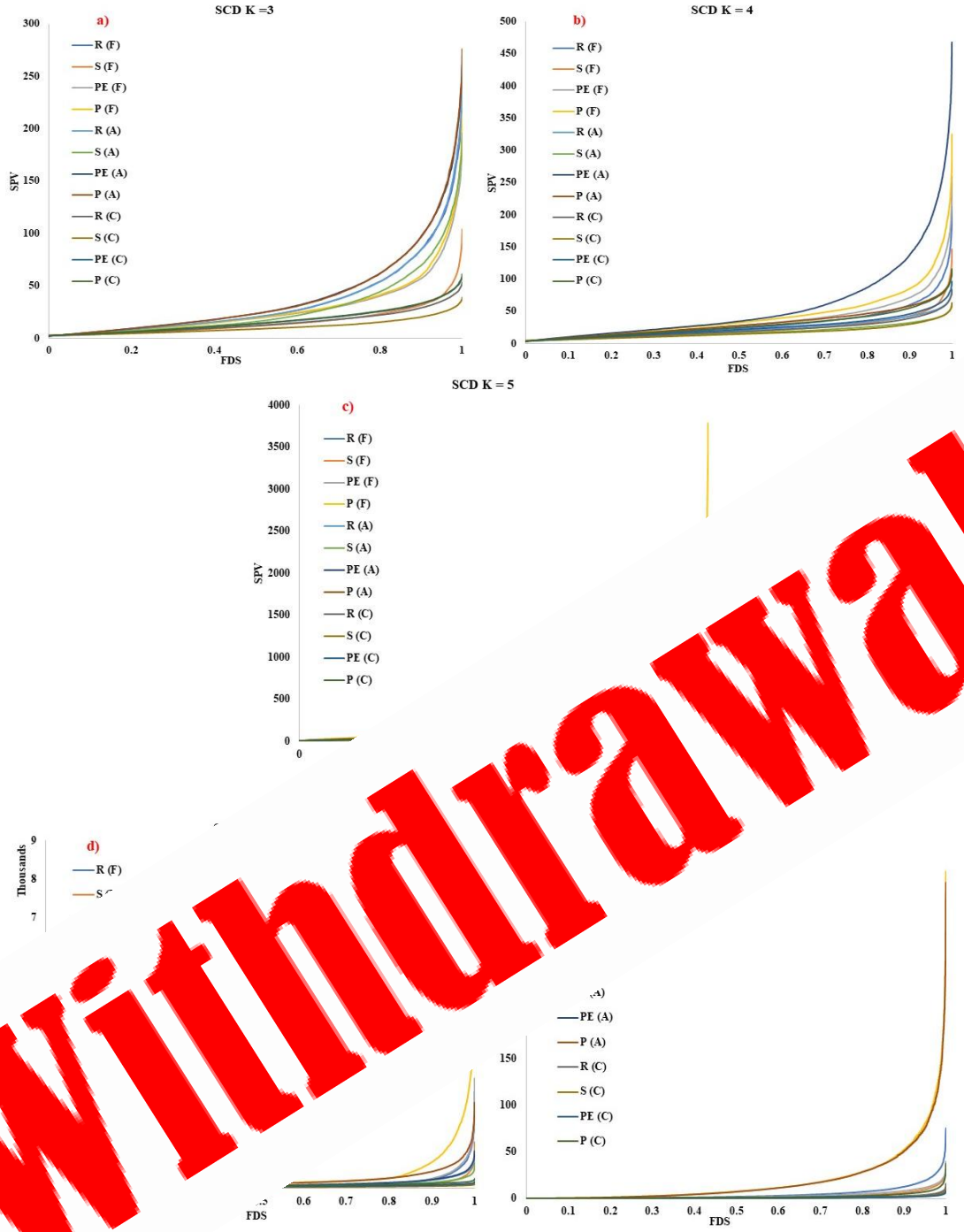
Withdrawal

Figure 3: *(a) FDS for CCD ¼ K = 8. (b) FDS for CCD ¼ K = 9 (a) FDS for CCD $^1/_8$ K = 10*

IV. SCD of factors K = 3 to 7

The FDS plots presented in Figures 4a, 4b, 4c, 4d, and 4e provide us with the means to assess the Small Composite Design. Unlike other increasing factors, the scaled prediction variance value is acceptable only when the factors are k = 3 and 4. Furthermore, for these factors (3 and 4), the performance is superior only when the alpha value is α_R and α_S, as opposed to other alpha values. However, for factors k = 6 and 7, there's a significant increase in the worst-case prediction variance and a decrease in G efficiency value. Therefore, it's concluded that opting for an SCD with higher

factors isn't the optimal choice for a robust experimental design.



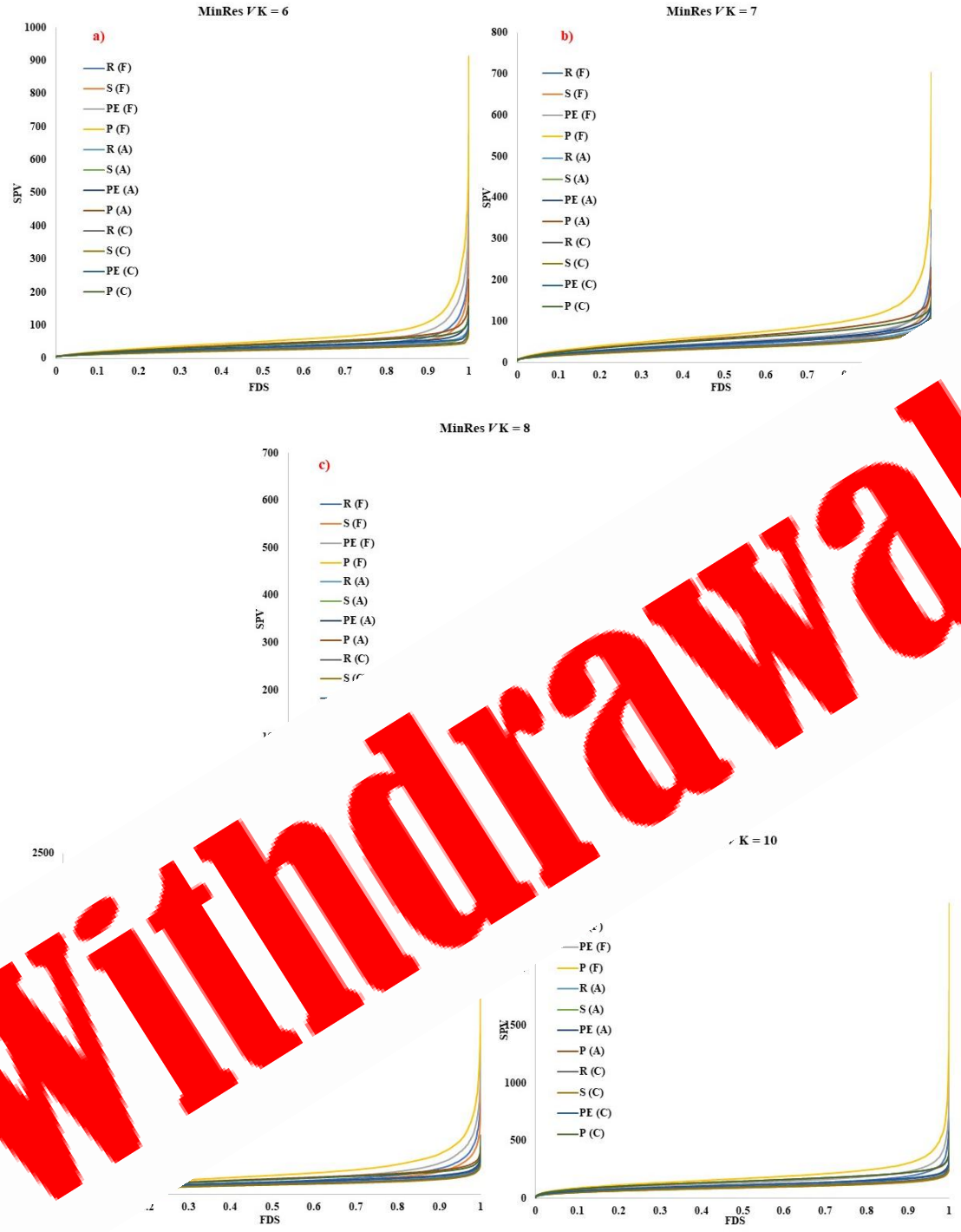
Withdrawal

SCD K = 3. (b) FDS for SCD K = 4. (c) FDS for SCD K = 5. (d) FDS for SCD K =6. (e) FDS for SCD K = 7.

V. MinRes *V* of factors K = 6 to 10

In the case of the MinRes *V* design from Figures 5a, 5b, 5c, 5d, and 5e, it's interesting to note that apart from the α_P value, all three different alpha values remain quite close until they reach a specific point in the FDS region. For instance, for k = 6, all design points travel up to 95% of the FDS region. For k = 7, all design points come closer to 90% of the FDS region, and for k = 8, all design points pass very close up to 85% of the FDS region. All these three factors have a more immediate destiny for below 100 SPV value. For higher factors k = 9 and 10 in the MinRes *V* design, the pattern of alpha

values clustering together remains consistent despite having a high SPV value. It's also worth noting that robust missing observations in the MinRes *V* design don't significantly impact using different alpha values, except for the α_P value.



Withdrawal

Figure 5: (a) *FDS for MinResV K = 6. (b) FDS for MinResV K = 7. (c) FDS for MinResV K = 8. (d) FDS for MinResV K =9. (e) FDS for MinResV K = 10.*

VI. Conclusion

In conclusion, CCD with various fractions results in missing a factorial observation in an experiment exhibiting superior efficiency in parameter estimation, significant relative G efficiency value, and a minimum SPV value compared to missing an axial observation. In the context of SCD, the absence

of a data point can lead to chaos in the experimental condition during analysis or interpretation due to poor estimation and a high SPV value. Notably, as the factor increases, the design performs poorly in all aspects. In the case of the Minimum run Resolution V design, the absence of an axial design point results in better predictive performance than missing a factorial design point. The lack of a center run does not degrade the performance of parameter estimation; it continues to uphold its reliability. The plot of the fraction of the design space indicates that the α_R and α_S values of CCD, SCD, and MinRes V have a moderate SPV for up to three quarters compared to α_{PE} and α_P values. Indeed, α_{PE} values outperform α_P values. We infer from our results that the newly introduced α_{PE} also provides superior parameter estimation and the lowest SPV in scenarios where an observation is missing.

References

[1] Khuri, A. I. and Mukhopadhyay, S. (2010). Response surface metho
interdisciplinary reviews: Computational statistics, 2(2):128-149.

[2] Drapery, N. R. (1961). Missing values in response surface design
398.

[3] Akhtar, M. and Prescott, P. (1986). Response surface desi
Communications in Statistics-Simulation and Computation, 15(

[4] Akhtar, M. (2001). Five-factor central com
observations. *Journal of Research (Science)*, 12(2):10

[5] Alrweili, H., Georgiou, S. and Styli
to missing data. *Quality and Reliability*

[6] Hayat, H., Akbar, A., Ahr ...sing
observation and optimalitie ...structure.
Communications in Stati

[7] Alanazi, K ...of central composite
designs wher ... *nternational*, 39(4):1143-
1171.

...hird order designs robust to one
... *Methods*, 48(24):6054-6062.

...ulya, C. K., Ebeneezar, S., Gills, R. and Jaggi,
...response surface design to missing observations.
...70-279.

...C. and Nwanya, J. C. (2009). On comparing the prediction
...ce designs in spherical regions: A review. *Statistica*, 69(4):285-298.

..., C. M., Anderson-Cook, C. and Montgomery, D. C. (2009). Graphical
... prediction variance performance for variations of the central composite
...actors. *Quality Technology & Quantitative Management*, 6(4):433-449.

...wuamaeze, C. U. (2021). Optimal prediction variance properties of some central
co... ...ce designs in the hypercube. *Communications in Statistics-Theory and Methods*, 50(8):1911-
1924.

[13] Gokul, A. R. and Pachamuthu, M. (2024). A graphical study on the missing data of central composite design within a spherical region. *Reliability: Theory & Applications*, 19(4(80)):228-239.

[14] Gokul, A. R. and Pachamuthu, M. (2024). Optimality prediction of second order boxbehnken design robust to missing observation. *Reliability: Theory & Applications*, 19(4(80)):635-647.

[15] Box, G. E., and Wilson, K. B. (1992). On the experimental attainment of optimum conditions. In Breakthroughs in statistics: methodology and distribution. *New York, NY: Springer New York,* 270-310.

Withdrawal

[16] Hartley, H. O. (1959). Smallest composite designs for quadratic response surfaces. *Biometrics*, 15(4), 611-624.

[17] Box, G. E., and Hunter, J. S. (1957). Multi-factor experimental designs for exploring response surfaces. *The Annals of Mathematical Statistics*, 195-241.

[18] Myers, R. H., Montgomery, D. C. and Anderson-Cook, C. M. Response surface methodology: process and product optimization using designed experiments. John Wiley & Sons, 2016.

[19] Anderson, M. J., and Whitcomb, P. J. RSM simplified: optimizing processes using response surface methods for design of experiments. Productivity press, 2016.

[20] Tarapara, B. and Divecha, J. (2023). Practical efficient alpha: Encompassing all alph of central composite design. *Communications in Statistics-Simulation and Computation*, 52(4):1

[21] Borkowski, J. J. (1995). Spherical prediction-variance properties of centra
Box—Behnken designs. *Technometrics*, 37(4), 399-410.

[22] Iwundu, M. P. (2017). Missing observations: The loss in relati
International journal of advanced mathematical sciences, 5(2), 43-9.

[23] Ozol-Godfrey, A., Anderson-Cook, C. M. and Montgo
space plots for examining model robustness. *Journal of qu*

[24] Zahran, A., Anderson-Cook, C. M. and M
assess prediction capability of response surface

[25] Herzberg, A. M., and Andrews, D
experiments in non-optimal situatio
Methodology, 38(3):284-289.

Withdrawal

RT&A, No 2 (84)
Volume 20, June 2025

FIXED POINT THEOREMS FOR SEQUENCE OF MAPPINGS IN TRICOMPLEX VALUED METRIC SPACE

Shivani Chourasiya[1], Kavita Shrivastava[2]

[1,2]Department of Mathematics and Statistics, Dr. Harisingh Gour Vishwavidyalaya, Sagar, M.P., 470003

[1]shivanichourasiya10@gmail.com, [2]kavita.rohit@rediffmail.com

Abstract

Fixed point theory plays a crucial role in mathematical analysis, with profound applications in differential equations, optimization, and dynamic systems. Over the years, researchers have extended classical fixed point results to more complex structures, such as tricomplex valued metric spaces, to address problems in multidimensional and hypercomplex settings. Despite the growing interest in tricomplex variable spaces, existing literature on fixed point theorems within these frameworks remains limited by various constraints. The existing literature exhibits several limitations; therefore, in this article, we utilize the concepts of continuity and weak compatibility to establish fixed point results for a novel class of contraction mappings by employing a sequence of mappings within a tricomplex valued metric space. If two mappings commute at their point of coincidence, they are said to be weakly compatible. These results not only deepen the comprehension of fixed point theory but also open up new possibilities for its application in more intricate and varied mathematical contexts. As a result, our research progresses in the field, providing a solid foundation for future investigations and potential uses across multiple scientific and engineering domains.

Keywords: fixed point, complex valued metric space, tricomplex valued metric space.

1. Introduction

In 1922 [1] introduced us to the Banach contraction principle that paved the way for developing the metric fixed-point theory. Consequently, many researchers generalize these results in different ways [see [2], [3], [4], [5], [6], [7], [8], etc.]. In 2011 [9] introduced complex valued metric space and proved the existence of common fixed points of a pair of mappings that satisfy contractive conditions. Extended and improved the contraction condition of these results and gave applications by [10]. In 2013, [11] generalized complex-valued b-metric spaces, providing the fixed point theorem. In 2013 [12] used the EA property and proved common fixed point theorems in complex valued metric spaces. Rational contraction in complex valued metric space and used point-dependent control functions by [13]. In 2018 [14] introduced us to multicomplex spaces that show the direction to the researcher to generalization of multicomplex spaces. Introduced the notion of complex valued metric spaces [15]. Bicomplex metric space given by [16]. In 2022 [17] proved common fixed point theorems in tricomplex valued metric spaces. By using the control function [18] proved fixed point results in tricomplex valued metric space. In 2023 [19] worked on bicomplex metric space with rational contractions using control functions of one variable. [see also [20], [21] etc.].

In 1982 [22] introduced weak commutativity, a pair of self-mappings (P, Q) on a metric space (Δ, Ω) if $\Omega(PQ\mu, QP\mu) \leq \Omega(Q\mu, P\mu), \forall \mu \in \Delta$. In 1986 [23] introduced compatible, if (P, Q) is a pair of self-mappings on a metric space (Δ, Ω) such that $\lim_{n\to\infty} \Omega(P\mu_n, Q\mu_n) = 0$ whenever

$\lim_{n\to\infty} P\mu_n = \lim_{n\to\infty} Q\mu_n = l \in \Delta$. According to [23] a pair of mappings is said to be weakly compatible if they commute at their coincidence point. In this paper, we generalize these results in a new type of contraction mapping by using a sequence of mappings.

2. Preliminaries

In this paper, we denote the families of real, complex, bicomplex, and tricomplex numbers, by C^0, C^1, C^2 and C^3 respectively. In 2018 [14] introduced us multicomplex spaces and define bicomplex and tricomplex numbers as follows:
Bicomplex numbers:

$$\sigma = \rho_1 + \rho_2 i_1 + \rho_3 i_2 + \rho_4 i_1 i_2,$$

where $\rho_1, \rho_1, \rho_1, \rho_1 \in C^0$ and independent units i_1 and i_2 such that $i_1^2 = i_2^2 = -1$ and $i_1 i_2 = i_2 i_1$ the set of bicomplex numbers define as

$$C^2 = \{\sigma : \sigma = \rho_1 + \rho_2 i_1 + \rho_3 i_2 + \rho_4 i_1 i_2, \rho_1, \rho_2, \rho_3, \rho_4 \in C^0\}$$

Or

$$C^2 = \{\sigma : \sigma = \Theta_1 + i_2\Theta_2, \Theta_1, \Theta_2 \in C^1\}$$

where $\Theta_1 = \rho_1 + \rho_2 i_1$ and $\Theta_2 = \rho_3 + \rho_4 i_1$.
Tricomplex numbers:

$$\Gamma = \rho_1 + \rho_2 i_1 + \rho_3 i_2 + \rho_4 j_1 + \rho_5 i_3 + \rho_6 j_2 + \rho_7 j_3 + \rho_8 i_4$$

Where $\rho_1, \rho_2, \rho_3, \rho_4, \rho_5, \rho_6, \rho_7, \rho_8 \in C^0$ and $i_1, i_2, i_3, i_4, j_1, j_2$ and j_3 are independent units such that $i_1^2 = i_2^2 = i_3^2 = i_4^2 = -1, j_1^2 = j_2^2 = j_3^2 = 1, j_1 = i_1 i_2 = i_2 i_1, j_2 = i_1 i_3 = i_3 i_1$ and, $i_4 = i_1 i_3 = i_1 i_2 i_3$. the set of tricomplex numbers define as

$$C^3 = \{\Gamma : \Gamma = \rho_1 + \rho_2 i_1 + \rho_3 i_2 + \rho_4 j_1 + \rho_5 i_3 + \rho_6 j_2 + \rho_7 j_3 + \rho_8 i_4, \rho_1, \rho_2, \rho_3, \rho_4, \rho_5, \rho_6, \rho_7, \rho_8 \in C^0\}$$

$$C^3 = \{\Gamma : \Gamma = \gamma_1 + i_3\gamma_2, \gamma_1, \gamma_2 \in C^2\}$$

where $\gamma_1 = \Theta_1 + \Theta_2 i_2 \in C^2$ and $\gamma_2 = \Theta_3 + \Theta_4 i_2 \in C^2$
Sum and Product of tricomplex numbers:
Let two tricomplex numbers Γ and Δ such that $\Gamma = \gamma_1 + i_3\gamma_2$ and $\Delta = \delta_1 + i_3\delta_2$ then

$$\Gamma \pm \Delta = (\gamma_1 + i_3\gamma_2) \pm (\delta_1 + i_3\delta_2) = \gamma_1 \pm \delta_1 + i_3(\gamma_2 \pm \delta_2)$$

and

$$\Gamma.\Delta = (\gamma_1 + i_3\gamma_2)(\delta_1 + i_3\delta_2) = (\gamma_1\delta_1 - \gamma_2\delta_2) + i_3(\gamma_1\delta_2 + \gamma_2\delta_1)$$

In the families of tricomplex numbers C^3 have four idempotent elements $0, 1, \varphi_1 = \frac{1+j_3}{2}$ and $\varphi_2 = \frac{1-j_3}{2}$ such that $\varphi_1 + \varphi_2 = 1$ and $\varphi_1.\varphi_2 = 0$. Hence φ_1, φ_2 are not trivial. Every tricomplex numbers can be expressed as an union of φ_1, φ_2 for example

$$\Gamma = \gamma_1 + i_3\gamma_2 = (\gamma_1 - i_2\gamma_2)\varphi_1 + (\gamma_1 + i_2\gamma_2)\varphi_2$$

Here Γ is idempotent of the tricomplex numeral and $\Gamma_1 = (\gamma_1 - i_2\gamma_2)\varphi_1$ and $\Gamma_2 = (\gamma_1 + i_2\gamma_2)\varphi_2$ are the idempotent components of the bicomplex numeral Γ.
A tricomplex number $\Gamma = \gamma_1 + i_3\gamma_2$ is invertible if there exists $\Delta \in C^3$ such that $\Gamma\Delta = 1$. A tricomplex number $\Gamma = \gamma_1 + i_3\gamma_2$ is nonsingular if $|\gamma_1^2 + \gamma_2^2| \neq 0$ and singular if $|\gamma_1^2 + \gamma_2^2| = 0$. Let $\Gamma = \gamma_1 + i_3\gamma_2$ and $\Delta = \delta_1 + i_3\delta_2$ be tricomplex numerals. Then the partial order relation $\leq_{i_3}$ on C^3 or $\Gamma \leq_{i_3} \Delta$ if the following axioms are satisfied.

1. $\gamma_1 = \delta_1, \gamma_2 = \delta_2$;
2. $\gamma_1 <_{i_3} \delta_1, \gamma_2 = \delta_2$;
3. $\gamma_1 = \delta_1, \gamma_2 <_{i_3} \delta_2$;
4. $\gamma_1 <_{i_3} \delta_1, \gamma_2 <_{i_3} \delta_2$;

If $\Gamma \lesssim_{i_3} \Delta$ implies $\Gamma <_{i_3} \Delta$ but $\Gamma \neq \Delta$
Having two tricomplex numerals $\Gamma, \Delta \in C^3$, the following conditions hold:

1. $\Gamma \leq_{i_3} \Delta$ if $||\Gamma|| \leq ||\Delta||$;
2. $||\Gamma + \Delta|| \leq ||\Gamma|| + ||\Delta||$;
3. $||\rho\Gamma|| = |\rho|||\Gamma||$, where $\rho \in C^0$
4. $||\Gamma\Delta|| \leq 2||\Gamma||||\Delta||$, If at least one of Γ, Δ nonsingular then equality holds.
5. $||\Gamma^{-1}|| = ||\Gamma||^{-1}$, if $||\Gamma|| \neq 0$
6. $||\frac{\Gamma}{\Delta}|| = \frac{||\Gamma||}{||\Delta||}$, if $||\Delta|| \neq 0$

Assume that $w_1, w_2 \in C^1$, where C^1 is the set of complex numbers that define a partial order $\leq$ on C^1 such that
$w_1 \leq w_2$ if and only if $Re(w_1) \leq Re(w_2)$ and $Im(w_1) \leq Im(w_2)$ that is $w_1 \leq w_2$. If any of the following is true [9]

1. $Re(w_1) = Re(w_2)$ and $Im(w_1) = Im(w_2)$;
2. $Re(w_1) < Re(w_2)$ and $Im(w_1) = Im(w_2)$;
3. $Re(w_1) = Re(w_2)$ and $Im(w_1) < Im(w_2)$;
4. $Re(w_1) < Re(w_2)$ and $Im(w_1) < Im(w_2)$;

we will write $w_1 \lesssim w_2$ if $w_1 \neq w_2$ and one of (2),(3), and (4) satisfiy. Take note of that

$$0 \leq w_1 \lesssim w_2 \Rightarrow |w_1| < |w_2|$$
$$w_1 \leq w_2 \text{ and } w_2 < w_3 \Rightarrow w_1 < w_3.$$

Let Δ be a set that is not empty. A mapping $\Omega : \Delta \times \Delta \to C^1$ is considered to be congruent with the following conditions [9]:

1. $0 \leq \Omega(\mu, \nu)$ for all $\mu, \nu \in \Delta$ and $\Omega(\mu, \nu) = 0$ if and only if $\mu = \nu$;
2. $\Omega(\mu, \nu) = \Omega(\nu, \mu)$, for all $\mu, \nu \in \Delta$;
3. $\Omega(\mu, \nu) \leq \Omega(\mu, z) + \Omega(z, \nu)$ for all $\mu, \nu, z \in \Delta$.

In this case, the metric Ω is referred to as a complex valued metric on the set Δ, and the space (Δ, Ω) is referred to as a complex valued metric space.
Let $\Delta \neq 0$. A mapping $\Omega : \Delta \times \Delta \longrightarrow C^3$ is called a tricomplex-valued metric if the following conditions are satisfied for all $\mu, \nu \in \Delta$ [18]:

1. $\Omega(\mu, \nu) \geq_{i_3} 0$ and $\Omega(\mu, \nu) = 0$ if and only if $\mu = \nu$
2. $\Omega(\mu, \nu) = \Omega(\nu, \mu)$
3. $\Omega(\mu, \nu) \leq_{i_3} \Omega(\mu, z) + \Omega(z, \nu)$, $\forall \mu, \nu, z \in \Delta$

Then (Δ, Ω) is called tricomplex-valued metric space.
Let C^3 is a tricomplex number and $\Omega : C^3 \times C^3 \longrightarrow C^3$ such that

$$\Omega(\rho_1, \rho_2) = |\mu_1 - \mu_2| + i_3|\nu_1 - \nu_2|$$

Where $\rho_1 = \mu_1 + i_3\nu_1$ and $\rho_2 = \mu_2 + i_3\nu_2$, then (C^3, Ω) is a tricomplex valued metric spaces. Let (Δ, Ω) is a tricomplex-valued metric space then [17]

1. A sequence $\{\nu_n\}$ is a Cauchy sequence if for all $\epsilon \in C^3$ and $\epsilon >_{i_3} 0$, there exist an integer N such that $\Omega(\nu_n, \nu_m) <_{i_3} \epsilon, \forall n, m > N$
2. A sequence $\{\nu_n\}$ is converges to if for all $\epsilon \in C^3$ and $\epsilon >_{i_3} 0$ there exist an integer N such that $\Omega(\nu_n, \nu) <_{i_3} \epsilon, \forall n > N$
3. tricomplex valued metric space (Δ, Ω) is complete if every Cauchy sequence in Δ converges in Δ.

Lemma 1. A sequence $\{\nu_n\}$ in a tricomplex-valued metric space (Δ, Ω) converges to ν if and only if $||\Omega(\nu_n, \nu)|| \longrightarrow 0$ as limit $\nu \longrightarrow \infty$ [18].

Lemma 2. A sequence $\{\nu_n\}$ in a tricomplex-valued metric space (Δ, Ω) is Cauchy if and only if $||\Omega(\nu_n, \nu_{n+m})|| \longrightarrow 0$ as limit $\nu \longrightarrow \infty$ [18].

3. Main Results

Theorem 1. Let (Δ, Ω) be a tricomplex valued metric space and P, Q be the self mappings and A^n and B^n are self mappings in Δ, such that for all $\mu, \nu \in \Delta$

1. $P(\Delta) \subset B^n(\Delta)$ and $Q(\Delta) \subset A^n(\Delta)$;
2. (A^n, P) and (B^n, Q) are commuting Pairs;
3. (A^n, P) and (B^n, Q) are commuting;
4. If $\Omega(Q\nu, A^n\mu) + \Omega(P\mu, B^n\nu) + \Omega(B^n\nu, A^n\mu) \neq 0$ then the following holds

$$\begin{aligned}\Omega(P\mu, Q\nu) \leq_{i_3} \delta_1 \left\{ \frac{\Omega(A^n\mu, P\mu) + \Omega(B^n\nu, Q\nu)\Omega(A^n\mu, P\mu)}{1 + \Omega(P\mu, Q\nu)} \right\} \\ +\delta_2 \max\{\Omega(A^n\mu, B^n\nu), \Omega(A^n\mu, P\mu), \Omega(B^n\nu, P\mu)\} \\ +\delta_3 \{\Omega(B^n\nu, Q\nu) + \Omega(Q\nu, A^n\mu) + \Omega(P\mu, B^n\nu)\} \\ +\delta_4 \left\{ \frac{\Omega(Q\nu, B^n\nu)\Omega(P\mu, A^n\mu)}{\Omega(Q\nu, A^n\mu) + \Omega(P\mu, B^n\nu) + \Omega(B^n\nu, A^n\mu)} \right\} \\ +\delta_5 \max\{\Omega(A^n\mu, P\mu), \Omega(B^n\nu, Q\mu), \Omega(A^n\mu, B^n\nu), \Omega(P\mu, Q\mu)\}\end{aligned}$$

Where $\delta_1 + \delta_2 + 2\delta_3 + \delta_4 + \delta_5 < 1$. For all $\mu, \nu \in \Delta$. Then A^n, B^n, P and Q have a unique common fixed point.

Proof. Let $\mu_0 \in \Delta$, since $P(\Delta) \subset B^n(\Delta)$ and $Q(\Delta) \subset A^n(\Delta)$ define for all $n \geq 0$ the sequence

$\{\nu_n\} \in \Delta$ by $\nu_{n+1} = P\mu_n = B^n\mu_{n+1}$ and $\nu_{n+2} = Q\mu_{n+1} = A^n\mu_{n+2}$ then,

$$\begin{aligned}
\Omega(\nu_{n+1}, \nu_{n+2}) &= \Omega(P\mu_n, Q\mu_{n+1}) \\
&\leq_{i_3} \delta_1 \left\{ \frac{\Omega(A^n\mu_n, P\mu_n) + \Omega(B^n\mu_{n+1}, Q\mu_{n+1})\Omega(A^n\mu_n, P\mu_n)}{1 + \Omega(P\mu_n, Q\mu_{n+1})} \right\} \\
&+\delta_2 \max\{\Omega(A^n\mu_n, B^n\mu_{n+1}), \Omega(A^n\mu_n, P\mu_n), \Omega(B^n\mu_{n+1}, P\mu_n)\} \\
&+\delta_3 \{\Omega(B^n\mu_{n+1}, Q\mu_{n+1}) + \Omega(Q\mu_{n+1}, A^n\mu_n) + \Omega(P\mu_n, B^n\mu_{n+1})\} \\
&+\delta_4 \left\{ \frac{\Omega(Q\mu_{n+1}, B^n\mu_{n+1})\Omega(P\mu_n, A^n\mu_n)}{\Omega(Q\mu_{n+1}, A^n\mu_n) + \Omega(P\mu_n, B^n\mu_{n+1}) + \Omega(B^n\mu_{n+1}, A^n\mu_n)} \right\} \\
&+\delta_5 \max\{\Omega(A^n\mu_n, P\mu_n), \Omega(B^n\mu_{n+1}, Q\mu_n), \Omega(A^n\mu_n, B^n\mu_{n+1}), \Omega(P\mu_n, Q\mu_n)\} \\
&\leq_{i_3} \delta_1 \left\{ \frac{\Omega(\nu_n, \nu_{n+1}) + \Omega(\nu_{n+1}, \nu_{n+2})\Omega(\nu_n, \nu_{n+1})}{1 + \Omega(\nu_{n+1}, \nu_{n+2})} \right\} \\
&+\delta_2 \max\{\Omega(\nu_n, \nu_{n+1}), \Omega(\nu_n, \nu_{n+1}), \Omega(\nu_{n+1}, \nu_{n+1})\} \\
&+\delta_3 \{\Omega(\nu_{n+1}, \nu_{n+2}) + \Omega(\nu_{n+2}, \nu_n) + \Omega(\nu_{n+1}, \nu_{n+1})\} \\
&+\delta_4 \left\{ \frac{\Omega(\nu_{n+2}, \nu_{n+1})\Omega(\nu_{n+1}, \nu_n)}{\Omega(\nu_{n+2}, \nu_n) + \Omega(\nu_{n+1}, \nu_{n+1}) + \Omega(\nu_{n+1}, \nu_n)} \right\} \\
&+\delta_5 \max\{\Omega(\nu_n, \nu_{n+1}), \Omega(\nu_{n+1}, \nu_{n+1}), \Omega(\nu_n, \nu_{n+1}), \Omega(\nu_{n+1}, \nu_{n+1})\}
\end{aligned}$$

$$\begin{aligned}
\implies ||\Omega(\nu_{n+1}, \nu_{n+2})|| &\leq |\delta_1|||\Omega(\nu_n, \nu_{n+1})|| + |\delta_2|||\Omega(\nu_n, \nu_{n+1})|| + |\delta_3|||\Omega(\nu_n, \nu_{n+1})|| \\
&+|\delta_4|||\Omega(\nu_n, \nu_{n+1})|| + |\delta_5|||\Omega(\nu_n, \nu_{n+1})|| \\
&\leq |(\delta_1 + \delta_2 + \delta_3 + \delta_4 + \delta_5)|||\Omega(\nu_n, \nu_{n+1})||
\end{aligned}$$

Now, $||\Omega(\nu_{n+1}, \nu_{n+2})|| \leq |(\delta_1 + \delta_2 + \delta_3 + \delta_4 + \delta_5)|||\Omega(\nu_n, \nu_{n+1})||$
Similarly,

$$\begin{aligned}
\Omega(\nu_{n+3}, \nu_{n+4}) &= \Omega(P\mu_{n+2}, Q\mu_{n+3}) \\
&\leq_{i_3} \delta_1 \left\{ \frac{\Omega(A^n\mu_{n+2}, P\mu_{n+2}) + \Omega(B^n\mu_{n+3}, Q\mu_{n+3})\Omega(A^n\mu_{n+2}, P\mu_{n+2})}{1 + \Omega(P\mu_{n+2}, Q\mu_{n+3})} \right\} \\
&+\delta_2 \max\{\Omega(A^n\mu_{n+2}, B^n\mu_{n+3}), \Omega(A^n\mu_{n+2}, P\mu_{n+2}), \Omega(B^n\mu_{n+3}, P\mu_{n+2})\} \\
&+\delta_3 \{\Omega(B^n\mu_{n+3}, Q\mu_{n+3}) + \Omega(Q\mu_{n+3}, A^n\mu_{n+2}) + \Omega(P\mu_{n+2}, B^n\mu_{n+3})\} \\
&+\delta_4 \left\{ \frac{\Omega(Q\mu_{n+3}, B^n\mu_{n+3})\Omega(P\mu_{n+2}, A^n\mu_{n+2})}{\Omega(Q\mu_{n+3}, A^n\mu_{n+2}) + \Omega(P\mu_{n+2}, B^n\mu_{n+3}) + \Omega(B^n\mu_{n+3}, A^n\mu_{n+2})} \right\} \\
&+\delta_5 \max\{\Omega(A^n\mu_{n+2}, P\mu_{n+2}), \Omega(B^n\mu_{n+3}, Q\mu_{n+2}), \Omega(A^n\mu_{n+2}, B^n\mu_{n+3}), \Omega(P\mu_{n+2}, Q\mu_{n+2})\} \\
&\leq_{i_3} \delta_1 \left\{ \frac{\Omega(\nu_{n+2}, \nu_{n+3}) + \Omega(\nu_{n+3}, \nu_{n+4})\Omega(\nu_{n+2}, \nu_{n+3})}{1 + \Omega(\nu_{n+3}, \nu_{n+4})} \right\} \\
&+\delta_2 \max\{\Omega(\nu_{n+2}, \nu_{n+3}), \Omega(\nu_{n+2}, \nu_{n+3}), \Omega(\nu_{n+3}, \nu_{n+3})\} \\
&+\delta_3 \{\Omega(\nu_{n+3}, \nu_{n+4}) + \Omega(\nu_{n+4}, \nu_{n+2}) + \Omega(\nu_{n+3}, \nu_{n+3})\} \\
&+\delta_4 \left\{ \frac{\Omega(\nu_{n+4}, \nu_{n+3})\Omega(\nu_{n+3}, \nu_{n+2})}{\Omega(\nu_{n+4}, \nu_{n+2}) + \Omega(\nu_{n+3}, \nu_{n+3}) + \Omega(\nu_{n+3}, \nu_{n+2})} \right\} \\
&+\delta_5 \max\{\Omega(\nu_{n+2}, \nu_{n+3}), \Omega(\nu_{n+3}, \nu_{n+3}), \Omega(\nu_{n+2}, \nu_{n+3}), \Omega(\nu_{n+3}, \nu_{n+3})\}
\end{aligned}$$

$$\begin{aligned}
\implies ||\Omega(\nu_{n+3}, \nu_{n+4})|| &\leq \delta_1||\Omega(\nu_{n+2}, \nu_{n+3})|| + \delta_2||\Omega(\nu_{n+2}, \nu_{n+3})|| + \delta_3||\Omega(\nu_{n+2}, \nu_{n+3})|| \\
&+\delta_4||\Omega(\nu_{n+2}, \nu_{n+3})|| + \delta_5||\Omega(\nu_{n+2}, \nu_{n+3})|| \\
&\leq (\delta_1 + \delta_2 + \delta_3 + \delta_4 + \delta_5)||\Omega(\nu_{n+2}, \nu_{n+3})||
\end{aligned}$$

Now, $||\Omega(\nu_{n+3}, \nu_{n+4})|| \leq (\delta_1 + \delta_2 + \delta_3 + \delta_4 + \delta_5)||\Omega(\nu_{n+2}, \nu_{n+3})||$
Preceding in this way we get

$||\Omega(\nu_n, \nu_{n+1})|| \leq (\delta_1 + \delta_2 + \delta_3 + \delta_4 + \delta_5)||\Omega(\nu_{n-1}, \nu_n)||$
$\implies ||\Omega(\nu_n, \nu_{n+1})|| \leq \delta||\Omega(\nu_{n-1}, \nu_n)||$
Where $(\delta = \delta_1 + \delta_2 + \delta_3 + \delta_4 + \delta_5) < 1$

$$\begin{aligned} ||\Omega(\nu_n, \nu_{n+1})|| &\leq \delta||\Omega(\nu_{n-1}, \nu_n)|| \\ &\leq \delta^2||\Omega(\nu_{n-2}, \nu_{n-1})|| \\ &\leq \delta^3||\Omega(\nu_{n-3}, \nu_{n-2})|| \\ &\ldots \leq \delta^n||\Omega(\nu_0, \nu_1)|| \end{aligned}$$

Now $\forall k > n$, we have,

$$||\Omega(\nu_k, \nu_n)|| \leq \delta^n||\Omega(\nu_0, \nu_1)|| + \delta^{n-1}||\Omega(\nu_0, \nu_1)|| + \delta^{n-2}||\Omega(\nu_0, \nu_1)|| + \ldots + \delta^{k-1}||\Omega(\nu_0, \nu_1)||$$

$\implies ||\Omega(\nu_k, \nu_n)|| \leq (\delta^n + \delta^{n-1} + \delta^{n-2} + \delta^{k-1})||\Omega(\nu_0, \nu_1)||$
$\implies ||\Omega(\nu_k, \nu_n)|| \leq \frac{\delta^n}{1-\delta}||\Omega(\nu_0, \nu_1)||$
as $\delta < 1$ then $\lim_{n,m\to\infty} ||\Omega(\nu_k, \nu_n)|| = 0$
Hence $\{\nu_n\}$ is a Cauchy sequence. Since Δ is complete then ν_n and its subsequences $\{P\mu_n\}, \{B^n\mu_{n+1}\}, \{Q\mu_{n+1}\}$ and $\{A^n\mu_{n+2}\}$ also converges to ω
$\lim_{n\to\infty} \nu_n = \lim_{n\to\infty} P\mu_n = \lim_{n\to\infty} B^n\mu_{n+1} = \lim_{n\to\infty} Q\mu_{n+1} = \lim_{n\to\infty} A^n\mu_{n+2} = \omega$
Let there exist $\omega* \in \Delta$ such that $\nu_n \to \omega*$ as $n \to \infty$, and $P\omega* = A^n\omega* = B^n\omega* = Q\omega* = \omega$
(A^n, P) and (B^n, Q) are weakly compatible, then they commute at their coincidence point.
Therefore
$P\omega = PA^n\omega* = A^nP\omega* = A^n\omega$ and $B^n\omega = B^nQ\omega* = QB^n\omega* = Q\omega$
Now we show that $P\omega = Q\omega$ then we have

$$\begin{aligned} \Omega(P\omega, Q\mu_{n+1}) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(A^n\omega, P\omega) + \Omega(B^n\mu_{n+1}, Q\mu_{n+1})\Omega(A^n\omega, P\omega)}{1 + \Omega(P\omega, Q\mu_{n+1})}\right\} \\ &+\delta_2 \max\{\Omega(A^n\omega, B^n\mu_{n+1}), \Omega(A^n\omega, P\omega), \Omega(B^n\mu_{n+1}, P\omega)\} \\ &+\delta_3\{\Omega(B^n\mu_{n+1}, Q\mu_{n+1}) + \Omega(Q\mu_{n+1}, A^n\omega) + \Omega(P\omega, B^n\mu_{n+1})\} \\ &+\delta_4\left\{\frac{\Omega(P\mu_{n+1}, B^n\mu_{n+1})\Omega(P\omega, A^n\omega)}{\Omega(Q\mu_{n+1}, A^n\omega) + \Omega(P\omega, B^n\mu_{n+1}) + \Omega(B^n\mu_{n+1}, A^n\omega)}\right\} \\ &+\delta_5 \max\{\Omega(A^n\omega, P\omega), \Omega(B^n\mu_{n+1}, Q\omega), \Omega(A^n\omega, B^n\mu_{n+1}), \Omega(P\omega, Q\omega)\} \end{aligned}$$

$$\begin{aligned} \implies \Omega(P\omega, \nu_{n+2}) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(P\omega, P\omega) + \Omega(\nu_{n+1}, \nu_{n+2})\Omega(P\omega, P\omega)}{1 + \Omega(P\omega, \nu_{n+2})}\right\} \\ &+\delta_2 \max\{\Omega(P\omega, \nu_{n+1}), \Omega(P\omega, P\omega), \Omega(\nu_{n+1}, P\omega)\} \\ &+\delta_3\{\Omega(\nu_{n+1}, \nu_{n+2}) + \Omega(\nu_{n+2}, P\omega) + \Omega(P\omega, \nu_{n+1})\} \\ &+\delta_4\left\{\frac{\Omega(\nu_{n+2}, \nu_{n+1})\Omega(P\omega, P\omega)}{\Omega(\nu_{n+2}, P\omega) + \Omega(P\omega, \nu_{n+1}) + \Omega(\nu_{n+1}, P\omega)}\right\} \\ &+\delta_5 \max\{\Omega(P\omega, P\omega), \Omega(\nu_{n+1}, P\omega), \Omega(P\omega, \nu_{n+1}), \Omega(P\omega, P\omega)\} \end{aligned}$$

If we take $n \to \infty$, then we have

$$\begin{aligned} \implies \Omega(P\omega, \omega) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(P\omega, P\omega) + \Omega(\omega, \omega)\Omega(P\omega, P\omega)}{1 + \Omega(P\omega, \omega)}\right\} \\ &+\delta_2 \max\{\Omega(P\omega, \omega), \Omega(P\omega, P\omega), \Omega(\omega, P\omega)\} \\ &+\delta_3\{\Omega(\omega, \omega) + \Omega(\omega, P\omega) + \Omega(P\omega, \omega)\} \\ &+\delta_4\left\{\frac{\Omega(\omega, \omega)\Omega(P\omega, P\omega)}{\Omega(\omega, P\omega) + \Omega(P\omega, \omega) + \Omega(\omega, P\omega)}\right\} \\ &+\delta_5 \max\{\Omega(P\omega, P\omega), \Omega(\omega, P\omega), \Omega(P\omega, \omega), \Omega(P\omega, P\omega)\} \end{aligned}$$

$\Longrightarrow ||\Omega(P\omega,\omega)|| \leq \delta_2||\Omega(P\omega,\omega)|| + 2\delta_3||\Omega(P\omega,\omega)|| + \delta_5||\Omega(P\omega,\omega)||$
$\Longrightarrow ||\Omega(P\omega,\omega)|| \leq (\delta_2 + 2\delta_3 + \delta_5)||\Omega(P\omega,\omega)||$
$\Longrightarrow ||\Omega(P\omega,\omega)|| \leq (\delta_2 + 2\delta_3 + \delta_5)||\Omega(P\omega,\omega)||$
This is contradiction that $(\delta_2 + 2\delta_3 + \delta_5) < 1$ therefore $P\omega = \omega$ and $A^n\omega = P\omega \Rightarrow A^n\omega = \omega$.
Now we will prove that $Q\omega = \omega$, putting $\mu = \nu = \omega$ then we have

$$\begin{aligned}\Omega(P\omega,Q\omega) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(A^n\omega,P\omega)+\Omega(B^n\omega,Q\omega)\Omega(A^n\omega,P\omega)}{1+\Omega(P\omega,Q\omega)}\right\}\\ &+\delta_2 \max\{\Omega(A^n\omega,B^n\omega),\Omega(A^n\omega,P\omega),\Omega(B^n\omega,P\omega)\}\\ &+\delta_3\{\Omega(B^n\omega,Q\omega)+\Omega(Q\omega,A^n\omega)+\Omega(P\omega,B^n\omega)\}\\ &+\delta_4\left\{\frac{\Omega(Q\omega,B^n\omega)\Omega(P\omega,A^n\omega)}{\Omega(Q\omega,A^n\omega)+\Omega(P\omega,B^n\omega)+\Omega(B^n\omega,A^n\omega)}\right\}\\ &+\delta_5 \max\{\Omega(A^n\omega,P\omega),\Omega(B^n\omega,Q\omega),\Omega(A^n\omega,B^n\omega),\Omega(P\omega,Q\omega)\}\end{aligned}$$

Taking $n \to \infty$ we get

$$\begin{aligned}\Omega(\omega,Q\omega) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(\omega,\omega)+\Omega(Q\omega,Q\omega)\Omega(\omega,\omega)}{1+\Omega(\omega,Q\omega)}\right\}\\ &+\delta_2 \max\{\Omega(\omega,Q\omega),\Omega(\omega,\omega),\Omega(Q\omega,\omega)\}\\ &+\delta_3\{\Omega(Q\omega,Q\omega)+\Omega(Q\omega,\omega)+\Omega(\omega,Q\omega)\}\\ &+\delta_4\left\{\frac{\Omega(Q\omega,Q\omega)\Omega(\omega,\omega)}{\Omega(Q\omega,\omega)+\Omega(\omega,Q\omega)+\Omega(Q\omega,\omega)}\right\}\\ &+\delta_5 \max\{\Omega(\omega,\omega),\Omega(Q\omega,Q\omega),\Omega(\omega,Q\omega),\Omega(\omega,Q\omega)\}\end{aligned}$$

$\Longrightarrow \Omega(\omega,Q\omega) \leq_{i_3} \delta_2\Omega(\omega,Q\omega) + 2\delta_3\Omega(\omega,Q\omega) + \delta_5\Omega(\omega,Q\omega)$
$\Longrightarrow ||\Omega(\omega,Q\omega)|| \leq (\delta_2 + 2\delta_3 + \delta_5)||\Omega(\omega,Q\omega)||$
Then
$\Longrightarrow ||\Omega(\omega,Q\omega)|| \leq (\delta_2 + 2\delta_3 + \delta_5)||\Omega(\omega,Q\omega)||$ Which is again a contradiction. Therefore $Q\omega = \omega$ and $B^n\omega = Q\omega \Longrightarrow B^n\omega = \omega$.
Uniqueness of fixed point
Let ω_0 be another fixed point of P, Q, A^n, B^n then we have

$$\begin{aligned}\Omega(P\omega,Q\omega_0) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(A^n\omega,P\omega)+\Omega(B^n\omega_0,Q\omega_0)\Omega(A^n\omega,P\omega)}{1+\Omega(P\omega,Q\omega_0)}\right\}\\ &+\delta_2 \max\{\Omega(A^n\omega,B^n\omega_0),\Omega(A^n\omega,P\omega),\Omega(B^n\omega_0,P\omega)\}\\ &+\delta_3\{\Omega(B^n\omega_0,Q\omega_0)+\Omega(Q\omega_0,A^n\omega)+\Omega(P\omega,B^n\omega_0)\}\\ &+\delta_4\left\{\frac{\Omega(Q\omega_0,B^n\omega_0)\Omega(P\omega,A^n\omega)}{\Omega(Q\omega_0,A^n\omega)+\Omega(P\omega,B^n\omega_0)+\Omega(B^n\omega_0,A^n\omega)}\right\}\\ &+\delta_5 \max\{\Omega(A^n\omega,P\omega),\Omega(B^n\omega_0,Q\omega),\Omega(A^n\omega,B^n\omega_0),\Omega(P\omega,Q\omega)\}\end{aligned}$$

Taking limit $n \to \infty$

$$\begin{aligned}\Omega(\omega,\omega_0) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(\omega,\omega)+\Omega(\omega_0,\omega_0)\Omega(\omega,\omega)}{1+\Omega(\omega,\omega_0)}\right\}\\ &+\delta_2 \max\{\Omega(\omega,\omega_0),\Omega(\omega,\omega),\Omega(\omega_0,\omega)\}\\ &+\delta_3\{\Omega(\omega_0,\omega_0)+\Omega(\omega_0,\omega)+\Omega(\omega,\omega_0)\}\\ &+\delta_4\left\{\frac{\Omega(\omega_0,\omega_0)\Omega(\omega,\omega)}{\Omega(\omega_0,\omega)+\Omega(\omega,\omega_0)+\Omega(\omega_0,\omega)}\right\}\\ &+\delta_5 \max\{\Omega(\omega,\omega),\Omega(\omega_0,\omega),\Omega(\omega,\omega_0),\Omega(\omega,\omega)\}\end{aligned}$$

$\Longrightarrow ||\Omega(\omega,\omega_0)|| \leq \delta_2||\Omega(\omega,\omega_0)|| + 2\delta_3||\Omega(\omega,\omega_0)|| + \delta_5||\Omega(\omega,\omega_0)||$
$\Longrightarrow ||\Omega(\omega,\omega_0)|| \leq (\delta_2 + 2\delta_3 + \delta_5)||\Omega(\omega,\omega_0)||$
$\Longrightarrow ||\Omega(\omega,\omega_0)|| \leq (\delta_2 + 2\delta_3 + \delta_5)||\Omega(\omega,\omega_0)||$
Which is again contradiction therefore $\omega = \omega_0$ is a unique common fixed point of P, Q, A^n and B^n ■

Theorem 2. Let (Δ, Ω) be a tricomplex valued metric space and P, Q be the self mappings and A^n and B^n are self mappings in Δ, such that for all $\mu, \nu \in \Delta$

1. $P(\Delta) \subset B^n(\Delta)$ and $Q(\Delta) \subset A^n(\Delta)$;
2. (A^n, P) are compatible, A^n or P is continuous and (B^n, Q) are weakly compatible;
3. (B^n, Q) are compatible, B^n or Q is continuous and (A^n, P) are weakly compatible;
4. For all $\mu, \nu \in \Delta$ and $\mu \neq \nu$ we have

$$\begin{aligned}\Omega(P\mu, Q\nu) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(A^n\mu, P\mu) + \Omega(B^n\nu, Q\nu)\Omega(A^n\mu, P\mu)}{1 + \Omega(P\mu, Q\nu)}\right\} \\ &+\delta_2 \max\{\Omega(A^n\mu, B^n\nu), \Omega(A^n\mu, P\mu), \Omega(B^n\nu, P\mu)\} \\ &+\delta_3\{\Omega(B^n\nu, Q\nu) + \Omega(Q\nu, A^n\mu) + \Omega(P\mu, B^n\nu)\} \\ &+\delta_4\left\{\frac{\Omega(Q\nu, B^n\nu)\Omega(P\mu, A^n\mu)}{\Omega(Q\nu, A^n\mu) + \Omega(P\mu, B^n\nu) + \Omega(B^n\nu, A^n\mu)}\right\} \\ &+\delta_5 \max\{\Omega(A^n\mu, P\mu), \Omega(B^n\nu, Q\mu), \Omega(A^n\mu, B^n\nu), \Omega(P\mu, Q\mu)\}\end{aligned}$$

Where $\delta_1 + \delta_2 + 2\delta_3 + \delta_4 + \delta_5 < 1$. For all $\mu, \nu \in \Delta$. Then A^n, B^n, P and Q have a unique common fixed point.

Proof. By theorem (1) $\{\nu_n\}$ is a Cauchy sequence and Δ is complete, then sequence $\{\nu_n\}$ converge to ω and subsequences $\{P\mu_n\}$, $\{B^n\mu_{n+1}\}$, $\{Q\mu_{n+1}\}$ and $\{A^n\mu_{n+2}\}$ also converges to ω.

$$\Longrightarrow \lim_{n\to\infty} \nu_n = \lim_{n\to\infty} P\mu_n = \lim_{n\to\infty} B^n\mu_{n+1} = \lim_{n\to\infty} A^n\mu_{n+2} = \lim_{n\to\infty} Q\mu_{n+1} = \omega. \tag{1}$$

Since (A^n, P) are compatible, let P is continuous then we have

$$\lim_{n\to\infty} A^nP\mu_{n+2} = \lim_{n\to\infty} PA^n\mu_{n+2} = P\omega \tag{2}$$

Now putting $\mu = \mu_{n+2}, \nu = \mu_{n+1}$ we have

$$\begin{aligned}\Omega(P\mu_{n+2}, Q\mu_{n+1}) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(A^n\mu_{n+2}, P\mu_{n+2}) + \Omega(B^n\mu_{n+1}, Q\mu_{n+1})\Omega(A^n\mu_{n+2}, P\mu_{n+2})}{1 + \Omega(P\mu_{n+2}, Q\mu_{n+1})}\right\} \\ &+\delta_2 \max\{\Omega(A^n\mu_{n+2}, B^n\mu_{n+1}), \Omega(A^n\mu_{n+2}, P\mu_{n+2}), \Omega(B^n\mu_{n+1}, P\mu_{n+2})\} \\ &+\delta_3\{\Omega(B^n\mu_{n+1}, Q\mu_{n+1}) + \Omega(Q\mu_{n+1}, A^n\mu_{n+2}) + \Omega(P\mu_{n+2}, B^n\mu_{n+1})\} \\ &+\delta_4\left\{\frac{\Omega(Q\mu_{n+1}, B^n\mu_{n+1})\Omega(P\mu_{n+2}, A^n\mu_{n+2})}{\Omega(Q\mu_{n+1}, A^n\mu_{n+2}) + \Omega(P\mu, B^n\mu_{n+1}) + \Omega(B^n\mu_{n+1}, A^n\mu_{n+2})}\right\} \\ &+\delta_5 \max\{\Omega(A^n\mu_{n+2}, P\mu_{n+2}), \Omega(B^n\mu_{n+1}, Q\mu_{n+2}), \Omega(A^n\mu_{n+2}, B^n\mu_{n+1}), \Omega(P\mu_{n+2}, Q\mu_{n+2})\}\end{aligned}$$

Taking limit $n \to \infty$ and using (1) and (2), we get

$$\begin{aligned}\Longrightarrow \Omega(P\omega, \omega) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(\omega, \omega) + \Omega(\omega, \omega)\Omega(\omega, \omega)}{1 + \Omega(\omega, \omega)}\right\} \\ &+\delta_2 \max\{\Omega(\omega, \omega), \Omega(\omega, \omega), \Omega(\omega, \omega)\} \\ &+\delta_3\{\Omega(\omega, \omega) + \Omega(\omega, \omega) + \Omega(\omega, \omega)\} \\ &+\delta_4\left\{\frac{\Omega(\omega, \omega)\Omega(\omega, \omega)}{\Omega(\omega, \omega) + \Omega(\omega, \omega) + \Omega(\omega, \omega)}\right\} \\ &+\delta_5 \max\{\Omega(\omega, \omega), \Omega(\omega, \omega), \Omega(\omega, \omega), \Omega(\omega, \omega)\}\end{aligned}$$

$\Longrightarrow ||\Omega(P\omega,\omega)|| \leq 0$
$\Longrightarrow ||\Omega(P\omega,\omega)|| \leq 0$. Therefore $P\omega = \omega$.
Now again putting $\mu = \omega$ and $\nu = \mu_{n+1}$ we have

$$\begin{aligned}\Omega(P\omega, Q\mu_{n+1}) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(A^n\omega, P\omega) + \Omega(B^n\mu_{n+1}, Q\mu_{n+1})\Omega(A^n\omega, P\omega)}{1+\Omega(P\omega, Q\mu_{n+1})}\right\}\\ &+\delta_2 \max\{\Omega(A^n\omega, B^n\mu_{n+1}), \Omega(A^n\omega, P\omega), \Omega(B^n\mu_{n+1}, P\omega)\}\\ &+\delta_3\{\Omega(B^n\mu_{n+1}, Q\mu_{n+1}) + \Omega(Q\mu_{n+1}, A^n\omega) + \Omega(P\omega, B^n\mu_{n+1})\}\\ &+\delta_4\left\{\frac{\Omega(Q\mu_{n+1}, B^n\mu_{n+1})\Omega(P\omega, A^n\omega)}{\Omega(Q\mu_{n+1}, A^n\omega) + \Omega(P\omega, B^n\mu_{n+1}) + \Omega(B^n\mu_{n+1}, A^n\omega)}\right\}\\ &+\delta_5 \max\{\Omega(A^n\omega, P\omega), \Omega(B^n\mu_{n+1}, Q\omega), \Omega(A^n\omega, B^n\mu_{n+1}), \Omega(P\omega, Q\omega)\}\end{aligned}$$

Let $n \to \infty$ and using (1) and (2), we have

$$\begin{aligned}\Longrightarrow \Omega(\omega,\omega) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(A^n\omega, \omega) + \Omega(\omega,\omega)\Omega(A^n\omega, P\omega)}{1+\Omega(\omega,\omega)}\right\}\\ &+\delta_2 \max\{\Omega(A^n\omega, \omega), \Omega(A^n\omega, \omega), \Omega(\omega,\omega)\}\\ &+\delta_3\{\Omega(\omega,\omega) + \Omega(\omega, A^n\omega) + \Omega(\omega,\omega)\}\\ &+\delta_4\left\{\frac{\Omega(\omega,\omega)\Omega(\omega, A^n\omega)}{\Omega(\omega, A^n\omega) + \Omega(\omega,\omega) + \Omega(\omega, A^n\omega)}\right\}\\ &+\delta_5 \max\{\Omega(A^n\omega, \omega), \Omega(\omega,\omega), \Omega(A^n\omega, \omega), \Omega(\omega,\omega)\}\end{aligned}$$

$\Longrightarrow (\delta_1 + \delta_2 + \delta_3 + \delta_5)||\Omega(A^n\omega,\omega)|| \geq 0$
$\Longrightarrow ||\Omega(A^n\omega,\omega)|| \geq 0$. Therefore $A^n\omega = \omega$
Since $P(\Delta) \subset B^n(\Delta)$, there exist $\omega* \in \Delta$ such that $P\omega = B^n\omega*$. Now to prove $B^n\omega* = Q\omega*$ putting $\mu = \omega$ and $\nu = \omega*$ we get

$$\begin{aligned}\Omega(P\omega, Q\omega*) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(A^n\omega, P\omega) + \Omega(B^n\omega*, Q\omega*)\Omega(A^n\omega, P\omega)}{1+\Omega(P\omega, Q\omega*)}\right\}\\ &+\delta_2 \max\{\Omega(A^n\omega, B^n\omega*), \Omega(A^n\omega, P\omega), \Omega(B^n\omega*, P\omega)\}\\ &+\delta_3\{\Omega(B^n\omega*, Q\omega*) + \Omega(Q\omega*, A^n\omega) + \Omega(P\omega, B^n\omega*)\}\\ &+\delta_4\left\{\frac{\Omega(Q\omega*, B^n\omega*)\Omega(P\omega, A^n\omega)}{\Omega(Q\omega*, A^n\omega) + \Omega(P\omega, B^n\omega*) + \Omega(B^n\omega*, A^n\omega)}\right\}\\ &+\delta_5 \max\{\Omega(A^n\omega, P\omega), \Omega(B^n\omega*, Q\omega), \Omega(A^n\omega, B^n\omega*), \Omega(P\omega, Q\omega)\}\end{aligned}$$

Here given that $P\omega = \omega = B^n\omega$. Then we get

$$\begin{aligned}\Omega(\omega, Q\omega*) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(\omega,\omega) + \Omega(\omega, Q\omega*)\Omega(\omega,\omega)}{1+\Omega(\omega, Q\omega*)}\right\}\\ &+\delta_2 \max\{\Omega(\omega,\omega), \Omega(\omega,\omega), \Omega(\omega,\omega)\}\\ &+\delta_3\{\Omega(\omega, Q\omega*) + \Omega(Q\omega*, \omega) + \Omega(\omega,\omega)\}\\ &+\delta_4\left\{\frac{\Omega(Q\omega*, \omega)\Omega(\omega,\omega)}{\Omega(Q\omega*, \omega) + \Omega(\omega,\omega) + \Omega(\omega,\omega)}\right\}\\ &+\delta_5 \max\{\Omega(\omega,\omega), \Omega(\omega, Q\omega), \Omega(\omega,\omega), \Omega(\omega, Q\omega)\}\end{aligned}$$

$\Longrightarrow ||\Omega(\omega, Q\omega*)|| \leq 2\delta_3||\Omega(\omega, Q\omega*)|| + \delta_5||\Omega(\omega, Q\omega*)||$
$\Longrightarrow ||\Omega(\omega, Q\omega*)|| \leq (2\delta_3 + \delta_5)||\Omega(\omega, Q\omega*)||$
$\Longrightarrow Q\omega* = \omega$
Since $B^n\omega*$ and Q are weakly compatible then we have $B^n\omega = B^nQ\omega* = Q\omega$. Therefore ω is a

coincidence point of B^n and P. Now to prove $Q\omega = \omega$, putting $\mu = \omega$ and $\nu = \omega$ we get

$$\begin{aligned}\Omega(P\omega, Q\omega) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(A^n\omega, P\omega) + \Omega(B^n\omega, Q\omega)\Omega(A^n\omega, P\omega)}{1 + \Omega(P\omega, Q\omega)}\right\} \\ &+\delta_2 \max\{\Omega(A^n\omega, B^n\omega), \Omega(A^n\omega, P\omega), \Omega(B^n\omega, P\omega)\} \\ &+\delta_3\{\Omega(B^n\omega, Q\omega) + \Omega(Q\omega, A^n\omega) + \Omega(P\omega, B^n\omega)\} \\ &+\delta_4\left\{\frac{\Omega(Q\omega, B^n\omega)\Omega(P\omega, A^n\omega)}{\Omega(Q\omega, A^n\omega) + \Omega(P\omega, B^n\omega) + \Omega(B^n\omega, A^n\omega)}\right\} \\ &+\delta_5 \max\{\Omega(A^n\omega, P\omega), \Omega(B^n\omega, Q\omega), \Omega(A^n\omega, B^n\omega), \Omega(P\omega, Q\omega)\}\end{aligned}$$

$$\begin{aligned}\implies \Omega(\omega, Q\omega) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(\omega, \omega) + \Omega(\omega, Q\omega)\Omega(\omega, \omega)}{1 + \Omega(\omega, Q\omega)}\right\} \\ &+\delta_2 \max\{\Omega(\omega, \omega), \Omega(\omega, \omega), \Omega(\omega, \omega)\} \\ &+\delta_3\{\Omega(\omega, Q\omega) + \Omega(Q\omega, \omega) + \Omega(\omega, \omega)\} \\ &+\delta_4\left\{\frac{\Omega(Q\omega, \omega)\Omega(\omega, \omega)}{\Omega(Q\omega, \omega) + \Omega(\omega, \omega) + \Omega(\omega, \omega)}\right\} \\ &+\delta_5 \max\{\Omega(\omega, \omega), \Omega(\omega, Q\omega), \Omega(\omega, \omega), \Omega(\omega, Q\omega)\}\end{aligned}$$

$\implies ||\Omega(\omega, Q\omega)|| \leq 2\delta_3||\Omega(\omega, Q\omega)|| + \delta_5||\Omega(\omega, Q\omega)||$
$\implies ||\Omega(\omega, Q\omega)|| \leq (2\delta_3 + \delta_5)||\Omega(\omega, Q\omega)||$
$\implies Q\omega = \omega$

Uniqueness of fixed point

Let ω_0 be another fixed point of P, Q, A^n, B^n then we have

$$\begin{aligned}\Omega(P\omega, Q\omega_0) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(A^n\omega, P\omega) + \Omega(B^n\omega_0, Q\omega_0)\Omega(A^n\omega, P\omega)}{1 + \Omega(P\omega, Q\omega_0)}\right\} \\ &+\delta_2 \max\{\Omega(A^n\omega, B^n\omega_0), \Omega(A^n\omega, P\omega), \Omega(B^n\omega_0, P\omega)\} \\ &+\delta_3\{\Omega(B^n\omega_0, Q\omega_0) + \Omega(Q\omega_0, A^n\omega) + \Omega(P\omega, B^n\omega_0)\} \\ &+\delta_4\left\{\frac{\Omega(Q\omega_0, B^n\omega_0)\Omega(P\omega, A^n\omega)}{\Omega(Q\omega_0, A^n\omega) + \Omega(P\omega, B^n\omega_0) + \Omega(B^n\omega_0, A^n\omega)}\right\} \\ &+\delta_5 \max\{\Omega(A^n\omega, P\omega), \Omega(B^n\omega_0, Q\omega), \Omega(A^n\omega, B^n\omega_0), \Omega(P\omega, Q\omega)\}\end{aligned}$$

Taking limit $n \to \infty$

$$\begin{aligned}\Omega(\omega, \omega_0) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(\omega, \omega) + \Omega(\omega_0, \omega_0)\Omega(\omega, \omega)}{1 + \Omega(\omega, \omega_0)}\right\} \\ &+\delta_2 \max\{\Omega(\omega, \omega_0), \Omega(\omega, \omega), \Omega(\omega_0, \omega)\} \\ &+\delta_3\{\Omega(\omega_0, \omega_0) + \Omega(\omega_0, \omega) + \Omega(\omega, \omega_0)\} \\ &+\delta_4\left\{\frac{\Omega(\omega_0, \omega_0)\Omega(\omega, \omega)}{\Omega(\omega_0, \omega) + \Omega(\omega, \omega_0) + \Omega(\omega_0, \omega)}\right\} \\ &+\delta_5 \max\{\Omega(\omega, \omega), \Omega(\omega_0, \omega), \Omega(\omega, \omega_0), \Omega(\omega, \omega)\}\end{aligned}$$

$\implies ||\Omega(\omega, \omega_0)|| \leq \delta_2||\Omega(\omega, \omega_0)|| + 2\delta_3||\Omega(\omega, \omega_0)|| + \delta_5||\Omega(\omega, \omega_0)||$
$\implies ||\Omega(\omega, \omega_0)|| \leq (\delta_2 + 2\delta_3 + \delta_5)||\Omega(\omega, \omega_0)||$
$\implies ||\Omega(\omega, \omega_0)|| \leq (\delta_2 + 2\delta_3 + \delta_5)||\Omega(\omega, \omega_0)||$

Which is contradiction since $(\delta_2 + 2\delta_3 + \delta_5) < 1$ therefore $\omega = \omega_0$ is a unique common fixed point of P, Q, A^n and B^n ■

Corollary 1. Let (Δ, Ω) be a tricomplex valued metric space and P, Q, A, B are self mappings in Δ, such that for all $\mu, \nu \in \Delta$

1. $P(\Delta) \subset B(\Delta)$ and $Q(\Delta) \subset A(\Delta)$;
2. (A, P) and (B, Q) are commuting Pairs;
3. (A, P) and (B, Q) are commuting;
4. If $\Omega(Q\nu, A\mu) + \Omega(P\mu, B\nu) + \Omega(B\nu, A\mu) \neq 0$ then the following holds

$$\begin{aligned}\Omega(P\mu, Q\nu) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(A\mu, P\mu) + \Omega(B\nu, Q\nu)\Omega(A\mu, P\mu)}{1 + \Omega(P\mu, Q\nu)}\right\} \\ &+\delta_2 \max\{\Omega(A\mu, B\nu), \Omega(A\mu, P\mu), \Omega(B\nu, P\mu)\} \\ &+\delta_3\{\Omega(B\nu, Q\nu) + \Omega(Q\nu, A\mu) + \Omega(P\mu, B\nu)\} \\ &+\delta_4\left\{\frac{\Omega(Q\nu, B\nu)\Omega(P\mu, A\mu)}{\Omega(Q\nu, A\mu) + \Omega(P\mu, B\nu) + \Omega(B\nu, A\mu)}\right\} \\ &+\delta_5 \max\{\Omega(A\mu, P\mu), \Omega(B\nu, Q\mu), \Omega(A\mu, B\nu), \Omega(P\mu, Q\mu)\}\end{aligned}$$

Where $\delta_1 + \delta_2 + 2\delta_3 + \delta_4 + \delta_5 < 1$. For all $\mu, \nu \in \Delta$. Then P, Q, A and B have a unique common fixed point.

Corollary 2. Let (Δ, Ω) be a tricomplex valued metric space and P, Q, A, B are self mappings in Δ, such that for all $\mu, \nu \in \Delta$

1. $P(\Delta) \subset B(\Delta)$ and $Q(\Delta) \subset A(\Delta)$;
2. (A, P) are compatible, A or P is continuous and (B, Q) are weakly compatible;
3. (B, Q) are compatible, B or Q is continuous and (A, P) are weakly compatible;
4. For all $\mu, \nu \in \Delta$ and $\mu \neq \nu$ we have

$$\begin{aligned}\Omega(P\mu, Q\nu) \leq_{i_3} \delta_1 &\left\{\frac{\Omega(A\mu, P\mu) + \Omega(B\nu, Q\nu)\Omega(A\mu, P\mu)}{1 + \Omega(P\mu, Q\nu)}\right\} \\ &+\delta_2 \max\{\Omega(A\mu, B\nu), \Omega(A\mu, P\mu), \Omega(B\nu, P\mu)\} \\ &+\delta_3\{\Omega(B\nu, Q\nu) + \Omega(Q\nu, A\mu) + \Omega(P\mu, B\nu)\} \\ &+\delta_4\left\{\frac{\Omega(Q\nu, B\nu)\Omega(P\mu, A\mu)}{\Omega(Q\nu, A\mu) + \Omega(P\mu, B\nu) + \Omega(B\nu, A\mu)}\right\} \\ &+\delta_5 \max\{\Omega(A\mu, P\mu), \Omega(B\nu, Q\mu), \Omega(A\mu, B\nu), \Omega(P\mu, Q\mu)\}\end{aligned}$$

Where $\delta_1 + \delta_2 + 2\delta_3 + \delta_4 + \delta_5 < 1$. For all $\mu, \nu \in \Delta$. Then P, Q, A and B have a unique common fixed point.

4. Conclusion

In this paper, we proved generalizations of fixed-point theorems in tricomplex-valued metric space for the sequence of mappings with rational type contraction and using commuting, compatible and weakly compatible pairs. our contributions have enhanced the field of tricomplex valued metric spaces.

References

[1] Banach, S. (1922). Sur les operations dans les ensembles abstraits et leur applications aux equations integrales. *Fund. Math.*, 3:133–181.
[2] Sedghi, S. Shobe, N. and Zhou, H. (2007). A common fixed point theorem in metric spaces. *Fixed point theory and Applications*, 2007:1-13.

[3] Long-Guang, H., Xian, Z. (2007). Cone metric spaces and fixed point theorems of contractive mappings. *J. Math. Anal. Appl*, 332(2):1468-1476.
[4] Amini-Harandi, A. (2012). Metric-like spaces, partial metric spaces and fixed points. *Fixed point theory and Applications*, 2012:1-10.
[5] Cho, Y. J. (2017). Survey on metric fixed point theory and applications. *Advances in Real and Complex Analysis with Applications*, 183–241.
[6] Dyshin, O., Maharramov, F. (2022). Multivariate Charts of Statistical Control of the Dynamic Process of Oil Field Development Management. *Reliability: Theory Applications*, 17(SI 3 (66)):158-163.
[7] Karlsson, A. (2024). A metric fixed point theorem and some of its applications. *Geometric and Functional Analysis*, 34(2):486–511.
[8] Chourasiya, S., and Shrivastava, K. (2024). Common Fixed Point Theorem in Controlled Metric Spaces. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(5):253–257.
[9] Azam, A., Fisher, B., and Khan, M. (2011). Common fixed point theorems in complex valued metric spaces. *Numerical Functional Analysis and Optimization*, 32(3):243-253.
[10] Sintunavarat, W., and Kumam, P. (2012). Generalized common fixed point theorems in complex valued metric spaces and applications. *Journal of inequalities and Applications*, 2012:1-12.
[11] Rao, K. P. R., Swamy, P. R., and Prasad, J. R. (2013). A common fixed point theorem in complex valued b-metric spaces. *Bulletin of Mathematics and Statistics research*, 1(1):1-8.
[12] Verma, R. K., and Pathak, H. K. (2013). Common fixed point theorems using property (EA) in complex-valued metric spaces. *Thai J. Math*, 11(2):347-355.
[13] Singh, N., Singh, D., Badal, A., and Joshi, V. (2016). Fixed point theorems in complex valued metric spaces. *Journal of the Egyptian Mathematical Society*, 24(3):402-409.
[14] Price, G. B. (2018). An introduction to multicomplex spaces and functions. *CRC Press*.
[15] Choi, J., Datta, S. K., Biswas, T., and Islam, M. N. (2017). Some fixed point theorems in connection with two weakly compatible mappings in bicomplex valued metric spaces. *Honam Mathematical Journal*, 39(1):115-126.
[16] Beg, I., Datta, S. K., Pal, D. (2021). Fixed point in bicomplex valued metric spaces. *International Journal of Nonlinear Analysis and Applications*, 12(2):717-727.
[17] Mani, G., Gnanaprakasam, A. J., Haq, A. U., Baloch, I. A., Jarad, F. (2022). Common fixed point theorems on tricomplex valued metric space. *Mathematical Problems in Engineering*, 2022(1):4617291.
[18] Ramaswamy, R., Mani, G., Gnanaprakasam, A. J., Abdelnaby, O. A. A., Radenovi‡, S. (2022). An application to fixed-point results in tricomplex-valued metric spaces using control functions. *Mathematics*, 10(18):3344.
[19] Tassaddiq, A., Ahmad, J., Al-Mazrooei, A. E., Lateef, D., Lakhani, F. (2023).On common fixed point results in bicomplex valued metric spaces with application. *AIMS Math*, 8: 5522-5539.
[20] Murali, A., Muthunagai, K. (2024). Some theorems on fixed points in bi-complex valued metric spaces with an application to integral equations. *Journal of the Nigerian Society of Physical Sciences*, 1750-1750.
[21] Siva, G. (2025). Fixed points in tricomplex valued S-metric spaces with applications. *International Journal of Nonlinear Analysis and Applications*, 16(3):115-123.
[22] Sessa, S. (1982). On a weak commutativity condition of mappings in fixed point considerations. *Publ. Inst. Math*, 32(46):149–153.
[23] Jungck, G. (1986). Compatible mappings and common fixed points. *International journal of mathematics and mathematical sciences*, 9(4):771–779.

DISCOUNT PRICING POLICY WITH STOCHASTIC DEMAND AND PRESERVATION TECHNOLOGY FOR DETERIORATING ITEMS

Ramkrishna Tiwari[1], Ram Kumar Tiwari[2], Lalji Kumar[3],
Pooran Lal Prajapati[4] , Uttam Kumar Khedlekar[5]

•

[1,3,5]Department of Mathematics and Statistics,
Dr. Harisingh Gour Vishwavidyalaya, Sagar, Madhya Pradesh, India - 470003
[2]Government College Makronia Bujurg, Sagar, Madhya Pradesh, India - 470004
[4]Sahodra Rai Government Polytechnic College, Sagar, Madhya Pradesh, India - 470001
[1]*r.krishnatiwari84@gmail.com*, [2]*shriram.adina@gmail.com*, [3]*laljikumar07@gmail.com*,
[4]*plprajapati16@gmail.com*, [5]*uvkkcm@yahoo.co.in*

Abstract

Managing inventory for deteriorating items presents significant challenges, especially in environments where demand is stochastic and price-sensitive. Retailers face the dual pressure of minimizing losses from product deterioration while ensuring sufficient stock availability to meet customer demand. Additionally, backlogged shortages, if not managed strategically, can erode customer satisfaction and profitability. These complexities are further compounded by the need to strike a balance between promotional efforts, preservation investments, and pricing strategies to remain competitive in dynamic markets. To address these challenges, this study develops a comprehensive mathematical model aimed at optimizing replenishment strategies for deteriorating items. The model integrates preservation technologies and strategic price discounts to enhance product demand while accounting for the intricate relationship between market potential, stock levels, promotional efforts, and selling prices. The objective is to determine optimal pricing policies, order quantities, promotional expenditures, preservation costs, and schedules for shortages and replenishment, ultimately maximizing total profit per unit time. The findings reveal that total profit per unit time is concave concerning price, shortage duration, and inventory holding period. Numerical analyses demonstrate the efficacy of preservation technologies in mitigating deterioration rates, enabling retailers to adopt competitive pricing and stimulate sales. Additionally, the demand pattern index is shown to significantly influence inventory policies, while promotional campaigns play a critical role in augmenting product sales. Quantity discounts, tied to demand and order volumes, are identified as effective tools for reducing procurement costs and enhancing retailer profitability. These insights highlight the strategic importance of investing in preservation technologies, adopting dynamic pricing strategies, and implementing targeted promotional campaigns. Together, these approaches provide a robust framework for optimizing inventory management, increasing sales, and maximizing profitability in highly competitive markets.

Keywords: Inventory, Deterioration, Stochastic Optimal Control, Preservation Technology Investment, Quantity Discounts

1. Introduction

The effective management of inventory systems has long been acknowledged as a critical concern in business operations, particularly in industries where products are prone to deterioration over

time. This challenge is especially prominent in the management of perishable goods such as pharmaceuticals, volatile liquids, and fresh produce. If left unaddressed, product deterioration can result in significant economic losses and a decline in customer service levels. Consequently, the formulation of sophisticated inventory models that account for deterioration dynamics has attracted substantial attention from researchers in operations research, computer science, and management studies.

In response to the complexities of real-world inventory systems, researchers have continually refined theoretical frameworks to reflect practical scenarios and improve decision-making processes. A key advancement in this domain has been the integration of preservation technologies, which aim to reduce the rate of product deterioration, thereby extending shelf life and minimizing economic losses. Simultaneously, the application of discount policies has emerged as an effective strategy for retailers to accelerate product turnover, enhance revenue streams, and maintain market competitiveness. Within this multidimensional context, the development of optimal inventory policies has become a focal point, requiring a careful combination of preservation technologies, pricing strategies, and promotional measures to achieve maximum profitability while ensuring satisfactory service levels.

Early contributions, such as the Economic Order Quantity (EOQ) model introduced by Harris, established the foundational principles of inventory management by balancing inventory holding and setup costs to determine optimal order quantities. While the EOQ model remains a cornerstone in inventory theory, contemporary research has advanced to address the complexities associated with deteriorating inventory systems. Extensions of the EOQ framework have incorporated diverse deterioration patterns, such as constant, Weibull, and time-dependent linear rates. Notable advancements in this area have been demonstrated in the works of several researchers, including [1], [2], and [3]. These studies have provided critical insights into the interaction between deterioration dynamics and inventory management strategies, offering a comprehensive basis for optimizing decision-making processes.

The integration of stochastic elements into inventory modeling has profoundly enriched theoretical frameworks, providing deeper insights into demand variability and its implications for optimal decision-making. Significant contributions to this domain include studies on inventory models characterized by constant deterioration rates, constant demand, and instantaneous replenishment, as examined by researchers such as [4]. These investigations have offered valuable perspectives on the dynamics of inventory systems operating under stochastic demand conditions. Expanding upon this foundation, subsequent research by [5] and [6] has focused on inventory models with fixed deterioration rates and exponentially time-fluctuating demand, emphasizing strategies to effectively manage uncertainties in both supply and demand dynamics.

In parallel, advancements in fuzzy inventory modeling have introduced novel approaches to addressing the ambiguity and uncertainty inherent in real-world inventory systems. The application of fuzzy concepts has led to the development of models that incorporate price-dependent demand and time-varying holding costs, as demonstrated in the works of [7] and [8]. These contributions present robust methodologies for making informed decisions in uncertain environments. Furthermore, models accounting for linearly time-dependent deterioration rates and partially backlogged shortages have been introduced by [9], highlighting the necessity of integrating sophisticated deterioration dynamics into inventory management frameworks.

Expanding beyond conventional inventory management approaches, recent research has increasingly emphasized multi-objective inventory optimization, particularly in scenarios involving incremental discounts and fuzzy decision-making frameworks. These methodologies have been instrumental in enhancing supplier selection and order allocation processes. Notable investigations, such as those conducted by [10], have explored strategies for minimizing purchasing risks while simultaneously maximizing total profit. This line of research underscores the significance of fostering sustainable buyer-supplier relationships, especially in competitive market environments.

Building on this extensive body of literature, the present study seeks to contribute to the field of stochastic optimal control in inventory modeling by introducing a comprehensive framework that integrates preservation technologies, pricing strategies, and promotional initiatives. By

employing stochastic optimization techniques, the proposed model aims to identify optimal inventory policies that effectively balance economic objectives with service level requirements. This approach is designed to enhance both profitability and competitiveness in dynamic market conditions. Additionally, through extensive numerical analysis and sensitivity testing, the study systematically evaluates the impact of various parameters on decision variables and overall profitability, thereby offering actionable insights for both industry practitioners and academic researchers.

In the modern business landscape, promotional strategies have emerged as a pivotal mechanism for stimulating consumer demand, differentiating brands, and increasing product visibility. Traditional promotional schemes—including price discounts, gifts, and coupons—have evolved in parallel with technological advancements. The rise of digital platforms, such as social networking websites and blogs, has further transformed the promotional landscape, enabling highly targeted marketing campaigns. Consequently, the integration of promotional strategies into inventory management research has gained substantial relevance, underscoring their critical role in shaping demand dynamics and optimizing operational efficiency.

A well-structured discount policy plays a pivotal role in fostering mutually beneficial relationships between consumers and retailers, generating economic advantages for both parties within the business ecosystem. The efficacy of quantity discount schedules, categorized into all-unit and incremental discounts, has been extensively documented in academic literature. Foundational contributions, such as the seminal EOQ model, have demonstrated the minimization of discounted cash flows through the determination of optimal order quantities under varying reorder point scenarios, as explored by researchers including [11].

Moreover, the intricate relationship between pricing strategies and inventory management is exemplified through dynamic lot-sizing problems formulated under discounting schemes. Research in this domain, such as the study presented in [12], leverages forecast data to guide speculative stocking decisions, reinforcing the strategic implications of discount-driven inventory control. Further investigations into the quantity discount-pricing problem, including the work of [13], highlight the complexities of pricing decisions in price-dependent demand environments. Expanding upon this foundation, subsequent research by [14] incorporates supplier credit terms, providing valuable insights into order decision-making under diverse financial constraints.

The strategic interplay between temporary price discounts and inventory policies has also garnered significant scholarly attention. Investigations by researchers such as [15], [16], and [17] underscore the deliberate use of pricing incentives to maximize savings and profitability within inventory systems characterized by time-dependent demand and transient price fluctuations. These contributions emphasize the necessity of informed decision-making in dynamic pricing environments, illustrating the critical role of adaptive pricing strategies in enhancing operational efficiency and financial performance.

The emerging domain of joint pricing and preservation decision-making underscores the synergy between pricing strategies, inventory control, and preservation technologies. Notable advancements, including models integrating stochastic demand, pricing, and preservation decisions, illustrate the intricate trade-offs involved in optimizing total profit under uncertain market conditions, as demonstrated in the work of [18]. Acknowledging the stochastic nature of demand and pricing dynamics, contemporary research has increasingly focused on optimizing pricing and production policies within stochastic inventory systems. Game-theoretic approaches to retail and wholesale pricing optimization, as explored by [20], highlight the strategic considerations underpinning competitive market pricing decisions. Additionally, pioneering insights into pricing and production policy optimization under fuzzy environments have been provided by [21], reinforcing the importance of dynamic pricing strategies in mitigating the adverse effects of uncertainty on inventory performance.

A sustainable production framework has been developed to address machine inefficiencies, environmental impacts, and demand-price uncertainties, offering versatile strategies applicable across greenhouse gas-emitting industries. This framework effectively balances economic and environmental objectives, contributing to the advancement of sustainable industrial practices, as

explored in studies such as [22]. Additionally, stochastic inventory models leveraging Particle Swarm Optimization (PSO) have demonstrated superior performance over deterministic models by effectively capturing market uncertainties. This approach enhances adaptive decision-making and risk mitigation within supply chains, reinforcing its applicability in complex and volatile market conditions, as highlighted in the research of [23].

Innovative approaches to inventory modeling have gained prominence, particularly in contexts characterized by seasonal sales and dynamic pricing. Research on stochastic demand and dynamic pricing strategies for seasonal sales has revealed that real-time pricing adjustments outperform traditional methods by improving inventory management and optimizing profit margins during demand fluctuations. This perspective is further substantiated by studies such as [24]. Similarly, an inventory model tailored for declining markets has integrated promotional efforts with pricing strategies to manage non-instantaneous deteriorating items. By dynamically adjusting pricing and replenishment schedules in response to time- and price-sensitive demand, this model has demonstrated its effectiveness in maximizing profitability, as examined in [25].

For deteriorating products, an advanced inventory management framework has been introduced, categorizing production life-cycle stages and incorporating rework processes to enhance product quality. This framework employs a hybrid approach combining Grey Wolf and Ant Colony Optimization algorithms, achieving significant cost reductions and improved profitability, as evidenced in studies like [26]. Furthermore, a conversion optimization model has been developed to determine the optimal timing, quantity, cost, and selling price for homogeneous and heterogeneous conversions. Particle Swarm Optimization has proven more effective than classical methods in maximizing profit by optimizing multiple variables, as demonstrated in research such as [27].

Comparative analyses of forecasting methods have also gained scholarly attention. A study evaluating Holt-Winters Exponential Smoothing (HWES) and ARIMA models concluded that ARIMA offers superior performance in minimizing lost sales and enhancing demand forecasting accuracy under varying economic conditions. These findings underscore the importance of advanced forecasting techniques in improving supply chain resilience and operational efficiency, as discussed in [28]. Additionally, an inventory management model integrating pricing fluctuations, advertising strategies, and preservation costs has demonstrated its potential to enhance supply chain efficiency while aligning inventory practices with sustainability goals in competitive markets, as explored in [29].

Finally, delayed deteriorating EOQ models have been developed to optimize inventory costs by accounting for time-varying demand and partial backlogging, thereby addressing critical challenges in managing inventory systems with deteriorating products ([30]).

Building on these scholarly advancements, this paper aims to address critical questions related to optimal pricing, discounting, and shortage management within stochastic inventory control systems. By examining the complexities of replenishment scheduling and order quantity determination under conditions of uncertain demand and dynamic promotional strategies, this study seeks to advance the discourse on stochastic inventory management. The findings are intended to provide actionable insights, thereby supporting both practitioners in decision-making and scholars in furthering theoretical developments in the field.

2. Notations & Assumption

Notations and assumption bearing usual tradition, utilized in the subsequent discussions, are laid down as follows:

Notations

Parameters:

$D(s)$ Non-linear discounted price and dependent demand function for products,

$I(t)$ On-hand inventory of product at time t,

s The selling price per unit,

θ The deterioration rate of defective products,

ξ The preservation technology cost per unit,

C_p Purchasing cost per unit,

C_h Holding cost per item per unit time,

C_b Backorder cost per unit per time unit,

C_o Ordering cost per order,

C_l The lost sale cost per unit,

ε The random variable of the demand function with $E(\varepsilon) = \mu$,

t_s The replenishment point, the duration of the shortage time period,

T The inventory period,

ρ The promotional effort, $\eta \geq 1$,

Q The ordering quantity per cycle,

$I_1(t)$ The level of negative inventory at time t where $0 \leq t \leq t_{(s)}$,

$I_2(t)$ The level of positive inventory at time t where $t_s \leq t \leq t_{(s)} + T$,

$\widehat{\Psi}$ The total profit.

3. Assumptions

1. The replenishment rate is infinite, but its order size is finite.
2. The basic demand function is $D(s) + \varepsilon$, where $D(s)$ is a decreasing and deterministic function of the selling price s, and ε is a non-negative and continuous random variable with $E(\varepsilon) = \mu$.
3. The promotional effort cost PC is considered as $PC = K(\rho - 1)^2 \left[\int_0^1 (D(s) + \varepsilon)dt\right]^{\alpha}$ where $K > 0$ and α is a constant.
4. Shortages occur at the beginning of the cycle in this model, and a fraction of the demand varying with waiting time is backlogged. Assume that the customer's impatient function is $B(x) = e^{-\delta x}$.
5. The demand function is influenced by promotional effort ρ and can be expressed by mark-up over the promotional effort.
6. The item deteriorates at a time-varying rate of deterioration θ, where $0 < \theta < 1$. Besides, there is no repair or replacement of deteriorated units during the replenishment cycle.
7. The proportion of reduced deterioration rate, $m(\xi) = 1 - e^{-\eta\xi}$, is a continuous, concave, and increasing function of capital investment, where η is a parameter and ξ is the preservation technology cost.

4. Mathematical Formulation

The inventory model commences with shortages and concludes without shortages, as illustrated graphically in Figure 1. Shortages are partially backlogged, with backlogged demand fulfilled at the replenishment point t_s, while the remaining inventory suffices to meet demand until time T. The depletion of inventory arises from the combined impact of demand and deterioration during the interval $[t_s, T]$, as depicted in Figure 1. Notably, preservation technology is employed to regulate the deterioration rate. The demand function is contingent on a non-linear discounted price, defined by:

$$D(s) = a(ds)^{-b_1}, \quad \text{where} \quad a,\, d,\, b_1 > 0.$$

With this understanding, the variation in inventory level at any instant $t \in [0, t_m]$ is governed by the following differential equation: The entire inventory system can be characterized by

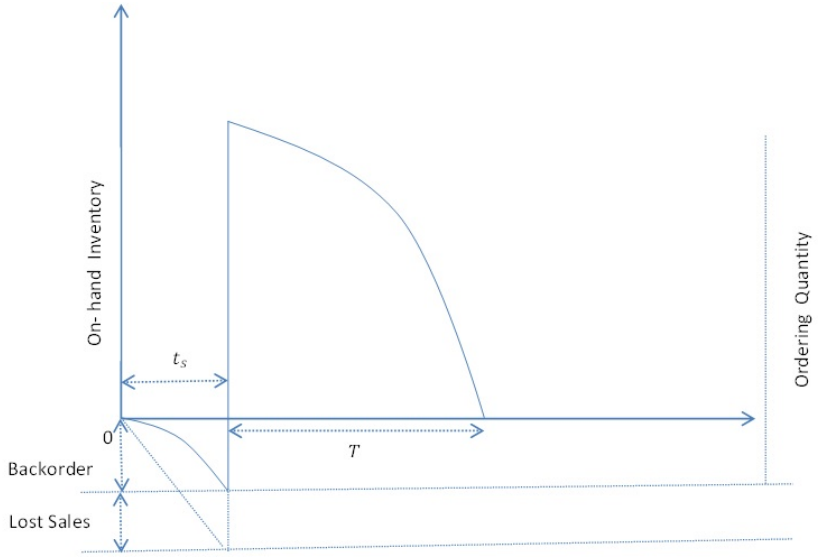


Figure 1: *Graphical Representation of the Inventory System*

leveraging the differential equation:

$$\frac{dI_1(t)}{dt} = -\rho(D(s) + \varepsilon)B(t - t_s), \quad 0 \le t \le t_s, \tag{1}$$

where $I_1(t)$ denotes the level of negative inventory, $B(t - t_s)$ represents the customer's impatient function capturing backlogged demand, and ε signifies the stochastic element of demand. This equation encapsulates the dynamic evolution of inventory levels during the shortage period.
with $I_1(t) = 0$, at $t = 0$

$$B(x) = e^{-\delta x}$$

Utilizing the conditions, the solutions of above differential equation

$$I_1(t) = -\frac{1}{\delta}\left[e^{t_s - t}(e^{\delta t} + 1)(D(s) + \varepsilon)\rho\right] \tag{2}$$

The change in inventory level at any instant of time $t \in [t_m, T]$, is represented by following differential equation

$$\frac{dI_2(t)}{dt} = -\theta(1 - m(\xi))I_2(t) - \rho(D(s) + \varepsilon), \quad 0 \le t \le t_m \tag{3}$$

with $I_2(T) = 0$ the solutions of above differential equation

$$I_2(t) = -\frac{1}{\theta\big(m(\xi)-1\big)}\left[e^{t_2\theta\left(m(\xi)-1\right)}\left\{e^{t_2\theta\left(m(\xi)-1\right)} + e^{t\theta\left(m(\xi)-1\right)}\right\}\rho(D(s)+\varepsilon)\right] \tag{4}$$

The lot sale quantity at time t, is

$$I_l(t) = (1 - e^{-(t-t_s)\delta})(D(s)+\varepsilon)\rho \tag{5}$$

The replenishment size with back logged is

$$Q = \left\{I_2(0 - I_1(t_s)\right\} \tag{6}$$

$$Q = \left[\frac{1}{\delta}\left[(e^{\delta t_s}+1) - \frac{1}{\theta(m(\xi)-1)}\left\{e^{T\theta\left(m(\xi)-1\right)}\left(e^{T\theta\left(m(\xi)-1\right)}+1\right)\right\}\right](D(s)+\varepsilon)\rho \tag{7}$$

The cost of lost sale during time interval $[0, t_s]$

$$\widehat{TLC} = E\left[C_l\int_0^{t_s} I_l(t)dt\right] \tag{8}$$

$$\widehat{TLC} = C_l\left(t_s + \frac{1-e^{t_s\delta}}{\delta}\right)(D(s)+\mu)\rho \tag{9}$$

The total cost of back logged for stock out during the time interval $[0, t_s]$

$$\widehat{TSC} = E\left[C_b\int_0^{t_s} -I_1(t)dt\right] \tag{10}$$

$$\widehat{TSC} = -\frac{C_h}{(\theta)^2\big(m(\xi)-1\big)^2}\left[e^{-t_2\theta\left(m(\xi)-1\right)}\left\{1 + e^{t_2\theta\left(m(\xi)-1\right)}\left(t_2\theta(m(\xi)-1)-1\right)\right\}\rho(D(s)+\varepsilon)\right] \tag{11}$$

The total holding cost during the time interval $[0, T]$

$$\widehat{THC} = E\left[C_h\int_0^{T} I_2(t)dt\right] \tag{12}$$

$$\widehat{THC} = \frac{C_h}{\delta^2}\left\{1 + e^{t_s\delta}(t_s\delta - 1)\right\}(D(s)+\mu)\rho \tag{13}$$

The total purchasing cost is

$$\widehat{TPC} = E(C_p.Q) = (D(s)+\mu)C_p\rho\left\{\frac{1-e^{t_s\delta}}{\delta} - \frac{1-e^{t_2\theta(m(\xi)-1)}}{\theta(m(\xi)-1)}\right\} \tag{14}$$

The total preservation technology cost is

$$\widehat{TPTC} = \xi(t_s + T) \tag{15}$$

The total promotional cost is

$$\widehat{T_pC} = E\left[K(\rho-1)^2\left\{\int_0^{t_s+T}(D(s)+\varepsilon)dt\right\}^{\alpha}\right] \tag{16}$$

$$\widehat{TPC} = K\left((t_s+T)(D(s)+\mu)\right)^{\alpha}(\rho-1)^2 \tag{17}$$

The total sales revenue is

$$\widehat{TSR} = E\left[s\left\{\rho(D(s)+\varepsilon)\int_0^T dt + \int_0^{t_s} B(t_s-t)dt\right\}\right] \tag{18}$$

$$\widehat{TSR} = s(D(s)+\mu)\rho\left(T + \frac{e^{t_s\delta}-1}{\delta}\right) \tag{19}$$

Now, the total profit

$$\widehat{\Psi} = \widehat{TSR} - \widehat{TC} \tag{20}$$

$$\begin{aligned}\widehat{\Psi}(t_s,T,s) =& s(D(s)+\mu)\rho\left(T+\frac{e^{t_s\delta}-1}{\delta}\right) - \left[A + C_l\left(t_s + \frac{1-e^{t_s\delta}}{\delta}\right)(D(s)+\mu)\rho\right.\\ &- \frac{C_h}{(\theta)^2\big(m(\xi)-1\big)^2}\left\{e^{-t_2\theta\left(m(\xi)-1\right)}\left\{1+e^{t_2\theta\left(m(\xi)-1\right)}\left(T(m(\xi)-1)-1\right)\right\}\rho(D(s)\right.\\ &\left.+\mu)\right\} + \frac{C_h}{\delta^2}\left(1+e^{t_s\delta}(t_s\delta-1)\right)(D(s)+\mu)\rho + (D(s)+\mu)C_p\rho\left\{\frac{1-e^{t_s\delta}}{\delta}\right.\\ &\left.\left.- \frac{1-e^{T\theta(m(\xi)-1)}}{\theta(m(\xi)-1)}\right\} + \xi(t_s+T) + K\left((t_s+T)(D(s)+\mu)\right)^{\alpha}(\rho-1)^2\right]\end{aligned} \tag{21}$$

In order to optimize the above non-linear optimization problem at first, the objective function is simplified and then some theorems are proposed. It is useful mentioning here that for a small value of x, the Taylor theorem (series) says that the exponential function has an approximation of $e^x \approx 1 + x + \frac{x^2}{2!}$. Using this result in Eq. (21), it is obtained the following:

$$\begin{aligned}\widehat{\Psi}(t_s,T,s) =& \Big[-4A - 4(t_s+T)\xi - 4K\big((t_s+T)(D(s)+\mu)\big)^{\alpha}(\rho-1)^2 + 2C_p t_s(2+t_s\delta)(D(s)+\mu)\rho\\ &+ 2C_l t_s^2\delta(D(s)+\mu)\rho - 2C_b t_s^2(1+t_s\delta)(D(s)+\mu)\rho + 2s\big(2(t_s+T)+t_s^2\delta\big)(D(s)+\mu)\rho\\ &- C_h T^2(D(p)+\mu)\rho\big(1-T\theta(1-m(\xi))\big)\Big\{2+T\theta(2+T\theta)\\ &+ T\theta m(\xi)\big(T\theta(m(\xi)-2)-2\big)\Big\} - C_p T(D(s)+\mu)\rho\big(T\theta(m(\xi)-1)+2\big)\big(2+T\theta(2\\ &+ T\theta) + T\theta m(\xi)\big(T\theta(m(\xi)-2)-2\big)\Big]\end{aligned} \tag{22}$$

Therefore, average total profit per unit time is

$$\widehat{\Psi_a}(t_s,T,s) = \frac{\widehat{\Psi}(t_s,T,s)}{t_s+T} \tag{23}$$

$$\begin{aligned}\widehat{\Psi_a}(t_s,T,s) =& \frac{1}{4(t_s+T)}\Big[-4A - 4(t_s+T)\xi - 4K\big((t_s+T)(D(s)+\mu)\big)^{\alpha}(\rho-1)^2 + 2C_p t_s(2+t_s\delta)(D(s)+\mu)\rho\\ &+ 2C_l t_s^2\delta(D(s)+\mu)\rho - 2C_b t_s^2(1+t_s\delta)(D(s)+\mu)\rho + 2s\big(2(t_s+T)+t_s^2\delta\big)(D(s)+\mu)\rho\\ &- C_h T^2(D(p)+\mu)\rho\big(1-T\theta(1-m(\xi))\big)\Big\{2+T\theta(2+T\theta)+T\theta m(\xi)\big(T\theta(m(\xi)-2)-2\big)\Big\}\\ &- C_p T(D(s)+\mu)\rho\big(T\theta(m(\xi)-1)+2\big)\big(2+T\theta(2+T\theta)+T\theta m(\xi)\big(T\theta(m(\xi)-2)-2\big)\Big]\end{aligned} \tag{24}$$

5. Solution Procedure

In this study, we aim to determine the optimal values of the inventory cycle length , shortage period , and selling price that maximize the average total profit per unit time . The objective function, derived using Taylor series approximations. Now, we seek to solve the following constrained optimization problem:

$$\max_{t_s,T,s} \quad \widehat{\Psi_a}(t_s,T,s) \tag{25}$$

To ensure the existence of a global maximum for the optimization problem, we examine the concavity of the profit function. A function is concave if its Hessian matrix is negative definite, i.e., all its leading principal minors have alternating signs. We state the following theorem:

Theorem 5.1. *The profit function $\widehat{\Psi_a}(t_s, T, s)$ is concave with respect to inventory period T shortage time t_s and selling price s if the corresponding Hessian matrix H of expected profit function is negative definite. where*

$$\text{H}= \begin{pmatrix} \frac{\partial^2 \widehat{\Psi_a}}{\partial T^2} & \frac{\partial^2 \widehat{\Psi}}{\partial s \partial T} & \frac{\partial^2 \widehat{\Psi}}{\partial T \partial t_s} \\ \frac{\partial^2 \widehat{\Psi}}{\partial s \partial T} & \frac{\partial^2 \widehat{\Psi}}{\partial s^2} & \frac{\partial^2 \widehat{\Psi}}{\partial t_s \partial s} \\ \frac{\partial^2 \widehat{\Psi}}{\partial T \partial t_s} & \frac{\partial^2 \widehat{\Psi}}{\partial t_s \partial s} & \frac{\partial^2 \widehat{\Psi}}{\partial t_s{}^2} \end{pmatrix}$$

Theorem 5.2. *When shortage time t_s is fixed, then the price s^* and the inventory period T^* obtained from Eq. (25) and Eq. (26) have absolute second order condition for profit objective maximization $\widehat{\Psi_a}$.*

At the value s^* and T^* obtained from the equating the necessary derivatives to zero:

$$H = \begin{bmatrix} \frac{\partial^2 \widehat{\Psi_a}}{\partial s^2} & \frac{\partial^2 \widehat{\Psi_a}}{\partial s \partial T} \\ \frac{\partial^2 \widehat{\Psi_a}}{\partial T \partial s} & \frac{\partial^2 \widehat{\Psi_a}}{\partial T^2} \end{bmatrix}$$

$$Det(H) = \frac{\partial^2 \widehat{\Psi_a}}{\partial s^2} * \frac{\partial^2 \widehat{\Psi_a}}{\partial T^2} - \left[\frac{\partial^2 \widehat{\Psi_a}}{\partial s \partial T}\right]^2 > 0 \tag{26}$$

The profit function will be concave with respect to selling price s and inventory period T, if the $Det(H) > 0$ or in other word corresponding Hessian matrix of the profit function is negative definite, i.e. if all the eigenvalues of the Hessian matrix H are negative. We proved it with the help of Mathematica Software.

Theorem 5.3. *When the inventory period T is fixed, then the price s^* and the shortage time $t_s{}^*$ obtained from Eq.(25) and Eq.(27) have absolute second order condition for profit objective maximization $\widehat{\Psi_a}$.*

Theorem 5.4. *When the selling price s^* is fixed, then the inventory period T^* and the shortage time $t_s{}^*$ obtained from Eq.(25) and Eq.(26) have absolute second order condition for profit objective maximization $\widehat{\Psi}$.*

Theorem 5.5. *When shortage time t_s, selling price s and inventory period T are fixed, then the profit function $\widehat{\Psi_a}(t_s, T, s)$ is concave with respect to promotional effort ρ.*

Proof. The first and second order partial derivatives of the total profit function $\widehat{\Psi}(t_s, T, s)$ given by Eq. (22) with respect to ρ are given below:

$$\frac{\partial^2 \widehat{\Psi_a}}{\partial \rho^2} = -\frac{2K\left[(t_s + T)\left(a(ds)^{-b_1} + \mu\right)\right]^\alpha}{t_s + T} < 0 \tag{27}$$

Hence, ρ^* is global optimal that optimizes the profit function $\widehat{\Psi_a}(t_s, T, s)$ for fixed values of selling price s, inventory cycle length T and shortage time t_s.

6. Numerical Example & Sensitivity Analysis

Using Mathematica, the example validates inventory management models by deriving optimal replenishment time, order quantity, and selling price. It demonstrates the framework's effectiveness in real-world scenarios, providing actionable insights for inventory optimization and strategic decision-making.

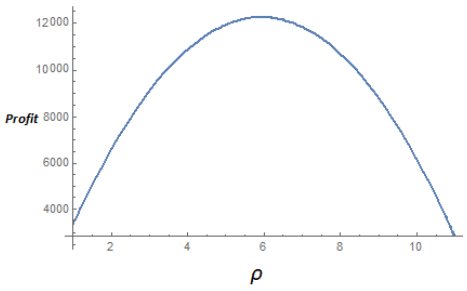


Figure 2: *Profit function w.r.t. promotional effort* ρ

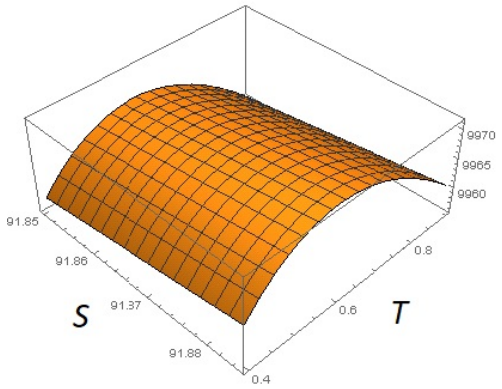


Figure 3: *Profit function w.r.t.* T *and* s

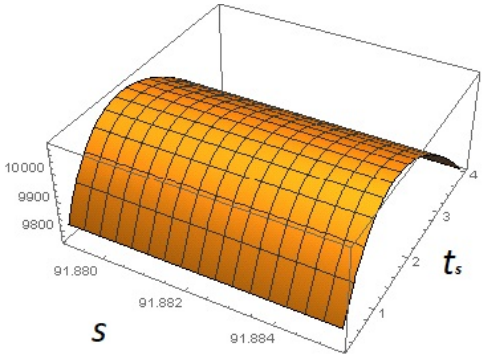


Figure 4: *Profit function w.r.t.* s *and* t_s

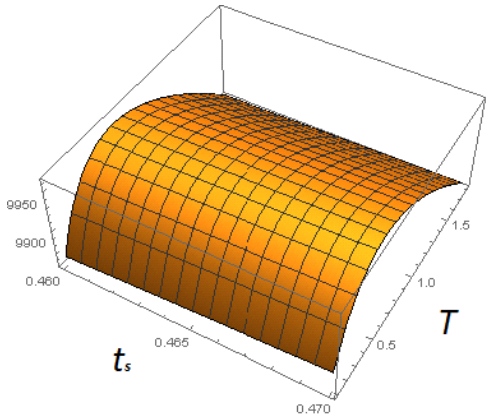


Figure 5: *Profit function w.r.t.* t_s *and* T

Example 1. *In this example, the demand function is given by* $D(s) = a(ds)^{-b_1}$*, where* a*,* d*, and* b_1 *are positive parameters. The system parameters are specified as follows: the ordering cost is* $A = 500$ *per cycle, the holding cost is* $C_h = 2.5$ *per unit per unit time, and the purchasing cost is* $C_p = 3$ *per unit. The backorder cost is* $C_b = 5$ *per unit, while the lost sale cost is* $C_l = 1.5$ *per unit, and the demand rate is* $\mu = 40$ *units per unit time. The coefficient for promotional effort is* $\rho = 3.4$*, and the customer impatience rate is* $\delta = 0.009$ *units per unit time. The deterioration rate is* $\theta = 0.02$ *per unit time, with a constant parameter* $\alpha = 1$*. Additional parameters include* $a = 100$*,* $d = 1$*, an adjustment factor* $a_1 = 0.001$*, and an elasticity parameter* $b_1 = 1.5$*. The cost of preservation technology is* $\xi = 1.2$ *per unit.*
To determine the optimal solutions, the following system of equations is solved:

$$\frac{\partial \widehat{\Psi_a}}{\partial s} = 0, \quad \frac{\partial \widehat{\Psi_a}}{\partial T} = 0, \quad \text{and} \quad \frac{\partial \widehat{\Psi_a}}{\partial t_s} = 0.$$

The optimal solutions are found to be $t_s^* = 0.4095$ *units of time,* $T^* = 1.4728$ *units of time, and* $s^* = 87$ *units. Correspondingly, the total average profit is calculated as:*

$$\widehat{\Psi_a}^* = 9102.46.$$

Through this example, we evaluate all eigenvalues of the Hessian matrix H, given by:

$$H = \begin{pmatrix} \frac{\partial^2 \widehat{\Psi_a}}{\partial T^2} & \frac{\partial^2 \widehat{\Psi_a}}{\partial s \partial T} & \frac{\partial^2 \widehat{\Psi_a}}{\partial T \partial t_s} \\ \frac{\partial^2 \widehat{\Psi_a}}{\partial s \partial T} & \frac{\partial^2 \widehat{\Psi_a}}{\partial s^2} & \frac{\partial^2 \widehat{\Psi_a}}{\partial t_s \partial s} \\ \frac{\partial^2 \widehat{\Psi_a}}{\partial T \partial t_s} & \frac{\partial^2 \widehat{\Psi_a}}{\partial t_s \partial s} & \frac{\partial^2 \widehat{\Psi_a}}{\partial {t_s}^2} \end{pmatrix}$$

The obtained eigenvalues -862.4011, -23.145, and -0.0000146 are negative, indicating that the profit function $\widehat{\Psi}_a(t_s, T, s)$ is concave with respect to t_s, T, and s.

Example 2. *Building on the dataset utilized in the previous example, a discount of* 30% *($d = 0.3$) is applied to the selling price. Under these conditions, the optimal solutions are determined as follows: the optimal time for stock depletion is $t_s^* = 0.4781$ units of time, the replenishment cycle time is $T^* = 1.4501$ units of time, and the optimal stock level is $s^* = 91.88$ units. Consequently, the total average profit is calculated as:*

$$\widehat{\Psi_a}^* = 9906.30.$$

Furthermore, the eigenvalues of the Hessian matrix are determined as follows:

$$H = \begin{pmatrix} \frac{\partial^2 \widehat{\Psi}_a}{\partial T^2} & \frac{\partial^2 \widehat{\Psi}_a}{\partial s \partial T} & \frac{\partial^2 \widehat{\Psi}_a}{\partial T \partial t_s} \\ \frac{\partial^2 \widehat{\Psi}_a}{\partial s \partial T} & \frac{\partial^2 \widehat{\Psi}_a}{\partial s^2} & \frac{\partial^2 \widehat{\Psi}_a}{\partial t_s \partial s} \\ \frac{\partial^2 \widehat{\Psi}_a}{\partial T \partial t_s} & \frac{\partial^2 \widehat{\Psi}_a}{\partial t_s \partial s} & \frac{\partial^2 \widehat{\Psi}_a}{\partial {t_s}^2} \end{pmatrix}$$

The obtained eigenvalues -785.3681, -22.015231, and -0.0000146 are negative, indicating the concavity of the profit function $\widehat{\Psi}_a(t_s, T, s)$ with respect to t_s, T, and s.

Example 3. *Consider the inventory system characterized by the demand function $D(s) = a(ds)^{-b_1}$, where $a, d, b_1 > 0$. The following parametric values are used: the ordering cost is $A = 600$ per cycle, the holding cost is $C_h = 1.5$ per unit per unit time, and the purchasing cost is $C_p = 3.2$ per unit. The backorder cost is specified as $C_b = 3$ per unit, while the lost sale cost is $C_l = 2$ per unit, and the mean demand is $\mu = 50$ units. The promotional effort parameter is $\rho = 2.4$, and the impatience parameter is $\delta = 0.009$ units per unit time. The deterioration rate is $\theta = 0.01$ per unit time, and the constant parameter is $\alpha = 1.2$. Additional parameters include $a = 200$, $d = 1.5$, $a_1 = 0.001$, $b_1 = 1.3$, and the preservation technology cost $\xi = 1.5$ per unit.*
Solving the equations $\frac{\partial \widehat{\Psi}_a}{\partial s} = 0$, $\frac{\partial \widehat{\Psi}_a}{\partial T} = 0$, and $\frac{\partial \widehat{\Psi}_a}{\partial t_s} = 0$ yields the optimal solutions $t_s^* = 0.2314$ units of time, $T^* = 0.9251$ units of time, and $s^* = 105$ units. Correspondingly, the total average profit is computed as:

$$\widehat{\Psi_a}^* = 10027.253.$$

Table 1: *Effect of changes of parameters on optimal solution*

Parameters	Change	t_s^*	T^*	s^*	$\widehat{\Psi_a}^*$
a	100	0.4780	1.4501	91.88	9906.30
	101	0.4761	1.4571	91.98	9919.54
	104	0.4712	1.4821	92.27	9959.33
b_1	1.50	0.4780	1.4501	91.88	9906.30
	1.53	0.4809	1.4339	87.59	9311.17
	1.54	0.4813	1.4295	86.28	9128.50
δ	0.009	0.4780	1.4501	91.88	9906.30
	0.006	0.4551	1.5101	92.72	10000.6
	0.005	0.4410	1.5413	92.96	10021.6
a_1	0.001	0.4780	1.4501	91.88	9906.3
	0.200	0.4725	1.4599	91.97	9917.14
	0.300	0.4701	1.4639	92.01	9921.77
K	9	0.4780	1.4501	91.88	9906.300
	12	0.8012	0.5231	99.74	10500.84
	13	0.8625	0.3921	103.4	10807.00

Table 2: *Effect of Changes for Different Values of ρ and Computational Results for Normal Distribution of ε*

Parameter	Value	t_s^*	T^*	s^*	$\widehat{\Psi_a}^*$
Effect of Promotional Effort (ρ)					
ρ	3.4	0.4780	1.4501	91.88	9906.3
	3.5	0.5321	1.3071	93.82	10400.20
	3.6	0.5510	1.1391	95.17	10825.21
	3.7	0.6532	0.9451	95.88	10931.00
	3.8	0.7251	0.7091	96.25	11320.20
Computational Results for Normal Distribution of ε					
$\varepsilon \sim N(30,1)$	-	0.5490	1.1586	88.50	7104.50
$\varepsilon \sim N(33,1)$	-	0.5386	1.2041	89.00	7875.86
$\varepsilon \sim N(36,1)$	-	0.5386	1.2825	89.91	8706.00
$\varepsilon \sim N(39,1)$	-	0.4917	1.3974	91.26	9586.61
$\varepsilon \sim N(42,1)$	-	0.4447	1.5898	93.51	10604.00

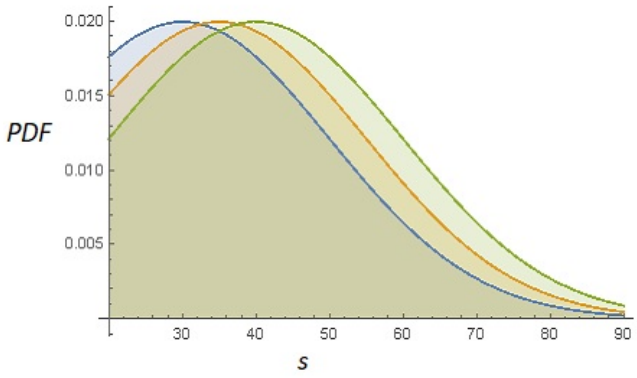


Figure 6: *Graphical representation of probability distribution function with price s*

7. Sensitivity Analysis

(i) From the analysis presented in Table 1, it is evident that an increase in the parameter a results in a corresponding increase in the optimal shortage time t_s^*, optimal inventory period T^*, optimal selling price s^*, and optimal profit $\widehat{\Psi_a}^*$. This trend signifies an enhancement in the market potential of the product, leading to improved optimal outcomes.

(ii) An escalation in the demand rate parameter b_1 leads to an increase in the optimal shortage time t_s^*. Consequently, the heightened shortage costs contribute to a decrease in the optimal profit. This phenomenon arises due to a reduction in demand resulting from the increased parameter b_1.

(iii) The observed results indicate that a decrease in the parameter δ corresponds to a decrease in the optimal shortage time t_s^*. This reduction in backlogging rates aids in minimizing total costs, consequently leading to an increase in optimal profit.

(iv) A rise in the parameter K correlates with a decrease in the optimal inventory period T^*, accompanied by increases in the optimal shortage time t_s^*, optimal selling price s^*, and optimal profit $\widehat{\Psi}^*$. This effect is attributed to the enhanced promotional efforts, resulting in heightened demand and subsequently increased profits.

The promotional effort parameter ρ exhibits a noteworthy influence on the optimal solution. As its value increases, there is a corresponding rise in the optimal shortage time t_s^*, optimal selling price s^*, and optimal profit $\widehat{\Psi}^*$, while the optimal replenishment period T^* decreases. The heightened promotional efforts lead to an augmented demand rate, resulting in a nonlinear increase in expected profit and subsequent decrease. This underscores the importance of judiciously implementing promotional strategies to ensure optimal outcomes.

Table 3 presents the anticipated results for the normal distribution parameters of a random variable. As the demand variable's value increases, so does the expected mean value. Correspondingly, the optimal selling price, preservation technology cost, total profit, and order quantity also increase, while the optimal shortage period decreases. This highlights the opportunity for retailers to escalate production and preservation technology costs in response to increased demand, thereby enhancing total profit and market potential.

Additionally, graphical analysis using three-dimensional (3D) plots for the profit function was conducted. Figures 3, 4, and 5 depict piecewise 3D plots of the profit function against two corresponding variables. The concave nature of the profit function concerning s and T (Figure 3), s and t_s (Figure 4), and T and t_s (Figure 5) is evident. Furthermore, Figure 2 illustrates the concave nature of the profit function with respect to promotional efforts ρ. Figure 6 showcases the probability distribution function with price s.

8. Conclusion

The proposed model offers inventory policies tailored for retailers to optimize their total profit within the current market landscape. Retailers strategically employ price discounts to stimulate product demand while leveraging preservation technology to mitigate deterioration rates. Promotional strategies are intermittently implemented to bolster sales, with market potential influenced by promotional efforts, inventory levels, and product pricing. A numerical example, coupled with sensitivity analysis, provides actionable managerial insights, elucidating the significance of promotional activities such as price discounts. Moreover, investing in preservation technology yields improved total profit, underscoring the pivotal role of prudent decision-making. Retailers and wholesalers alike stand to benefit from strategic promotional efforts, which amplify product demand and subsequently enhance total profit. Additionally, offering price discounts incentivizes larger order sizes, mutually benefiting retailers and customers.

Numerous avenues for future research exist, including extensions of this model to accommodate fuzzy environments, production policies, credit-linked demand, stock-dependent demand, and payment delay facilities. Furthermore, considerations such as supply disruptions, bargaining dynamics, campaign efforts, and return strategies warrant exploration. Additionally, the model's applicability can be extended to encompass multi-item EOQ models and other pertinent inventory management scenarios.

Acknowledgement

Declaration: We declare that this research paper represents our own work and has not been submitted elsewhere for publication. All sources used have been properly cited.

Data Availability Statement:
The datasets generated and analyzed during the current study are not applicable as there was no specific data collected or used for this research.

Funding:
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] Goyal, S. K., & Giri, B. C. (2001). Recent trends in modeling of deteriorating inventory. *European Journal of Operational Research*, 134(1): 1-16.

[2] Goswami A., Chaudhuri K. (1991). EOQ model for an inventory with a linear trend in demand and finite rate of replenishment considering shortages.*International Journal of Systems Science*, 22(1): 181-187.

[3] Chakrabarti T., Chaudhuri K. S. (1997).An EOQ model for deteriorating items with a linear trend in demand and shortages in all cycles. *International Journal of Production Economics*, 49(3): 205-213.

[4] Jalan A. K., Chaudhuri K. S. (1999). Structural properties of an inventory system with deterioration and trended demand.*International Journal of Systems Science*, 30(6): 627-633.

[5] Hariga M. A., Benkherouf L. (1994).Optimal and heuristic inventory replenishment models for deteriorating items with exponential time-varying demand. *European Journal of Operational Research*, 79(1): 123-137.

[6] Wee H. M. (1995). A deterministic lot-size inventory model for deteriorating items with shortages and a declining market. *Computers & Operations Research*, 22(3): 345-356.

[7] Garai T., Chakraborty D., Roy T. K. (2019).Fully fuzzy inventory model with price-dependent demand and time varying holding cost under fuzzy decision variables. *Journal of Intelligent & Fuzzy Systems*, 36(4): 3725-3738.

[8] Mahapatra G. S., Adak S., Kaladhar K. (2019).A fuzzy inventory model with three-parameter Weibull deterioration with reliant holding cost and demand incorporating reliability. *Journal of Intelligent & Fuzzy Systems*, 36(6): 5731-5744.

[9] Olalekan A. A., Bin M. O. (2020).An inventory model for linearly time-dependent deteriorating rate and time-varying demand with shortages partially backlogged. *Journal of Intelligent & Fuzzy Systems*, 38(4): 4545-4557.

[10] Safaeian M., Fathollahi-Fard A. M., Tian G., Li Z., Ke H. (2019).A multi-objective supplier selection and order allocation through incremental discount in a fuzzy environment. *Journal of Intelligent & Fuzzy Systems*, 37(1): 1435-1455.

[11] Aucamp D. C., Kuzdrall P. J. (1986). Lot sizes for one-time-only sales. Journal of the Operational Research Society, 37(1): 79-86.

[12] Bastian M. (1992). A perfect lot-tree procedure for the discounted dynamic lot-size problem with speculation. *Naval Research Logistics (NRL)*, 39(5): 651-668.

[13] Martin, G. E. (1993).A buyer-independent quantity discount pricing alternative. *Omega*, 21(5): 567-572.

[14] Cardenas-Barron L. E., Sana S. S. (2014).A production-inventory model for a two-echelon supply chain when demand is dependent on sales teams' initiatives. *International Journal of Production Economics*, 155: 249-258.

[15] Wee H. M. (1995).Joint pricing and replenishment policy for deteriorating inventory with declining market. *International Journal of Production Economics*, 40(2-3): 163-171.

[16] Shah N. H. (2015).Retailer's replenishment and credit policies for deteriorating inventory under credit period-dependent demand and bad-debt loss. *Top*, 23(1): 298-312.

[17] Tripathi R. P., Tomar S. S. (2015). Optimal order policy for deteriorating items with time-dependent demand in response to temporary price discount linked to order quantity. *International Journal of Mathematical Analysis*, 9(23): 1095-1109.

[18] Soni H. N., Chauhan A. D. (2018).Joint pricing, inventory, and preservation decisions for deteriorating items with stochastic demand and promotional efforts. *Journal of Industrial Engineering International*, 14(4): 831-843.

[19] Liu, P., Hendalianpour, A., & Hamzehlou, M. (2021).Pricing model of two-echelon supply chain for substitutable products based on double-interval grey-numbers. *Journal of Intelligent & Fuzzy Systems*, 40(5): 8939-8961.

[20] Long L. N. B., Kim H. S., Cuong T. N., You S. S. (2023). Intelligent decision support tool for optimizing stochastic inventory systems under uncertainty. *Journal of Intelligent & Fuzzy Systems*, 44(4): 6609-6629.

[21] Arcelus F. J., Bhadury J., Srinivasan G., Arcelus F. J., Bhadury J. (1996). Buyer-seller quantity discount strategies under profit maximizing objectives: A game theoretical analysis. *International Journal of Operations & Quantitative Management (IJOQM)*, 2: 49-72.

[22] Khedlekar, U. K., Kumar, L., Sharma, K., & Dwivedi, V. (2024). An Optimal Sustainable Production Policy for Imperfect Production System with Stochastic Demand, Price, and Machine Failure with FRW Policy Under Carbon Emission. *Process Integration and Optimization for Sustainability*, 8: 919–938.

[23] Kumar, L., Ghoshi, P. S., Saxena, S., & Sharma, K. (2024). A Comparative Study of Inventory Modelling: Deterministic Over Stochastic Approach. *Reliability: Theory & Applications*, 19(1 (77)): 804–818.

[24] Kumar, L., Sharma, K., & Khedlekar, U. K. (2024). Dynamic pricing strategies for efficient inventory management with auto-correlative stochastic demand forecasting using exponential smoothing method. *Results in Control and Optimization*, 15: 100432.

[25] Keswani, M., Dwivedi, V., Kumar, L., Kumar, B., & Khedlekar, K. U. (2024). Efficiency of managing a stochastic inventory system in a declining market for non-instantaneous deteriorating items under partial backlogs. *Yugoslav Journal of Operations Research*, (00): 47–47.

[26] Tiwari, R., Khedlekar, S., Tiwari, R. K., Kumar, L., & Khedlekar, U. K. (2024). Lifecycle product inventory optimization: A hybrid approach with Grey-Wolf and Ant-Colony methods. *Journal of Remanufacturing*, 14: 219–252.

[27] Khedlekar, U. K., & Kumar, L. (2024). Mathematical modelling for convertible items with rework using particle swarm optimisation. *International Journal of Systems Science: Operations & Logistics*, 11(1): 2306222.

[28] Kumar, L., Khedlekar, S., & Khedlekar, U. K. (2024). A comparative assessment of Holt-Winters Exponential Smoothing and Autoregressive Integrated Moving Average for inventory optimization in supply chains. *Supply Chain Analytics*, 8: 100084.

[29] Sharma, K., Kumar, L., & Khedlekar, U. K. (2024). Stochastic optimization of perishable inventory incorporating preservation, freshness index, expiry date, and optimizing promotional strategies for effective management. *Reliability: Theory & Applications*, 19(4 (80)): 666–685.

[30] Adaraniwon A. O., Omar M. B. (2019). An inventory model for delayed deteriorating items with power demand considering shortages and lost sales. *Journal of Intelligent & Fuzzy Systems*, 36(6): 5397-5408.

CAUSES OF ELECTROMECHANICAL RESONANCE IN WIND TURBINES AND METHODS FOR ITS ELIMINATION

N.S. Mammadov

•

Azerbaijan State Oil and Industry University, Baku, Azerbaijan
nicat.memmedov.sa@asoiu.edu.az

Abstract

Electromechanical resonance in wind turbines is a complex and multifaceted phenomenon that significantly affects the performance, service life, and reliability of wind turbines. This problem is especially relevant for modern high-power wind turbines operating under variable loads and high air turbulence. The article provides a detailed analysis of the mechanisms of electromechanical resonance, including the influence of design parameters, aerodynamic loads, electromagnetic interactions, and operating conditions. The physical foundations of resonance phenomena and their impact on the dynamic behavior of wind turbines are considered, which allows for a deeper understanding of the mechanisms of their occurrence and the development of methods for their prevention. Particular attention is paid to modern methods of diagnostics and prediction of resonance effects. The existing technologies for suppressing resonance oscillations are analyzed, such as the use of damping elements, active vibration dampers, adaptive control systems, and digital twins, which allow for real-time prediction and correction of generator behavior. In addition, the article proposes innovative engineering solutions aimed at increasing the stability of wind turbines to resonance phenomena. These solutions include optimizing design parameters, selecting new high-strength and damping materials, and improving automated control systems. Research into this problem and developing effective measures to eliminate it are key aspects of further development of wind energy, allowing for an increase in the service life of equipment, its reliability, and a reduction in operating costs.

Keywords: wind turbine, electromechanical resonance, aerodynamic loads, damping elements, reliability.

I. Introduction

Modern wind power plays a crucial role in the global transition toward renewable energy sources. It offers a sustainable, environmentally friendly, and cost-effective means of electricity generation, reducing dependence on fossil fuels and lowering greenhouse gas emissions. However, the increasing scale and complexity of wind turbines introduce numerous engineering challenges that must be addressed to ensure their efficiency, reliability, and longevity. Among these challenges, electromechanical resonance stands out as a critical issue that can significantly impact the stability and operational performance of wind energy systems. Electromechanical resonance is a phenomenon characterized by the interaction of mechanical and electrical oscillations within a system. This complex interplay can lead to excessive vibrations, component overheating, and accelerated wear and tear, ultimately threatening the integrity of wind turbine structures. If not properly managed, electromechanical resonance can cause severe mechanical failures and substantial financial losses due to maintenance, repairs, and operational downtime.

Electromechanical resonance occurs when the natural frequencies of a turbine's mechanical components coincide with the characteristic frequencies of the electrical system. The mechanical elements of a wind turbine—such as the rotor, blades, bearings, and tower—exhibit specific vibrational modes that can be excited under certain operating conditions. Similarly, the electrical system, which includes generators, converters, and power electronics, possesses inherent resonance frequencies associated with circuit dynamics and electromagnetic interactions.

When these mechanical and electrical frequencies align, resonance amplifies oscillatory motion, leading to increased vibrations. This excessive movement exerts additional mechanical stress on the turbine structure, accelerating fatigue and leading to critical component failures. Key components that are particularly vulnerable to resonance-induced damage include:

- Rotor and Blades: Excessive vibrations can cause cracks, deformation, and eventual structural failure.

- Bearings: Increased mechanical load results in premature wear, overheating, and reduced operational lifespan.

- Supporting Structures: Towers and nacelles may experience structural fatigue, leading to stability issues.

- Generators and Power Electronics: Resonance can cause thermal stress, insulation breakdown, and component degradation [1-3].

External Factors Contributing to Resonance. While the fundamental causes of electromechanical resonance lie in the interaction of mechanical and electrical dynamics, external environmental conditions further exacerbate the issue. Wind turbines, especially high-capacity models operating in turbulent atmospheric conditions, are particularly susceptible to fluctuations in operating loads. Some of the key external factors that can amplify resonance effects include:

1. Sudden Wind Gusts: Rapid changes in wind speed generate abrupt variations in aerodynamic forces, triggering structural oscillations.
2. Temperature Fluctuations: Thermal expansion and contraction alter material properties, shifting natural frequencies and resonance characteristics.
3. Voltage Instability in the Grid: Variations in electrical load and grid conditions can induce frequency mismatches in the power system, contributing to resonance effects.

Given these external influences, mitigating electromechanical resonance is one of the top priorities for wind turbine developers, engineers, and maintenance specialists.

The study of electromechanical resonance in wind power systems dates back to the mid-20th century, when the wind energy industry was in its infancy. Early wind turbine designs, though relatively simple, faced operational instabilities due to unanticipated resonance effects. As turbine technology advanced and power ratings increased, the impact of resonance phenomena became more pronounced.

Today, with the widespread deployment of multi-megawatt wind turbines—some exceeding 10 MW in capacity—the challenge of electromechanical resonance has grown significantly. Modern turbine designs incorporate advanced power electronics, variable-speed generators, and sophisticated control algorithms, making the dynamics of resonance even more complex. According to industry statistics, approximately 30% of wind turbine failures are linked to excessive vibrations caused by resonance effects. This underscores the urgency of developing new mitigation strategies to improve the resilience and reliability of wind power systems [4-6].

Electromechanical resonance is inherently a multidisciplinary issue, arising at the intersection of mechanical engineering, electrodynamics, materials science, and automated control systems. This multifaceted nature makes its detection and elimination particularly challenging. Some of the main difficulties include: complex system interactions, variable operating conditions, scaling issues. To effectively address these challenges, researchers and engineers have developed a combination of advanced diagnostic tools, predictive modeling techniques, and control strategies.

II. Formulation of the problem

Relevance of the Problem. The significance of electromechanical resonance has increased with the advancement of modern wind energy technologies. High-power wind turbines, which operate under variable aerodynamic and electromagnetic loads, are particularly susceptible to resonance phenomena. The continuous changes in wind speed, direction, and turbulence, combined with the influence of electrical grid fluctuations, create a dynamic environment where resonance conditions can arise unexpectedly. These effects not only reduce the efficiency of energy conversion but also accelerate wear and tear on the system, leading to increased maintenance costs and reduced service life.

Fundamental Causes of Electromechanical Resonance. The emergence of electromechanical resonance in wind turbines is influenced by several key factors:

1. Structural and Design Parameters: The mass distribution, stiffness, and damping characteristics of turbine components play a crucial role in determining the natural frequencies of the system. Poorly optimized designs may result in frequency coincidences, leading to resonance.
2. Aerodynamic Loads: Wind turbulence, gusts, and rapid changes in airflow exert fluctuating forces on the rotor blades, contributing to mechanical oscillations that can align with the system's resonance frequencies.
3. Electromagnetic Interactions: The generator and power electronics introduce electrical oscillations that can couple with mechanical vibrations. Harmonic distortions and grid instabilities further amplify these interactions, increasing the risk of resonance.
4. Operating Conditions: Variations in temperature, wind shear, and grid voltage fluctuations affect material properties and electrical system stability, potentially triggering resonance phenomena under specific conditions.

Consequences of Electromechanical Resonance. The impact of resonance in wind turbines can be severe, affecting both performance and reliability:

- Increased Vibrations: Excessive oscillations accelerate material fatigue, leading to cracks, wear, and eventual failure of critical components.
- Structural Damage: Bearings, gearbox components, and rotor blades are particularly vulnerable to long-term resonance exposure.
- Efficiency Reduction: Energy losses due to uncontrolled oscillations result in suboptimal power generation.
- Operational Downtime: Frequent failures require costly repairs and increase turbine downtime, reducing overall energy output and financial viability [7-9].

Challenges in Mitigation. Addressing electromechanical resonance requires a multidisciplinary approach, combining mechanical engineering, aerodynamics, electromagnetics, and control systems. The complexity of wind turbine dynamics makes it difficult to predict and eliminate resonance effects entirely. Existing mitigation methods include:

- Passive Damping Systems: Integration of structural damping materials and mechanical absorbers to reduce vibration amplitude.

- Active Vibration Control: Implementation of real-time adaptive control strategies using sensors and actuators to counteract resonance effects.

- Digital Twin Technology: Advanced simulation and predictive maintenance approaches that model turbine behavior under different operating conditions to anticipate and prevent resonance occurrences.

Research and Future Directions. The development of innovative solutions for resonance suppression is essential for the continued advancement of wind energy. Future research should focus on:

- Optimization of Structural Design: Tailoring turbine geometries to minimize resonance

risks through computational modeling.

• Development of High-Strength Materials: Exploration of new composite materials with enhanced damping properties to improve structural resilience.

• Intelligent Control Algorithms: Implementation of AI-driven predictive maintenance and adaptive control strategies for real-time resonance mitigation.

By addressing electromechanical resonance through advanced engineering approaches, wind turbine manufacturers and operators can improve the reliability, efficiency, and longevity of their systems, ultimately contributing to the sustainable growth of renewable energy [10-11].

III. Problem solution

Electromechanical resonance can lead to fatigue damage, power loss and failure of key components such as blades, drives and generator sets. To cope with this challenge, it is necessary to implement an integrated approach that includes both engineering solutions and innovative diagnostic and control methods.

One of the most effective ways to combat electromechanical resonance is to optimize the design characteristics of the turbine. In particular, the blade design is a critical element. Incorrectly selected blade geometry or angle of attack can lead to significant vibrations and increased resonance effects. Therefore, it is important to carefully design and test blades taking into account aerodynamic characteristics, as well as their dynamic behavior in different wind conditions. Modern approaches include the use of blades with variable geometry, which allows their characteristics to be adapted depending on the wind speed and reduce the likelihood of resonance.

Another important aspect is the design of the turbine tower. The tower vibration frequencies must be designed in such a way that they do not coincide with the frequencies of external influences, such as blade vibrations or wind flows. Optimization of the tower rigidity and mass, as well as the use of materials that can reduce the amplitude of vibrations, helps reduce mechanical loads and prevent damage [12-15].

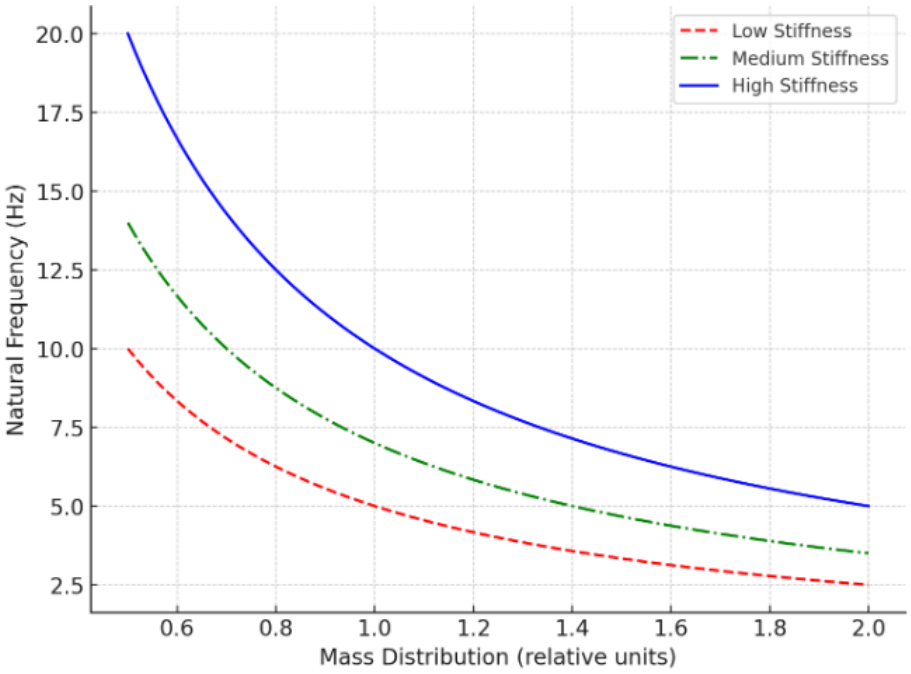


Figure 1: *Dependence of natural frequency on mass distribution and stifness*

Figure 1 shows the dependence of the natural frequency of the structure on the parameters of the mass distribution and rigidity of the elements. A clear shift in resonant frequencies is visible with an increase in the rigidity of the tower and the use of reinforced materials in the blades. In addition, an important aspect is the selection of materials that can increase the system's resistance to resonance effects. Modern materials, such as carbon composites or aluminum alloys, provide high strength with less weight, which helps reduce resonant vibrations and improve the overall dynamics of the turbine. The use of materials with improved damping properties is also of great importance. Materials such as special silicate coatings or rubber-like inserts absorb vibration energy and reduce its transmission to other components of the system. For example, composite materials with an integrated damping layer can significantly reduce vibration levels at critical frequencies, as shown in Figure 2.

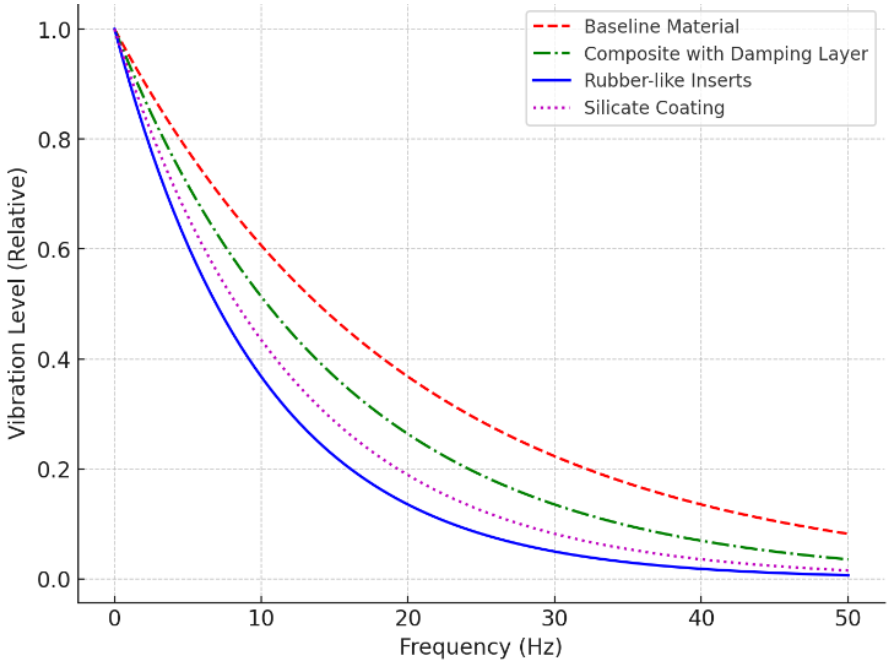


Figure 2: *Effect of damping materials on vibration reduction*

Aerodynamic stabilization is one of the most important factors in the process of suppressing resonance in wind turbines. Wind turbines, due to their design and operation, are exposed to turbulent air flows, which plays a key role in exciting mechanical vibrations that can lead to the occurrence of resonance phenomena. Wind turbulence, especially under conditions of variable wind speed and direction, can significantly amplify vibrations if turbine components, such as blades or a tower, are not properly adjusted [16-20]. This process can lead to unwanted vibrations that increase the load on the structure and reduce operating efficiency. In order to minimize the likelihood of resonance, modern technologies offer advanced blade designs that are able to operate effectively in various aerodynamic conditions. The design process for such blades involves not only selecting the optimal geometry and material, but also implementing mechanisms that allow the blades to dynamically change their angle of attack depending on the wind speed and direction. These real-time angle of attack control mechanisms allow for more precise control of wind flow effects, which reduces the load on the turbine structure and prevents resonant oscillations. Thus, aerodynamic stabilization helps reduce the impact of wind turbulence, increasing the operational stability and durability of turbines. Figure 3 shows the effect of active angle of attack control on the amplitude of blade vibrations. It can be seen that, when optimally configured, the system effectively dampens oscillations even with sharp changes in wind speed.

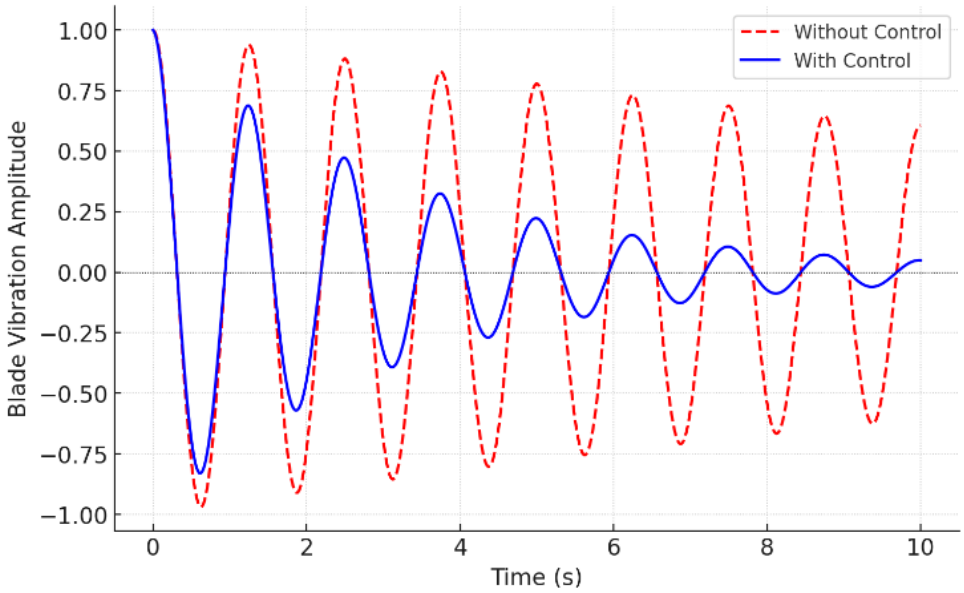


Figure 3: *Effect of active pitch control on blade oscillations*

An important addition to this process is the use of electromechanical control systems, which significantly enhance the turbine's resistance to resonance phenomena. Smart power electronics integrated with grid stabilization systems allow for effective control of electrical loads, preventing unwanted oscillations between the mechanical and electrical components of the system. Ideally, such systems can actively regulate the flow of energy in the grid, maintaining turbine stability even when external conditions change. In addition, modern turbines widely use real-time feedback mechanisms, which play a crucial role in adapting turbine operation to changes in external conditions. Intelligent controllers integrated with a monitoring system are capable of continuously monitoring changes during operation, such as an increase or decrease in wind load, and adjusting turbine operating parameters accordingly. These control systems allow the turbine to adapt to changes in resonance conditions and prevent its occurrence by adjusting factors such as the pitch of the blades or their rotation speed. As a result, these high-tech solutions not only prevent possible component failures and degradation, but also significantly improve the overall efficiency of turbines, especially under variable wind loads [21-24].

Aerodynamic stabilization and electromechanical control systems work closely together, ensuring a synchronized response of the turbine to changing environmental conditions and reducing the risks associated with resonant oscillations. Innovative approaches to design and control have significantly improved the performance of turbines, ensuring their reliable operation over a long period of time in different weather conditions.

One of the significant achievements in the field of resonance suppression in wind turbines has been the introduction of digital twin technology. This approach involves creating a virtual copy of a physical object, in this case a wind turbine, which functions as an active model of the real device. The digital twin receives a continuous stream of data on the actual operation of the turbine, including information on its vibrations, temperature, pressure and other parameters. This data is used to simulate various operating and external conditions that can affect the behavior of the turbine. Thus, a digital twin not only reflects the current state of the turbine, but also allows you to predict possible problems and scenarios that the device may face in the future. One of the key aspects of using digital twins is predictive maintenance. Using virtual models, you can analyze the behavior of the turbine in various situations, which helps to identify potential anomalies associated with resonance before they become a threat to the system. This makes it possible to take

measures in advance to prevent catastrophic breakdowns and extend the service life of the device. By predicting possible failures and problems, you can optimize maintenance schedules and minimize turbine downtime, which is important for the efficient operation of wind turbines. In addition to digital twins, monitoring and diagnostic systems that are implemented in the designs of modern wind turbines play an important role in suppressing resonant oscillations. These systems are equipped with a wide range of sensors that continuously collect data on vibrations, temperature, rotation speed and other indicators of the turbine's operation. Sensor networks provide round-the-clock monitoring of the device's condition, which allows you to quickly respond to any changes in its operation. Machine learning and artificial intelligence algorithms process the collected information and analyze it to detect characteristic patterns that may indicate the presence of resonance. This enables the system to diagnose anomalies at an early stage of their occurrence, before they lead to serious damage. Thus, the integration of digital twins and modern sensor technologies creates a powerful platform for predicting, monitoring and eliminating resonance-related problems. The use of these technologies can significantly improve the reliability and safety of wind turbines, reducing the risk of their breakdowns and optimizing maintenance and repair costs [25-28].

IV. Conclusions

Electromechanical resonance remains a critical issue in modern wind power, particularly with the increasing capacity of wind turbines exceeding 10 MW. Recent advancements in turbine design, including adaptive blade geometries, high-damping composite materials, and optimized tower structures, have significantly improved resistance to resonance-induced failures.

A key innovation in this field is the integration of digital twin technology, which enables real-time predictive analysis of turbine behavior under varying conditions. By continuously processing operational data, digital twins help preemptively identify resonance risks, allowing for proactive adjustments before structural damage occurs. Additionally, AI-driven control systems and active vibration damping mechanisms have introduced new levels of adaptability, ensuring optimal turbine performance even in turbulent wind conditions.

The combination of advanced structural engineering, real-time monitoring, and intelligent control algorithms marks a new era in wind turbine reliability. These innovations not only extend the lifespan of critical components but also enhance energy output efficiency, reinforcing wind power as a cornerstone of the global renewable energy transition.

References

[1] N.S. Mammadov, N.A. Ganiyeva, G.A. Aliyeva, “Role of Renewable Energy Sources in the World”. Journal of Renewable Energy, Electrical, and Computer Engineering, Vol. 2, №2, pp. 63-67, Indonesia, 30 September 2022

[2] Shuai Z., Liu D., Shen J., Tu C., Luo, A., “Series and parallel resonance problem of wideband frequency harmonic and its elimination strategy”, IEEE Trans. Power Electron., 2014, 29, pp. 1941–1952

[3] Song Y., Ebrahimzadeh E., Blaabjerg F., “Analysis of high-frequency resonance in DFIG-based offshore wind farm via long transmission cable”, IEEE Trans. Energy Convers., 2018, 33, pp. 1036–1046

[4] Mammadov N.S., Rzayeva S.V., Ganiyeva N.A., “Analisys of synchronized asynchronous generator for a wind electric installation”, Przegląd Elektrotechniczny, №5, 2023, pp.37-40, doi:10.15199/48.2023.05.07

[5] Nijat Mammadov, “Analysis of systems and methods of emergency braking of wind turbines”. International Science Journal of Engineering & Agriculture, Vol. 2, № 2, pp. 147-152, Ukraine, April 2023

[6] Zhang Y., Christian Klabunde C., Wolter M., “Harmonic Filtering in DFIG-based Offshore Wind Farm through Resonance Damping”, In Proceedings of the 2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), Bucharest, Romania, 29 September–2 October 2019

[7] Nijat Mammadov, Ilkin Marufov, Saadat Shikhaliyeva, Gulnara Aliyeva, Saida Kerimova, “Research of methods power control of wind turbines”, Przeglad elektrotechniczny, R. 100 NR 5/2024, pp. 236-239

[8] Bossanyi E., Wright A., "Field testing of individual pitch control on the NREL CART-2 wind turbine", In Proceedings of the European Wind Energy Conference, Marseille, France, 16–19 March 2009
[9] Thakur S., Saha N., "Load Reduction on Offshore Wind Turbines by Aerodynamic Flaps", In Proceedings of the ASME 2017 36th International Conference on Ocean, Offshore and Arctic Engineering, American Society of Mechanical Engineers, Trondheim, Norway, 25–30 June 2017
[10] Ilkin Marufov, Aynura Allahverdiyeva, Nijat Mammadov, "Study of application characteristics of cylindrical structure induction levitator in general and vertical axis wind turbines", Przegląd elektrotechniczny, R. 99 NR 10/2023, pp.196-199
[11] Mammadov N.S., "Vibration research in wind turbines", XV International Scientific and Practical Conference «The main directions of the development of scientific research», Helsinki, Finland, 2023, pp. 345-346
[12] N.S. Mammadov, "Methods for improving the energy efficiency of wind turbines at low wind speeds", Vestnik nauki journal, Issue 2, Vol. 61, №4, Russia, April 2023
[13] Bons J.P., Sondergaard R., Rivir R.B., "Turbine separation control using pulsed vortex generator jets", In Proceedings of the ASME Turbo Expo 2000: Power for Land, Sea, and Air, American Society of Mechanical Engineers, Munich, Germany, 8–11 May 2000.
[14] Ilkin Marufov, Najiba Piriyeva, Nijat Mammadov, Shukufa Ismayilova, "Calculation of induction levitation vertical axis wind generator-turbine system parameters, levitation and influence loop" Przegląd elektrotechniczny, 2024 , No.2, pp.135-139
[15] Cheah-Mane M., Liang J., Jenkins N., Sainz L., "Electrical resonance instability study in HVDC-connected Offshore Wind Power Plants", In Proceedings of the 2016 IEEE Power and Energy Society General Meeting (PESGM), Boston, MA, USA, 17–21 July 2016
[16] Mammadov N.S., Aliyeva G.A. "Energy efficiency improving of a wind electric installation using a thyristor switching system for the stator winding of a two-speed asynchronous generator", IJTPE, Issue 55, Volume 55, Number 2 , pp. 285-290, June 2023
[17] Brown, R.E. Electric Power Distribution Reliability; CRC Press: Boca Raton, FL, USA, 2017
[18] Maldonado V., Farnsworth J., Gressick W., Amitay M., "Active control of flow separation and structural vibrations of wind turbine blades. Wind", Energy Int. J. Prog. Appl. Wind Power Convers. Technol., 2010, 13, pp. 221–237
[19] I.M. Marufov, N.S. Mammadov, K.M. Mukhtarova, N.A. Ganiyeva, G.A. Aliyeva "Calculation of main parameters of induction levitation device used in vertical axis wind generators". International Journal on technical and Physical Problems of Engineering", Issue 54, Volume 15, Number 1, pp. 184-189, March 2023
[20] Mammadov N. S., Mukhtarova K. M., "METHODOLOGY FOR ASSESSING THE RELIABILITY OF AGS BASED ON RENEWABLE ENERGY SOURCES", Reliability: Theory & Applications, Volume 19, No 4 (80), pp.648-653, 2024
[21] Mammadov N. S., "SAFETY IN THE OPERATION OF WIND TURBINES: ENSURING RELIABILITY AND PROTECTION", Vestnik nauki, Issue 1, №.7, Vol. 76, pp. 883-888, 2024
[22] Nadir Aliyev, Elbrus Ahmedov, Samira Khanahmedova, Sona Rzayeva, "Synthesis of the Exact Parameters of the Electromagnetic Brake of a Wind Electric Installation", Przeglad Elektrotechniczny journal, № 10, Poland, 2023
[23] I.N. Rahimli, S.V. Rzayeva, E.E. Umudov, "DIRECTION OF ALTERNATIVE ENERGY", Vestnik nauki, Issue 2, Vol. 61, №4, April 2023
[24] Nijat Mammadov, Najiba Piriyeva, Shukufa Ismayilova, "Research of lightning protection systems for wind electric installations", Przeglad Elektrotechniczny journal,
[25] Rzayeva S.V., Piriyeva N.M., Guseynova I.A. Analysis of reliability of typical power supply circuits. Reliability: Theory and Applications, RTA, №3 (79), Volume 19, pp. 173-178, September 2024
[26] Nijat Mammadov, "PROSPECTS FOR THE DEVELOPMENT OF RENEWABLE ENERGY SOURCES", The 29th International scientific and practical conference "Modern scientific trends and youth development"(July 25–28, 2023) Warsaw, Poland, International Science Group, pp. 211-213, 2023
[27] Elshad Safiyev, Ilham Rahimli, Nijat Mammadov, "Method of qualitative impregnation of electric motor windings." Przegląd Elektrotechniczny, R. 100 NR 8/2024, pp.238 -240
[28] Carcangiu C.E., Pineda I., Fischer T., Kuhnle B., Scheu M., Martin M., "Wind turbine structural damping control for tower load reduction", In Civil Engineering Topics; Springer: Berlin/Heidelberg, Germany, 2011; Volume 4, pp. 141–153

MINIMIZATION OF ACTIVE POWER LOSSES IN ALTERNATING CURRENT MAGNETIC SYSTEMS

G.V.Mamedova
•
Azerbaijan State Oil and Industry University, Baku, Azerbaijan
gulaya68@mail.ru

Abstract

In magnetic systems of electromagnetic alternating current devices, active power losses occur in steel and copper, reducing which is an urgent task. In order to minimize the loss of active power, a method based on the geometric optimization of the magnetic system has been developed. The optimization criteria are the coefficients of the proportionality principle. The last requirement is necessary for the proper coordination of the designed device with other mechanisms and for improving the technical and economic performance of the devices. One of the proportionality principles for electromagnetic devices is the construction of designs with a given coefficient $n_0 = h_0/c_0$, where h_0/c_0, and are the height and thickness of the winding, respectively, as the overall dimensions of the device are mainly determined by these parameters. As a rule, the value of the coefficient n_0 can be ensured after the preliminary calculation of the magnetic circuit. On the other hand, there are complex interrelations between the coefficient n_0 ampere-turns, and the temperature rise of the winding. Excessive increase or decrease of the coefficient n_0 disrupts this principle, reduces the current density in the winding, increases the cross-section of the material and the winding capacity, and fails to achieve the specified temperature rise of the winding, etc. In this case, the main task of optimizing the winding dimensions is to ensure the maximum current density in the winding while maintaining the given coefficient n_0, temperature rise, and the ampere-turn values obtained from the preliminary calculation. In this study, for the most common electromagnetic design, considering the given ratios of winding height to its thickness, temperature rise, and ampere-turns, the optimal parameters and dimensions of the magnetic system have been determined. Loss minimization has been carried out for electromagnetic devices with voltage winding and current winding.

Key words: electromechanical device, control system, frequency converter, control device, power part, modulation method, output voltage, three-phase autonomous voltage inverter, semiconductor device, analog-to-digital converter, controller, comporator.

I. Introduction

It is well-known that reducing losses of active power in electromagnetic devices operating with alternating current of various purposes (chokes, measurement transformers, relays, displacement and force transducers, etc.) is absolutely necessary, as it enhances the technical and economic performance of these devices. The research conducted in this direction has led to the development of electrical steels and conductive materials with improved characteristics. However, reducing losses of active power in alternating current magnetic systems (ACMS) is possible through the proper selection of the geometry of the magnetic system (MS) design.

In this study, the active resistances of the excitation winding (EW) r, the losses of active power in the EW windings P_{10}, the short-circuited turn (SCT) P_s, and the steel sections of the magnetic circuit P_m of alternating current magnetic systems (ACMS) are minimized using geometric optimization methods.

The optimization criteria are dimensionless coefficients of the proportionality principle: $n_a=b/a$, $n_c=b/c$, $n_0=h_0/c_0$ [3]. In this case, the core area of the magnetic circuit $S_c=ab$ and the area of the EW $S_c=c_0h_0$ are considered as constant parameters, as they are respectively defined by the known values of magnetic induction in the core B_{max} and the ampere-turns of the EW F=IW [1,5].

In electromagnetic devices with current windings, the current value I in the winding remains constant when the air gap of the armature changes (I=const). However, in electromagnetic devices with voltage windings (U=const), increasing the armature air gap leads to an increase in the coil current because the coil inductance L decreases [1,2]. Due to this, the method of determining active power losses differs between I=const and U=const cases. In the former case $P_{10} \sim I^2$, while in the latter case, $P_{10} \sim U^2$.

The same applies to the losses of active power P_m caused by losses in the steel sections of the magnetic circuit. Additionally, some ACMS contain short-circuited turns (SCT), which also contribute to losses P_s. As a result, the total losses P are determined as the sum of P_{10}, P_s and P_m. This further complicates the determination of active power losses in ACMS [5-9].

The diagrams of the most common alternating current magnetic systems are shown in Fig. 1-3, where the main dimensions of the structures are indicated.

Taking into account the above, solutions to the following tasks are considered here:

1. Determination of the main dimensions and parameters of the MS using dimensionless coefficients of the proportionality principle $n_0=h_0/c_0$, $n_a=b/a$, $n_c=b/c$.
2. Selection of methods for determining active power losses for various alternating current magnetic systems [3].
3. Minimization of active power losses in the current winding (I=const), voltage winding (U=const) and closed loop.
4. Determination of the optimal values of the geometric dimensions of the magnetic circuit and EW.
5. Minimization of active power losses in the steel of the split pole of a single-phase electromagnet [8].

II. Statement and solution of the problem.

In this context, we utilize expressions for the electrical resistances of the excitation winding and the short-circuited turn.

$$Z = \frac{U}{I} = r_{10} + r_{\sim} + j(x_{1s} + x_m); \tag{1}$$

$$r_{10} = \rho_1 \frac{l_1 w_1^2}{k_{з1} S_{01}}; \quad r_k = \rho_k \frac{l_k w_k^2}{k_k S_k} \tag{2}$$

$$x = x_s + x_m = \sigma x_m \tag{3}$$

for the ampere-turns of the excitation winding (EW) and the short-circuited turn (SCT).

$$F_1 = j_1 k_{31} S_{01}; \quad F_2 = j_k k_k S_{02}; \quad F_2 = b_2 F_1, \tag{4}$$

for magnetic resistances of steel and air sections:

$$\dot{Z}_m = \frac{\dot{F}_1}{\dot{\Phi}_0} = R_m + jX_m \tag{5}$$

$$R_m = \rho_R \frac{l_m}{S_c}; \quad X_m = \rho_x \frac{l_m}{S_c}; \quad R_\delta = \frac{\delta}{\mu_0 S_c}, \tag{6}$$

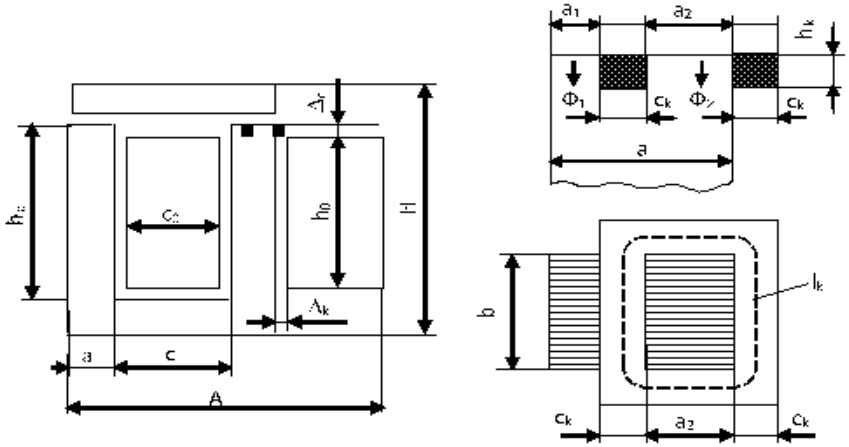
for losses of active capacities in the excitation winding P_{10}, the short-circuited coil P_k and the steel sections P_c:

$$P_{10} = I^2 r_{10}; \quad P_k = I_k^2 r_k; \quad P_c = I^2 r_{\sim}, \tag{7}$$

for the temperature rise of the excitation winding (EW) and the short-circuited turn (SCT).:

$$\tau_1 = \frac{P_{10} + P_k + P_c}{k_T S_{ox}}; \quad \tau_k = \frac{P_k}{k_T S_{ox}} \quad , \tag{8}$$

where r_{10} and r_k are the active resistances of the excitation winding (EW) and the short-circuited turn (SCT); $r_{\sim}$ is the equivalent active resistance, attributed to losses in the steel and SCT; x_{1s} and x_m are the leakage inductive resistance and the main inductive resistance of the EW, corresponding to leakage fluxes and the working magnetic flux Φ_0 ; ρ_1 and $\rho_к$ are the specific electrical resistances of the EW and SCT conductors; ρ_R and ρ_x are the specific active and reactive magnetic resistances of the steel; $к_{31}$ and $к_к$ are the fill factors of copper in the EW and SCT; σ is the coefficient of magnetic flux dispersion; j_1 and j_2 are the current densities in the EW and SCT; $l_{ср1}$ and l_k are the average lengths of the winding turns in the EW and SCT; l_m is the average length of the path of the working magnetic flux Φ_0 ; R_m and x_m are the active and reactive magnetic resistances of the steel sections; R_δ is the magnetic resistance of the air gap δ through which the magnetic flux Φ_0 passes; τ_1 and $\tau_к$ are the temperatures of overheating of the EW and SCT.



a) b)

Figure 1. *Magnetic system of a single-phase electromagnet with a split pole.*

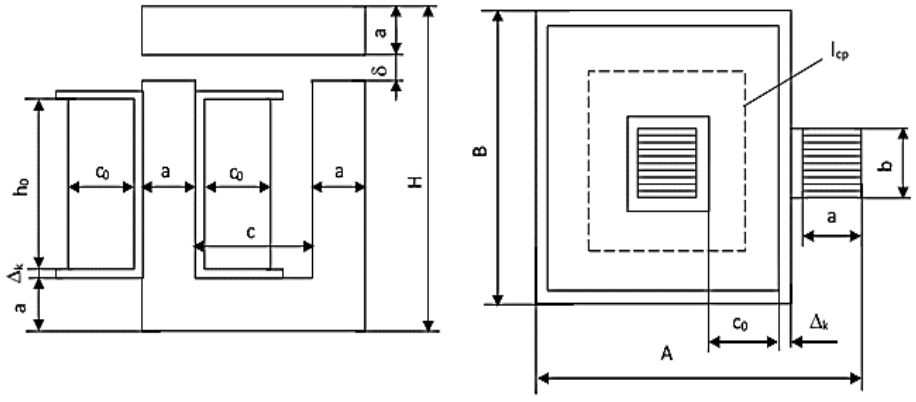


Figure 2. *Magnetic system of an inductive displacement transducer.*

The subscripts "1" and "k" in the formulas correspond respectively to the EW (excitation winding) and SCT (short-circuited turn). Formulas with the subscript "k" relate to Figures 1 and 2b.

The notations in the provided expressions are well-known [1-5] and do not require explanations. Equations (2) and (7) take into accou nt the overheating temperatures τ_1 and τ_k of the windings [3,5].

III. Definition of linear dimensions and key parameters in terms of dimensionless quantities.

To simplify optimization problem solutions, we will express linear dimensions and fundamental parameters in terms of dimensionless quantities.
Using the notations provided in Figures 1-2, we determine the overall dimensions of the magnetic system for a single-phase electromagnet and an inductive displacement transducer.

$$A = 2a + 2c_0 + 4\Delta_k \approx 2a + 2c = 2c\left(1 + \frac{n_c}{n_a}\right), \tag{9}$$

$$B = b + 2c_0 + 4\Delta_k \approx b + 2c = c(2 + n_c) \tag{9,a}$$

$$H = \delta_{mak} + 2a + h_0 + 2\Delta_k \approx \delta_{mak} + c\left(n_0 + 2\frac{n_c}{n_a}\right) \tag{9, b}$$

where

$$n_a = \frac{b}{a};\ \ n_c = \frac{b}{c};\ \ n_0 = \frac{h_0}{c_0};\ b = cn_c;\ \frac{a}{c} = \frac{n_c}{n_a}. \tag{10}$$

$4\Delta_к << (2a + 2c_0)$; $4\Delta_к \le (b + 2c_0)$; $2\Delta_к << (2a + h_0)$; $c = c_0$

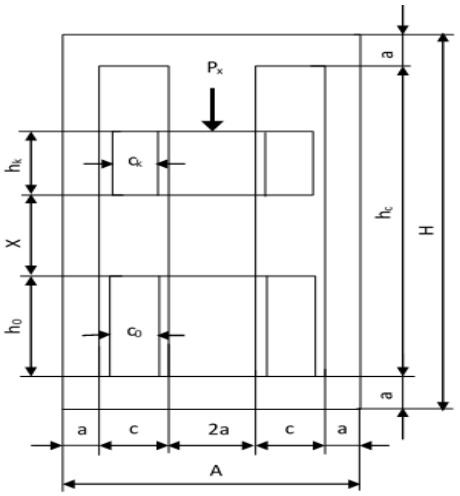


Figure 3. *Magnetic system of a force transducer.*

For the force transducer magnetic system (see Figure 3), we obtain:

$$A = 4a + 2c = 2c\left(1 + \frac{n_c}{n_a}\right) \tag{10, a}$$

$$B = b + 2c_0 + 4\Delta_k \approx b + 2c = c(2 + n_c) \tag{10, b}$$

$$H = x_{мак} + h_0 + h_k + 2a = x_{mak} + n_0 c_0 + \\ + n_{e2} c_k + 2a = x_{mak} + c\left(n_0 + n_{e2}\frac{c_k}{c} + 2\frac{n_c}{n_a}\right) \tag{10, c}$$

where

$$x = x_{mak};\ h_c = x_{mak} + h_0 + h_k;\ c_0 \approx c;$$
$$n_{e2} = h_k / c_k;\ 4\Delta_k << (b + 2c_0).$$

The perimeters of the cross-sectional areas of the magnetic circuit rods M_c and the excitation winding M_0:

$$M_c = 2(a + b) = 2a(1 + n_a) = 2N_c\sqrt{S_c}; \tag{11}$$

$$M_0 = 2(c_0 + h_0) = 2c_0(1 + n_0) = 2N_0\sqrt{S_0}, \tag{12}$$

where

$$N_c = \frac{1+n_a}{\sqrt{n_a}}; \quad N_0 = \frac{1+n_0}{\sqrt{n_0}} \ . \tag{13}$$

The cross-sectional areas of the excitation winding S_0 and the magnetic circuit rod S_s

$$S_0 = c_0 h_0 = n_0 c_0^2 \ ; \quad S_s = a \cdot b = n_a a^2 = c_0^2 \frac{n_c^2}{n_a} \ . \tag{14}$$

According to equation (6), the active and reactive magnetic resistances of the steel sections through M_0 can be determined as:

$$R_m = \frac{\rho_R}{S_c}(M_0 + a'); \quad X_m = \frac{\rho_x}{S_c}(M_0 + a') \ , \tag{15}$$

where

$$l_m = 2(2a + h_c + c) = M_0 + a'; \tag{15,a}$$

$$a' = 4a + 10\Delta_k + 2\Delta_c = \Delta + 4\sqrt{\frac{S_c}{n_a}}; \quad \Delta = 10\Delta_k + 2\Delta_c; \quad c = c_0 + 3\Delta_k; \quad h_c = h_0 + \Delta_c + 2\Delta_k;$$

$$a = \sqrt{\frac{S_c}{n_a}}; \quad c_0 = \sqrt{\frac{S_0}{n_0}} \ .$$

The thickness of the excitation winding frame wall is approximately $\Delta_k \approx (0.5 \div 1.0)10^{-3}$ m. The active resistances of the excitation winding and the short-circuited turn are determined by:

$$r_{10} = m_0' \left(M_c + 8\Delta_k + 4\sqrt{\frac{S_0}{n_0}} \right); \tag{16}$$

$$r_k = m_k' \left(2a_2 + 4c_k + 2\sqrt{n_a S_c} \right), \tag{17}$$

where

$$m_0' = \frac{\rho_1 w_1^2}{k_{30} S_0}; \quad m_k' = \frac{b_k^2 w_k^2 \rho_k}{k_k S_k}; \quad l_{cp} = l_a + 4c_0; \quad l_a = 2(a + b + 4\Delta_k) = M_c + 8\Delta_k; \tag{18}$$

$$l_k = 2(a_2 + b + 2c_k) = 2a_2 + 4c_k + 2\sqrt{n_a S_c}; \tag{19}$$

The cooling area of the excitation winding, determined by the coefficient n_0 and the perimeter M_c, is calculated as:

$$S_{ox} = n_0 c_0 (l_a + 8c_0) = n_0 c_0 (M_c + 8\Delta_k + 8c_0) \tag{20}$$

In reference [3], the main linear dimensions and parameters of the direct current magnetic system are expressed in terms of dimensionless quantities in order to determine the dimensions of the magnetic system while considering the principle of proportionality. Explanations for the symbols provided above are also given.

From equations (11)–(16), it is easy to conclude that the magnetic resistances (R_m and X_m) and the electrical resistance (r_{10}) of the magnetic system are directly proportional to the perimeters M_c and M_0. Since these dimensions depend respectively on the dimensionless coefficients N_c and N_0, or on the optimization criteria n_a and n_0, the minimization of active power losses P_{10} should be carried out based on expressions (7) and (11)–(16). These expressions take into account the temperatures of overheating τ_1 and τ_k (or the dependence of specific electrical resistances ρ_1 and ρ_2 respectively on τ_1 and τ_k) and the influences of the coefficients n_a, n_c, and n_0 on the fundamental parameters and dimensions of the magnetic system [10-18].

IV. Conclusions

Active power losses in the steel sections of the magnetic circuit, the excitation winding conductors, and the short-circuited turn depend on the parameters of the electrical, magnetic, and thermal circuits of alternating current electromagnetic devices. Initial analytical expressions based on the equations of these circuits were derived, allowing linear dimensions and key parameters of ACMS to be expressed in terms of dimensionless quantities. The minimum values of the active electrical resistances of the excitation winding and the short-circuited turn, as well as active power

losses, depend on the dimensional sizes of the cross-sectional areas of the excitation winding M_0 and the core.

In AC electromagnetic devices, the values of active power losses in the excitation winding conductors and the steel sections of the magnetic circuit vary significantly depending on the design of the ACMS. Through a comparative analysis of methods for determining active power losses, it was established that the method of equivalent electrical resistance allows for simultaneous consideration of all types of active power losses. However, in this case, the analytical expressions of parameters are quite cumbersome and do not allow for separate determination and comparison of all types of active power losses.

This drawback does not apply to the energy method and geometric optimization of the dimensions of the excitation winding and magnetic circuit. Optimization is carried out for electromagnetic devices with voltage and current windings, where the optimization criteria are the coefficients of proportionality principle n_a, n_c and n_0.

References

[1] G.V.Mamedova, G.S. Kerimzade, N.M.Piriyeva. "Electromagnetic calculation of tension devices for winding wires of small cross sections", International Journal on technical and Physical Problems of Engineering (IJTPE), Issue 53, Vol. 14, No.4, pp. 80-85, December 2022.

[2] G.V.Mamedova, G.S. Kerimzade, N.M.Piriyeva."Issues of electrical devices with levitation elements", International Journal on technical and Physical Problems of Engineering (IJTPE), Issue 56, Vol. 15, No.3, pp.120-125, September 2023.

[3] N.M.Piriyeva,G.S. Kerimzade. "Methods for increasing electromagnetic efficiency in induction levitator" , PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT.ISSN0033-2097, R.99 NR 10/2023. Warszawa. pp192-195.

[4] N.M.Piriyeva,G.S. Kerimzade." Electromagnetic efficiency in induction levitators and ways to improve it", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT.ISSN0033-2097,R.99.NR 06/2023. Warszawa. pp204-207.

[5] N.M.Piriyeva,G.S. Kerimzade." Mathematical model for the calculation of electrical devices based on induction levitators", International Journal on technical and Physical Problems of Engineering (IJTPE), Issue 55, Vol.15, No.2, pp.274-280, Yune 2023.

[6] Ya.R. Abdullaev, G.S. Kerimzade, "Design of Electric Devices with LE", Electrical Engineering, No. 5, pp. 16-22, 2015.

[7] G.S. Kerimzade, G.V. Mamedova, "Working Modes for Designing Electrical Apparatuses with Induction-mi with Levitation Elements", Bulletin Buildings, Vol. 17, No. 1, pp. 42-46, Baku, Azerbaijan, 2015.

[8] G.S. Kerimzade, G.V. Mamedova, "Analysis of the Parameters of Electric Devices with Levitational Elements", News of Universities Instrumentation, No. 12, Vol. 61, pp. 67-71, Sant Petersburg, Russia, 2018.

[9] G.S. Kerimzade, "Analytical Connections of the Parameters and sizes of the Precision Stabilizer of Alternating Current Using the Effect of Inductive Levitation", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 52, Vol. 14, No. 3, pp. 175-184, September 2022.

[10] G.S. Kerimzade, "Analysis of the Methodology for Calculation Current Stabilizer with Induction Levitation", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 53, Vol. 14, No. 4, pp. 170-174, December 2022.

[11]G.S. Kerimzade, "Analytical Expressions of the Relationship to the Calculation of the AC Stabilizer with Induction Levitation", International Journal on technical and Physical Problems of Engineering (IJTPE), Issue 54, Vol. 15, No. 1, pp. 135-139, March 2023.

[12] G.S. Kerimzade, "High Precision Current Stabilizers with Induction Levitation", Journal of Renewable Energy, Electrical, and Computer Engineering, e-ISSN: 2776-0049, Vol. 3, No. 1, pp.

26-31, Indonesia, March 2023.

[13] G.S. Kerimzade, "Choice of Expressions of the Method of Calculation of the Current Stabilizer on the Principle of Induction Levitation", News of Azerbaijan High Technical Educational Institutions, Vol. 4, pp. 71-76, 2023.

[14] G.S. Kerimzade, "Optimization of parameters of AC stabilizer on the principle of induction levitation", The 19th International Conference on "TPE". ICTPE-2023. Number 06. Code 01PES06. 31 October 2023, pp.32-36.

[15] G.S. Kerimzade." Calculation of parameters of control induction support", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT.ISSN0033-2097,R.99 NR 03/2024. Warszawa. pp192-195.

[16] G.S. Kerimzade "Calculation of parameters of control induction support", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT.ISSN0033-2097,R.99 NR 03/2024. Warszawa. pp192-195.

[17] G.V.Mamedova, G.S. Kerimzade."Design parameters for electromechanical devices with a levitation element", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT.ISSN0033-2097, 2024. Warszawa.(print in edidion)

[18] G.V.Mamedova Characteristic features of control methods in electromechanical devices Journal "Reliability: Theory & Applications". ISSN 1932-2321. Volume 19, № 4 (80).2024. pp.958-966.

[19] G.V.Mamedova, "Determining the leakage inductances in magnetic systems with movable screens.", The 20th International Conference on "TPE". ICTPE-2024. Number06. Code 01PES06. 31 October 2024, pp.32-36.

RT&A, No 2 (84)
Volume 20, June 2025

STEADY STATE ANALYSIS OF BULK ARRIVAL, FIXED BATCH SERVICE QUEUE WITH INSPECTION, REWORK AND MULTIPLE VACATIONS

[1]S. Karpagam, [2]R. Lokesh

•

[1,2]Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology,
Tamil Nadu, India.
[1]karpagammaths19@gmail.com, [2]rlokesh1020@gmail.com

Abstract

A non-Markovian bulk queueing system receives bulk arrivals through a single server that handles fixed-size batch services. The server conducts quality checks after each service to identify defective items, which are then sent to rework for quality control purposes. The server takes random-length vacations until new arrivals reach the threshold "K" before returning to service. Through the supplementary variable technique, the key performance metrics are computed together with steady-state governing equations. Using graphical representations and numerical tables, the model demonstration shows essential information about system performance through calculated data.

Keywords: bulk arrival, batch service, inspection, rework, performance measures

1. Introduction

Service industries, together with modern manufacturing, adopt batch service systems to maximize operational efficiency while maintaining high-quality control standards. A system that operates with fixed batch sizes enables better resource management by designing server operations to reduce operational expenses. Product quality maintenance reaches optimal levels through inspections that reveal defective items that require rework after each batch service. The systematic implementation of this approach results in lower manufacturing faults and waste output while achieving better production dependability rates. Batch processing systems that integrate inspection procedures and rework activities serve to maximize quality control for automated production environments as well as additive manufacturing and assembly lines. The batch service framework helps industries achieve both productivity growth and defect-free output, which results in better customer satisfaction and production cost decreases.

Lee et al. [16] investigated fixed-batch service queues with Poisson arrival patterns and vacation periods. The study used decomposition methods to generate queue length distributions both for basic vacation systems and multiple vacation systems. Ayyappan et al. [1] applied fixed batch service together with multiple vacations and gated service to study transportation and manufacturing processes. By combining generating functions and Rouche's theorem, they developed precise solutions for queue probabilities as well as performance metrics. Goswami et al. [8] established a novel method to study sojourn-time distributions in $M/G^a/1$ queues with batch service through their derivation of probability density function expressions using partial fractions and the residue theorem. Yuvarani and Vijayalakshmi [24] studied a multistage batch arrival, batch service queue system that incorporated multiple vacation policies and server breakdowns

and delayed repairs. They developed solutions for transient and steady-state conditions by using the supplementary variable technique.

Koyama et al. [12] conducted their study based on services from sharing economies, including ride-sharing and shared-taxi services. Their model presented a $M + M(K)/M/1$ queue system that allowed customers to decide between individual or combined services. Ghosh et al. [7] conducted a thorough analysis of a finite-buffer queueing system that incorporates correlated batch arrivals and batch-size-dependent bulk service. Van Ommeren et al. [23] studied bulk-service multi-server queues featuring heterogeneous server capacities that perfectly fit passenger transport and package delivery scenarios. Their decomposition-based method using renewal theory principles delivered a major breakthrough in analyzing multi-type server systems operating in complex service environments.

Chen et al. [4] investigated a modified Markovian queue system that used bulk arrivals and bulk services with general state-dependent control protocols. The study began with analyzing a bulk-arrival and bulk-service queue system in a resting state to reveal its relationship with broader queueing models. Kumar [13] investigated a single-server queue slideshow network that incorporated discouragement features during multiple vacation periods. D'Arienzo et al. [5] investigated a single-server retrial queue system that utilized MAP flow and PH service times and featured a finite capacity pool for customer accumulation and group service. The model implemented a threshold system that started the next group service only after reaching a predetermined customer count in the pool. Their main contribution involved demonstrating that group service durations can be shorter than the combined lengths of individual services because of batch processing advantages.

Pradhan et al. [20] investigated an infinite-buffer batch-service queue with Markovian arrivals and batch-size-dependent service times to derive performance metrics while examining arrival process correlation effects. Germs and Van Foreest [6] investigated state-dependent finite-buffer bulk queues through semi-regenerative analysis to derive queue length distributions. Kumar [14] examined a population model under catastrophic conditions, which he modeled through renewal growth processes along with Poisson catastrophes, utilizing geometric patterns to represent batch-sized population effects. Somasudaram et al. [21] investigated a batch arrival single service queue that included optional service and working breakdown. Chaudhry et al. [3] explored the queueing system with Poisson arrival patterns along with bulk service characteristics within limited capacity constraints.

Niranjan and Indhira [17] developed a comprehensive analysis of the batch service model. In more recent surveys, Panta et al. [19] and Upadhyaya [22] discussed additional types of models for vacation queueing systems, each with a different service policy. Niranjan et al. [18] studied a bulk queueing model with vacations and queue length and service request-based load adjustments in data center architecture to optimize cloud computing traffic distribution through power business intelligence and queueing theory. Haridass and Arumuganathan [9] developed a new batch service queuing framework that includes three essential features: multiple server vacations as well as setup durations followed by server-controlled re-service entry decisions.

Karpagam [11] developed a batch arrival bulk service queue model that incorporates unreliable servers and service resumption and rework for faulty items and multiple vacations with threshold policies for manufacturing systems. Ayyappan and Karpagam [2] developed a non-Markovian model that integrates batch arrivals with general bulk service and server breakdowns and multiple vacations and re-service and a standby server. Jeyakumar and Arumuganathan [10] developed a non-Markovian single-server queueing model featuring Poisson batch arrivals and general bulk service with controlled optional re-service based on probabilities and queue length-dependent vacation policies that trigger multiple vacations until reaching a threshold.

This paper focuses on analyzing the $M^X/G^K/1$ queueing system with inspection, rework, and multiple vacations where bulk arrivals adhere to a compound Poisson process denoted by M^X and fixed-size batch service durations follow a general distribution marked by G^K alongside a single server identified by 1.

The rest of the paper has been divided into eight sections. Section 2 gives practical application for the suggested model. Section 3 presents the model description, which includes Figure 1 schematic diagram and the model notation. Section 4 explores the steady-state queue size distribution, and Section 5 introduces the probability-generating function that describes queue-size distribution in steady state. Section 6 establishes key performance metrics. Numerical results together with graphical comparisons appear in Section 7. Finally, Section 8 concludes the paper.

2. Real-world application of the proposed model

The implementation of a fixed batch processing system in industrial 3D printing (Additive Manufacturing - AM) ensures printing automation through waiting for at least K parts to prepare while both optimizing machine performance and decreasing energy consumption. The system remains in vacation mode when the queue contains fewer than (K-1) parts. The AI driving inspection system ensures that the quality checks of the printing are met by detecting misalignment of the layer and other problems through CT analysis and optical scanning methods. The defective components need rework that includes laser remelting for metal parts and UV curing for resin prints together with heat treatment for thermoplastics to minimize reprinting requirements. The system processes final finishing and packaging that follows quality inspections of the high-quality outputs before they move to shipment. Manufacturing through this process benefits sectors including aerospace (turbine blades), medical (implants, prosthetics), automotive (3D-printed brackets), and fashion (customized footwear), while enhancing both production speed and waste reduction and providing accurately precise results.

3. Detailed model description

In this section, the proposed system is modeled using the supplementary variable technique, incorporating the remaining service time of the batch in service and the remaining vacation time of the server as supplementary variables at an arbitrary time. The schematic representation of the model is shown in Figure 1.

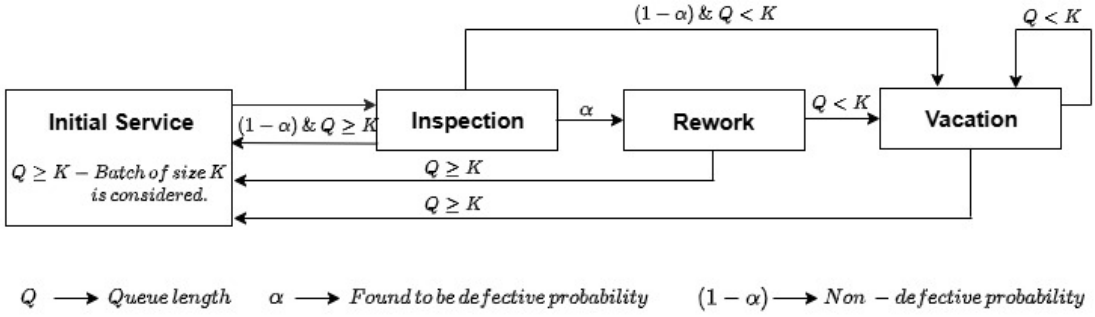


Figure 1: *Pictorial depiction of the model*

The server operates with batch services of a fixed size 'K'. The initial service begins only when the queue reaches 'K'. Upon completing a service, the batch undergoes an inspection. If defects are found, the batch moves to rework with probability 'α' If no defects are detected, the system checks the queue length:

- If the queue length meets or exceeds 'K', the next service starts immediately with probability '$(1-\alpha)$'.
- If the queue length is less than 'K', the server goes on vacation with probability '$(1-\alpha)$'.

After rework, a similar decision process follows:

- If the queue length reaches or exceeds 'K', the next service starts.
- If the queue length remains below 'K', the server enters a vacation.

This vacation cycle continues until the queue accumulates 'K' parts, at which point service starts.

3.1. Notations

The model parameters include the arrival rate λ, the group size random variable X with probability $g_i = \Pr(X = i)$, and its probability generating function $X(z)$. The mean batch size is denoted as $E(X) = X_1$. Let $\aleph_s(t)$ and $\aleph_q(t)$ denote the number of products in service and queue at time 't'. Let $\chi(t)$ denote the server state at time 't'. The basic notations are listed in Table 1:

Table 1: *The list of notations used in this paper*

Function Type	**Initial service time**	**Inspection time**	**Rework time**	**Vacation Time**
Cumulative Density Function (CDF)	$M_b(\hbar)$	$H_1(\hbar)$	$H_2(\hbar)$	$V(\hbar)$
Probability Density Function (PDF)	$m_b(\hbar)$	$h_1(\hbar)$	$h_2(\hbar)$	$v(\hbar)$
Laplace-Stieltjes Transform (LST)	$\tilde{M}_b(\mho)$	$\tilde{H}_1(\mho)$	$\tilde{H}_2(\mho)$	$\tilde{V}(\mho)$
Remaining Service Time	$M_b^0(t)$	$H_1^0(t)$	$H_2^0(t)$	$V^0(t)$

1. $\chi(t) = 1, 2, 3$ and 4 denote the server is on initial service, inspection, rework and vacation, respectively.
2. $\pounds(t) = l$, if the server is on l^{th} vacation.

We define the following state probabilities as follows :

$$F_{K,j}(\hbar, t)\Delta t = Pr\{\aleph_s(t) = K,\ \aleph_q(t) = j,\ \hbar \le {M_b}^0(t) < \hbar + \Delta t,\ \chi(t) = 1\},\ j \ge 0;$$
$$Q_n(\hbar, t)\Delta t = Pr\{\aleph_q(t) = n,\ \hbar \le H_1^0(t) < \hbar + \Delta t,\ \chi(t) = 2\}, \text{where } n \ge 0;$$
$$R_n(\hbar, t)\Delta t = Pr\{\aleph_q(t) = n,\ \hbar \le {H_2}^0(t) < \hbar + \Delta t,\ \chi(t) = 3\}, \text{where } n \ge 0;$$
$$W_{l,j}(\hbar, t)\Delta t = Pr\{\aleph_q(t) = j,\ \hbar \le V^0(t) < \hbar + \Delta t,\ \chi(t) = 4, \pounds(t) = l\}, \text{where } j \ge 0, l \ge 1.$$

4. Steady state queue size distribution

This section develops the probability-generating function (PGF) for the queue size at an arbitrary time point to determine essential performance metrics. The first step involves developing steady-state difference-differential equations. The Laplace - Stieltjes transform applied to these equations combined with the method from Lee et al. [15] leads to a PGF solution through algebraic steps.

Server is in initial service

$$-F'_{K,0}(\hbar) = -\lambda F_{K,0}(\hbar) + R_K(0)m_b(\hbar) + (1-\alpha)Q_K(0)m_b(\hbar) + \sum_{l=1}^{\infty} W_{n,K}(0)m_b(\hbar), \tag{1}$$

$$-F'_{K,j}(\hbar) = -\lambda F_{K,j}(\hbar) + R_{K+j}(0)m_b(\hbar) + (1-\alpha)Q_{K+j}(0)m_b(\hbar) + \sum_{l=1}^{\infty} W_{n,K+j}(0)m_b(\hbar) + \sum_{i=1}^{j} F_{K,j-i}(\hbar)\lambda g_i,\ j \geq 1, \tag{2}$$

Server is in inspection (3)

$$-Q_0(\hbar) = -\lambda Q_0(\hbar) + F_{K,0}(0)h_1(\hbar), \tag{4}$$

$$-Q_n(\hbar) = -\lambda Q_n(\hbar) + F_{K,n}(0)h_1(\hbar) + \sum_{i=1}^{n} Q_{n-i}(\hbar)\lambda g_i,\ \ n \geq 1, \tag{5}$$

Server is in rework (6)

$$-R_0(\hbar) = -\lambda R_0(\hbar) + \alpha Q_0(0)h_2(\hbar), \tag{7}$$

$$-R_n(\hbar) = -\lambda R_n(\hbar) + \alpha Q_n(0)h_2(\hbar) + \sum_{i=1}^{n} R_{n-i}(\hbar)\lambda g_i,\ \ n \geq 1. \tag{8}$$

Server is in vacation

$$-W'_{1,0}(\hbar) = -\lambda W_{1,0}(\hbar) + R_0(0)v(\hbar) + (1-\alpha)Q_0(0)v(\hbar), \tag{9}$$

$$-W'_{1,n}(\hbar) = -\lambda W_{1,n}(\hbar) + R_n(0)v(\hbar) + (1-\alpha)Q_n(0)v(\hbar) + \sum_{i=1}^{n} W_{1,n-i}(\hbar)\lambda g_i, \quad 1 \leq n \leq K-1, \tag{10}$$

$$-W'_{1,n}(\hbar) = -\lambda W_{1,n}(\hbar) + \sum_{i=1}^{n} W_{1,n-i}(\hbar)\lambda g_i,\ n \geq K \tag{11}$$

$$-W'_{l,0}(\hbar) = -\lambda W_{l,0}(\hbar) + W_{l-1,0}(0)v(\hbar),\ l \geq 2, \tag{12}$$

$$-W'_{l,n}(\hbar) = -\lambda W_{l,n}(\hbar) + W_{l-1,0}(0)v(\hbar) + \sum_{i=1}^{n} W_{l,n-i}(\hbar)\lambda g_i,\ 1 \leq n \leq K-1,\ l \geq 2, \tag{13}$$

$$-W'_{l,n}(\hbar) = -\lambda W_{l,n}(\hbar) + \sum_{i=1}^{n} W_{l,n-i}(\hbar)\lambda g_i,\ n \geq K,\ l \geq 2. \tag{14}$$

Using the LST on both sides of equations (1) to (14), we get

$$\upsilon\tilde{F}_{K,0}(\upsilon) - F_{K,0}(0) = \lambda\tilde{F}_{K,0}(\upsilon) - R_K(0)\tilde{M}_b(\upsilon) - (1-\alpha)Q_K(0)\tilde{M}_b(\upsilon) - \sum_{l=1}^{\infty} W_{l,K}(0)\tilde{M}_b(\upsilon), \tag{15}$$

$$\upsilon\tilde{F}_{K,j}(\upsilon) - F_{K,j}(0) = \lambda\tilde{F}_{K,j}(\upsilon) - R_{K+j}(0)\tilde{M}_b(\upsilon) - (1-\alpha)Q_{K+j}(0)\tilde{M}_b(\upsilon) - \sum_{l=1}^{\infty} W_{l,K+j}(0)\tilde{M}_b(\upsilon) - \sum_{i=1}^{j} \tilde{B}_{K,j-i}(\upsilon)\lambda g_i,\ j \geq 1, \tag{16}$$

$$\upsilon\tilde{Q}_0(\upsilon) - Q_0(0) = \lambda\tilde{Q}_0(\upsilon) - F_{K,0}(0)\tilde{H}_1(\upsilon), \tag{17}$$

$$\upsilon\tilde{Q}_n(\upsilon) - Q_n(0) = \lambda\tilde{Q}_n(\upsilon) - F_{K,n}(0)\tilde{H}_1(\upsilon) - \sum_{i=1}^{n} \tilde{Q}_{n-i}(\upsilon)\lambda g_i, n \geq 1, \tag{18}$$

$$\upsilon\tilde{R}_0(\upsilon) - R_0(0) = \lambda\tilde{R}_0(\upsilon) - \alpha Q_0(0)\tilde{H}_2(\upsilon), \tag{19}$$

$$\upsilon\tilde{R}_n(\upsilon) - R_n(0) = \lambda\tilde{R}_n(\upsilon) - \alpha Q_n(0)\tilde{H}_2(\upsilon) - \sum_{i=1}^{n} \tilde{R}_{n-i}(\upsilon)\lambda g_i, n \geq 1, \tag{20}$$

$$\upsilon\tilde{W}_{1,0}(\upsilon) - W_{1,0}(0) = \lambda\tilde{W}_{1,0}(\upsilon) - R_0(0)\tilde{V}(\upsilon) - (1-\alpha)Q_0(0)\tilde{V}(\upsilon) \tag{21}$$

$$\upsilon\tilde{W}_{1,n}(\upsilon) - W_{1,n}(0) = \lambda\tilde{W}_{1,n}(\upsilon) - R_n(0)\tilde{V}(\upsilon) - (1-\alpha)Q_n(0)\tilde{V}(\upsilon) - \sum_{i=1}^{n}\tilde{W}_{1,n-i}(\upsilon)\lambda g_i, \quad 1 \le n \le K-1, \tag{22}$$

$$\upsilon\tilde{W}_{1,n}(\upsilon) - W_{1,n}(0) = \lambda\tilde{W}_{1,n}(\upsilon) - \sum_{i=1}^{n}\tilde{W}_{1,n-i}(\upsilon)\lambda g_i,\ n \ge K, \tag{23}$$

$$\upsilon\tilde{W}_{l,0}(\upsilon) - W_{l,0}(0) = \lambda\tilde{W}_{l,0}(\upsilon) - W_{l-1,0}(0)\tilde{V}(\upsilon),\ l \ge 2, \tag{24}$$

$$\upsilon\tilde{W}_{l,0}(\upsilon) - W_{l,0}(0) = \lambda\tilde{W}_{l,0}(\upsilon) - W_{l-1,0}(0)\tilde{V}(\upsilon) - \sum_{i=1}^{n}\tilde{W}_{l,n-i}(\upsilon)\lambda g_i,\ 1 \le n \le K-1, l \ge 2, \tag{25}$$

$$\upsilon\tilde{W}_{l,n}(\upsilon) - W_{l,n}(0) = \lambda\tilde{W}_{l,n}(\upsilon) - \sum_{i=1}^{n}\tilde{W}_{l,n-k}(\upsilon)\lambda g_i,\ n \ge K,\ l \ge 2. \tag{26}$$

The following probability-generating functions (PGFs) are defined in order to derive the PGF for the queue size at an arbitrary time epoch:

$$\left.\begin{aligned}
&\tilde{F}_K(z,\upsilon) = \sum_{j=0}^{\infty}\tilde{F}_{K,j}(\upsilon)z^j, \quad F_K(z,0) = \sum_{j=0}^{\infty}F_{K,j}(0)z^j,\\
&\tilde{Q}(z,\upsilon) = \sum_{j=0}^{\infty}\tilde{Q}_j(\upsilon)z^j, \quad Q(z,0) = \sum_{j=0}^{\infty}Q_j(0)z^j,\\
&\tilde{R}(z,\upsilon) = \sum_{j=0}^{\infty}\tilde{R}_j(\upsilon)z^j, \quad R(z,0) = \sum_{j=0}^{\infty}R_j(0)z^j,\\
&\tilde{W}_l(z,\upsilon) = \sum_{j=0}^{\infty}\tilde{W}_{l,j}(\upsilon)z^j, \quad W_l(z,0) = \sum_{j=0}^{\infty}W_{l,j}(0)z^j, \quad l \ge 1.
\end{aligned}\right\} \tag{27}$$

By multiplying equations (15) to (26) with suitable powers of z^j and summing over $'j'$ ($j = 0$ to ∞), then by using (27) and after some mathematical manipulations, we obtain

$$(\upsilon - u(z))\tilde{W}_1(z,\upsilon) = \big(\tilde{V}(u(z)) - \tilde{V}(\upsilon)\big)\sum_{n=0}^{K-1}\Big[R_n(0) + (1-\alpha)Q_n(0)\Big] \tag{28}$$

$$(\upsilon - u(z))\tilde{W}_n(z,\upsilon) = \big(\tilde{V}(u(z)) - \tilde{V}(\upsilon)\big)\sum_{n=0}^{K-1}W_{l-1,n}(0),\ l \ge 2, \tag{29}$$

$$(\upsilon - u(z))\tilde{Q}(z,\upsilon) = \big(\tilde{H}_1(u(z)) - \tilde{H}_1(\upsilon)\big)F_K(z,0), \tag{30}$$

$$(\upsilon - u(z))\tilde{R}(z,\upsilon) = \big(\tilde{H}_2(u(z)) - \tilde{H}_2(\upsilon)\big)\alpha Q(z,0)\tilde{H}_3(u(z)), \tag{31}$$

$$y(z)(\upsilon - u(z))\tilde{F}_K(z,\upsilon) = \big(\tilde{M}_b(u(z)) - \tilde{M}_b(\upsilon)\big)\Bigg[R(z,0) - \sum_{n=0}^{K-1}R_n(0)z^n + (1-\alpha)Q(z,0) - (1-\alpha)\sum_{n=0}^{K-1}Q_n(0)z^n + \sum_{l=1}^{\infty}\Big(W_l(z,0) - \sum_{n=0}^{K-1}W_{l,n}(0)z^n\Big)\Bigg] \tag{32}$$

5. Probability generating function of the queue size

Let P(z) be the PGF of the queue size at an arbitrary time epoch. Then,

$$P(z) = \tilde{F}_K(z,0) + \tilde{R}(z,0) + \tilde{Q}(z,0) + \sum_{l=1}^{\infty}\tilde{W}_l(z,0). \tag{33}$$

Substituting $\mho = 0$ in equations (28) - (32), then (33) becomes

$$P(z) = \frac{\sum_{n=0}^{K-1} c_n z^n (\tilde{V}(u(z)) - 1) a(z)}{u(z) y(z)} \tag{34}$$

where

$$r_n = R_n(0),\ q_n = Q_n(0), v_n = \sum_{l=1}^{\infty} W_{l,n}(0),\ c_n = r_n + (1-\alpha) q_n + v_n,\ u(z) = \lambda - \lambda X(z),$$
$$y(z) = z^K - (1-\alpha)\tilde{M}_b(u(z))\tilde{H}_1(u(z)) - \alpha \tilde{M}_b(u(z))\tilde{H}_1(u(z))\tilde{H}_2(u(z)) \text{ and}$$
$$a(z) = z^K + 1 - 2(1-\alpha)\tilde{M}_b(u(z))\tilde{H}_1(u(z)) - 2\alpha \tilde{M}_b(u(z))\tilde{H}_1(u(z))\tilde{H}_2(u(z)).$$

Remark. The PGF must obtain $P(1) = 1$. This criteria is satisfied by the result obtained by equating a term to 1 and applying the L'Hopital's rule in (34). Then $\rho = (M_{b1} + H_{11} + \alpha H_{21}/K) < 1$ is the condition for the steady state for the model under consideration existence.

5.1. Computational aspects

Rouche's theorem on complex variables can be utilized to determine these constants. $y(z)$ has one on the unit circle $|z| = 1$ and $(K-1)$ zeroes within, according to Rouche's theorem. $P(z)$ is analytic both on and within the unit circle; therefore, at these points the numerator must vanish, which gives 'K' equations in 'K' unknowns. These equations can be solved using any suitable numerical technique.

6. Performance measures

Some useful results of our model are listed below.

6.1. Expected queue length

Mean number of customers in the queue is the equation (34) is obtained by using

$$E(Q) = \left(\frac{dP(z)}{dz}\right)_{z=1}$$
$$E(Q) = \frac{6\sum_{n=0}^{K-1} c_n \left[W_1(2na_1 + a_2) + W_2 a_1\right](-\lambda X1) y_1 - 6W_1 \sum_{n=0}^{K-1} c_n a_1 \left[(-\lambda X1) y_2 - (\lambda X2) y_1\right]}{3\left(3\left[(-\lambda X1) y_2 - (\lambda X2) y_1\right]\right)^2} \tag{35}$$

where

$$y_1 = K - M_{b1} - H_{11} - \alpha H_{21}, M_{b1} = \lambda X_1 E(M_b), M_{b2} = \lambda X_2 E(M_b) + (\lambda X_1)^2 E({M_b}^2),$$
$$y_2 = K(K-1) - M_{b2} - H_{12} - 2M_{b1}H_{11} - \alpha(2H_{21}H_{11} + 2H_{21}M_{b1} + H_{21}),$$
$$a_1 = K - 2(M_{b1} + H_{11} + \alpha H_{21}), H_{11} = \lambda X_1 E(H_1), H_{21} = \lambda X_1 E(H_2)),$$
$$a_2 = K(K-1) - 2[M_{b2} + H_{12} + 2M_{b1}H_{11} + \alpha(2H_{21}H_{11} + 2H_{21}M_{b1} + H_{21})],$$
$$H_{12} = \lambda X_2 E(H_1) + (\lambda X_1)^2 E({H_1}^2), H_{22} = \lambda X_2 E(H_2) + (\lambda X_1)^2 E({H_2}^2),$$
$$W_1 = \lambda X_1 E(V), W_2 = \lambda X_2 E(V) + (\lambda X_1)^2 E(V^2).$$

6.2. Expected waiting time

Using Little's formula, we get expected waiting time at arbitrary epoch is

$$E(W) = \frac{E(Q)}{\lambda X_1} \tag{36}$$

7. Numerical Illustration

This section presents a numerical illustration of the proposed queueing model by analyzing parameter variations using MATLAB software. The computed results are displayed in both tabular and graphical formats, assuming that the batch size distribution of arrivals follows a geometric distribution with a mean of 2. For the numerical demonstration, we assume that the initial service time, inspection, rework, and vacation times follow general distributions with parameters μ_s, μ_i, μ_r, and γ, respectively. Numerical values satisfying the stability criteria are selected randomly.

λ	**Exponential**		**Erlang-2**		**Hyperexponential**	
	E(Q)	**E(W)**	**E(Q)**	**E(W)**	**E(Q)**	**E(W)**
5.00	28.0325	2.8032	27.8501	2.7850	25.4377	2.5438
5.25	31.8792	3.0361	31.8001	3.0286	29.3675	2.7969
5.50	37.9381	3.4489	38.1386	3.4671	35.7476	3.2498
5.75	48.4498	4.2130	49.5277	4.3068	47.6078	4.1398
6.00	70.4681	5.8723	75.6091	6.3008	76.9288	6.4107

Table 2: *Performance metrics for different service time distributions*

The expected queue length $E(Q)$ and expected waiting time $E(W)$ are computed for various values of λ and presented in Table 2. The default parameter values for numerical analysis are:

- Exponential distribution: $\mu_s = 6$, $\mu_i = 5$, $\mu_r = 4$, $\alpha = 0.7$, $K = 7$, $\gamma = 3$.
- Erlang-2 distribution: $\mu_s = 9$, $\mu_i = 7$, $\mu_r = 5$, $\alpha = 0.1$, $K = 7$, $\gamma = 3$.
- Hyperexponential distribution: $\mu_s = 9$, $\mu_i = 7$, $\mu_r = 5$, $\alpha = 0.1$, $K = 7$, $c = 0.9$, $d = 0.7$, $e = 0.6$, $p = 6$, $q = 4$, $r = 4$, $f_1 = 0.3$, $s = 2$, $\gamma = 3$.

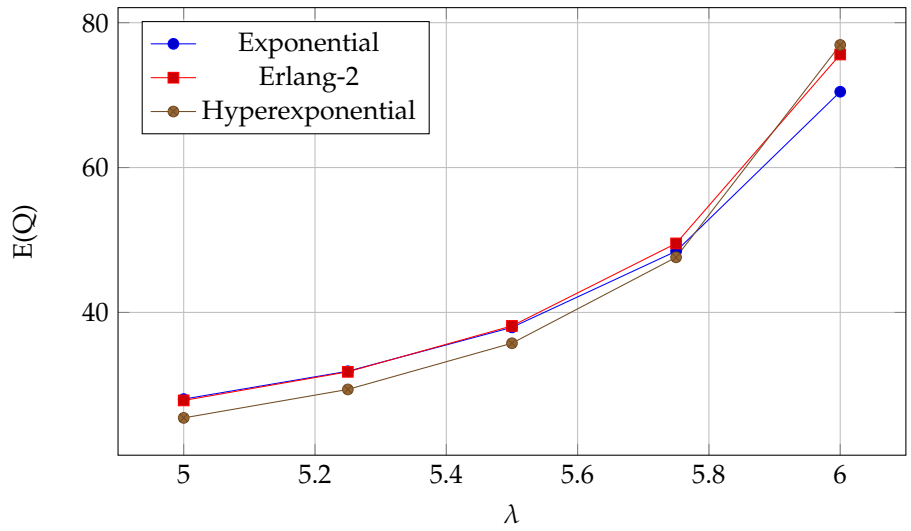


Figure 2: *Mean queue length E(Q) vs. arrival rate λ for different service time distributions.*

Figure 2 shows how system workload and queue length (E(Q)) react when production demand (λ) rises using various service time distributions. The system workload E(Q) experiences a substantial increase when demand rises, which suggests potential manufacturing, inspection, and rework stage congestion. Increased service time variability produces longer item waiting lines, which creates processing inefficiencies as well as delays in the fulfillment of orders.

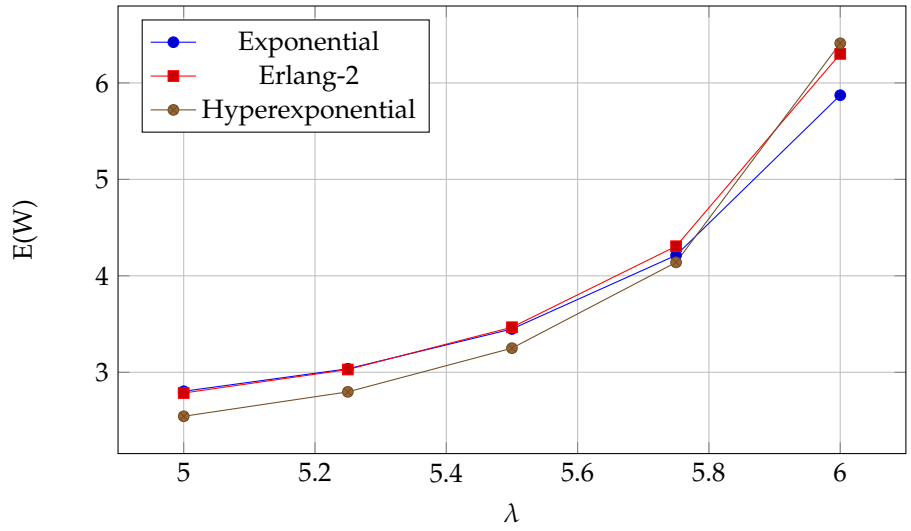


Figure 3: *Mean waiting time E(W) vs. arrival rate λ for different service time distributions.*

Figure 3 illustrates how raising demand levels prolongs waiting time (E(W)). Raising the λ value produces increased waiting durations mostly in systems whose service times show high variation. Long waiting times affect system productivity through disrupted scheduling operations. Optimizing scheduling combined with predictive maintenance systems and adaptable batch processing enables better control of operational difficulties to improve performance quality.

μ_s	$\lambda = 5$		$\lambda = 6$	
	E(Q)	**E(W)**	**E(Q)**	**E(W)**
9	24.0400	2.4040	35.3044	2.9420
10	11.0944	1.1094	22.7451	1.8954
11	10.5894	1.0589	18.0331	1.5028
12	10.3115	1.0312	15.5947	1.2996
13	10.1810	1.0181	14.1277	1.1773

Table 3: *E(Q) and E(W) for different service rates (μ_s)*

Table 3 employs Erlang-2 distributions for μ_s initial service rate along with $\mu_i = 7$ inspection rate and $\mu_r = 5$ rework rates while the vacation rate follows an exponential distribution with parameters $\gamma = 3$, $K = 7$, and $\alpha = 0.1$. The queue performance indicators E(Q) and E(W) appear in Table 3, where different service rate assumptions demonstrate system dynamic variances.

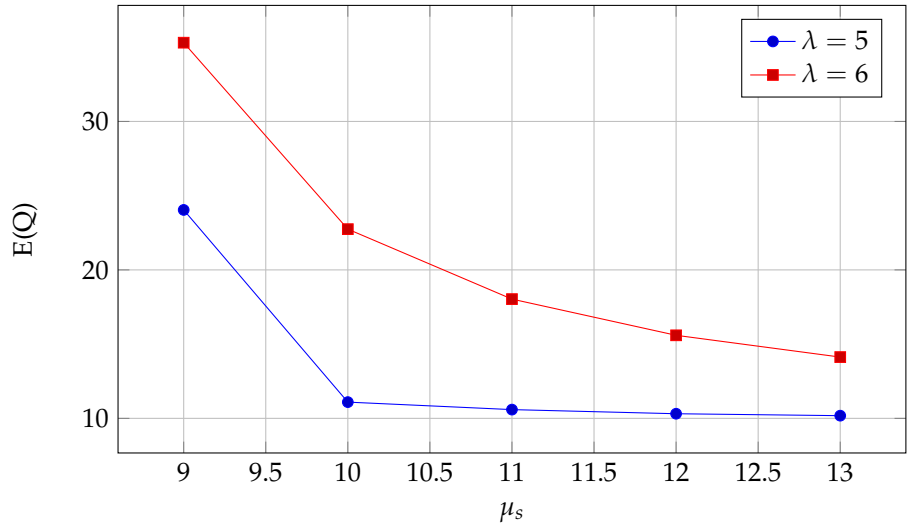


Figure 4: *E(Q) vs. Service Rate (μ_s)*

Figure 4 depicts the relationship between E(Q) and μ_s while λ varies. The E(Q) decreases when μ_s increases because this action minimizes queue congestion in all stages. E(Q) decreases through decreased queues, which improves part and order processing efficiency, especially when demand is high ($\lambda = 6$).

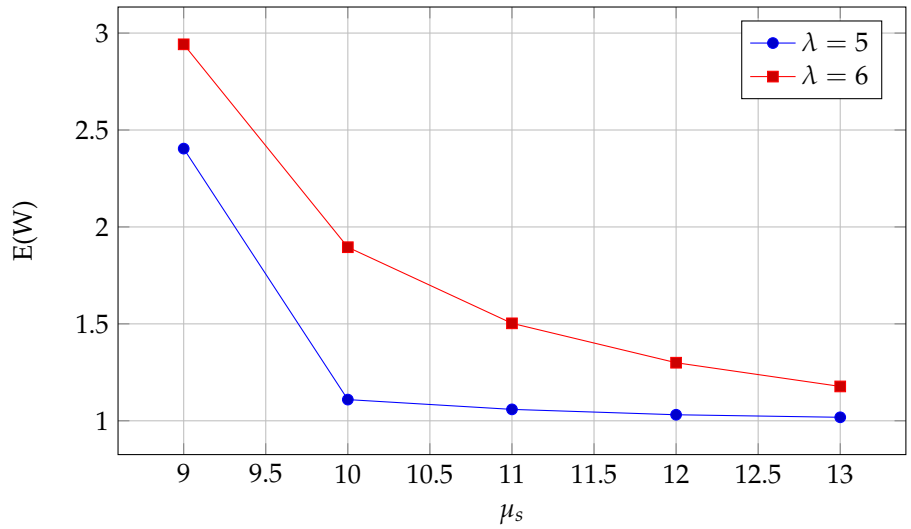


Figure 5: *E(W) vs. Service Rate (μ_s)*

Figure 5 shows how processing efficiency (μ_s) relates to the mean waiting time (E(W)) for orders. The mean waiting time decreases when processing efficiency (μ_s) rises because faster processing times generate fewer delays in order completion. Service rates play a vital role in reducing delays because higher demand levels ($\lambda = 6$) produce more pronounced effects.

Table 4 shows system parameters which contain the defect rate probability (α), service rate (μ_s), expected queue length ($E(Q)$), and expected waiting time ($E(W)$). The system adopts a fixed batch size $K = 7$ with an arrival rate $\lambda = 5$. Initial service time follows an Erlang-2 distribution with mean service rate at μ_s and inspection and rework services use Erlang-2 distributions at $\mu_i = 7$ and $\mu_r = 6$ respectively. The vacation period follows an exponential distribution with rate $f = 3$.

α	$\mu_s = 10$		$\mu_s = 11$		$\mu_s = 12$	
	E(Q)	**E(W)**	**E(Q)**	**E(W)**	**E(Q)**	**E(W)**
0.1	21.4589	2.1459	20.5613	2.0561	20.1177	2.0118
0.2	24.5557	2.4556	22.8812	2.2881	21.8684	2.1868
0.3	29.7274	2.9727	26.7080	2.6708	24.8457	2.4846
0.4	39.3315	3.9331	33.3550	3.3355	29.9064	2.9906
0.5	62.6701	6.2670	46.9792	4.6979	39.3757	3.9376
0.6	195.4627	19.5463	88.9730	8.8973	62.4679	6.2468

Table 4: *Queueing metrics for different service rates (μ_s) and defect rates (α)*

Figure 6 demonstrates how defect probability rates (α) affect expected queue length ($E(Q)$) when processing rates ($\mu_s = 10, 11, 12$) are used. An increase in the α value produces substantial growth in $E(Q)$ especially when processing rates remain low. The system experiences congestion because defect rates which increase cause more rework combined with longer inspection times. Processing rates that run efficiently work to reduce the accumulation of work in queue thereby sustaining a smooth production flow while minimizing batch manufacturing bottlenecks.

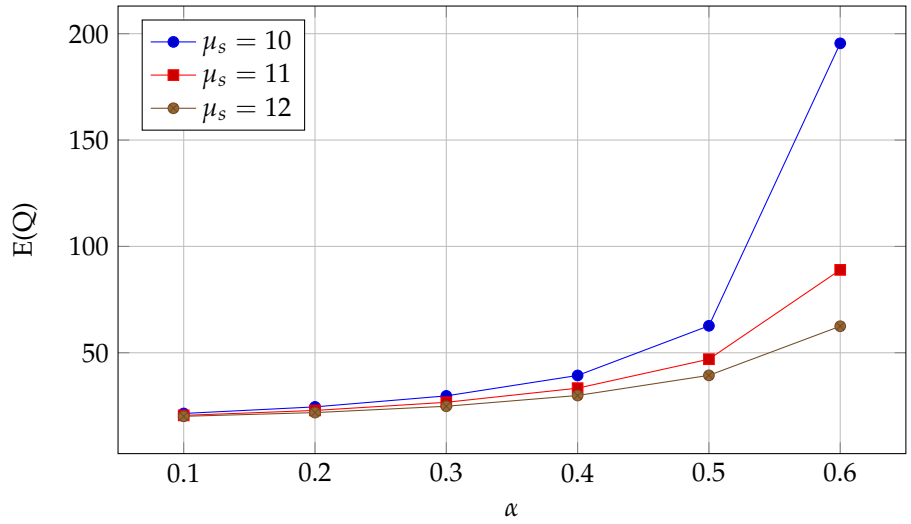


Figure 6: *E(Q) vs. Defect probability rate* (α)

Figure 7 displays how defect probabilities influence the expected waiting time ($E(W)$). The processing requirements along with rework increase because of the value increment in α. The combination of reduced service rates with μ_s intensifies these delays because it indicates the need to maintain high processing efficiency while implementing effective quality control measures. Higher quality inspection management as well as better rework procedures enhance both production speed and system operational effectiveness.

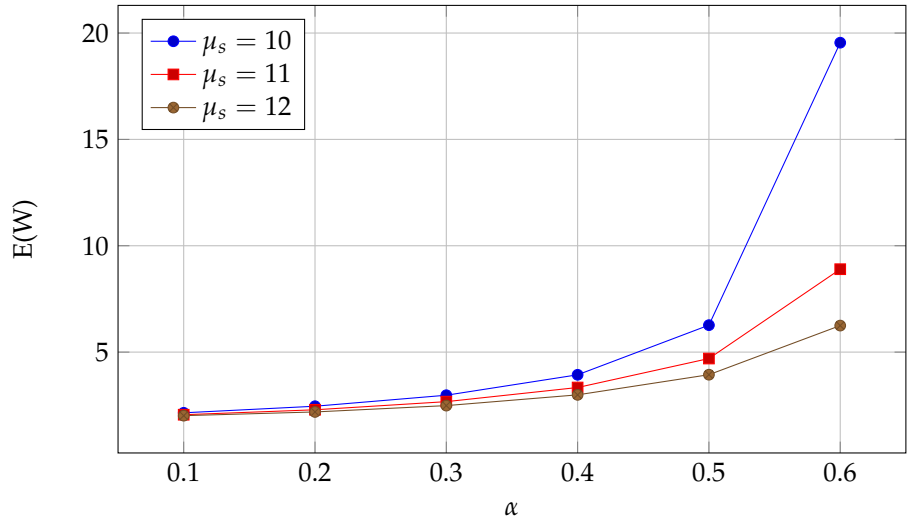


Figure 7: *E(W) vs. Defect probability rate* (α)

8. Conclusion and further scope for extension of work

The proposed model investigates a single-server non-Markovian bulk queueing system that employs bulk arrivals accompanied by fixed-size batch services and includes inspection capabilities and vacation features and rework operations. The supplementary variable technique enables steady-state analysis of the system to derive essential performance metrics. The combination of batch service policies with post-service inspection and rework operations leads to substantial improvements in both production efficiency and product quality. Numerical calculations and graphical presentation tools help reveal system behavior, allowing users to understand how batch sizes combine with arrival rates and service rates and how rework policies affect system results. Future research should expand this model through investigations of multi-server systems and dynamic batch size adaptation and priority-based rework strategies to achieve maximum efficiency. The implementation of machine learning techniques alongside adaptive vacation policies would boost the practical use of this system. The integration of cost-based optimization methods shows promise to reduce operational expenses while achieving maximum throughput and quality levels.

References

[1] Ayyappan, G., Devipriya, G., and Subramanian, A. M. G. (2014). Analysis of single server fixed batch service queueing system under multiple vacations with gated service. *International Journal of Computer Applications*, 89(5), 15–19.

[2] Ayyappan, G., and Karpagam, S. (2018). An $M^{[X]}/G(a,b)/1$ queueing system with breakdown and repair, stand-by server, multiple vacation and control policy on request for re-service. *Mathematics*, 6(6), 101.

[3] Chaudhry, M., Datta Banik, A., Barik, S., and Goswami, V. (2023). A novel computational procedure for the waiting-time distribution (in the queue) for bulk-service finite-buffer queues with Poisson input. *Mathematics*, 11(5), 1142.

[4] Chen, A., Wu, X., and Zhang, J. (2020). Markovian bulk-arrival and bulk-service queues with general state-dependent control. *Queueing Systems*, 95(3), 331–378.

[5] D'Arienzo, M. P., Dudin, A. N., Dudin, S. A., and Manzo, R. (2020). Analysis of a retrial queue with group service of impatient customers. *Journal of Ambient Intelligence and Humanized Computing*, 11:2591–2599.

[6] Germs, R., and Van Foreest, N. (2013). Analysis of finite-buffer state-dependent bulk queues. *OR Spectrum*, 35(3), 563–583.

[7] Ghosh, S., Banik, A. D., Walraevens, J., and Bruneel, H. (2022). A detailed note on the finite-buffer queueing system with correlated batch-arrivals and batch-size-/phase-dependent bulk-service. *4OR*, 20(2), 241–272.

[8] Goswami, V., Chaudhry, M., and Banik, A. D. (2022). Sojourn-time Distribution for $M/G^a/1$ Queue with Batch Service of Fixed Size-Revisited. *Methodology and Computing in Applied Probability*, 24(4), 2897–2912.

[9] Haridass, M., and Arumuganathan, R. (2012). A batch service queueing system with multiple vacations, setup time and server's choice of admitting reservice. *International Journal of Operational Research*, 14(2), 156–186.

[10] Jeyakumar, S., and Arumuganathan, R. (2011). A non-Markovian bulk queue with multiple vacations and control policy on request for re-service. *Quality Technology & Quantitative Management*, 8(3), 253–269.

[11] Karpagam, S. (2022). Analysis of bulk service queuing system with rework, unreliable server, resuming service and two kinds of multiple vacation. *Yugoslav Journal of Operations Research*, 33(1), 17–39.

[12] Koyama, Y., Nakamura, A., and Phung-Duc, T. (2024). Sojourn Time Analysis of a Single-Server Queue with Single-and Batch-Service Customers. *Mathematics*, 12(18), 2820.

[13] Kumar, A. (2023). Single server multiple vacation queue with discouragement solved by confluent hypergeometric function. *Journal of Ambient Intelligence and Humanized Computing*, 14(5), 6411–6422.

[14] Kumar, N. (2024). Analysis of a renewal arrival process subject to geometric catastrophe with random batch killing. *RAIRO-Operations Research*, 58(1), 933–946.

[15] Lee, H. W, Lee, S. S, Park, J. O., and Chae, K. C. (1994). Analysis of the $M^x/G/1$ queue by N-policy and multiple vacations. *Journal of Applied Probability*, 31(2), 476–496.

[16] Lee, H. W., Lee, S. S., and Chae, K. C. (1996). A fixed-size batch service queue with vacations. *International Journal of Stochastic Analysis*, 9(2), 205–219.

[17] Niranjan, S. P., and Indhira, K. (2016). A review on classical bulk arrival and batch service queueing models. *International Journal of Pure and Applied Mathematics*, 106(8), 45–51.

[18] Niranjan, S. P., Latha, S. D., Vlase, S., and Scutaru, M. L. (2024). Analysis of Bulk Queueing Model with Load Balancing and Vacation. *Axioms*, 14(1), 18.

[19] Panta, A. P., Ghimire, R. P., Panthi, D., and Pant, S. R. (2021). A review of vacation queueing models in different frameworks. *Ann. Pure Appl. Math*, 24, 99–121.

[20] Pradhan, S., and Gupta, U. C. (2019). Analysis of an infinite-buffer batch-size-dependent service queue with Markovian arrival process. *Annals of Operations Research*, 277, 161–196.

[21] Somasundaram, B., Karpagam, S., Lokesh, R., and Mary, A. K. S. (2023). An $M^{[X]}/G/1$ queue with optional service and working breakdown. *Ural Mathematical Journal*, 9(1), 162–175.

[22] Upadhyaya, S. (2016). Queueing systems with vacation: an overview. *International Journal of Mathematics in Operational Research*, 9(2), 167–213.

[23] Van Ommeren, J. K., Baer, N., Mishra, N., and Roy, D. (2020). Batch service systems with heterogeneous servers. *Queueing Systems*, 95(3), 251–269.

[24] Yuvarani, C., and Vijayalakshmi, C. (2020). Analysis of multistage $M^{[X]}/G^K/1$ queue with different server's interruptions and its application in energy consumption. *International Journal of Ambient Energy*, 41(6), 691–698.

ANALYSIS OF THE PERFORMANCE ON SINGLE SERVER BATCH SERVICE MULTIPLE WORKING VACATION QUEUING MODEL WITH COMPULSORY AND EXTENDED REPAIR

Lidiya P[1*], K Julia Rose Mary[2]

•

[1,2]Department Of Mathematics, Nirmala College For Women, Coimbatore
[*]lidiyap481@gmail.com
juliakulandaisamy@gmail.com

Abstract

In this paper, the concept of single-server multiple-working vacations queuing model with compulsory and extended repair is analyzed. In this model, customers arrive at a service facility and form a queue to be served by a single server. The arrival follows the Poisson distribution, and the service follows the exponential distribution. Batches of customers are served under the General Bulk Service Rule. In GBSR, rather than the individual customer arriving in a queue one by one, the customers arrive in groups or batches. Thus, each batch of services contains a minimum of $'a'$ units and a maximum of $'b'$ units of customers. In this study, two types of repairs, that is, compulsory and extended repair, are considered. The steady-state equation, the performance of measures for the system, and particular cases of described model are derived. Finally, in the form of tables and graphs, numerical results have been analyzed.

Keywords: Repair,Batch Service, performance measures, Multiple Working vacation (MWV), queue length

1. Introduction

Erlang developed queuing theory while working for the Telephone Company to analyze the behavior of telephone traffic and optimize the capacity of telephone exchanges. His work laid the foundation for the study of waiting lines and has since been widely applied in various fields to improve system performance and efficiency. The main objective of queuing theory is to understand and optimize the performance of systems that involve waiting lines. By studying factors such as arrival rates, service rates, queue lengths, and waiting times, queuing theory provides insights into how to improve efficiency and reduce waiting times.

In queuing theory, a vacation queuing model is a type of queuing system where the server may take breaks or go on vacation, when service is not available. This type of model is often used in scenarios where service providers have scheduled breaks, such as in customer service centers, healthcare facilities, or manufacturing processes. Analyzing and optimizing vacation queuing models involves factors of the duration and frequency of vacations, the impact on service during vacation periods, and strategies to minimize the effects of downtime on customer satisfaction. The concept of "multiple working vacations" refers to a scenario where the server in a queuing system takes several breaks or vacations during their work.

The concept of the GBS rule was indeed introduced by Neuts. The GBS rule, which he introduced, is used to analyze queuing systems where customers arrive in batches and are served

as a single entity with a fixed service time for the entire batch. In a batch arrival process, rather than individual customers arriving at a queue one by one, customers arrive in groups or batches.

The breakdown refers to a situation where a service facility or a server becomes unavailable or inoperable, leading to a disruption in the service process. Breakdowns can occur due to various reasons, such as equipment failure, system malfunction, or scheduled maintenance. When a breakdown occurs, it can lead to increase waiting times for entities in the queue and can impact the overall performance of the system. Understanding the behavior of a queuing system under breakdown conditions help to improve system performance.

The repair process refers to the activities involved in restoring a failed or broken component back to a functional state. When a breakdown occurs in a system, the repair process is triggered to address the issue and restore the system to normal operation. Planning and scheduling the repair activities based on factors such as repair priorities, resource availability, and system constraints. The compulsory and extended repair model is a concept within queuing theory that is used to analyze and optimize repair processes in systems where maintenance and repair activities are essential. This model considers two types of repair scenarios: compulsory repairs, which are necessary to restore a failed component to working condition, and extended repairs, which involve more complex or time-consuming maintenance tasks. Overall, the compulsory and extended repair model provides a framework for understanding and managing repair processes in various systems and industries.

This paper analyzes the queuing system that combines multiple working vacations, batch service, a single server, compulsory and extended repair. For this model, we obtained steady state equations, measures of performances, and analyzed the particular cases.

2. Review of Literature

A single server feedback retrial queuing system with multiple working vacations, balking customer, breakdown and vacation interruption studied by [24]. Using the supplementary variable method they obtained steady state probability generating function for the system size. The queue caused by a lack of servers, (M / G / I) with single working vacation and vacation interruption presented by [14]. The results of that model analysis show that the server is not busy or does not need to add a server. The multiple working vacations queuing-inventory model with compliment analyzed by [19]. Also studied steady state, expected total cost rate and measures of system performance. A discrete-time queuing system with multiple working vacations and reneging analyzed by [22]. For that steady-state probabilities are obtained recursively from the difference equations that represent the model. In addition, they obtain some other performance measures and a cost model is formulated to determine the optimal service rate during working vacation. $M^{[x]}/G(a,b)/1$ queuing system with additional optional service, multiple vacation and setup time analyzed by [2]. Using supplementary variable technique, the probability generating function of the performance measures is developed. The non-Markovian single server batch arrival queuing system of compulsory three stages of services with fourth optional stage service, service interruptions and deterministic server vacations developed by [29]. A non-Markovian batch arrival general bulk service single server queuing system with server breakdown and repair, single vacation and stand-by server discussed by [4]. The probability generating function of queue size at an arbitrary time and some performance measures of the system are derived.

The backup server, breakdown, repair, and vacation queuing system derived by [7]. They also prove decomposition results for the rate matrix and the mean number in the system under some special cases. A renewal input single server batch service queue with impatient customers and multiple working vacations presented by [11]. Applying the embedded Markov chain approach and the displacement operator method, they find the probability distribution of the queue-length. $M^{[X]}/G(a,b)/1$ queuing model with multiple vacation controllable arrival during multiple vacation and two phases of repair with delay analyzed by [3]. Using supplementary variable technique, the probability generating function of the queue size. the modeling the steady-state behavior of a retrial queuing system with phases of service analyzed by [1] . The first

phase of service is essential for all customers, any customer has three options after the completion of the phase, after each phase, the probabilistic failure, delay, repair and vacation are considered. A single server queuing model with setup process, close-down process, and standby server are analyzed by using the Matrix-Analytic method helped to obtain steady-state probability are investigated by [5]. $M^X/G/1$ retrial queue with impatient customers, modified vacation policy and Bernoulli feedback analyzed by[28] . Supplementary variable technique and probability generating function method is used to derive the system size distribution. An algorithm for determining optimal parameters of the queuing system that allows server vacation proposed by [18]. Also they first propose an efficient algorithm to determine steady-state probabilities of the system, and second, a non-linear integer programming problem is formulated to determine optimal system parameters.

M/M/1 retrial queuing model with a waiting server subject to breakdown and repair under working vacation, vacation interruption presented by [12] . Steady-state solutions are obtained using the probability generating function technique. A mathematical queuing model for multi-phase essential repair (MPER) systems with unreliable servers developed by [26]. Moreover, the service of the customer is interjected due to unavailability of the server. Also time-dependent probabilities are calculated by Runge-Kutta. The M/M/2 machine repair problem with operating machines and two removable servers are investigated described by [17]. The failed machines are repaired under the multiple working vacations and triadic policy. A retrial queuing system with working vacation, interruption, feedback, and setup time with the perfect repair studied by [13]. Also to save power, the concept of setup time is included in that model. The Bernoullis vacation on each service includes additional service after breakdowns and repairs by the method of supplementary variable techniques analyzed by [25]. To implement these concepts we build a mathematical model using the complementary variable method and obtained a function that gives the probabilities for the number of customers in the system when the server is free or on repair or on vacation.

Open network queuing system problems with breakdowns and its alternative repair strategies proposed by [9]. Queuing model with working vacation, balking, unreliable server with multi-phase repair discussed by [16]. The steady state queue size distribution and mean operational system metrics have been obtained by using probability generating function and supplementary variable methodology. Also cost optimization using particle swarm optimization technique have been drawn to explore the practical utility of the model. An unreliable bulk arrival single server queue rendering two-heterogeneous optional repeated service with delayed repair, under Bernoulli Vacation Schedule and N-policy analyzed by [6].

The queuing system with phase type vacations, breakdowns, and repairs developed by [8]. In the phase type they consider, first the service is restored to normalcy immediately after the server becomes idle or a fixed number of services is offered. In the second one, the restoration takes a random time that is modeled using a phase type distribution. The single server queuing model, working vacation, working breakdown and two-phase repair examined by [27]. They considered the repair process in two phases to recover the server from breakdown to working state with improved service rate phase-wise. Matrix geometric solution method is used to obtain stationary probability vectors. Single server multiple working vacation queuing system under breakdown, balking and encouraged arrival, and heterogeneous arrival analyzed by [20, 23, 21] . Batch arrival queuing system with a random vacation policy and optional second service. They consider the servers are unreliable and can be repaired immediately demonstrated by [15]. The queuing model for machine repair problem with two failure modes and working vacation analyzed by [30]. The steady-state system-size distribution is derived via the matrix analysis method and matrix decomposition. The repair personnel of the enterprise using the theory of queuing was developed by [10]. The proposed methodology can be applied to evaluate the effectiveness of production equipment repair operations.

3. Methodology

In this paper, $M/M(a,b)/1/MWV$ queuing model with compulsory and extended repair analyzed. Instead of the server being fully idle during the vacation period, server serves at a different rate during multiple working vacations. The service rate varies depending on the arrival state. Customers arrive to the system with parameter λ_v follows Poisson distribution where the server provides service in regular busy period with parameter μ_{rb} and under the multiple working vacations server provides service with parameter μ_{wv} with exponential distribution. In this model batches of customers are served under General Bulk Service Rule. Thus each batch of service contains minimum $'a'$ units and maximum $'b'$ units. Suppose the customers waiting in the queue is less than $'a'$ server begins a vacation random variable V with parameter ξ.

In this model, we consider two types of repairs-compulsory repair ψ and extended repair ϕ. It is also assumed that breakdown happens only in busy period and denoted with the parameter β_v. When a breakdown occurs during busy period, a compulsory repair process begins instantly as aimed to quickly restore the breakdown condition. The immediate repair is essential to minimize disruptions in service and maintain system's performance during peak demand times. Also, in this model the extended repair process is considered in working vacations periods. The repair process during working vacations may be extended for various reasons. This could be due to limited availability of repair resources, or extended maintenance schedules.

Consider the customers in the queue at time t denote as $N_c(t)$, and L(t)=0,1 or 2 according to the server is idle, busy or regular busy on vacation state respectively. Let

$$\begin{aligned} R_n^I(t) &= Pr\{N_c(t)=n, L(t)=0\} \qquad 0 \le n \le a-1 \\ Q_n^V(t) &= Pr\{N_c(t)=n, L(t)=1\} \qquad n \ge 0 \\ P_n^B(t) &= Pr\{N_c(t)=n, L(t)=2\} \qquad n \ge 0 \end{aligned}$$

When the server is idle: $L(t)=0$, the size of the queue and system are the same.
When the server is busy or in working start during the vacation: $L(t)=1$ *or* 2, denote the total number of customers in the system is the sum of the number of customers in queue and the size of the service batches that contains particular $a \le x \le b$ customers.
Probabilities of the steady state are ,

$$Q_n^V = \lim_{t\to\infty} Q_n^V(t); \qquad R_n^I = \lim_{t\to\infty} R_n^I(t); \qquad P_n^B = \lim_{t\to\infty} P_n^B(t);$$

exist and the Chapman Kolmogrove equations satisfied by them in the steady state are given by,

$$\lambda_v R_0^I = \mu_{rb} P_0^B + \mu_{wv} Q_0^V \tag{1}$$

$$\lambda_v R_n^I = \lambda_v R_{n-1}^I + \mu_{rb} P_n^B + \mu_{wv} Q_n^V; \quad 1 \le n \le a-1 \tag{2}$$

$$(\lambda_v + \mu_{wv} + \phi + \xi) Q_0^V = \lambda_v R_{a-1}^I + \mu_{wv} \sum_{n=a}^{b} Q_n^V \tag{3}$$

$$(\lambda_v + \mu_{wv} + \phi + \xi) Q_n^V = (\lambda_v + \phi) Q_{n-1}^V + \mu_{wv} Q_{n+b}^V; \qquad n \ge 1 \tag{4}$$

$$(\lambda_v + \mu_{rb} + \psi + \beta_v) P_0^B = \mu_{rb} \sum_{n=a}^{b} P_n^B + \xi Q_0^V \tag{5}$$

$$(\lambda_v + \mu_{rb} + \psi + \beta_v) P_n^B = (\lambda_v + \psi + \beta_v) P_{n-1}^B + \mu_{rb} P_{n+b}^B + \xi Q_n^V; \qquad n \ge 1 \tag{6}$$

4. Steady state solution

To solve the steady state equation, the forward shifting operator E on P_n^B and Q_n^V are introduced then,

$$E(P_n^B) = P_{n+1}^B; \qquad E(Q_n^V) = Q_{n+1}^V \qquad for\ n \ge 0$$

Thus the equation (4) gives homogeneous difference equation as,

$$[\lambda_v + \phi + \mu_{wv}E^{b+1} - (\lambda_v + \phi + \xi + \mu_{wv})E]Q_n^V = 0 \tag{7}$$

The characteristics equation of (7) is obtained as,

$$z(u) = \lambda_v + \phi + \mu_{wv}u^{b+1} - (\lambda_v + \phi + \xi + \mu_{wv})u = 0 \tag{8}$$

by taking $x(u) = (\lambda_v + \phi + \xi + \mu_{wv})u$ and $y(u) = \lambda_v + \phi + \mu_{wv}u^{b+1}$, it is found that $|y(u)| < |x(u)|$ on $|u| = 1$. By Rouche's theorem $z(u)$ has unique root r_v inside the contour $|u| = 1$. Equation (7) has a homogeneous solution as,

$$Q_n^V = r_v^n Q_0^V \tag{9}$$

From equation (6) we get,

$$[(\lambda_v + \psi + \beta_v) + \mu_{rb}E^{b+1} - (\lambda_v + \mu_{rb} + \psi + \beta_v)E]P_n^B = -\xi r_v^{n+1}Q_0^V \tag{10}$$

By applying Rouches theorem to equation (10) as,

$$[(\lambda_v + \psi + \beta_v) + \mu_{rb}E^{b+1} - (\lambda_v + \mu_{rb} + \psi + \beta_v)E]P_n^B = 0$$

The above equation has unique root r with $|r| < 1$. Also equation (10) gives a non-homogeneous solution as,

$$P_n^B = \left[Zr^n - \frac{\xi r_v^{n+1}}{[(\lambda_v + \psi + \beta_v) + \mu_{rb}r_v^{b+1} - (\lambda_v + \mu_{rb} + \psi + \beta_v)r_v]}\right]Q_0^V \tag{11}$$

$$P_n^B = (Zr^n + Z^*r_v^n)Q_0^V \tag{12}$$

where

$$Z^* = \frac{\xi r_v}{[\lambda_v(r_v - 1) + \psi(r_v - 1) + \beta_v(r_v - 1) + \mu_{rb}r_v(1 - r_v^b)]} \tag{13}$$

The expression for R_n^I is obtained by adding equations (1) & (2) and substitute P_n^B and Q_n^V values,

$$R_n^I = \left[\frac{\mu_{rb}}{\lambda_v}\left(\frac{Z(1 - r^{n+1})}{(1 - r)} + \frac{Z^*(1 - r_v^{n+1})}{(1 - r_v)}\right) + \frac{\mu_{wv}}{\lambda_v}\frac{(1 - r_v^{n+1})}{(1 - r_v)}\right]Q_0^V$$

Now to calculate Z, consider equation (5) and substitute P_n^B and Q_n^V value,

$$Z\left[(\lambda_v + \mu_{rb} + \psi + \beta_v) - \frac{\mu_{rb}(r^a - r^{b+1})}{(1 - r)}\right] = \xi - Z^*\left((\lambda_v + \mu_{rb} + \psi + \beta_v) - \frac{\mu_{rb}(r_v^a - r_v^{b+1})}{(1 - r_v)}\right) \tag{14}$$

the above expression can be simplified as,

$$\frac{Z\mu_{rb}(1 - r^a)}{(1 - r)} = \frac{\xi}{(1 - r_v)} - \frac{Z^*\mu_{rb}(1 - r_v^a)}{(1 - r_v)} \tag{15}$$

Hence the probability queue size of the steady-state equation in terms of Q_0^V are obtained,

$$Q_n^V = (r_v^n)Q_0^V \qquad n \geq 0 \tag{16}$$

$$P_n^B = (Zr^n + Z^*r_v^n)Q_0^V \qquad n \geq 0 \tag{17}$$

where

$$Z = \frac{(1 - r)}{\mu_{rb}(1 - r^a)}\left[\frac{\xi}{(1 - r_v)} - \frac{Z^*\mu_{rb}(1 - r_v^a)}{(1 - r_v)}\right] \tag{18}$$

$$Z^* = \frac{\xi r_v}{[\lambda_v(r_v-1)+\psi(r_v-1)+\beta_v(r_v-1)+\mu_{rb}r_v(1-r_v^b)]} \tag{19}$$

and

$$R_n^I = \left[\frac{\mu_{rb}}{\lambda_v}\left(\frac{Z(1-r^{n+1})}{(1-r)}+\frac{Z^*(1-r_v^{n+1})}{(1-r_v)}\right)+\frac{\mu_{wv}}{\lambda_v}\frac{(1-r_v^{n+1})}{(1-r_v)}\right]Q_0^V \tag{20}$$

by using normalizing condition and calculated the value of Q_0^V

$$\sum_{n=0}^{\infty} Q_n^V + \sum_{n=0}^{\infty} P_n^B + \sum_{n=0}^{a-1} R_n^I = 1$$

By substituting P_n^B, Q_n^V and R_n^I is observe that,

$$\sum_{n=0}^{\infty} r_v^n Q_0^V + \sum_{n=0}^{\infty}(Zr^n + Z^* r_v^n)Q_0^V + \sum_{n=0}^{a-1}\left[\frac{\mu_{rb}}{\lambda_v}\left(\frac{Z(1-r^{n+1})}{(1-r)}+\frac{Z^*(1-r_v^{n+1})}{(1-r_v)}\right)+\frac{\mu_{wv}}{\lambda_v}\frac{(1-r_v^{n+1})}{(1-r_v)}\right]Q_0^V = 1$$

Hence $(Q_0^V)^{-1}$, is given by,

$$(Q_0^V)^{-1} = \omega(r_v,\mu_{wv}) + Z\omega(r,\mu_{rb}) + Z^*\omega(r_v,\mu_{rb}) \tag{21}$$

where

$$\omega(x,y) = \frac{1}{(1-x)}\left(1+\frac{y}{\lambda_v}(c-\frac{x(1-x^a)}{(1-x)})\right)$$

5. Performance Measures

In this section, we derive measures of performances for multiple working vacations model with compulsory and extended repair, the expected queue length, expected waiting time of the queue, expected system length and expected waiting time of the queue.

5.1. Mean queue length

The expected queue length of multiple working vacations model with reverse balking is given by;

$$L_q = \sum_{n=1}^{\infty} n(Q_n^V + P_n^B) + \sum_{n=1}^{a-1} nR_n^I$$

By substituting P_n^B, Q_n^V and R_n^I we observe that,

$$L_q = \sum_{n=1}^{\infty} n(r_v^n Q_0^V) + \sum_{n=1}^{\infty} n(Zr^n + Z^* r_v^n)Q_0^V + \sum_{n=1}^{a-1} n\left[\frac{\mu_{rb}}{\lambda_v}\left(\frac{Z(1-r^{n+1})}{(1-r)}+\frac{Z^*(1-r_v^{n+1})}{(1-r_v)}\right)+\frac{\mu_{wv}}{\lambda_v}\frac{(1-r_v^{n+1})}{(1-r_v)}\right]Q_0^V$$

$$L_q = Z\omega^*(r,\mu_{rb}) + Z^*\omega^*(r_v,\mu_{rb}) + \omega^*(r_v,\mu_{wv}) \tag{22}$$

where

$$\omega^*(x,y) = \frac{x}{(1-x)^2} + \frac{y}{(\lambda_v)(1-x)}\left[\frac{a(a-1)}{2} + \frac{ax^{a+1}(1-x)-x^2(1-x^a)}{(1-x)^2}\right]$$

and Z & Z^* are given by equations (18) & (19).

5.2. Mean System Length

The expected system length for the model,

$$L_s = L_q + \rho$$

$$L_s = \{Z\omega^*(r, \mu_{rb}) + Z^*\omega^*(r_v, \mu_{rb}) + \omega^*(r_v, \mu_{wv})\} + \rho$$

where

$$\omega^*(x, y) = \frac{x}{(1-x)^2} + \frac{y}{(\lambda_v)(1-x)}\left[\frac{a(a-1)}{2} + \frac{ax^{a+1}(1-x) - x^2(1-x^a)}{(1-x)^2}\right]$$

5.3. Mean Waiting Time of the Queue

The expected waiting time of the queue of our model is given by,

$$W_q = \frac{L_q}{\lambda_v}$$

$$\text{i.e., } W_q = \frac{Z\omega^*(r, \mu_{rb}) + Z^*\omega^*(r_v, \mu_{rb}) + \omega^*(r_v, \mu_{wv})}{\lambda_v}$$

where

$$\omega^*(x, y) = \frac{x}{(1-x)^2} + \frac{y}{(\lambda_v)(1-x)}\left[\frac{a(a-1)}{2} + \frac{ax^{a+1}(1-x) - x^2(1-x^a)}{(1-x)^2}\right]$$

5.4. Mean Waiting Time of the System

The expected waiting time of the system of above specifed model,

$$W_s = \frac{L_s}{\lambda_v}$$

$$W_s = \frac{\{Z\omega^*(r, \mu_{rb}) + Z^*\omega^*(r_v, \mu_{rb}) + \omega^*(r_v, \mu_{wv})\} + \rho}{\lambda_v}$$

where

$$\omega^*(x, y) = \frac{x}{(1-x)^2} + \frac{y}{(\lambda_v)(1-x)}\left[\frac{a(a-1)}{2} + \frac{ax^{a+1}(1-x) - x^2(1-x^a)}{(1-x)^2}\right]$$

If $Pr_{(wv)}$, $Pr_{(busy)}$ *and* $Pr_{(idle)}$ denote the probability that the server in idle, regular busy and busy vacation period then

$$Pr_{(idle)} = \sum_{n=0}^{a-1} R_n^I \tag{23}$$

where the R_n^I is given by equation (19).

$$Pr_{(busy)} = \sum_{n=0}^{\infty} P_n^B = \left(\frac{Z}{(1-r)} + \frac{Z^*}{(1-r_v)}\right) Q_0^V \tag{24}$$

$$Pr_{(wv)} = \sum_{n=0}^{\infty} Q_n^V = \frac{Q_0^V}{(1-r_v)} \tag{25}$$

6. Particular cases

6.1. Classical $M/M(a,b)/1/MWV$ model

By letting $\psi = \phi = 0$ in equations (16) to (22) and we obtain,

$$Q_n^V = (r_v^n)Q_0^V \qquad n \geq 0$$

$$P_n^B = (Zr^n + Z^* r_v^n)Q_0^V \qquad n \geq 0$$

$$R_n^I = \left[\frac{\mu_{rb}}{\lambda_v}(Zg_n(r) + Z^* g_n(r_v) + g_n(r_v)\right]Q_0^V \qquad 0 \leq n \leq a-1$$

where

$$Z = \frac{(1-r)}{\mu_{rb}(1-r^a)}\left[\frac{\xi}{(1-r_v)} - \frac{Z^*\mu_{rb}(1-r_v^a)}{(1-r_v)}\right]$$

$$Z^* = \frac{\xi r_v}{\lambda_v(r_v-1) + \mu_{rb}r_v(1-r_v^b)}$$

Further we observe that L_q of multiple working vacations queuing model with compulsory and extended repair by $\psi = \phi = 0$ we have;

$$L_q = Z\omega^*(r, \mu_{rb}) + Z^*\omega^*(r_v, \mu_{rb}) + \omega^*(r_v, \mu_{wv})$$

where

$$\omega^*(x,y) = \frac{x}{(1-x)^2} + \frac{y}{\lambda_v(1-x)}\left[\frac{a(a-1)}{2} + \frac{ax^{a+1}(1-x) - x^2(1-x^a)}{(1-x)^2}\right]$$

Thus, observed that our specified model coincides with the $M/M(a,b)/1/MWV$ queuing model analyzed by Julia Rose Mary and Aftab Begum (2011).

7. Numerical Analysis

In the numerical section, we consider various parameters like the the arrival rate (λ_v), regular service rate (μ_{rb}), the service rate during multiple working vacations (μ_{wv}), vacation parameter (ξ) & the compulsory and extended repair rate (ψ & ϕ). With the aid of the above parameters numerical analysis is carried out to analyze the performance measures of the specified model. For that purpose, consider the (7) & (10), where r_v and r are the roots of the following characteristic equations,

$$[\lambda_v + \phi + \mu_{wv}E^{b+1} - (\lambda_v + \phi + \xi + \mu_{wv})E] = 0$$

$$[(\lambda_v + \psi + \beta_v) + \mu_{rb}E^{b+1} - (\lambda_v + \mu_{rb} + \psi + \beta_v)E] = 0$$

respectively, and that lie in the interval [0,1].

The system performance measures and probabilities of each state is calculated and tabulated below. For the different values for λ_v, μ_{rb}, μ_{wv}, ξ, ψ and ϕ. Table 1 to Table 3 show various performance measures and Table 4 shows the probabilities of the idle state, the working vacation state and busy state.

Table 1: *L_q, L_s, W_q, and W_s with respect to ξ and μ_{wv}*

μ_{wv}	r_v	ξ	L_q	L_s	W_q	W_s
	0.9985	0.01	664.9178	665.2881	132.9836	133.0576
	0.9961	0.03	254.9934	255.3638	50.9987	51.0728
0.05	0.9938	0.05	160.2198	160.5902	32.0440	32.1180
	0.9915	0.07	116.9284	117.2988	23.3857	23.4598
	0.9892	0.09	92.2283	92.5987	18.4457	18.5197
	0.9981	0.01	524.6106	524.9806	104.9220	104.9961
	0.9956	0.03	225.8886	226.2590	45.1777	45.2518
0.10	0.9934	0.05	150.4587	150.8290	30.0917	30.1658
	0.9908	0.07	108.0308	108.4011	21.6062	21.6802
	0.9881	0.09	83.7757	84.1461	16.7551	16.8292
	0.9978	0.01	452.8352	453.2055	90.5670	90.6411
	0.9951	0.03	202.7179	203.0883	40.5436	40.6177
0.15	0.9927	0.05	135.9779	136.3482	27.1956	27.2696
	0.9894	0.07	93.8078	94.1781	18.7616	18.8356
	0.9871	0.09	77.3475	77.7179	15.4695	15.5436
	0.9967	0.01	301.4009	301.7713	60.2802	60.3543
	0.9945	0.03	180.4764	180.8467	36.0953	36.1693
0.20	0.9919	0.05	122.4987	122.8690	24.4997	24.5738
	0.9886	0.07	87.2368	87.6072	17.4474	17.5214
	0.9856	0.09	69.4125	69.7829	13.8825	13.9566

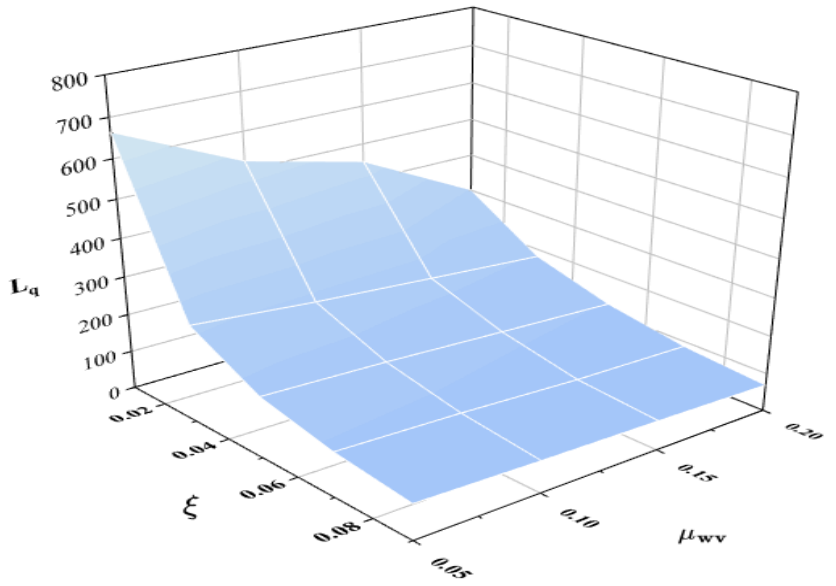


Figure 1: *Expected length of the queue with respect to ξ and μ_{wv}*

From Table (1) and Figure (1), we consider the parameters for arrival rate as $\lambda_v = 5$, service rate during regular busy period $\mu_{rb} = 0.9$, compulsory and extended repair rate $\psi = 4$ & $\phi = 4.5$, and then by varying vacation parameter ξ & service rate during vacation period μ_{wv}. Using these parameters, we calculated various measures of performances like expected queue length, & system length and theirs waiting time also. For a certain value for μ_{wv} (i.e $\mu_{wv} = 0.05$) noted that as vacation parameter (ξ) increases, L_s, L_q, W_s & W_q decrease gradually shown in Table (1).

From Figure (1) we observe that, the service rate during vacation period (μ_{wv}) and vacation

parameter (ξ) increases the length of the queue decreases, i.e., $\mu_{wv} = 0.05$ and $\xi = 0.01$ the length of the queue reaches the maximum value (664.9178) and $\mu_w v = 0.20$ & $\xi = 0.09$ length of the queue reaches the minimum value (69.4125).

Table 2: *L_q, L_s, W_q, W_s with respect to λ_v and μ_{wv}*

μ_{wv}	r_v	λ_v	L_q	L_s	W_q	W_s
	0.9985	6.0	665.9296	666.3741	110.9883	111.0623
	0.9961	5.8	255.5344	255.9640	44.0577	44.1317
0.05	0.9938	5.6	160.3357	160.7505	28.6314	28.7054
	0.9915	5.4	116.6397	117.0397	21.6000	21.6740
	0.9892	5.2	91.5454	91.9306	17.6049	17.6790
	0.9981	6.0	525.5545	525.9989	87.5924	87.6665
	0.9956	5.8	226.3512	226.7808	39.0261	39.1001
0.10	0.9934	5.6	150.4929	150.9077	26.8737	26.9478
	0.9908	5.4	107.6203	108.0203	19.9297	20.0038
	0.9881	5.2	82.9210	83.3062	15.9464	16.0204
	0.9978	6.0	453.7440	454.1884	75.6240	75.6981
	0.9951	5.8	203.1062	203.5358	35.0183	35.0924
0.15	0.9927	5.6	135.9050	136.3198	24.2688	24.3428
	0.9894	5.4	93.2164	93.6164	17.2623	17.3364
	0.9871	5.2	76.3332	76.7183	14.6795	14.7535
	0.9967	6.0	302.1914	302.6359	50.3652	50.4393
	0.9945	5.8	180.7857	181.2153	31.1699	31.2440
0.20	0.9919	5.6	122.3141	122.7289	21.8418	21.9159
	0.9886	5.4	86.5216	86.9216	16.0225	16.0966
	0.9856	5.2	68.1961	68.5813	13.1146	13.1887

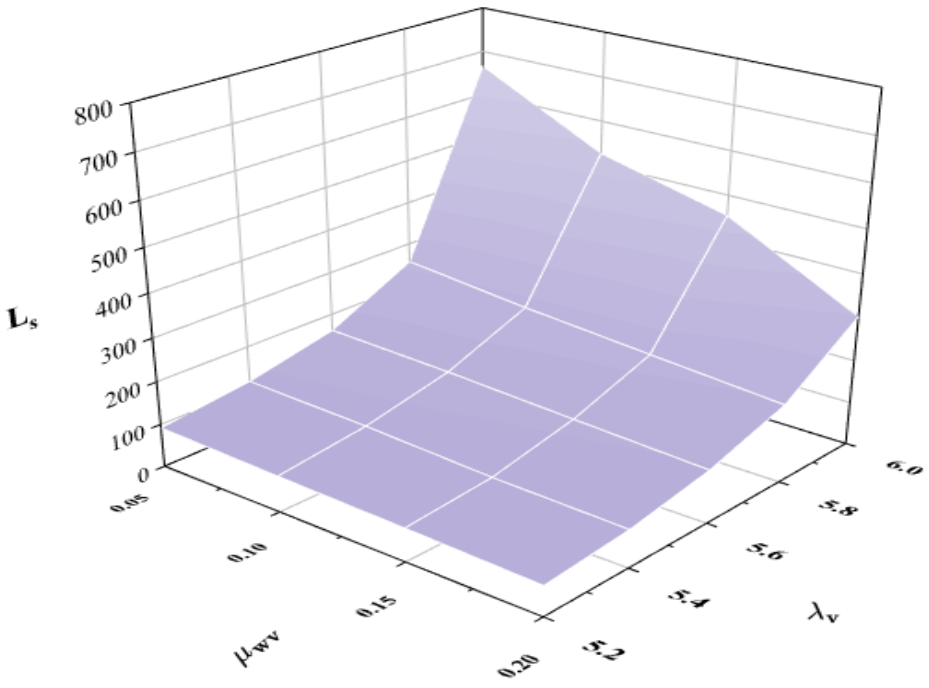


Figure 2: *Expected length of the system with respect to λ_v and μ_{wv}*

For the fixed values for the parameters $\mu_{rb} = 0.9$, vacation parameter $\xi = 0.01$, $\psi = 4$ & $\phi = 4.5$, and by varying service rate during vacation period (μ_{wv}) with respect to arrival rate (λ_v) we calculated other characteristics of the model shown in Table (2) & Figure (2). We consider a fixed value for μ_{wv} (i.e. $\mu_{wv} = 0.05$) and by varying the arrival rate (λ_v), we notice that as the arrival rate increase, then the various operating characteristics also increase. Also, for a fixed value for arrival rate λ_v (i.e. $\lambda_v = 6$) and by varying service rate of working vacation (μ_{wv}), we observe that μ_{wv} increases, the operating characteristics are decreasing. For example, $\lambda_v = 6$, $\mu_{wv} = 0.05$ the characteristic values reached their highest point ($L_q = 665.93$, $L_s = 666.37$, $W_q = 110.99$ & $W_s = 111.06$) and $\lambda_v = 6$, $\mu_{wv} = 0.20$ the characteristic values reached lowest point ($L_q = 302.19$, $L_s = 302.64$, $W_q = 50.37$ & $W_s = 50.44$)

Table 3: *L_q, L_s, W_q, W_s with respect to ψ and μ_{wv}*

μ_{wv}	r_v	ψ	L_q	L_s	W_q	W_s
	0.9975	4.1	398.0205	398.3909	79.6041	79.6782
	0.9949	4.3	194.7185	195.0889	38.9437	39.0178
0.05	0.9927	4.5	135.9419	136.3123	27.1884	27.2625
	0.9903	4.7	102.2985	102.6689	20.4597	20.5338
	0.9879	4.9	82.0785	82.4488	16.4157	16.4898
	0.9972	4.1	355.1622	355.5326	71.0324	71.1065
	0.9943	4.3	174.0373	174.4077	34.8075	34.8815
0.10	0.9920	4.5	123.8886	124.2590	24.7777	24.8518
	0.9895	4.7	94.3656	94.7360	18.8731	18.9472
	0.9870	4.9	76.2777	76.6480	15.2555	15.3296
	0.9969	4.1	320.5837	320.9541	64.1167	64.1908
	0.9938	4.3	159.8289	160.1992	31.9658	32.0398
0.15	0.9912	4.5	112.4559	112.8263	22.4912	22.5653
	0.9882	4.7	83.8072	84.1776	16.7614	16.8355
	0.9855	4.9	68.2585	68.6289	13.6517	13.7258
	0.9959	4.1	241.9815	242.3518	48.3963	48.4704
	0.9934	4.3	149.9802	150.3506	29.9960	30.0701
0.20	0.9901	4.5	99.7657	100.1360	19.9531	20.0272
	0.9871	4.7	76.5057	76.8761	15.3011	15.3752
	0.9835	4.9	59.8556	60.2260	11.9711	12.0452

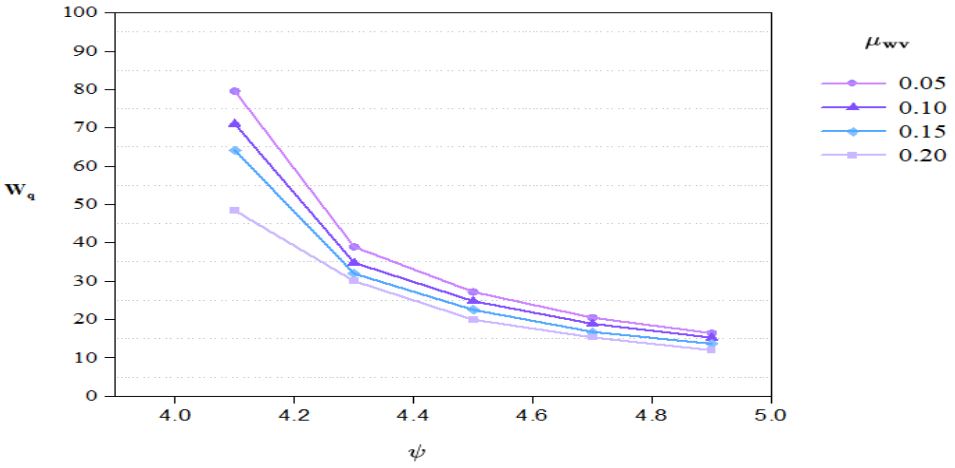


Figure 3: *Expected waiting time of the queue with respect to ψ and μ_{wv}*

Various performance measures of the compulsory and extended repair model are calculated and tabulated in Table (3). For that, we consider the values for $\lambda_v = 5$, $\mu_{rb} = 0.9$, $\xi = 0.01$, & $\phi = 4.5$ and different values for μ_{wv} & ψ.

Table (3), shows that for a desired value for μ_{wv} (i.e. $\mu_{wv} = 0.05$) and by different values for (increasing value) compulsory repair rate (ψ), the performance measures of the model L_s, L_q, W_s & W_q are decreasing. As μ_{wv} increases, the system operators are decreasing for a particular values for ψ. Figure (3) explains that as the expected waiting time of the queue with parameters like μ_{wv} & ψ values increase, the waiting time of the queue is decreasing step by step. That shows as the compulsory repair rate increases then waiting time of the queue decreases significantly.

Table 4: *$Pr(busy)$, $Pr(wv)$, and $Pr(idle)$ with respect to ξ and μ_{wv}*

μ_{wv}	r_v	ξ	$Pr(busy)$	$Pr(wv)$	$Pr(idle)$
0.05	0.9985	0.01	0.6325	0.3653	0.0022
	0.9975	0.02	0.6780	0.3182	0.0039
	0.9961	0.03	0.6760	0.3183	0.0058
	0.9949	0.04	0.6857	0.3069	0.0074
	0.9938	0.05	0.6966	0.2946	0.0088
0.10	0.9981	0.01	0.5784	0.4189	0.0027
	0.9972	0.02	0.6544	0.3413	0.0043
	0.9956	0.03	0.6518	0.3418	0.0064
	0.9943	0.04	0.6646	0.3273	0.0082
	0.9934	0.05	0.6853	0.3053	0.0094
0.15	0.9978	0.01	0.5441	0.4528	0.0031
	0.9969	0.02	0.6327	0.3625	0.0047
	0.9951	0.03	0.6298	0.3630	0.0071
	0.9938	0.04	0.6484	0.3427	0.0089
	0.9927	0.05	0.6671	0.3226	0.0104
0.20	0.9967	0.01	0.4498	0.5459	0.0043
	0.9959	0.02	0.5717	0.4223	0.0060
	0.9945	0.03	0.6060	0.3860	0.0080
	0.9934	0.04	0.6362	0.3542	0.0096
	0.9919	0.05	0.6482	0.3404	0.0114

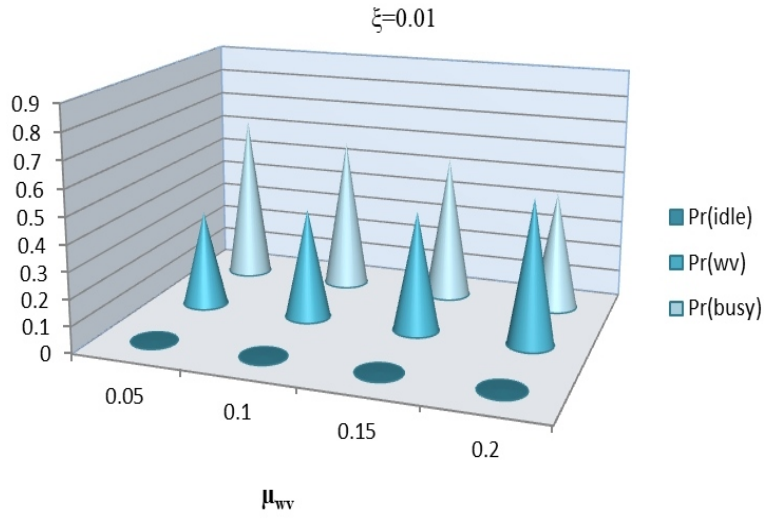


Figure 4: *$Pr(busy)$, $Pr(wv)$, and $Pr(idle)$ with respect to ξ and μ_{wv}*

The probability for various states, $Pr(idle)$, $Pr(wv)$ and $Pr(busy)$ are analyzed and the characteristic values are calculated and tabulated in (Table 4) by using the equations (23) to (25). The chosen parameters are $\lambda_v = 5$, $\mu_{rb} = 0.9$, $\psi = 4$ & $\phi = 4.5$, and by varying Service rate during working vacation (μ_{wv}) & vacation parameter (ξ), from Figure (4) we observe that as particular value for ξ (i.e., $\xi = 0.01$) and μ_{wv} values increase, $Pr(busy)$ decreases significantly and $Pr(idle)$ & $Pr(wv)$ increase gradually.

Conclusion

In this study, $M/M(a,b)/1/MWV$ queuing model for compulsory and extended repair are analyzed. In this model, GBSR is followed. The steady-state solution, the various performance measures for the system and particular cases are calculated. Finally, in the form of tables and graphs, numerical results have been analyzed.

Analyzing a batch arrival multiple working vacations model with compulsory and extended repair processes helps in understanding the complex system. In the compulsory repair model, failed components or entities are required to undergo repair immediately to ensure the system's continued operation. These repairs are essential to prevent system failures. An extended repair process refers to a situation where a failed entity or component undergoes a repair process that takes longer than usual repairs. While extended repairs may result in longer downtime. By combining compulsory and extended repair processes, organizations can optimize maintenance schedules, allocate resources effectively, and enhance system performance. The model helps in analysing repair queues, and determining optimal repair strategies to minimize downtime, and improve service quality. Overall, the compulsory and extended repair model serves as a valuable tool for organizations to enhance system performance. In future the model maybe extended to concept of breakdown situation.

References

[1] Abdollahi, S. & Rad, M. S. (2020). Reliability and sensitivity analysis of a batch arrival retrial queue with k-phase services, feedback, vacation, delay, repair and admission. *International Journal of Reliability, Risk and Safety: Theory and Application*, 3(2), 27-40.

[2] Ayyappan, G. & Deepa, T. (2019). Analysis of batch arrival bulk service queue with additional optional service multiple vacation and setup time. *International Journal of Mathematics in Operational Research*, 15(1), 1-25.

[3] Ayyappan, G. & Deepa, T. (2020). Analysis of an m [x]/g (a, b)/1 queueing model with multiple vacation controllable arrival during multiple vacation and two phases of repair with delay. *International Journal of Industrial and Systems Engineering*, 34(3), 409-431.

[4] Ayyappan, G. & Karpagam, S. (2019). An m [x]/g (a, b)/1 queueing system with server breakdown and repair, stand-by server and single vacation. *International Journal of Mathematics in Operational Research*, 14(2), 221-235.

[5] Ayyappan, G. & Thilagavathy, K. (2020). Analysis of map/ph/1 queueing model with setup, closedown, multiple vacations, standby server, breakdown, repair and reneging. *Reliability: Theory & Applications*, 15(2), 104-143.

[6] Begum, A. & Choudhury, G. (2022). Analysis of a bulk arrival n-policy queue with two-service genre, breakdown, delayed repair under bernoulli vacation and repeated service policy. *RAIRO-Operations Research*, 56(2), 979-1012.

[7] Chakravarthy, S. R., Kulshrestha, R., et al. (2020). A queueing model with server breakdowns, repairs, vacations, and backup server. *Operations research perspectives*, 7, 100131.

[8] Choudhary, A., Chakravarthy, S. R., & Sharma, D. C. (2023). Impact of the degradation in service rate in map/ph/1 queueing system with phase type vacations, breakdowns, and repairs. *Annals of Operations Research*, 331(2), 1207-1248.

[9] Elliriki, M., Reddy, C., Anand, K., & Saritha, S. (2022). Multi server queuing system with crashes and alternative repair strategies. *Communications in Statistics-Theory and Methods*, 51(23), 8173-8185.

[10] Gerasimova, M., Firer, A., Khramov, I., & Egarmina, A. (2024). Optimization of the number of repair personnel of the enterprise using the theory of queuing. *In ITM Web of Conferences*, volume 59 (pp. 01008).: EDP Sciences.

[11] Goswami, V. & Mund, G. (2020). Analysis of renewal input batch service queue with impatient customers and multiple working vacations. *International Journal of Management Science and Engineering Management*, 15(2), 96-105.

[12] Gupta, P. (2021). Study of feedback retrial queueing system with working vacation, setup time and perfect repair. *Ratio Mathematica*, 41, 291-307.

[13] Gupta, P. & Kumar, N. (2021). Performance analysis of retrial queueing model with working vacation, interruption, waiting server, breakdown and repair.*Journal of scientific research*, 13(3), 833-844.

[14] Hertini, E., Harisbaya, C., Nahar, J., et al. (2019). Queuing model using sojourn time distribution with single working vacation and vacation interruption. *In IOP Conference series: Materials science and engineering*, volume 567 (pp. 012036).: IOP Publishing.

[15] Huang, K.-B. (2023). Analysis and application of a batch arrival queueing model with the second optional service and randomized vacation policy. *In International Conference on Human-Computer Interaction* (pp. 320-333).: Springer.

[16] Jain, M. & Kumar, A. (2022). Unreliable server m [x]/g/1 queue with working vacation and multi-phase repair with delay in verification. *International Journal of Applied and Computational Mathematics*, 8(4), 210.

[17] Ketema, T., Demie, S., & Belachew, M. T. (2021). Controllable m/m/2 machine repair problem with multiple working vacations and triadic (0, q, n, m) policy. *International Journal of Management Science and Engineering Management*, 16(3), 184-196.

[18] Kumari, U. & Sharma, D. (2020). Optimal parameter selection for a machine repair system with servers vacation and controlling f-policy. *Jordan Journal of Mathematics and Statistics* (JJMS), 12(4), 625-642.

[19] Lakshmanan, K., Padmasekaran, S., & Jeganathan, K. (2019). Mathematical analysis of queueing-inventory model with compliment and multiple working vacations. *International Journal of Engineering and Advanced Technology, 8(6), 4239-4240*

[20] Lidiya, P. & Mary (2023). Performance study on heterogeneous arrival of batch service for multiple working vacations queuing system with breakdowns in the busy period. In 2023 First International Conference on Advances in Electrical, Electronics and Computational Intelligence (ICAEECI) *(pp. 1-6).: IEEE.*

[21] Lidiya, P. & Mary, K. J. R. (2024). Performance study of the m/m(a,b)/1/mwv queuing system with types of breakdowns. Journal of Applied Mathematics, Statistics and Informatics*, 19(2), 23-38.*

[22] Pikkala, V. L. & Kassahun, W. T. (2019). Discrete time analysis of multi-server queueing system with multiple working vacations and reneging of customers International Journal of Mathematical Modelling & Computations, 9(1 (WINTER)), 1-15.

[23] Prakati, P. & Mary, K. J. R. (2023). Evaluation of batch service mwv heterogeneous queuing system under balking and encouraged arrival. In 2023 First International Conference on Advances in Electrical, Electronics and Computational Intelligence (ICAEECI) *(pp. 1-7).: IEEE.*

[24] Rajadurai, P., Saravanarajan, M., & Chandrasekaran, V. (2018). A study on m/g/1 feedback retrial queue with subject to server breakdown and repair under multiple working vacation policy. Alexandria Engineering Journal*, 57(2), 947-962.*

[25] Rajam, V. et al. (2021). Retrial queue m(x)/g/c with bernoullis vacation, repair, service and re-service in orbit. Turkish Journal of Computer and Mathematics Education (TURCOMAT), *12(12), 4535-4540.*

[26] Sharma, R.& Kumar, G. (2021). Multi-phase essential repair analysis for multi-server queue under multiple working vacation. In Intelligent Systems: Proceedings of SCIS 2021 *(pp. 345-354).: Springer.*

[27] Thakur, S., Jain, A., & Ahuja, A. (2023). Analysis of map/ph/1 model with working vacation, working breakdown and two-phase repair. Arabian Journal for Science and Engineering, *(pp. 1-21).*

[28] Upadhyaya, S. & Kushwaha, C. (2020). Performance prediction and anfis computing for unreliable retrial queue with delayed repair under modified vacation policy. International Journal of Mathematics in Operational Research, *17(4), 437-466.*

[29] Vignesh, P., Srinivasan, S., & Sundari, S. M. (2019). Analysis of a non-markovian single server batch arrival queueing system of compulsory three stages of services with fourth optional stage service, service interruptions and deterministic server vacations. International Journal of Operational Research, *34(1), 28-53.*

[30] Wu, C.H., Yang, D.Y., & Ko, M.H. (2024). Performance sensitivity analysis for machine repair problem with two failure modes and working vacation. International Journal of Reliability, Quality and Safety Engineering, *(pp. 2350040).*

A NEW EXTENSION OF TWO PARAMETER ARADHANA DISTRIBUTION WITH COMPREHENSIVE STATISTICAL PROPERTIES AND APPLICATIONS

Rashid A. Ganaie[1,*], Manzoor A. Khanday[2], Prem Lata Gautam[3], R. Shenbagaraja[4] and T. Vivekanandan[5]

•

[1,*]Department of Statistics, Annamalai University, Tamil Nadu-608002, India
[2, 3]Department of Statistics, Lovely Professional University, Phagwara, Punjab, India
[4]Department of Statistics, Dr. M. G. R. Government Arts and Science College for Women, Villupuram, Tamil Nadu-605602, India
[5]Department of Science and Humanities, Vel Tech Multi Tech Dr.Rangarajan Dr.Sagunthala Engineering College, Chennai, Tamil Nadu-600062, India
[1,*]rashidau7745@gmail.com, [2]manzoorstat@gmail.com, [3]premgautam61@gmail.com, [4]statraja@gmail.com, [5]mtvivek2017@gmail.com

Abstract

In this study, we introduce a modified extension of standard two parameter Aradhana distribution called as the length biased two parameter Aradhana distribution. The length-biased distributions are a subclass of weighted distributions that adjust the probabilities based on lengths or sizes of observed phenomena, making them suitable for modelling certain practical scenarios. The length biased version is identified as a specific instance of a broader category of weighted distributions, which are widely used in statistical modelling when data collection mechanisms introduce some inherent bias. The distribution is thoroughly examined and analyzed with new structural properties and parameter estimation is performed using a robust and widely used technique of maximum likelihood estimation. The distribution's superiority and effectiveness are evaluated through a comparative analysis involving two real-world datasets. These datasets are used to highlight the practical applicability of the new distribution, demonstrate its flexibility and ability to fit real-world data better than existing models. This comprehensive approach highlights both theoretical advancements and real-world relevance, emphasizing the distribution's versatility and utility.

Keywords: length biased distribution, two parameter Aradhana distribution, reliability measures, order statistics

I. Introduction

In probability and statistics, the inception of weighted distributions plays a key role because the standard distributions may not provide the best fit from different data. Since then a situation arised to introduce an additional parameter to existing classical distribution which brings more flexibility in their nature. This additional parameter can be approached through several famous techniques i.e.

weighted technique. The idea of weighted distributions was suggested by Fisher [7] which later on modified by Rao [16] in unifying theory for problems where observations fall in non-experimental, non-replicated and non-random manner. The weighted distribution reduces to the length biased distribution especially if weight function considers only length of units of interest. The concept of length biased distribution was introduced by Cox [6] in renewal theory. Length biased sampling situation may occur in clinical trials, reliability theory and population studies were proper sampling frame is absent.

Many authors contributed extraordinarily to describe some length biased probability distributions along with applications. Ganaie et al. [8], Shanker and Prodhani [19], Alzoubi [4], Mustafa and Khan [13], Saghir et al. [18], Rashwan [15], Modi and Gill [11], Mudasir and Ahmad [12], Alidamat and Al-Omari [1], Ganaie and Rajagopalan [9], Al-Omari and Alsmairan [3], Almakhareez and Alzoubi [2], Chaito and Khamkong [5].

A two-parameter Aradhana distribution is a recently introduced new distribution which was proposed by Shanker et al. [20] which is a particular case of one parameter Aradhana distribution. It's important statistical properties have been described and derived. Further, its parameters are estimated by using the maximum likelihood estimation.

II. Length Biased Two Parameter Aradhana (LBTPA) Distribution

The probability density function of two parameter Aradhana distribution is given by

$$f(x;\theta,\alpha) = \frac{\theta^3}{\left(\theta^2\alpha^2 + 2\theta\alpha + 2\right)}(\alpha + x)^2 e^{-\theta x};\ x > 0, \theta > 0, \alpha > 0 \tag{1}$$

and cumulative distribution function of two parameter Aradhana distribution is given by

$$F(x;\theta,\alpha) = 1 - \left(1 + \frac{\theta x(\theta x + 2\theta\alpha + 2)}{\left(\theta^2\alpha^2 + 2\theta\alpha + 2\right)}\right)e^{-\theta x};\ x > 0, \theta > 0, \alpha > 0 \tag{2}$$

Consider the random variable X which represents non-negative condition with probability density function $f(x)$. Suppose its non-negative weight function $w(x)$, then probability density function of weighted random variable X_w is given by

$$f_w(x) = \frac{w(x)f(x)}{E(w(x))},\quad x > 0.$$

Here $w(x)$ is the weight function which is non - negative, then $E(w(x)) = \int w(x) f(x)dx < \infty$.

In this paper, we have considered the weight function as $w(x) = x$ to obtain the length biased version of two parameter Aradhana distribution and then its probability density function is given by

$$f_1(x;\theta,\alpha) = \frac{x f(x;\theta,\alpha)}{E(x)} \tag{3}$$

Here $E(x) = \int_0^\infty x f(x;\theta,\alpha)\,dx$

$$E(x) = \frac{\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}{\theta\left(\theta^2\alpha^2 + 2\theta\alpha + 2\right)} \tag{4}$$

After substituting equation (1) and (4) in equation (3), we will get the probability density function of length biased two parameter Aradhana distribution which is given by

$$f_l(x;\theta,\alpha) = \frac{x\theta^4}{\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}(\alpha + x)^2 e^{-\theta x} \tag{5}$$

The cumulative distribution function of length biased two parameter Aradhana distribution should be determined as

$$F_l(x;\theta,\alpha) = \int_0^x f_l(x;\theta,\alpha)dx$$

$$= \int_0^x \frac{x\theta^4}{\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}(\alpha + x)^2 e^{-\theta x}dx \tag{6}$$

After simplification of equation (6), we will get the cumulative distribution function of length biased two parameter Aradhana distribution which is given by

$$F_l(x;\theta,\alpha) = \frac{1}{\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}\left(\alpha^2\theta^2\gamma(2,\ \theta x) + \gamma(4,\ \theta x) + 2\alpha\theta\,\gamma(3,\ \theta x)\right) \tag{7}$$

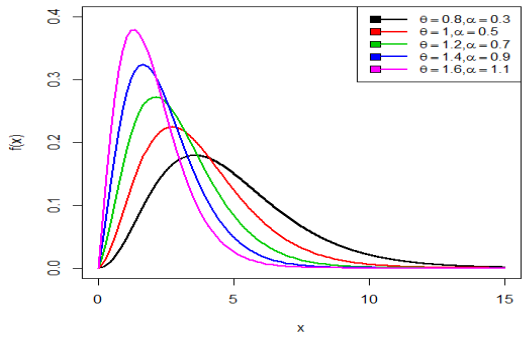


Figure 1: *Pdf plot of LBTPA distribution*

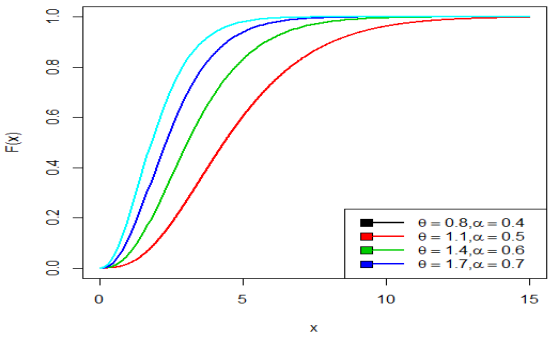


Figure 2: *Cdf plot of LBTPA distribution*

III. Reliability measures

In this section, here we have derived the reliability function, hazard rate function, reverse hazard rate function and mills ratio of length biased two parameter Aradhana distribution. The reliability function of length biased two parameter Aradhana distribution is given by

$$R(x) = 1 - F_l(x;\theta,\alpha)$$

$$R(x) = 1 - \frac{1}{\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}\left(\alpha^2\theta^2\gamma(2,\ \theta x) + \gamma(4,\ \theta x) + 2\alpha\theta\,\gamma(3,\ \theta x)\right)$$

The hazard rate of length biased two parameter Aradhana distribution is given by

$$h(x) = \frac{f_l(x;\theta,\alpha)}{1 - F_l(x;\theta,\alpha)}$$

$$h(x) = \frac{x\theta^4(\alpha + x)^2 e^{-\theta x}}{(\alpha^2\theta^2 + 6 + 4\alpha\theta) - (\alpha^2\theta^2\gamma(2,\ \theta x) + \gamma(4,\ \theta x) + 2\alpha\theta\,\gamma(3,\ \theta x))}$$

The reverse hazard rate function of length biased two parameter Aradhana distribution is given by

$$h_r(x) = \frac{f_1(x;\theta,\alpha)}{F_1(x;\theta,\alpha)}$$

$$h_r(x) = \frac{x\theta^4(\alpha + x)^2 e^{-\theta x}}{(\alpha^2\theta^2\gamma(2,\ \theta x) + \gamma(4,\ \theta x) + 2\alpha\theta\,\gamma(3,\ \theta x))}$$

The Mills Ratio of length biased two parameter Aradhana distribution is given by

$$M.R = \frac{1}{h_r(x)} = \frac{(\alpha^2\theta^2\gamma(2,\ \theta x) + \gamma(4,\ \theta x) + 2\alpha\theta\,\gamma(3,\ \theta x))}{x\theta^4(\alpha + x)^2 e^{-\theta x}}$$

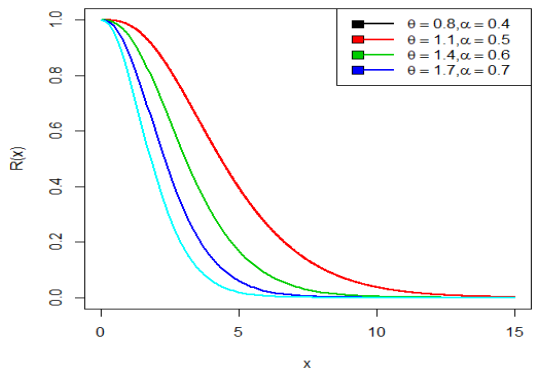


Figure 3:*Reliability plot of LBTPA distribution*

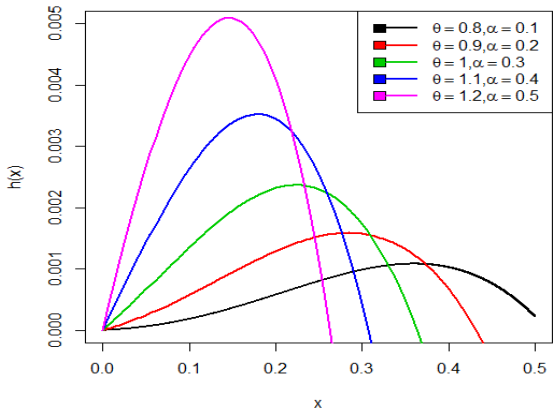


Figure 4:*Hazard plot of LBTPA distribution*

IV. Order Statistics

Order statistics is a basic concept in statistics which deals with arrangement of independent and identically distributed random variables. Consider the order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ of a random sample $X_1, X_2, \ldots, X_n$ from a continuous distribution with probability density function $f_x(x)$ and cumulative distribution function $F_x(x)$, then probability density function of r^{th} order statistics $X_{(r)}$ is given by

$$f_{x(r)}(x) = \frac{n!}{(r-1)!(n-r)!} f_X(x) \left(F_X(x)\right)^{r-1} \left(1 - F_X(x)\right)^{n-r} \qquad (8)$$

After substituting equations (5) and (7) in equation (8), we will get the probability density function of r^{th} order statistics $X_{(r)}$ of length biased two parameter Aradhana distribution as

$$f_{x(r)}(x) = \frac{n!}{(r-1)!(n-r)!}\left(\frac{x\theta^4(\alpha+x)^2 e^{-\theta x}}{\left(\alpha^2\theta^2+6+4\alpha\theta\right)}\right)\left(\frac{1}{\left(\alpha^2\theta^2+6+4\alpha\theta\right)}\left(\alpha^2\theta^2\gamma(2,\theta x)+\gamma(4,\theta x)+2\alpha\theta\gamma(3,\theta x)\right)\right)^{r-1}$$

$$\times\left(1-\frac{1}{\left(\alpha^2\theta^2+6+4\alpha\theta\right)}\left(\alpha^2\theta^2\gamma(2,\theta x)+\gamma(4,\theta x)+2\alpha\theta\gamma(3,\theta x)\right)\right)^{n-r}$$

For $r = n$, we will get the probability density function of higher order statistics $X_{(n)}$ of length biased two parameter Aradhana distribution as

$$f_{x(n)}(x) = \frac{n x\theta^4}{\left(\alpha^2\theta^2+6+4\alpha\theta\right)}(\alpha+x)^2 e^{-\theta x}\left(\frac{1}{\left(\alpha^2\theta^2+6+4\alpha\theta\right)}\left(\alpha^2\theta^2\gamma(2,\theta x)+\gamma(4,\theta x)+2\alpha\theta\gamma(3,\theta x)\right)\right)^{n-1}$$

For $r = 1$, we will get the probability density function of first order statistics $X_{(1)}$ of length biased two parameter Aradhana distribution as

$$f_{x(1)}(x) = \frac{n x\theta^4}{\left(\alpha^2\theta^2+6+4\alpha\theta\right)}(\alpha+x)^2 e^{-\theta x}\left(1-\frac{1}{\left(\alpha^2\theta^2+6+4\alpha\theta\right)}\left(\alpha^2\theta^2\gamma(2,\theta x)+\gamma(4,\theta x)+2\alpha\theta\gamma(3,\theta x)\right)\right)^{n-1}$$

V. Entropy

The term entropy is a fundamental concept which leads to measure the amount of uncertainty or disorder in a system.

I. Renyi Entropy

The Renyi entropy is a quantity which formulates the basis of concept of generalized dimensions. The Renyi entropy also generalizes several forms of entropy

$$R_T(\gamma) = \frac{1}{1-\gamma}\log\left\{\int f_1^{\gamma}(x;\theta,\alpha)\,dx\right\}$$

$$= \frac{1}{1-\gamma}\log\left(\int_0^{\infty}\left(\frac{x\theta^4}{\left(\alpha^2\theta^2+6+4\alpha\theta\right)}(\alpha+x)^2 e^{-\theta x}\right)^{\gamma}dx\right) \qquad (9)$$

After the simplification of equation (9), we get

$$R_T(\gamma) = \frac{1}{1-\gamma}\log\left(\left(\frac{\theta^{3\gamma-k-1}}{\left(\alpha^2\theta^2+6+4\alpha\theta\right)^{\gamma}}\right)\sum_{k=0}^{2\gamma}\binom{2\gamma}{k}\alpha^{2\gamma-k}\frac{\Gamma(\gamma+k+1)}{\gamma^{\gamma+k+1}}\right)$$

Obvious, the Tsallis entropy of proposed distribution will be determined by using the expression

$$T_s(\beta) = \frac{1}{\beta-1}\left(1-\int_0^{\infty} f_l^{\beta}(x;\theta,\alpha)\,dx\right)$$

$$= \frac{1}{\beta-1}\left(1-\int_0^{\infty}\left(\frac{x\theta^4}{\left(\alpha^2\theta^2+6+4\alpha\theta\right)}(\alpha+x)^2 e^{-\theta x}\right)^{\beta}dx\right) \tag{10}$$

After the simplification of above equation (10), we get

$$T_S(\beta) = \frac{1}{\beta-1}\left(1-\left(\frac{\theta^{3\beta-i-1}}{\left(\alpha^2\theta^2+6+4\alpha\theta\right)^{\beta}}\right)\sum_{i=0}^{2\beta}\binom{2\beta}{i}\alpha^{2\beta-i}\frac{\Gamma(\beta+i+1)}{\beta^{\beta+i+1}}\right)$$

VI. Statistical Properties

In this section, we have described and derived several structural properties of length biased two parameter Aradhana distribution those are

I. Moments

Consider random variable X constitutes the length biased two parameter Aradhana distribution, then r^{th} order moment of presented distribution will be determined as

$$\mu_r' = E(X^r) = \int_0^{\infty} x^r f_l(x;\theta,\alpha)\,dx$$

$$= \int_0^{\infty} x^r \frac{x\theta^4}{\left(\alpha^2\theta^2+6+4\alpha\theta\right)}(\alpha+x)^2 e^{-\theta x}dx \tag{11}$$

After the simplification of equation (11), we get

$$\mu_r' = E(X^r) = \frac{\left(\alpha^2\theta^2\Gamma(r+2)+\Gamma(r+4)+2\alpha\theta(r+3)\right)}{\theta^r\left(\alpha^2\theta^2+6+4\alpha\theta\right)} \tag{12}$$

Now substituting $r = 1$ and 2 in equation (12), we will get the first two moments of length biased two parameter Aradhana distribution

$$\mu_1' = \frac{\left(2\alpha^2\theta^2 + 24 + 12\alpha\theta\right)}{\theta\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}$$

$$\mu_2' = \frac{\left(6\alpha^2\theta^2 + 120 + 48\alpha\theta\right)}{\theta^2\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}$$

$$\text{Variance} = \frac{\left(2\alpha^4\theta^4 + 24\alpha^3\theta^3 + 108\alpha^2\theta^2 + 192\alpha\theta + 144\right)}{\theta^2\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)^2}$$

$$\text{S.D} = \frac{\sqrt{\left(2\alpha^4\theta^4 + 24\alpha^3\theta^3 + 108\alpha^2\theta^2 + 192\alpha\theta + 144\right)}}{\theta\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}$$

II. Harmonic Mean

The harmonic mean of executed distribution will be determined by applying following expression

$$\text{H.M} = E\left(\frac{1}{x}\right) = \int_0^\infty \frac{1}{x} f_l(x;\theta,\alpha)dx$$

$$= \int_0^\infty \frac{1}{x} \frac{x\theta^4}{\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}(\alpha + x)^2 e^{-\theta x}dx \qquad (13)$$

After the simplification of equation (13), we get

$$\text{H.M} = \frac{\theta(\alpha^2\theta + 2 + 2\alpha\theta)}{\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}$$

III. Moment generating function and characteristic function

The moment generating function of proposed length biased two parameter Aradhana distribution will be determined as

$$M_X(t) = E\left(e^{tx}\right) = \int_0^\infty e^{tx} f_l(x;\theta,\alpha)dx$$

$$= \int_0^\infty e^{tx} \frac{x\theta^4}{\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}(\alpha + x)^2 e^{-\theta x}dx \qquad (14)$$

After the simplification of equation (14), we get

$$M_X(t) = \frac{\theta^4}{(\theta - t)^4\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}\left(\alpha^2(\theta - t)^2 + 6 + 4\alpha(\theta - t)\right)$$

Obviously, the characteristic function of length biased two parameter Aradhana distribution is

$$M_X(it) = \frac{\theta^4}{(\theta - it)^4 \left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}\left(\alpha^2(\theta - it)^2 + 6 + 4\alpha(\theta - it)\right)$$

VII. Bonferroni and Lorenz Curves

The bonferroni and Lorenz curves were applied in economics to reveal the graphical representation of distribution of income or wealth. The proposed curves are defined as

$$B(p) = \frac{1}{p\mu_1'}\int_0^q x f_1(x;\theta,\alpha)\,dx$$

$$L(p) = \frac{1}{\mu_1'}\int_0^q x f_1(x;\theta,\alpha)\,dx$$

$$\text{Here } \mu_1' = \frac{\left(2\alpha^2\theta^2 + 24 + 12\alpha\theta\right)}{\theta\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}$$

$$B(p) = \frac{\theta\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}{p\left(2\alpha^2\theta^2 + 24 + 12\alpha\theta\right)}\int_0^q x\frac{x\theta^4}{\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}(\alpha + x)^2 e^{-\theta x}dx \qquad (15)$$

After the simplification of equation (15), we get

$$B(p) = \frac{1}{p\left(2\alpha^2\theta^2 + 24 + 12\alpha\theta\right)}\left(\alpha^2\theta^2\gamma(3,\ \theta q) + \gamma(5,\ \theta q) + 2\alpha\theta\gamma(4,\ \theta q)\right)$$

$$L(p) = \frac{1}{\left(2\alpha^2\theta^2 + 24 + 12\alpha\theta\right)}\left(\alpha^2\theta^2\gamma(3,\ \theta q) + \gamma(5,\ \theta q) + 2\alpha\theta(4,\ \theta q)\right)$$

VIII. Maximum Likelihood Estimation

In this section, we estimate the parameters of length biased two parameter Aradhana distribution by using maximum likelihood estimation. Consider the random sample $X_1, X_2,\ldots,X_n$ of size n from the length biased two parameter Aradhana distribution, then the likelihood function can be defined as

$$L(x) = \prod_{i=1}^n f_1(x;\theta,\alpha)$$

$$= \prod_{i=1}^n \left(\frac{x_i\,\theta^4}{\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}(\alpha + x_i)^2 e^{-\theta x_i}\right)$$

$$= \frac{\theta^{4n}}{\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)^n}\prod_{i=1}^n \left(x_i(\alpha + x_i)^2 e^{-\theta x_i}\right)$$

The log likelihood function should be stated as

$$\log L = 4n\log\theta - n\log\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right) + \sum_{i=1}^{n}\log x_i + 2\sum_{i=1}^{n}\log(\alpha + x_i) - \theta\sum_{i=1}^{n}x_i \quad (16)$$

Now differentiating log-likelihood equation (16) with respect to parameters θ and α. The following normal equations must be established

$$\frac{\partial \log L}{\partial \theta} = \frac{4n}{\theta} - n\left(\frac{\left(2\alpha^2\theta + 4\alpha\right)}{\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}\right) - \sum_{i=1}^{n}x_i = 0$$

$$\frac{\partial \log L}{\partial \alpha} = -n\left(\frac{\left(2\alpha\theta^2 + 4\theta\right)}{\left(\alpha^2\theta^2 + 6 + 4\alpha\theta\right)}\right) + 2\sum_{i=1}^{n}\left(\frac{1}{(\alpha + x_i)}\right) = 0$$

The above likelihood equations are too complicated to solve it algebraically. Therefore, we use the numerical technique like R and wolfram mathematics for estimating the required parameters.

IX. Application

In this section, we have performed the goodness of fit of length biased two parameter Aradhana distribution by analyzed the two real lifetime data sets and then comparison has been developed in order to reveal that the length biased two parameter Aradhana distribution provides a better fit in comparison over two parameter Aradhana, Garima, Pranav, Pratibha, Akash and Uma distributions.
Data set 1:The following real data set given below which represents the waiting times (in minutes) before service of 100 bank customers that was examined and analyzed by Ghitany et al. [10] for fitting the lindley distribution and the data set is given as under
0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9, 3.1, 3.2, 3.3, 3.5, 3.6, 4, 4.1, 4.2, 4.2, 4.3, 4.3, 4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9, 5, 5.3, 5.5, 5.7, 5.7, 6.1, 6.2, 6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1, 7.1, 7.1, 7.4, 7.6, 7.7, 8, 8.2, 8.6, 8.6, 8.6, 8.8, 8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11, 11, 11.1, 11.2, 11.2, 11.5, 11.9, 12.4, 12.5, 12.9, 13, 13.1, 13.3, 13.6, 13.7, 13.9, 14.1, 15.4, 15.4, 17.3, 17.3, 18.1, 18.2, 18.4, 18.9, 19, 19.9, 20.6, 21.3, 21.4, 21.9, 23, 27, 31.6, 33.1, 38.5
Data set 2:The second real symmetric data set constitutes the failure times of windshields which was reported by Murthy et al. [14] and the real data set is given below as
0.04, 0.3, 0.31, 0.557, 0.943, 1.07, 1.124, 1.248, 1.281, 1.281, 1.303, 1.432, 1.48, 1.51, 1.51, 1.568, 1.615, 1.619, 1.652, 1.652, 1.757, 1.795, 1.866, 1.876, 1.899, 1.911, 1.912, 1.9141, 0.981, 2.010, 2.038, 2.085, 2.089, 2.097, 2.135, 2.154, 2.190, 2.194, 2.223, 2.224, 2.23, 2.3, 2.324, 2.349, 2.385, 2.481, 2.610, 2.625, 2.632, 2.646, 2.661, 2.688, 2.823, 2.89, 2.9, 2.934, 2.962, 2.964, 3, 3.1, 3.114, 3.117, 3.166, 3.344, 3.376, 3.385, 3.443, 3.467, 3.478, 3.578, 3.595, 3.699, 3.779, 3.924, 4.035, 4.121, 4.167, 4.240, 4.255, 4.278, 4.305, 4.376, 4.449, 4.485, 4.570, 4.602, 4.663, 4.694

To enumerate the model comparison criterions the unknown parameters are estimated thoroughly by applying R software. To predict the capability and effectiveness of length biased two parameter Aradhana distribution over two parameter Aradhana, Garima, Pranav, Pratibha, Akash and Uma distributions, we consider criterions like *AIC* (Akaike Information Criterion), *CAIC* (Consistent Akaike Information Criterion), *BIC* (Bayesian Information Criterion), *AICC* (Corrected Akaike Information Criterion), Shannon's entropy *H(X)* and *-2logL*. It is quite obvious that the distribution is

better if it has the smaller criterion values of *AIC, BIC, AICC, CAIC, H*(*X*) and *-2logL* in comparison over other compared distributions. The formulae for computing the criterion value are as given

$$AIC = 2k - 2\log L, \qquad BIC = k\log n - 2\log L, \qquad AICC = AIC + \frac{2k(k+1)}{n-k-1}$$

$$CAIC = \frac{2kn}{n-k-1} - 2\log L, \qquad H(X) = -\frac{2\log L}{n}$$

Where *n* is sample size, *k* is the number of parameters in statistical model and –2log*L* is maximized value of log-likelihood function under the considered model.

Table 1: *Shows MLE and S.E of Data set 1 and Data set 2 of performed distributions*

Data set 1			Data set 2		
Distribution	MLE	S.E	Distribution	MLE	S.E
LBTPA	$\hat{\alpha} = 5.011230$ $\hat{\theta} = 2.028885$	$\hat{\alpha} = 1.103495$ $\hat{\theta} = 1.430256$	LBTPA	$\hat{\alpha} = 0.330304$ $\hat{\theta} = 1.444080$	$\hat{\alpha} = 0.239622$ $\hat{\theta} = 0.103652$
TPA	$\hat{\alpha} = 1.79978517$ $\hat{\theta} = 0.25969442$	$\hat{\alpha} = 1.04516631$ $\hat{\theta} = 0.02479628$	TPA	$\hat{\alpha} = 0.0946144$ $\hat{\theta} = 1.1265797$	$\hat{\alpha} = 0.088121$ $\hat{\theta} = 0.078480$
Garima	$\hat{\theta} = 0.15154412$	$\hat{\theta} = 0.01309641$	Garima	$\hat{\theta} = 0.5571774$	$\hat{\theta} = 0.050554$
Pranav	$\hat{\theta} = 0.4047845$	$\hat{\theta} = 0.0200951$	Pranav	$\hat{\theta} = 1.2094683$	$\hat{\theta} = 0.052114$
Pratibha	$\hat{\theta} = 0.28916795$	$\hat{\theta} = 0.01658711$	Pratibha	$\hat{\theta} = 0.8937320$	$\hat{\theta} = 0.050786$
Akash	$\hat{\theta} = 0.29527837$	$\hat{\theta} = 0.01683424$	Akash	$\hat{\theta} = 0.9320800$	$\hat{\theta} = 0.055209$
Uma	$\hat{\theta} = 0.39680518$	$\hat{\theta} = 0.01958624$	Uma	$\hat{\theta} = 1.2098061$	$\hat{\theta} = 0.059669$

Table 2: *shows comparison, performance and analysis of fitted distributions of data set 1*

Distribution	-2logL	AIC	BIC	AICC	CAIC	H(X)
LBTPA	634.6014	638.6014	643.8117	638.7251	638.7251	6.3460
TPA	637.5075	641.5075	646.7179	641.6312	641.6312	6.3750
Garima	649.6375	651.6375	654.2427	651.6783	651.6783	6.4963
Pranav	665.9148	667.9148	670.52	667.9556	667.9556	6.6591
Pratibha	640.3329	642.3329	644.9381	642.3737	642.3737	6.4033
Akash	641.9292	643.9292	646.5344	643.9700	643.9700	6.4192
Uma	659.1849	661.1849	663.7901	661.2257	661.2257	6.5918

Table 3: *shows comparison, performance and analysis of fitted distributions of data set 2*

Distribution	-2logL	AIC	BIC	AICC	CAIC	H(X)
LBTPA	282.5138	286.5138	291.4685	286.6549	286.6549	3.2103
TPA	286.6318	290.6318	295.5865	290.7729	290.7729	3.2571
Garima	331.4615	333.4615	335.9388	333.5080	333.5080	3.7666
Pranav	300.3776	302.3776	304.8549	302.4241	302.4241	3.4133
Pratibha	303.18	305.18	307.6573	305.2265	305.2265	3.4452
Akash	306.3466	308.3466	310.824	308.3931	308.3931	3.4812
Uma	298.3461	300.3461	302.8234	300.3926	300.3926	3.3902

It is quite clear and illustrated from the results given above in table 2 and table 3 that the length biased two parameter Aradhana distribution has smaller *AIC*, *BIC*, *AICC*, *CAIC*, *H(X)* and *-2logL* values in comparison over two parameter Aradhana, Garima, Pranav, Pratibha, Akash and Uma distributions which indicates that the length biased two parameter Aradhana distribution provides a better fit over two parameter Aradhana, Garima, Pranav, Pratibha, Akash and Uma distributions.

X. Conclusion

In this paper, we have incorporated a new extension of two parameter Aradhana distribution termed as length biased two parameter Aradhana distribution has been established. The formulated new distribution has been comprehensively introduced and explored by applying the length biased technique to two parameter Aradhana distribution. Some of its fundamental statistical features like moments, shape of behavior of pdf and cdf, variance and its mean, coefficient of variation, reliability function, hazard function, reverse hazard function, moment generating function, bonferroni and lorenz curves have been described. Further its order statistics and Renyi entropy has been discussed. Additionally its parameters are estimated by using maximum likelihood estimation. Furthermore, the predictability and potentiality of length biased two parameter Aradhana distribution has been examined and highlighted by applying the two real lifetime data sets and hence it is carefully explored from the result that length biased two parameter Aradhana distribution provides a quite satisfactory fit as compared over two parameter Aradhana, Garima, Pranav, Pratibha, Akash and Uma distributions.

References

[1] Alidamat, A. J. and Al-Omari, A. I. (2021). The extended length biased two parameters Mirra distribution with an application to engineering data. *Advanced Mathematical Models & Applications*, 6(2):113-127.

[2] Almakhareez, S. and Alzoubi, L. (2024). Length-biased Benrabia distribution with applications. *Electronic Journal of Applied Statistical Analysis*, 17(02):255-277.

[3] Al-Omari, A. I. and Alsmairan, I. K. (2019). Length-biased Suja distribution and its applications. *Journal of Applied Probability and Statistics*, 14(3):95-116.

[4] Alzoubi, L. (2024). Length-biased Loai distribution: statistical properties and application. *Electronic Journal of Applied Statistical Analysis*, 17(02):278-301.

[5] Chaito, T. and Khamkong, M. (2021). The length-biased Weibull-Rayleigh distribution for application to hydrological data. *Lobachevskii Journal of Mathematics*, 42:3253-3265.

[6] Cox, D. R. (1962). Renewal theory, *Barnes and Noble*, New York.

[7] Fisher, R. A. (1934). The effects of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics*, 6:13-25.

[8] Ganaie, R. A., Shenbagaraja, R. and Soumya, V. P. (2024). A generalization and some new contribution to Pratibha distribution with statistical properties and its applications. *International Journal of Statistics and Reliability Engineering*, 11(1):108-116.

[9] Ganaie, R. A. and Rajagopalan, V. (2023). The weighted power quasi lindley distribution with properties and applications of lifetime data. *Pak.j.stat.oper.res*., 19(2):279-298.

[10] Ghitany, M. E., Atieh, B. and Nadarajah, S. (2008). Lindley distribution and its Applications. *Mathematics Computing and Simulation*, 78:493-506.

[11] Modi, K. and Gill, V. (2015). Length-biased weighted Maxwell distribution. *Pak.j.stat.oper.res*., 11(4):465-472.

[12] Mudasir, S. and Ahmad, S. P. (2018). Characterization and estimation of the length biased Nakagami distribution. *Pak.j.stat.oper.res*., 14(3):697-715.

[13] Mustafa, A. and Khan, M. I. (2023). Length biased extended Rayleigh distribution and its applications. *Journal of Modern Applied Statistical Methods*, 22(2):1-19.

[14] Murthy, D. N. P., Xie, M and Jiang, R. (2004). *Weibull Models*, John Wiley & Sons Inc, Hoboken.

[15] Rashwan, N. I. (2013). A length-biased version of the generalized gamma distribution. *Advances and Applications in Statistics*, 32(2):119-137.

[16] Rao, C. R. (1965). On discrete distributions arising out of method of ascertainment, in classical and Contagious Discrete, G.P. Patiled; Pergamum Press and Statistical publishing Society, Calcutta. 320-332.

[17] R core team. R version 3.5.3 (2019): A language and environment for statistical Computing. R foundation for statistical computing, Vienna, Austria. URL http://www.R-project.org/.

[18] Saghir, A., Saleem, M., Khadim, A. and Tazeem, S. (2015). The modified double weighted exponential distribution with properties. *Mathematical Theory and Modeling*, 5(8):78-91.

[19] Shanker, R. and Prodhani, H. R. (2024). On weighted Amarendra distribution with properties and applications. *Biometrics & Biostatistics International Journal*, 13(3):76-88.

[20] Shanker, R., Soni, N. K., Shanker, R., Ray, M. and Prodhani, H. R. (2024). A two-Parameter Aradhana distribution with applications to Reliability engineering. *Reliability: Theory & Applications*, 19, 3(79):757-774.

RT&A, No 2 (84)
Volume 20, June 2025

LAKIBUL - G FAMILY OF DISTRIBUTIONS: SPECIAL MODELS, PROPERTIES AND APPLICATIONS

Idzhar A. Lakibul

•

Department of Mathematics
College of Arts and Sciences
Mindanao State University - Sulu, Jolo, Sulu, Philippines
idzhar.lakibul@msusulu.edu.ph

Abstract

In this study, a new generated family of distributions is introduced, called the Lakibul - G Family of Distributions. The T-X family of distributions is used in the derivation of the proposed class. Three special models of the proposed family are derived, such as the Lakibul - Uniform, Lakibul-Kumaraswamy, and Lakibul-Rayleigh distributions. Some properties of the proposed special models such as moments, mean, variance, moment generating function, survival, and hazard functions are studied. The Maximum likelihood approach is used to estimate the parameters of proposed models. Results show the following: (i) for Lakibul-Uniform distribution, it is found that this proposed probability distribution produces better estimate for electronic dataset as compared with the Cubic Transmuted Uniform distribution; (ii) for Lakibul - Kumaraswamy distribution is found to give a good fit for the COVID-19 dataset compared with the Marshall-Olkin reduce Kies and the Transmuted Kumaraswamy distributions; and (iii) for the Lakibul - Rayleigh distribution, it is observed that this distribution provides a better fit for precipitation dataset as compared with the Exponentiated Transform Inverse Rayleigh and the Exponentiated Inverse Rayleigh distributions.

Keywords: Uniform distribution, T-X family of distributions, Rayleigh distribution, Kumaraswamy distribution.

1. Introduction

The probability distribution plays an important role in describing the behavior of real-life data. The Weibull, log-logistic, and log-normal distributions are popular classical probability distributions for modeling non-negative or lifetime data. For bimodal behavior, bathtub behavior, and other complex behaviors of the data, the classical distributions are no longer flexible to describe the said behaviors of the data. Because of that, the extension and generalization of any classical distributions have become an area of interest for some researchers.

Azzalini [3] proposed a skewed family of distributions to generate a distribution with an additional skewed parameter. Eugene et al. [7] introduced the Beta-G family of distributions using the beta distribution. Cordeiro and de Castro [6] developed a Kumaraswamy - G family as an alternative to the Beta-G family using Kumaraswamy distribution. Other identified family of distributions are the Transmuted - G family [16], the exponentiated generalized class of distributions [5], the generalized transmuted - G family of distributions [13] and the exponentiated generalized Kumaraswamy - G class [17].

Rahman et al. [14] proposed a cubic transmuted family of distributions by adding an additional parameter to the baseline distribution. In addition, an extended version of the uniform distribution called the cubic transmuted uniform distribution was derived, and some of its statistical properties

such as its moments, mean, variance, moment generating function, characteristic function, mean absolute deviation, quantile function, median, reliability function, hazard function, and Shannon entropy were studied.

Lakibul and Tubo [11] established a new generated family of distributions called the T-extended Standard U-quadratic - G family of distributions (TeSU-G). In addition, a bimodal version of the Weibull distribution was derived using the TeSU-G family of distributions. An extension of the TeSU-G family of distributions can be found in the study by Lakibul and Tubo [10]. Furthermore, the bivariate version of the extended Standard U-quadratic distribution can be seen in the paper by Lakibul, Polestico, and Supe [12].

In this paper, the main objectives are the following: (i) to introduce the Lakibul-G family of distributions; (ii) to derive some special distributions to be derived such as the Lakibul-Uniform distribution, Lakibul-Kumaraswamy distribution and the Lakibul-Rayleigh distribution using the proposed class; (iii) to derive some properties of each of the proposed distributions such as moment, mean, variance, moment generating function, survival and hazard functions; (iv) to estimate the parameters of each of the proposed distributions using the maximum likelihood method; and (v) to apply the proposed distributions on real data and compare with some existing distributions.

The remainder of the paper is organized as follows. Lakibul - G family of distributions is introduced in Section 2. In Section 3, Lakibul - Uniform distribution is derived and its properties are studied. Lakibul-Kumaraswamy distribution and its properties are presented in Section 4. In Section 5, the Lakibul-Rayleigh distribution and its properties are introduced and derived. The applicability of each of the proposed distributions is discussed in Section 6. In Section 7, a concluding remark is presented.

2. Lakibul-G Family of Distributions

Alzaatreh et al. [2] introduced the T-X family of distributions to extend any continuous distribution to a generalized distribution. The cumulative distribution function of the special case of the T-X family of distribution is given by

$$F_{T-X}(x) = \int_0^{G(x)} f(t)dt, \tag{1}$$

where $G(x)$ is any baseline cumulative distribution function and $f(t)$ is any probability density function (pdf) with support on the interval $[0, 1]$. Lakibul - G family of distributions is obtained by using

$$f(t) = 2(1-\lambda)(1-t) + 3\lambda(4t^2 - 4t + 1), 0 \leq \lambda \leq 1. \tag{2}$$

in (1), and the cumulative distribution function (cdf) of the proposed family of distribution is

$$F_L(x) = (2+\lambda)G(x) - (1+5\lambda)G^2(x) + 4\lambda G^3(x) \tag{3}$$

with corresponding probability density function given by

$$f_L(x) = g(x)[2 + \lambda - 2(1+5\lambda)G(x) + 12\lambda G^2(x)]$$

, where $g(x)$ is the probability density function associated with the cdf $G(x)$.

3. Lakibul-Uniform Distribution (LU) and Its Properties

Let X be a random variable follows a uniform distribution with cumulative distribution function given by

$$G_U(x) = x, x \in [0, 1]. \tag{4}$$

The cdf of Lakibul-Uniform distribution (LU) is obtained by inserting (4) into (3) and is

$$F_{LU}(x) = (2+\lambda)x - (1+5\lambda)x^2 + 4\lambda x^3 \tag{5}$$

with corresponding pdf given by

$$f_{LU}(x) = 2 + \lambda - 2(1+5\lambda)x + 12\lambda x^2, x \in [0,1], \lambda \in [0,1]. \tag{6}$$

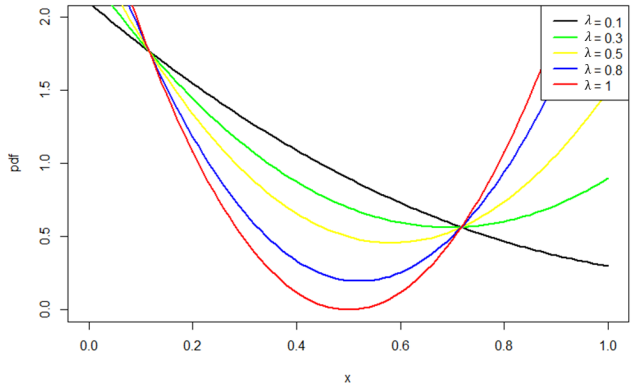


Figure 1: *pdf plots of the LU distribution for varied values of λ.*

Figure 1 shows some possible density shapes of the Lakibul-Uniform distribution. It reveals that the Lakibul-Uniform distribution can model a data with bathtub density shape.

3.1. Moments of LU

Theorem 1. The *rth* raw moment of LU distribution with density (6) and is

$$\mu'_r = \frac{r(3\lambda r + 2 + \lambda) + 6}{(r+1)(r+2)(r+3)}. \tag{7}$$

The mean and variance are, respectively, $\frac{\lambda+2}{6}$ and $\frac{22\lambda - 5\lambda^2 - 10}{180}$.

Proof. The *rth* raw moment is defined by

$$\begin{aligned}
\mu'_r &= E[X^r] \\
&= \int_0^\infty x^r f(x) dx \\
&= \int_0^1 x^r [2 + \lambda - 2(1+5\lambda)x + 12\lambda x^2] dx \\
&= (2+\lambda)\int_0^1 x^r dx - 2(1+5\lambda)\int_0^1 x^{r+1} dx + 12\lambda \int_0^1 x^{r+2} dx \\
&= \frac{2+\lambda}{r+1} - \frac{2(1+5\lambda)}{r+2} + \frac{12\lambda}{r+3} \\
&= \frac{r(3\lambda r + 2 + \lambda) + 6}{(r+1)(r+2)(r+3)}.
\end{aligned}$$

The mean of *LU* distribution is obtained by taking $r = 1$ in (7) and is

$$\mu'_1 = \frac{\lambda + 2}{6}.$$

The 2*nd* raw moment of LU distribution is obtained from (7) by taking $r = 2$ and is

$$\mu'_2 = \frac{7\lambda + 5}{30}.$$

The variance of LU distribution is obtained as

$$\begin{aligned}\sigma^2 &= \mu_2' - (\mu_1')^2 \\ &= \frac{7\lambda+5}{30} - \left(\frac{\lambda+2}{6}\right)^2 \\ &= \frac{22\lambda - 5\lambda^2 - 10}{180}.\end{aligned}$$

3.2. Moment Generating Function of LU

Theorem 2. Let X follows the LU distribution then the moment generating function $M_X(t)$ is given by

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r[r(3\lambda r + 2 + \lambda) + 6]}{r!(r+1)(r+2)(r+3)}, t \in \mathbb{R}.$$

Proof. By definition of moment generating function and pdf (6), we have

$$M_X(t) = \mathbb{E}(e^{tX}) = \int_0^1 e^{tX} f_{LU}(x)dx.$$

Recall that $e^{tX} = \sum_{r=0}^{\infty} \frac{t^r}{r!}x^r$ then we have

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!}\mu_r'.$$

Hence,

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r[r(3\lambda r + 2 + \lambda) + 6]}{r!(r+1)(r+2)(r+3)}.$$

3.3. Survival and Hazard Functions of LU

Let X be a random variable with cdf (5) and pdf (6) then the survival $S_{LU}(x)$ and hazard $h_{LU}(x)$ functions of LU distribution are respectively, given by

$$S_{LU}(x) = 1 - (2+\lambda)x - (1+5\lambda)x^2 + 4\lambda x^3$$

and

$$h_{LU}(x) = \frac{2 + \lambda - 2(1+5\lambda)x + 12\lambda x^2}{1 - (2+\lambda)x - (1+5\lambda)x^2 + 4\lambda x^3},$$

where $\lambda \in [0, 1]$.

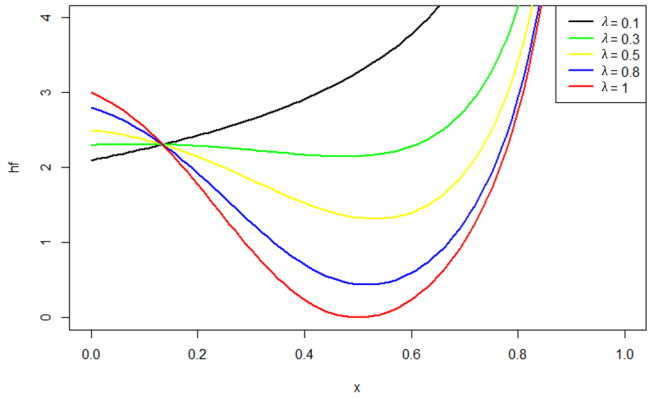


Figure 2: *hf plots of the LU distribution for varied values of λ.*

Figure 2 shows some possible shapes of the hazard function of the LU distribution. It reveals that the hazard function of the LU distribution can model the behavior of the bathtub hazard rate of the data.

3.4. Maximum Likelihood Estimates of LU

Let X_1, X_2,..., X_n be a random sample of size n from LU distribution, then the likelihood function is

$$L = \prod_{i=1}^{n} [2 + \lambda - 2(1 + 5\lambda)x_i + 12\lambda x_i^2].$$

The log-likelihood function is

$$l = \sum_{i=1}^{n} log[2 + \lambda - 2(1 + 5\lambda)x_i + 12\lambda x_i^2]. \tag{8}$$

Taking the derivative of (8) with respect to parameter λ then, we have

$$\frac{\partial l}{\partial \lambda} = \sum_{i=1}^{n} \frac{1 - 10x_i + 12x_i^2}{2 + \lambda - 2(1 + 5\lambda)x_i + 12\lambda x_i^2}. \tag{9}$$

Equating (9) to 0, we can obtain the numerical maximum likelihood estimate of the LU distribution parameter.

4. Lakibul-Kumaraswamy Distribution (LK) and Its Properties

Let X be a random variable follows a Kumaraswamy distribution with cdf given by

$$G_K(x) = 1 - (1 - x^a)^b, x \in (0, 1), a, b > 0. \tag{10}$$

The cdf of Lakibul - Kumaraswamy distribution (LK) is obtained by inserting (10) into (3) and is

$$F_{LK}(x) = 1 - 3\lambda(1 - x^a)^b + (7\lambda - 1)(1 - x^a)^{2b} - 4\lambda(1 - x^a)^{3b} \tag{11}$$

with corresponding pdf given by

$$f_{LK}(x) = abx^{a-1}(1 - x^a)^{b-1} \left[3\lambda + 2(1 - 7\lambda)(1 - x^a)^b + 12\lambda(1 - x^a)^{2b}\right], \tag{12}$$

where $x \in (0, 1)$, a, $b > 0$ and $\lambda \in [0, 1]$.

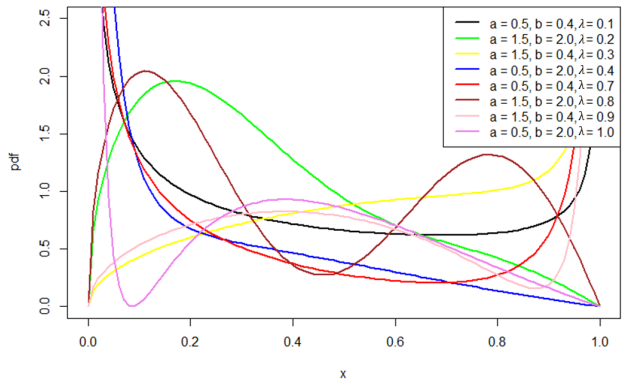


Figure 3: *pdf plots of the LK distribution for some values of a and b and varied values of λ.*

Figure 3 presents some possible density shapes of the LK distribution. It reveals that the LK distribution can model data with either of the following shapes: exponential, unimodal, bimodal, and bathtub density shapes.

4.1. Moments of LK

Theorem 3. The *rth* raw moment of LK distribution with density (12) and is

$$\mu'_r = 3b\lambda\mathbb{B}\left(\frac{r}{a}+1,b\right)+2b(1-7\lambda)\mathbb{B}\left(\frac{r}{a}+1,2b\right)+12b\lambda\mathbb{B}\left(\frac{r}{a}+1,3b\right). \tag{13}$$

The mean and variance of LK distribution are respectively, given by

$$\mu = 3b\lambda\mathbb{B}\left(\frac{1}{a}+1,b\right)+2b(1-7\lambda)\mathbb{B}\left(\frac{1}{a}+1,2b\right)+12b\lambda\mathbb{B}\left(\frac{1}{a}+1,3b\right)$$

and

$$\begin{aligned}Var(X) =&3b\lambda\mathbb{B}\left(\frac{2}{a}+1,b\right)+2b(1-7\lambda)\mathbb{B}\left(\frac{2}{a}+1,2b\right)+12b\lambda\mathbb{B}\left(\frac{2}{a}+1,3b\right)\\ &-b^2\left(3b\lambda\mathbb{B}\left(\frac{1}{a}+1,b\right)+2b(1-7\lambda)\mathbb{B}\left(\frac{1}{a}+1,2b\right)+12b\lambda\mathbb{B}\left(\frac{1}{a}+1,3b\right)\right)^2.\end{aligned}$$

Proof. The *rth* raw moment is defined by

$$\begin{aligned}\mu'_r =&E[X^r]\\ =&\int_0^\infty x^r f(x)dx\\ =&\int_0^1 x^r abx^{a-1}(1-x^a)^{b-1}\left[3\lambda+2(1-7\lambda)(1-x^a)^b+12\lambda(1-x^a)^{2b}\right]dx\\ =&3\lambda ab\int_0^1 x^r x^{a-1}(1-x^a)^{b-1}dx+2(1-7\lambda)ab\int_0^1 x^r x^{a-1}(1-x^a)^{2b-1}dx\\ &+12\lambda ab\int_0^1 x^r x^{a-1}(1-x^a)^{3b-1}dx\\ =&3b\lambda\mathbb{B}\left(\frac{r}{a}+1,b\right)+2b(1-7\lambda)\mathbb{B}\left(\frac{r}{a}+1,2b\right)+12b\lambda\mathbb{B}\left(\frac{r}{a}+1,3b\right).\end{aligned}$$

The mean of LK distribution is obtained by taking $r=1$ in (13) and is

$$\mu = \mu'_1 = 3b\lambda\mathbb{B}\left(\frac{1}{a}+1,b\right)+2b(1-7\lambda)\mathbb{B}\left(\frac{1}{a}+1,2b\right)+12b\lambda\mathbb{B}\left(\frac{1}{a}+1,3b\right).$$

The 2nd raw moment of LK distribution is obtained by taking $r=2$ in (13) and is

$$\mu'_2 = 3b\lambda\mathbb{B}\left(\frac{2}{a}+1,b\right)+2b(1-7\lambda)\mathbb{B}\left(\frac{2}{a}+1,2b\right)+12b\lambda\mathbb{B}\left(\frac{2}{a}+1,3b\right).$$

The variance of LK distribution is obtained as

$$\begin{aligned}Var(X) =&\mu'_2-(\mu'_1)^2\\ =&3b\lambda\mathbb{B}\left(\frac{2}{a}+1,b\right)+2b(1-7\lambda)\mathbb{B}\left(\frac{2}{a}+1,2b\right)+12b\lambda\mathbb{B}\left(\frac{2}{a}+1,3b\right)\\ &-b^2\left(3b\lambda\mathbb{B}\left(\frac{1}{a}+1,b\right)+2b(1-7\lambda)\mathbb{B}\left(\frac{1}{a}+1,2b\right)+12b\lambda\mathbb{B}\left(\frac{1}{a}+1,3b\right)\right)^2.\end{aligned}$$

4.2. Moment Generating Function of LK

Theorem 4. Let X follows the *LK* distribution, then the moment generating function $M_X(t)$ is given by

$$M_X(t)=\sum_{r=0}^{\infty}\frac{t^r b}{r!}\left[3\lambda\mathbb{B}\left(\frac{r}{a}+1,b\right)+2(1-7\lambda)\mathbb{B}\left(\frac{r}{a}+1,2b\right)+12\lambda\mathbb{B}\left(\frac{r}{a}+1,3b\right)\right], t\in\mathbb{R}.$$

The proof is similar to the proof of Theorem 2.

4.3. Survival and Hazard Functions of LK

Let X be a random variable with cdf (11) and pdf (12), then the survival $S_{LK}(x)$ and hazard $h_{LK}(x)$ functions of LK distribution are respectively, given by

$$S_{LK}(x) = 3\lambda(1-x^a)^b - (7\lambda-1)(1-x^a)^{2b} + 4\lambda(1-x^a)^{3b}$$

and

$$h_{LK}(x) = \frac{abx^{a-1}(1-x^a)^{b-1}\left[3\lambda + 2(1-7\lambda)(1-x^a)^b + 12\lambda(1-x^a)^{2b}\right]}{3\lambda(1-x^a)^b - (7\lambda-1)(1-x^a)^{2b} + 4\lambda(1-x^a)^{3b}},$$

where $x \in (0,1)$, a, $b > 0$ and $\lambda \in [0,1]$.

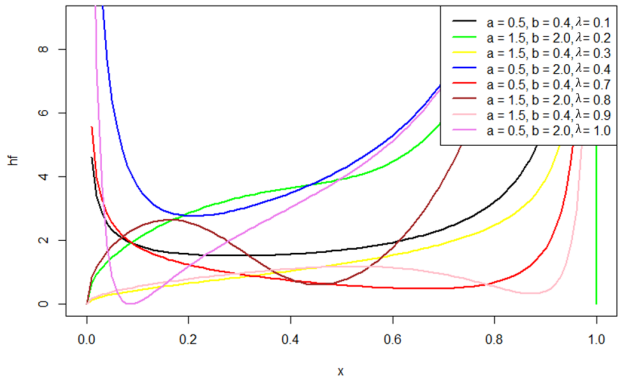


Figure 4: *hf plots of the LK distribution for some values of a and b and varied values of λ.*

Figure 4 shows some possible hazard rate behaviors of the LK distribution. It shows that the hazard rate of the LK distribution can model a data with bathtub, monotonic increasing, or non-monotonic increasing hazard behaviors.

4.4. Maximum Likelihood Estimates of LK

Let X_1, X_2,..., X_n be a random sample of size n from LK distribution then the likelihood function is

$$L = \prod_{i=1}^{n} abx_i^{a-1}(1-x_i^a)^{b-1}\left[3\lambda + 2(1-7\lambda)(1-x_i^a)^b + 12\lambda(1-x_i^a)^{2b}\right].$$

The log-likelihood function is

$$l = nlog(a) + nlog(b) + (a-1)\sum_{i=1}^{n} log(x_i) + (b-1)\sum_{i=1}^{n} log(1-x_i^a)$$

$$+ \sum_{i=1}^{n} log\left[3\lambda + 2(1-7\lambda)(1-x_i^a)^b + 12\lambda(1-x_i^a)^{2b}\right]. \tag{14}$$

Taking the derivative of (14) with respect to the parameters a, b and λ then, we have

$$\frac{\partial l}{\partial a} = \frac{n}{a} - 2b\sum_{i=1}^{n} \frac{x_i^a(1-x_i^a)^{b-1}log(x_i)\left[1-7\lambda+12\lambda(1-x_i^a)^b\right]}{3\lambda + 2(1-7\lambda)(1-x_i^a)^b + 12\lambda(1-x_i^a)^{2b}}$$

$$+ \sum_{i=1}^{n} log(x_i) - (b-1)\sum_{i=1}^{n} \frac{x_i^a log(x_i)}{1-x_i^a}, \tag{15}$$

$$\frac{\partial l}{\partial b} = \frac{n}{b} + \sum_{i=1}^{n} log(1-x_i^a) + 2\sum_{i=1}^{n} \frac{(1-x_i^a)^b log(1-x_i^a)\left[1-7\lambda+12\lambda(1-x_i^a)^b\right]}{3\lambda + 2(1-7\lambda)(1-x_i^a)^b + 12\lambda(1-x_i^a)^{2b}} \tag{16}$$

and

$$\frac{\partial l}{\partial \lambda} = \sum_{i=1}^{n} \frac{3 - 14(1 - x_i^a)^b + 12(1 - x_i^a)^{2b}}{3\lambda + 2(1 - 7\lambda)(1 - x_i^a)^b + 12\lambda(1 - x_i^a)^{2b}}. \tag{17}$$

The numerical maximum likelihood estimates of the LK distribution parameters can be obtained by equating (15), (16) and (17) to 0, respectively.

5. Lakibul-Rayleigh Distribution (LR) and Its Properties

Let X be a random variable follows a Rayleigh distribution with cdf given by

$$G_R(x) = 1 - e^{-\frac{x^2}{2\sigma^2}}, x > 0, \sigma > 0. \tag{18}$$

The cdf of Lakibul-Rayleigh distribution (LR) is derived by inserting (18) into (3) and is

$$F_{LR}(x) = 1 - 3\lambda e^{-\frac{x^2}{2\sigma^2}} + (7\lambda - 1)e^{-\frac{x^2}{\sigma^2}} - 4\lambda e^{-\frac{3x^2}{2\sigma^2}} \tag{19}$$

with corresponding pdf given by

$$f_{LR}(x) = \frac{xe^{-\frac{x^2}{2\sigma^2}}}{\sigma^2}\left[3\lambda - 2(7\lambda - 1)e^{-\frac{x^2}{2\sigma^2}} + 12\lambda e^{-\frac{x^2}{\sigma^2}}\right], \tag{20}$$

where $x > 0$, $\sigma > 0$ and $\lambda \in [0, 1]$.

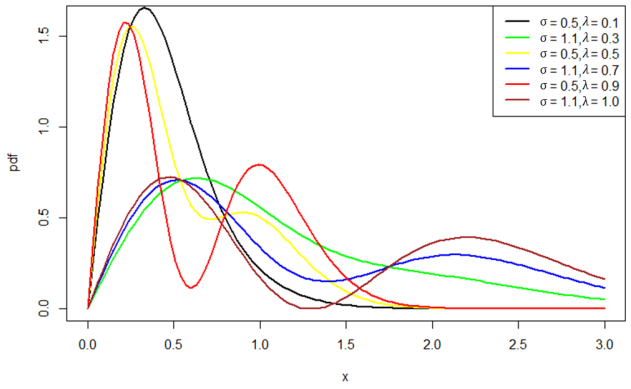


Figure 5: *pdf plots of the LR distribution for some values of σ and varied values of λ.*

Figure 5 presents some possible density shapes of the LR distribution. It reveals that the LR distribution can model data with right-tailed unimodal or bimodal behaviors.

5.1. Moments of LR

Theorem 5. Let X be a random variable that follows Lakibul-Rayleigh distribution, then the *rth* raw moment of LR distribution is given by

$$\mu_r' = \sigma^r \left[1 + \lambda\left(2^{\frac{r}{2}}3 + 2^{\frac{r}{2}+2}3^{-\frac{r}{2}} - 7\right)\right]\Gamma\left(\frac{r}{2} + 1\right). \tag{21}$$

The mean and variance of LR distribution are respectively, given by

$$\mu = \frac{\sigma\sqrt{\pi}}{2}\left[1 + \lambda\left(3\sqrt{2} + \frac{4\sqrt{2}}{\sqrt{3}} - 7\right)\right]$$

and

$$Var(X) = \sigma^2\left\{1 + \frac{5\lambda}{3} - \frac{\pi}{4}\left[1 + \lambda\left(3\sqrt{2} + \frac{4\sqrt{2}}{\sqrt{3}} - 7\right)\right]^2\right\}.$$

Proof. The *rth* raw moment is defined by

$$\begin{aligned}
\mu_r' &= E[X^r] \\
&= \int_0^\infty x^r f(x)dx \\
&= \int_0^\infty x^r \frac{xe^{-\frac{x^2}{2\sigma^2}}}{\sigma^2}\left[3\lambda - 2(7\lambda-1)e^{-\frac{x^2}{2\sigma^2}} + 12\lambda e^{-\frac{x^2}{\sigma^2}}\right]dx \\
&= 3\lambda\int_0^\infty x^r \frac{xe^{-\frac{x^2}{2\sigma^2}}}{\sigma^2}dx - 2(7\lambda-1)\int_0^\infty x^r \frac{xe^{-\frac{x^2}{\sigma^2}}}{\sigma^2}dx + 12\lambda\int_0^\infty x^r \frac{xe^{-\frac{3x^2}{2\sigma^2}}}{\sigma^2}dx \\
&= 3\lambda 2^{\frac{r}{2}}\sigma^r\Gamma\left(\frac{r}{2}+1\right) - (7\lambda-1)\sigma^r\Gamma\left(\frac{r}{2}+1\right) + 2^{\frac{r}{2}+2}3^{-\frac{r}{2}}\lambda\sigma^2\Gamma\left(\frac{r}{2}+1\right) \\
&= \sigma^r\left[1 + \lambda\left(2^{\frac{r}{2}}3 + 2^{\frac{r}{2}+2}3^{-\frac{r}{2}} - 7\right)\right]\Gamma\left(\frac{r}{2}+1\right).
\end{aligned}$$

The mean of LR distribution is obtained by taking $r = 1$ in (21) and is

$$\begin{aligned}
\mu = \mu_1' &= \sigma\left[1 + \lambda\left(2^{\frac{1}{2}}3 + 2^{\frac{1}{2}+2}3^{-\frac{1}{2}} - 7\right)\right]\Gamma\left(\frac{1}{2}+1\right) \\
&= \sigma\left[1 + \lambda\left(3\sqrt{2} + \frac{4\sqrt{2}}{\sqrt{3}} - 7\right)\right]\Gamma\left(\frac{3}{2}\right) \\
&= \sigma\left[1 + \lambda\left(3\sqrt{2} + \frac{4\sqrt{2}}{\sqrt{3}} - 7\right)\right]\frac{1}{2}\Gamma\left(\frac{1}{2}\right) \\
&= \frac{\sigma\sqrt{\pi}}{2}\left[1 + \lambda\left(3\sqrt{2} + \frac{4\sqrt{2}}{\sqrt{3}} - 7\right)\right].
\end{aligned}$$

The 2*nd* raw moment of LR distribution is obtained by taking $r = 2$ in (21) and is

$$\begin{aligned}
\mu_2' &= \sigma^2\left[1 + \lambda\left(2^{\frac{2}{2}}3 + 2^{\frac{2}{2}+2}3^{-\frac{2}{2}} - 7\right)\right]\Gamma\left(\frac{2}{2}+1\right) \\
&= \frac{\sigma^2}{3}(3+5\lambda)\Gamma(2) \\
&= \frac{\sigma^2}{3}(3+5\lambda).
\end{aligned}$$

The variance of LR distribution is obtained as

$$\begin{aligned}
Var(X) &= \mu_2' - (\mu_1')^2 \\
&= \frac{\sigma^2}{3}(3+5\lambda) - \frac{\sigma^2\pi}{4}\left[1 + \lambda\left(3\sqrt{2} + \frac{4\sqrt{2}}{\sqrt{3}} - 7\right)\right]^2 \\
&= \sigma^2\left\{1 + \frac{5\lambda}{3} - \frac{\pi}{4}\left[1 + \lambda\left(3\sqrt{2} + \frac{4\sqrt{2}}{\sqrt{3}} - 7\right)\right]^2\right\}.
\end{aligned}$$

5.2. Moment Generating Function of LR

Theorem 6. Let X follows the *LR* distribution, then the moment generating function $M_X(t)$ is given by

$$M_X(t) = \sum_{r=0}^{\infty}\frac{t^r\sigma^r\left[1 + \lambda\left(2^{\frac{r}{2}}3 + 2^{\frac{r}{2}+2}3^{-\frac{r}{2}} - 7\right)\right]}{r!}\Gamma\left(\frac{r}{2}+1\right), t \in \mathbb{R}.$$

The proof is similar to the proof of Theorem 2.

5.3. Survival and Hazard Functions of LR

Let X be a random variable with cdf (19) and pdf (20), then the survival $S_{LR}(x)$ and hazard $h_{LR}(x)$ functions of LR distribution are respectively, given by

$$S_{LR}(x) = 3\lambda e^{-\frac{x^2}{2\sigma^2}} - (7\lambda - 1)e^{-\frac{x^2}{\sigma^2}} + 4\lambda e^{-\frac{3x^2}{2\sigma^2}}$$

and

$$h_{LR}(x) = \frac{xe^{-\frac{x^2}{2\sigma^2}}\left[3\lambda - 2(7\lambda - 1)e^{-\frac{x^2}{2\sigma^2}} + 12\lambda e^{-\frac{x^2}{\sigma^2}}\right]}{\sigma^2\left[3\lambda e^{-\frac{x^2}{2\sigma^2}} - (7\lambda - 1)e^{-\frac{x^2}{\sigma^2}} + 4\lambda e^{-\frac{3x^2}{2\sigma^2}}\right]},$$

where $x > 0$, $\sigma > 0$ and $\lambda \in [0, 1]$.

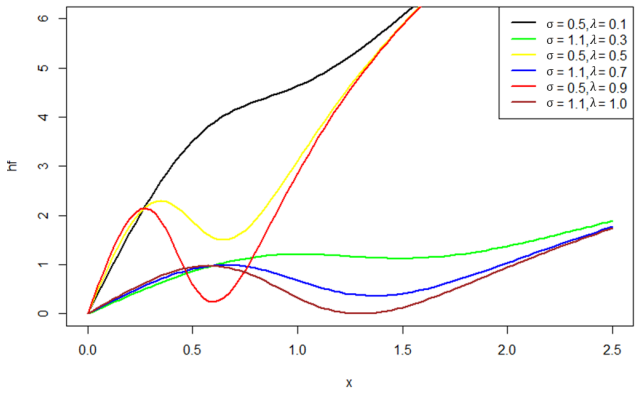


Figure 6: *hf plots of the LR distribution for some values of σ and varied values of λ.*

Figure 6 shows some possible shapes of the hazard rate function of the LR distribution. It reveals that the hazard rate function of the LR distribution can model data with monotonic or non-monotonic increasing hazard behaviors.

5.4. Maximum Likelihood Estimates of LR

Let X_1, X_2,..., X_n be a random sample of size n from LR distribution then the likelihood function is

$$L = \prod_{i=1}^{n} \frac{x_i e^{-\frac{x_i^2}{2\sigma^2}}}{\sigma^2}\left[3\lambda - 2(7\lambda - 1)e^{-\frac{x_i^2}{2\sigma^2}} + 12\lambda e^{-\frac{x_i^2}{\sigma^2}}\right].$$

The log-likelihood function is

$$l = \sum_{i=1}^{n} log(x_i) - \sum_{i=1}^{n} \frac{x_i^2}{2\sigma^2} - 2nlog(\sigma) + \sum_{i=1}^{n} log\left[3\lambda - 2(7\lambda - 1)e^{-\frac{x_i^2}{2\sigma^2}} + 12\lambda e^{-\frac{x_i^2}{\sigma^2}}\right]. \tag{22}$$

Taking the derivative of (22) with respect to the parameters σ and λ then, we have

$$\frac{\partial l}{\partial \sigma} = \frac{1}{\sigma}\sum_{i=1}^{n} x_i^2 - \frac{2n}{\sigma} + \frac{2}{\sigma}\sum_{i=1}^{n} \frac{x_i^2 e^{-\frac{x_i^2}{2\sigma^2}}\left(1 - 7\lambda + 12\lambda e^{-\frac{x_i^2}{2\sigma^2}}\right)}{3\lambda - 2(7\lambda - 1)e^{-\frac{x_i^2}{2\sigma^2}} + 12\lambda e^{-\frac{x_i^2}{\sigma^2}}} \tag{23}$$

and

$$\frac{\partial l}{\partial \lambda} = \sum_{i=1}^{n} \frac{3 - 14e^{-\frac{x_i^2}{2\sigma^2}} + 12e^{-\frac{x_i^2}{\sigma^2}}}{3\lambda - 2(7\lambda - 1)e^{-\frac{x_i^2}{2\sigma^2}} + 12\lambda e^{-\frac{x_i^2}{\sigma^2}}}. \tag{24}$$

Equating (23) and (24) to 0, respectively, we can obtain numerical maximum likelihood estimates of the LR distribution parameters.

6. Application

In this section, we apply the LU, LK and LR distributions on the real datasets and compare with some existing distributions, respectively. The applications of each of the proposed distributions are presented in the following subsections. In the analysis of data, we use a package "fitdistrplus" in R software. For comparison of the distributions, we use the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

6.1. Application of LU Distribution

In this subsection, the LU distribution is applied to real data and compared with the Cubic Transmuted Uniform (CTU) distribution proposed by Rahman et al. [14]. The probability density function of the CTU distribution is given by

$$f_{CTU}(x) = 1 - \lambda + 6\lambda x - 6\lambda x^2, x \in [0,1], \lambda \in [-1,1].$$

The data set used in this application is about the lifetimes of 30 electronic devices in days. This set of data was used by Rahman et al. [14]. They fitted the CTU distribution to this dataset and compared it with the Kumaraswamy, Beta, and Transmuted Uniform distributions, and they found that the CTU distribution had the best fit for this dataset. The data set is given as follows: 0.020, 0.029, 0.034, 0.044, 0.057, 0.096, 0.106, 0.139, 0.156, 0.164, 0.167, 0.177, 0.250, 0.326, 0.406, 0.607, 0.650, 0.672, 0.676, 0.736, 0.817, 0.838, 0.910, 0.931, 0.946, 0.953, 0.961, 0.981, 0.982 and 0.990.

Table 1 lists the MLE estimates with the corresponding standard errors of the models fitted to the electronic data set. Table 2 indicates that the LU distribution gives a better fit for this dataset since it has the smallest AIC and BIC values compared to the CTU distribution. Moreover, the estimated pdf of the models fitted to the electronic dataset is presented in Figure 7, revealing that the LU distribution can mimic the behavior of the electronic dataset compared to the CTU distribution.

Table 1: *MLEs and Standard Error (SE) of the fitted models for the electronic dataset.*

Distribution	$\hat{\lambda}$	*SE*
LU	0.8239858	0.159316
CTU	-1	0.4436516

Table 2: *Numerical values of log-likelihood (logLike), AIC, and BIC of the fitted models for electronic dataset.*

Distribution	*logLike*	*AIC*	*BIC*
LU	**7.020164**	**-12.04033**	**-10.63913**
CTU	6.196299	-10.3926	-8.991401

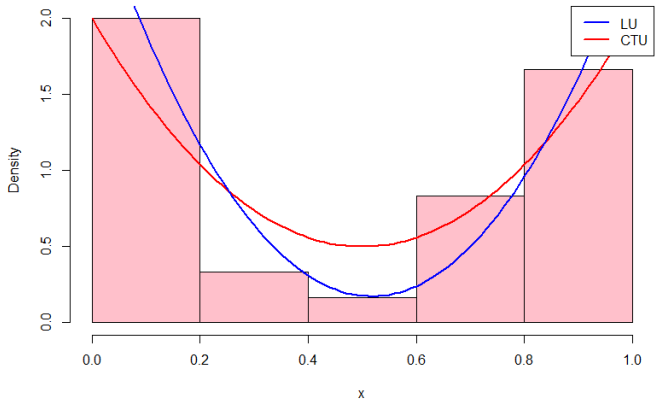


Figure 7: *Estimated pdf of the fitted models for electronic dataset.*

6.2. Application of LK Distribution

In this subsection, the LK distribution is applied to the COVID-19 dataset. This dataset covers the death rates due to COVID-19 infectious in Spain from 3 March to 7 May 2020. This dataset was used by Afify et al. [1] and fitted the Marshall-Olkin reduced Kies (MORKI) distribution, and compared to the Kumaraswamy, Beta, reduced Kies and G-beta distributions. They found that the MORKI distribution was the best fit for this dataset. The dataset is given as follows: 0.3330, 0.5000, 0.5000, 0.5714, 0.2500, 0.3469, 0.4839, 0.2105, 0.2311, 0.3127, 0.4800, 0.2749, 0.3625, 0.3922, 0.3414, 0.3711, 0.4288, 0.4077, 0.3939, 0.4076, 0.4079, 0.4408, 0.4046, 0.3836, 0.3545, 0.3275, 0.3162, 0.3150, 0.3053, 0.2930, 0.2790, 0.2685, 0.2588, 0.2492, 0.2481, 0.2453, 0.2355, 0.2285, 0.2241, 0.2193, 0.2162, 0.2153, 0.2129, 0.2098, 0.2037, 0.2066, 0.2087, 0.2038, 0.2029, 0.2023, 0.1993, 0.1962, 0.1711, 0.1678, 0.1646, 0.1629, 0.1613, 0.1544, 0.1510, 0.1484, 0.1465, 0.1453, 0.1436, 0.1420, 0.1396 and 0.1372.

In this application, we compare the LK distribution with the MORKI and transmuted Kumaraswamy distributions (TK) [9]. The probability density functions of the MORKI and TK are, respectively, given by

$$f_{MORKI}(x) = \frac{\varphi\delta x^{\delta-1}(1-x)^{-\delta-1}e^{-\left(\frac{x}{1-x}\right)^{\delta}}}{\left\{1-(1-\varphi)e^{-\left(\frac{x}{1-x}\right)^{\delta}}\right\}^{2}}, x \in (0,1), \varphi > 0, \delta > 0;$$

$$f_{TK}(x) = abx^{a-1}(1-x^{a})^{b-1}\left[1-\lambda+2\lambda(1-x^{a})^{b}\right], x \in (0,1), a > 0, b > 0, \lambda \in [-1,1].$$

Table 3 lists the MLE estimates of the models fitted to the COVID-19 dataset. Table 4 reveals that the LK distribution is chosen as the best model for this dataset compared to MORKI and TK since it has the lowest AIC and BIC statistics values. Furthermore, the plots of the fitted models in the COVID-19 dataset is presented in Figure 8 and shows that the LK distribution provides a good fit for this dataset.

Table 3: *MLEs of the fitted models for the second COVID-19 dataset.*

Distribution	$\hat{a}$	$\hat{b}$	$\hat{\lambda}$	$\hat{\varphi}$	$\hat{\delta}$
LK	3.6039558	47.6246377	0.4579338		
MORKI				0.04649913	2.95909586
TK	2.884895	21.050565	0.479449		

Table 4: *Numerical values of the AIC and BIC of the models fitted to the COVID-19 dataset.*

Distribution	*AIC*	*BIC*
LK	**-111.994**	**-106.4251**
MORKI	−110.6726	−106.2933
TK	−105.86	−99.29104

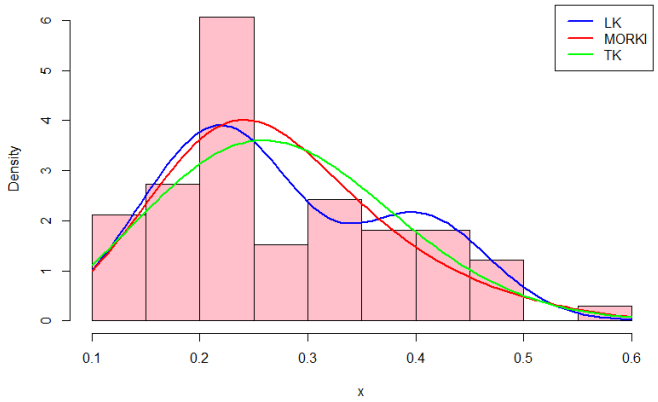


Figure 8: *Estimated pdf of the fitted models for the COVID-19 dataset.*

6.3. Application of LR Distribution

In this subsection, we apply the LR distribution on real data and compare with the following distributions:

- Exponentiated Transformed Inverse Rayleigh distribution (ETIR) [4]

$$f_{ETIR}(x) = \left(\frac{2\sigma^2}{e-1}\right)\frac{1}{x^3}e^{e^{-\left(\frac{\sigma}{x}\right)^2}}, x > 0, \sigma > 0;$$

- Exponentiated Inverse Rayleigh dsitribution (EIRD) [15]

$$f_{EIRD}(x) = \frac{2\alpha\sigma^2}{x^3}e^{-\left(\frac{\sigma}{x}\right)^2}\left(1-e^{-\left(\frac{\sigma}{x}\right)^2}\right)^{\alpha-1}, x \geq 0, \sigma > 0, \alpha > 0.$$

The dataset used in this application is from the study of Hinkley [8]. This data consists of thirty successive observations of the March precipitation (in inches). The dataset is given as follows: 0.77, 1.74, 0.81, 1.2, 1.95, 1.2, 0.47, 1.43, 3.37, 2.2, 3, 3.09, 1.51, 2.1, 0.52, 1.62, 1.31, 0.32, 0.59, 0.81, 2.81, 1.87, 1.18, 1.35, 4.75, 2.48, 0.96, 1.89, 0.9 and 2.05. This data was used by [4] and fitted the ETIR distribution, and compared with the exponentiated inverse Rayleigh, inverse Rayleigh, transmuted inverse Rayleigh and the Rayleigh distributions. They found that the ETIR was the best fit for this dataset.

The results of this analysis are given in Tables 5 and 6. Table 5 presents the MLE estimates of the models fitted to this dataset. Table 6 reveals that the LR distribution gives a better estimate for this dataset since it has the lowest values of AIC and BIC compared to the EIRD and ETIR distributions. In addition, the same result is observed in Figure 9.

Table 5: *MLEs of the fitted models for the Precipitation dataset.*

Distribution	$\hat{\sigma}$	$\hat{\lambda}$	$\hat{\alpha}$
LR	1.7814893	0.1136674	
ETIR	0.8293138		
EIRD	0.8287479		0.7315789

Table 6: *Numerical values of the AIC and BIC of the fitted models for the Precipitation dataset.*

Distribution	*AIC*	*BIC*
LR	**80.59769**	**83.40009**
ETIR	86.05263	87.45383
EIRD	90.40234	93.20474

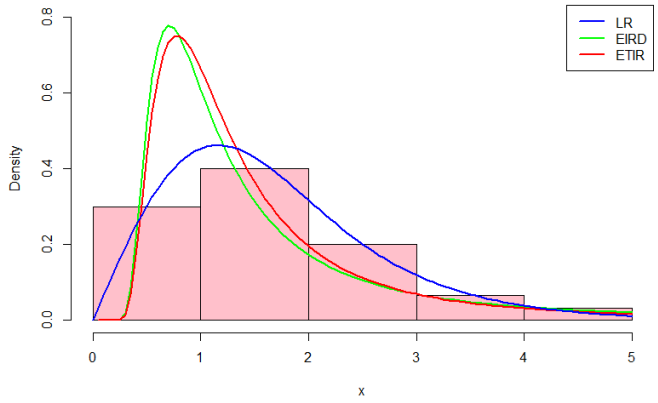


Figure 9: *Estimated pdf of the fitted models for the Precipitation dataset.*

7. Concluding Remarks

In this paper, a new generated family of distributions called the Lakibul-G Family of Distributions has been introduced. Three special distributions of the proposed class, such as the Lakibul-Uniform, Lakibul-Kumaraswamy, and Lakibul-Rayleigh distributions, were derived. Some properties of each of the proposed models, such as its moments, mean, variance, and moment generating function, were derived. The maximum likelihood method was used to estimate the parameters of each of the proposed distributions. Real data sets were used to examine the flexibility and applicability of the proposed distributions. For the Lakibul-Uniform distribution, it was found that this distribution provides a better estimate for the electronic dataset compared to the Cubic Transmuted Uniform distribution. Next, for the Lakibul - Kumaraswamy distribution, it was observed that the LK distribution produces a better fit for the COVID-19 dataset as compared with the Marshall - Olkin reduce Kies and the Transmuted Kumaraswamy distributions. Lastly, for the Lakibul - Rayleigh distribution, it was found that the LR distribution gives better estimates than the Exponentiated Transform Inverse Rayleigh and the Exponentiated Inverse Rayleigh distributions.

References

[1] Afify, A. J., Nassar, M., Kumar, D. and Cordeiro, G. M. (2022). A new unit distribution: properties, imference, and applications. *Electronic Journal of Applied Statistical Analysis*, 15:438–462.

[2] Alzaatreh, A., Lee, C. and Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, 71:63–79.

[3] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12:171–178.

[4] Banerjee, P. and Bhunia, S. (2022).Exponentiated Transformed Inverse Rayleigh Distribution: Statistical Properties and Different Methods of Estimation. *Austrian Journal of Statistics*, 51:60–75.

[5] Cordeiro, G., Ortega, E. M. M. and Cunha, D. C. C. (2013). The exponentiated generalized class of distributions. *Journal of Data Science*, 11:1–27.

[6] Cordeiro, G. M. and de Castro, M. (2011). A new family of generalized distributions. *Journal of Statistica; Computation and Simulation*, 81:883–893.

[7] Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*, 31:497–512.

[8] Hinkley, D. (1977). On Quick Choice of Power Transformation. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 26:67–69.

[9] Khan, M. S., King, R. and Hudson, I. L. (2016). Transmuted Kumaraswamy Distribution. *Statistics in Transition*, 17:183–210.

[10] Lakibul, I. A. and Tubo, B. F. (2023). On the Four-Parameter T-extended Standard U-quadratic Exponentiated Weibull Distribution. *The Mindanawan journal of Mathematics*, 5:17–33.

[11] Lakibul, I. A. and Tubo, B. F. (2023). On the TeSU-G family of distributions applied to lifedata analysis. *Reliability: Theory and Applications*, 18:24–38.

[12] Lakibul, I. A., Polestico, D. L. and Supe, A. P. (2024). On the Bivariate Extension of the extended Standard U-quadratic Distribution. *European Journal of Pure and Applied Mathematics*, 17:790-809.

[13] Nofal, Z. M., Afify, A. Z., Yousof, H. M. and Cordeiro, G. M. (2017). The generalized transmuted - G family of distributions. *Communication Statistics Theory and Methods*, 49:4119–4136.

[14] Rahman, M. M., Al-Zahrani, B. and Shahbaz, S. H. (2019). Cubic Transmuted Uniform Distribution: An Alternative to Beta and Kumaraswamy Distributions. *Journal of Pure and Applied Mathematics*, 12:1106–1121.

[15] Rao, G. S. and Mbwambo, S. (2019). Exponentiated Inverse Rayleigh Distribution and an Application to Coating Weights of Iron Sheets Darta. *Journal of Probability and Statistics*, https://doi.org/10.1155/2019/7519429.

[16] Shaw, W. T. and Buckley, I. R. C. (2007). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *Research report*.

[17] Silva, R., Silva, F. G., Ramos, M., Cordeiro, G. M., Marinho, P. and de Andrade, T. A. N. (2019). The Exponentiated Kumaraswamy-G Class: General Properties and Application. *R. C de Estadistica*, 42:1–33.

RT&A, No 2 (84)
Volume 20, June 2025

MULTIVARIATE IMPUTATION BY MAHALANOBIS DISTANCE OPTIMIZATION (MIMDO)

Geovert John D. Labita[1], Altea S. Labita[2], Bernadette F. Tubo[3]

•

[1,2]University of Science and Technology of Southern Philippines
[3]Mindanao State University - Iligan Institute of Technology
[1]geovertjohn.labita@ustp.edu.ph, [2]altea.labita@ustp.edu.ph, [3]bernadette.tubo@g.msuiit.edu.ph

Abstract

This paper introduces a new method for missing data imputation based on an optimization approach and is now available as an R package called "mimdo". This method deals with imputing the missing values by computing the values that minimize the Mahalanobis distance between an observation and the overall mean. The effectivity of mimdo was demonstrated in both classification and regression tasks using popular benchmark datasets. From all experiments, it was found out that using mimdo for imputing the missing values in the dataset, on the average, the classification rate is more than 80% and an R-squared of more than 50%. Furthermore, the consistency of the results were validated through simulation studies.

Keywords: missing data, cluster analysis, regression analysis, optimization, mahalanobis

1. Introduction

Real-world datasets always are accompanied by missing data which is one major factor affecting data quality [9]. As a result, missing data pose a challenge to most data analysis techniques because a substantial proportion of the data may be missing and predictions must be made for cases with missing inputs. Thus, missing data must be handled properly in order to obtain quality knowledge and also because many statistical models and machine learning algorithms rely on complete datasets.

The risk of missing data is evident in many applications. For example, in cluster analysis, missing values can complicate the application of clustering algorithms. Specifically, in astronomy, imaging sensors have limited sensitivity and may fail to detect light intensities below a minimal threshold frustrating the clustering of celestial bodies [10]. Similarly, according to [2], meteorological or weather stations are riddled with missing climatic data. Thus, the regional grouping of these climatic elements are greatly affected.

In regression analysis, as the amount of data that is missing increases, there can be a substantial reduction of sample size and a resulting loss of power which means that there is a potential for biases in the regression estimates and their standard errors. For example, the problem of non-response to one or more questions in budgetary studies may be very troublesome when the data are to be used in regression analysis [6].

A common practice for dealing with missing values in the context of clustering is to first impute the missing values, and then apply the clustering algorithm on the completed data [5]. However, some of the data imputation methods, in some way, have their drawbacks. For example, the basic data imputation technique like the mean imputation is fundamentally changing the structure of the underlying data. As a consequence, the final clusters obtained are to some extent, a consequence of the decision to replace the missing values by the means rather than the data itself

[4]. Moreover, highly correlated variables is also an issue in data imputation with multivariate imputation by chained equations (mice) as an example. In most cases, mice algorithm will leave these variables out of the imputation process [7].

Inspired from the work of [3], this paper introduces a new methodology for data imputation, which imputes the missing data by computing the values that will minimize the Mahalanobis distance between an observation and the overall mean. By utilizing the Mahalanobis distance, this novel method is also suitable for datasets with high correlation since Mahalanobis distance can handle non-spherical datasets which are commonly caused from highly correlated variables.

This paper is arranged as follows. Methodology is introduced and discussed in section 2. The model solution is presented and derived in section 3 while section 4 illustrates the some applications of *mimdo*. Section 5 presents the concluding remarks.

2. Methods

This section presents the derivation of the optimization model with schematic diagram and *R* installation of *mimdo*. It begins by providing a brief introduction to optimization before proceeding to give a detailed discussion of the methodology.

Definition 1. [1] Let $\boldsymbol{v} \in \mathcal{X} \subset \mathbb{R}^d$ be a vector of decision variables v_i, $i = 1, 2, \ldots, d$ where $\mathcal{X}$ is the ground set of vectors of decision variables. Also, let $f : \mathbb{R}^d \to \mathbb{R} \cup \{\pm\infty\}$ be the objective function and $g_j : \mathbb{R}^d \to \mathbb{R}$ be the constraint function defining restriction on $\boldsymbol{v}$. Then the optimization problem is to

$$\begin{aligned} \min \quad & f(\boldsymbol{v}) && (1)\\ \text{subject to} \quad & g_j(\boldsymbol{v}) \geq 0, \quad && j \in \mathcal{I} \text{ (inequality constraints)} \\ & g_j(\boldsymbol{v}) = 0, \quad && j \in \mathcal{E} \text{ (equality constraints).} \end{aligned}$$

For a maximization problem, the sign of f is changed. In this paper, f can be interpreted as the distance between two observations. In an unconstrained optimization, $\mathcal{I} \cup \mathcal{E} = \emptyset$ and $\mathcal{X} = \mathbb{R}^d$. The set of vectors $S \subset \mathbb{R}^d$ satisfying the constraints in model (1) is called the set of *feasible solutions* to the problem.

To derive the desired optimization model, we first specify the missing and known values for a given dataset $\boldsymbol{X} = \{\boldsymbol{x}_i\}_{i=1}^n$ with p variables, that is,

$$\mathcal{M} = \left\{(i,q) : x_{iq} \text{ is missing, } 1 \leq q \leq p\right\},$$

$$\mathcal{N} = \left\{(i,q) : x_{iq} \text{ is known, } 1 \leq q \leq p\right\}.$$

Also, let J be the set of indices of all incomplete observations given by

$$J = \left\{i \ : \boldsymbol{x}_i \text{ has at least 1 missing coordinate}\right\}.$$

Now, let $\mathbf{W} \in \mathbb{R}^{n \times p}$ be the matrix with imputed values, where w_{jq} is the imputed value for entry x_{jq} for $(j,q) \in \mathcal{M}$. The full imputation for observation $\boldsymbol{x}_j$ is referred to as $\boldsymbol{w}_j$. The idea is to consider the missing data problem as an optimization problem so that the key decision variables are the missing values $\left\{w_{jq} : (j,q) \in \mathcal{M}\right\}$.

Let $\boldsymbol{\mu} = \left\{\mu_1, \ldots, \mu_p\right\}$ be the mean vector and $\boldsymbol{\Sigma}$ be the covariance matrix of a complete dataset, then the squared Mahalanobis distance of the observation $\boldsymbol{w}_i$ from the mean $\boldsymbol{\mu}$ is given by

$$M_i(\boldsymbol{w}_i, \boldsymbol{\mu}) = \begin{bmatrix} w_{i1} - \mu_1 & \cdots & w_{ip} - \mu_p \end{bmatrix} \boldsymbol{\Sigma}^{-1} \begin{bmatrix} w_{i1} - \mu_1 \\ \vdots \\ w_{ip} - \mu_p \end{bmatrix}$$

assuming that the inverse covariance matrix exists. In the presence of missing data, our goal is to solve for the values of $\left\{w_{jq} : (j,q) \in \mathcal{M}\right\}$ that would minimize M_j for each $j \in J$. Thus, if we

let D_j to be the set of indices of variables containing missing values for each $j \in J$, we have the following optimization problem,

$$\min f\left(w_{jd}\right) = M_j\left(\boldsymbol{w}_j, \boldsymbol{\mu}\right) \tag{2}$$

with respect to w_{jd} where $(j,d) \in \mathcal{M}$, that is, $d \in D_j$ for $\left|D_j\right| \leq p$ and $w_{jq} = x_{jq}$ for all $(j,\ q) \in \mathcal{N}$.

Model (2) will be run for a number of iterations with $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ being updated from each iteration. For the first iteration, the initial values of the decision variables $\left\{w_{jq} : (j,q) \in \mathcal{M}\right\}$ are the mean values from all observed data in each variable so that $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ will be computed from this initial imputed dataset. For the second iteration onwards, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ will be updated from the solutions of model (2). As an illustration, the schematic diagram of the procedure is given in Figure 1.

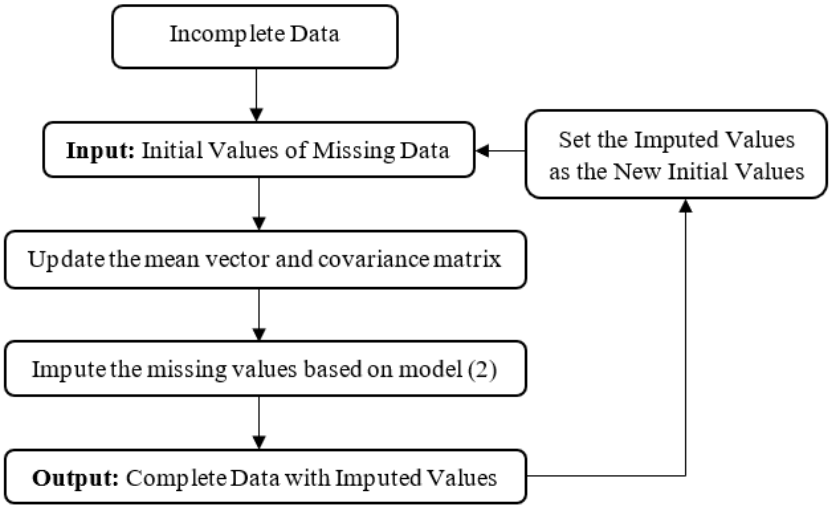


Figure 1: *Schematic Diagram of mimdo*

The default number of iterations is set to 30, but can be adjusted for large datasets to avoid long running time. After the desired number of iterations is performed, the final complete imputed dataset is then obtained based on the average imputed values. The proposed algorithm is now available on the Comprehensive *R* Archive Network (CRAN) as an *R* package called "mimdo: Multivariate Imputation by Mahalanobis Distance Optimization". The implementation is given as follows:

```
install.packages("mimdo")
library(mimdo)
incomplete_data<-read.csv("file_path.csv")
#the first  row of the csv file must be the variable names

complete_data<-mimdo(as.data.frame(incomplete_data),inverse=TRUE)
```

3. Results

This section presents the solution and properties of *mimdo*.

Proposition 1. For the case $\left|D_j\right| = 1$, let $\boldsymbol{X} = \{\boldsymbol{x}_i\}_{i=1}^{n}$ be a given incomplete dataset with p variables. Also, let $\boldsymbol{W} \in \mathbb{R}^{n \times p}$ be the matrix with imputed values where w_{iq} is the imputed value for entry x_{iq}. If $(j,d) \in \mathcal{M}$, then the solution of the optimization problem (2) is given by

$$w_{jd} = \mu_d - \frac{1}{\sigma^*_{dd}} \sum_{a:a \neq d}^{p} \sigma^*_{da}\left(w_{ja} - \mu_a\right) \tag{3}$$

where $\mu_a, \mu_d, \sigma^*_{da} \in \mathbb{R}$ and $\sigma^*_{dd} > 0$.

Proof. Let $(j,d) \in \mathcal{M}$ and consider the optimization model (2). Now, we have to solve for w_{jd} that would minimize

$$M_j = \begin{bmatrix} w_{j1}-\mu_1 & \cdots & w_{jq}-\mu_q & \cdots & w_{jp}-\mu_p \end{bmatrix} \mathbf{\Sigma}^{-1} \begin{bmatrix} w_{j1}-\mu_1 \\ \vdots \\ w_{jq}-\mu_q \\ \vdots \\ w_{jp}-\mu_p \end{bmatrix}.$$

Suppose that

$$\mathbf{\Sigma}^{-1} = \begin{bmatrix} \sigma^*_{11} & \cdots & \sigma^*_{1q} & \cdots & \sigma^*_{1p} \\ \vdots & & \vdots & & \vdots \\ \sigma^*_{q1} & \cdots & \sigma^*_{qq} & \cdots & \sigma^*_{qp} \\ \vdots & & \vdots & & \vdots \\ \sigma^*_{p1} & \cdots & \sigma^*_{pq} & \cdots & \sigma^*_{pp} \end{bmatrix}$$

where the diagonal elements are positive, then

$$M_j = \sum_{b=1}^{p} \sum_{a=1}^{p} \sigma^*_{ab} \left(w_{ja}-\mu_a\right) \Big(w_{jb}-\mu_b\Big).$$

To differentiate M_j, we have to separate the terms containing w_{jd} and using the symmetric property of the inverse covariance matrix, we have,

$$M_j = \sigma^*_{dd}\Big(w_{jd}-\mu_d\Big)^2 + 2\sum_{a:a\neq d}^{p} \sigma^*_{da}\Big(w_{jd}-\mu_d\Big)\left(w_{ja}-\mu_a\right) + \sum_{b:b\neq d}^{p}\sum_{a:a\neq d}^{p} \sigma^*_{ab}\left(w_{ja}-\mu_a\right)\Big(w_{jb}-\mu_b\Big)$$

$$\Rightarrow D_{w_{jd}}\left(M_j\right) = 2\sigma^*_{dd}\Big(w_{jd}-\mu_d\Big) + 2\sum_{a:a\neq d}^{p} \sigma^*_{da}\left(w_{ja}-\mu_a\right).$$

Finally, equating the derivative to 0 will solve for the imputed value, that is,

$$2\sigma^*_{dd}\Big(w_{jd}-\mu_d\Big) + 2\sum_{a:a\neq d}^{p} \sigma^*_{da}\left(w_{ja}-\mu_a\right) = 0$$

$$\Rightarrow\ w_{jd} = \mu_d - \frac{1}{\sigma^*_{dd}} \sum_{a:a\neq d}^{p} \sigma^*_{da}\left(w_{ja}-\mu_a\right).$$

■

For the case when $|D_j| > 1$, we take the partial derivative with respect to w_{jd} for each $d \in D_j$, set to zero and obtain a system of equations of the form

$$\sigma^*_{d1}w_{j1} - \sigma^*_{d1}\mu_1 + \cdots + \sigma^*_{dd}w_{jd} - \sigma^*_{dd}\mu_d + \cdots + \sigma^*_{d,|D_j|}w_{j,|D_j|} - \sigma^*_{d,|D_j|}\mu_{|D_j|} + \sum_{q\notin D_j} \sigma^*_{dq}\left(w_{jq}-\mu_q\right) = 0$$

where we rewrite the subscripts as $\{1,2,\ldots,|D_j|\}$ corresponding to the indices with missing values. Thus, in matrix form, we have,

$$\begin{bmatrix} \sigma^*_{11} & \sigma^*_{12} & \cdots & \sigma^*_{1,|D_j|} \\ \sigma^*_{21} & \sigma^*_{22} & & \sigma^*_{2,|D_j|} \\ \vdots & & \ddots & \vdots \\ \sigma^*_{|D_j|,1} & \sigma^*_{|D_j|,2} & \cdots & \sigma^*_{|D_j|,|D_j|} \end{bmatrix} \begin{bmatrix} w_{j1} \\ w_{j2} \\ \vdots \\ w_{j,|D_j|} \end{bmatrix}$$

$$= \begin{bmatrix} \sigma^*_{11}\mu_1 + \sum_{q\in D_j\setminus\{1\}} \sigma^*_{1q}\mu_q - \sum_{q\notin D_j} \sigma^*_{1q}\left(w_{jq}-\mu_q\right) \\ \sigma^*_{22}\mu_2 + \sum_{q\in D_j\setminus\{2\}} \sigma^*_{2q}\mu_q - \sum_{q\notin D_j} \sigma^*_{2q}\left(w_{jq}-\mu_q\right) \\ \vdots \\ \sigma^*_{|D_j|,|D_j|}\mu_{|D_j|} + \sum_{q\in D_j\setminus\{|D_j|\}} \sigma^*_{|D_j|,q}\mu_q - \sum_{q\notin D_j} \sigma^*_{|D_j|,q}\left(w_{jq}-\mu_q\right) \end{bmatrix}$$

in which the solution are the imputed values for w_{jd}, $d \in D_j$.

To show global optimality of w_{jd} in (3), we need to show that f is convex. This can easily be seen from the proof of Proposition 1 where

$$f = \sigma^*_{dd}\left(w_{jd}-\mu_d\right)^2 + 2\sum_{a:a\neq d}^{p} \sigma^*_{da}\left(w_{jd}-\mu_d\right)\left(w_{ja}-\mu_a\right) + \sum_{b:b\neq d}^{p}\sum_{a:a\neq d}^{p} \sigma^*_{ab}\left(w_{ja}-\mu_a\right)\left(w_{jb}-\mu_b\right)$$

which is a parabola. To verify this, the abscissa of the vertex of the parabola must contain the solution w_{jd}. Let $T = w_{jd}-\mu_d$, $\alpha = \sigma^*_{dd}$, $\beta = 2\sum_{a:a\neq d}^{p} \sigma^*_{da}\left(w_{ja}-\mu_a\right)$, and $\gamma = \sum_{b:b\neq d}^{p}\sum_{a:a\neq d}^{p} \sigma^*_{ab}\left(w_{ja}-\mu_a\right)\left(w_{jb}-\mu_b\right)$, then

$$f = \alpha T^2 + \beta T + \gamma.$$

Now, the abscissa of the vertex of the parabola can be solved using the formula

$$T = -\frac{\beta}{2\alpha} = -\frac{2\sum_{a:a\neq d}^{p} \sigma^*_{da}\left(w_{ja}-\mu_a\right)}{2\sigma^*_{dd}} = -\frac{1}{\sigma^*_{dd}}\sum_{a:a\neq d}^{p} \sigma^*_{da}\left(w_{ja}-\mu_a\right)$$

$$\Rightarrow w_{jd}-\mu_d = -\frac{1}{\sigma^*_{dd}}\sum_{a:a\neq d}^{p} \sigma^*_{da}\left(w_{ja}-\mu_a\right)$$

$$\Rightarrow \qquad w_{jd} = \mu_d - \frac{1}{\sigma^*_{dd}}\sum_{a:a\neq d}^{p} \sigma^*_{da}\left(w_{ja}-\mu_a\right)$$

which is the exact solution from Proposition 1. Thus, f is convex having w_{jd} as the global minimum. Figure 2 illustrates the model f with its global minimum in two-dimensional space assuming $\left|D_j\right| = 1$.

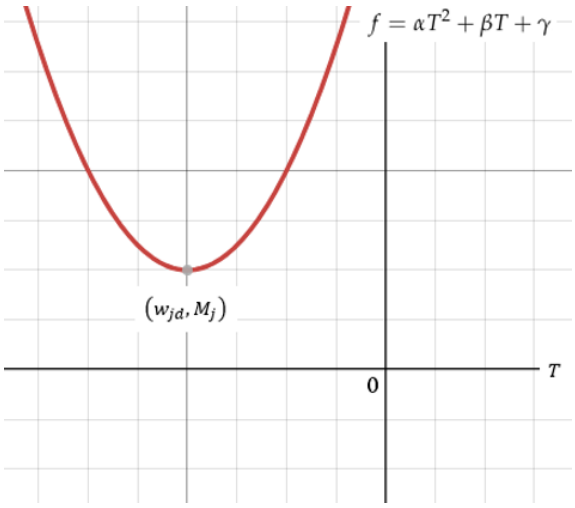


Figure 2: *Graph of Model (2) for the case* $\left|D_j\right| = 1$

As we can see from Figure 2, the vertex contains the global minimum which is the solution w_{jd} and the corresponding minimum Mahalanobis distance M_j. The next result shows the convergence of the solution of *mimdo*.

Conjecture 1. Let f be the objective function given in (2). If $f^{(t)}$ is the objective function value at iteration t, then

$$\lim_{t\to\infty}\left(f^{(t)}-f^{(t-1)}\right)=0.$$

Conjecture 1 implies that as the number of iteration increases, the objective function in model (2) will likely to converge to a specific value. This means that for each $j \in J$, the imputed values $\left\{w_{jd}\right\}, d \in D_j$ converges to specific values. This conjecture will be shown in the application in the next section.

4. Application

For demonstration, *mimdo* is applied on popular datasets taken from the UCI (University of California Irvine) Machine Learning Repository which have been widely used by students, educators, and researchers all over the world. These datasets are the "Iris Flower" and "Yacht Hydrodynamics" downloaded at *https://archive.ics.uci.edu/datasets* which will be used for the empirical analysis of data imputation algorithms and is a good point of reference for assessment or comparison of different imputation methods.

In doing the experiment, incomplete datasets are generated (5%-30%) from complete datasets using MCAR (missing completely at random) and MAR (missing at random) missing data mechanisms from the *R* Package *missMethods*. For classification (Iris Flower), the percentage of missing data is generated per cluster while for regression (Yacht Hydrodynamics), the percentage of missing data is generated directly from the whole dataset. Specifically, 100 incomplete datasets will then be generated per missing rate. For assessment purposes, the missing elements among the data are then imputed using *mimdo*, *mice*, and *knn* (k-nearest neighbor) where *mice* is set to one single imputation with 30 maximum iterations (*mice* - *R* Package *mice*, *knn* - *R* Package *bnstruct*).

For classification task, the traditional K-means algorithm is applied into the imputed dataset and obtain the clustering accuracy (classification rate) based on the original cluster labels. For regression analysis, the imputed dataset is regressed with respect to the dependent variable and obtain the R-squared. The average result from 100 incomplete datasets is then reported. Figure 3 shows the average classification rate results using Iris dataset while Figure 4 shows the average R-squared results using Yacht dataset.

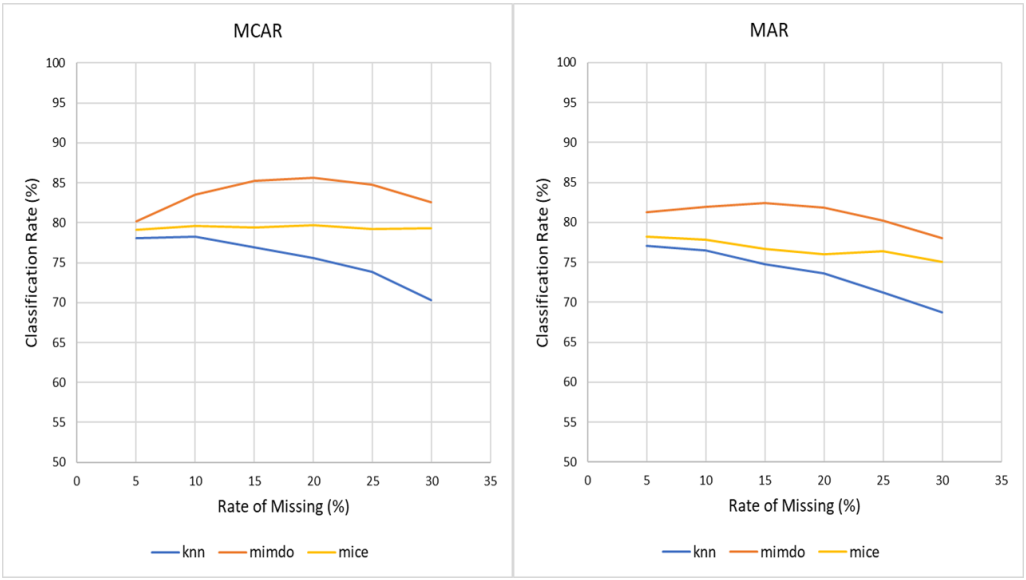


Figure 3: *Classification Rate Results from Iris Dataset*

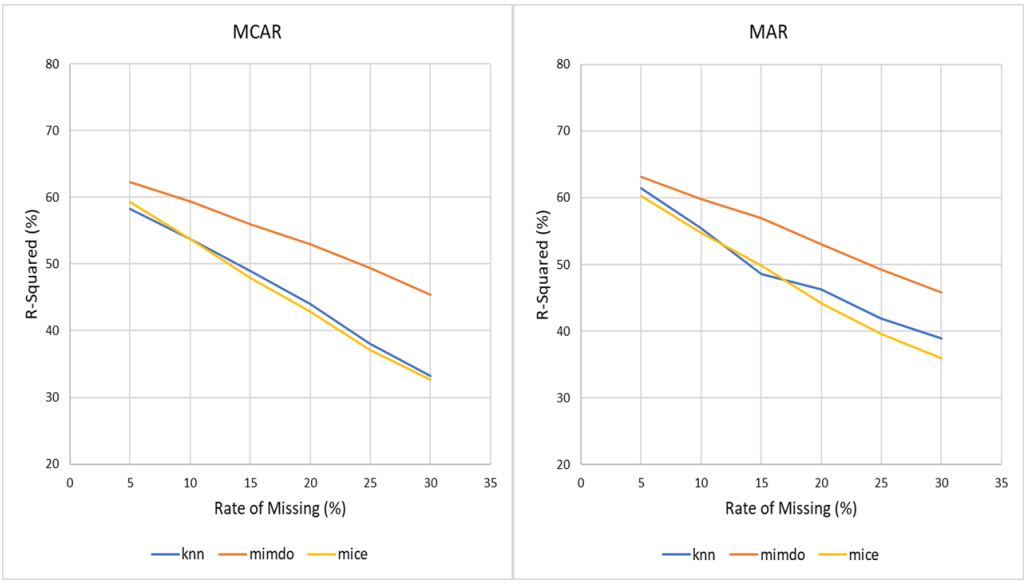


Figure 4: *Regression Results from Yacht Dataset*

It can be observed from Figures 3 & 4 that the results from *mimdo* are higher especially with MCAR assumption. Consistency of this result will be explored in section 4.2 using simulation studies.

4.1. Convergence of Solutions

This section shows the corresponding results of the convergence of the solutions of *mimdo*. This means that as the iteration increases, the objective function in model (2) will likely to converge to a specific value. For example, Figure 5 shows the average objective value differences between each consecutive iterations.

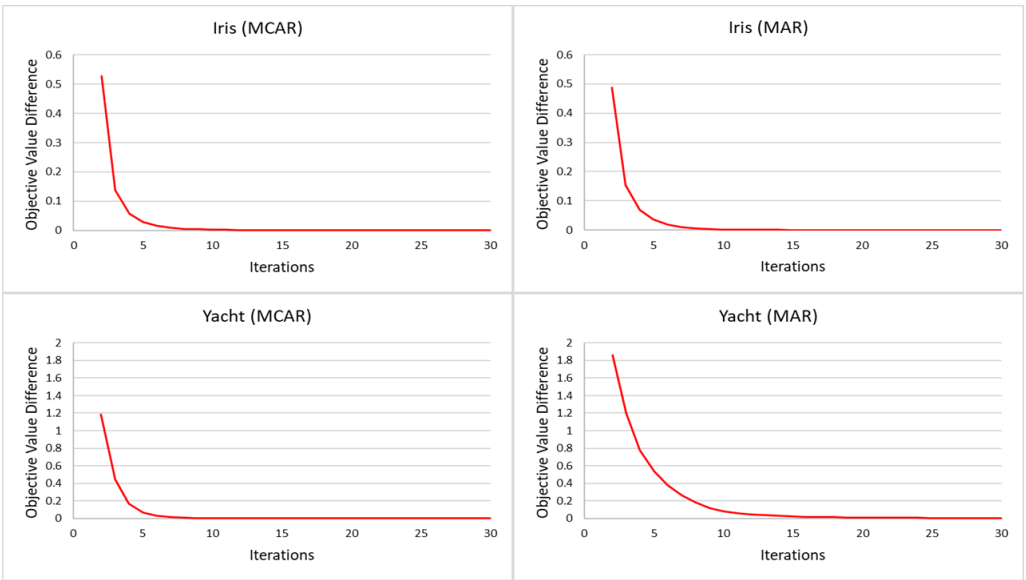


Figure 5: *Overall Average Variation*

It can be observed from Figure 5 that the graph decreases eventually to zero which indicates the consistency of the solution (imputed values) of *mimdo* as stated in Conjecture 1.

4.2. Simulated Datasets

In this section, we generate simulated datasets that have the same distributions and correlations as the existing datasets considered using the *simdf()* function from the *R* Package *faux*. Specifically, 100 simulated datasets were generated for each dataset considered and the average results were reported. Figures 6 & 7 show the average results using simulated Iris and simulated Yacht datasets respectively.

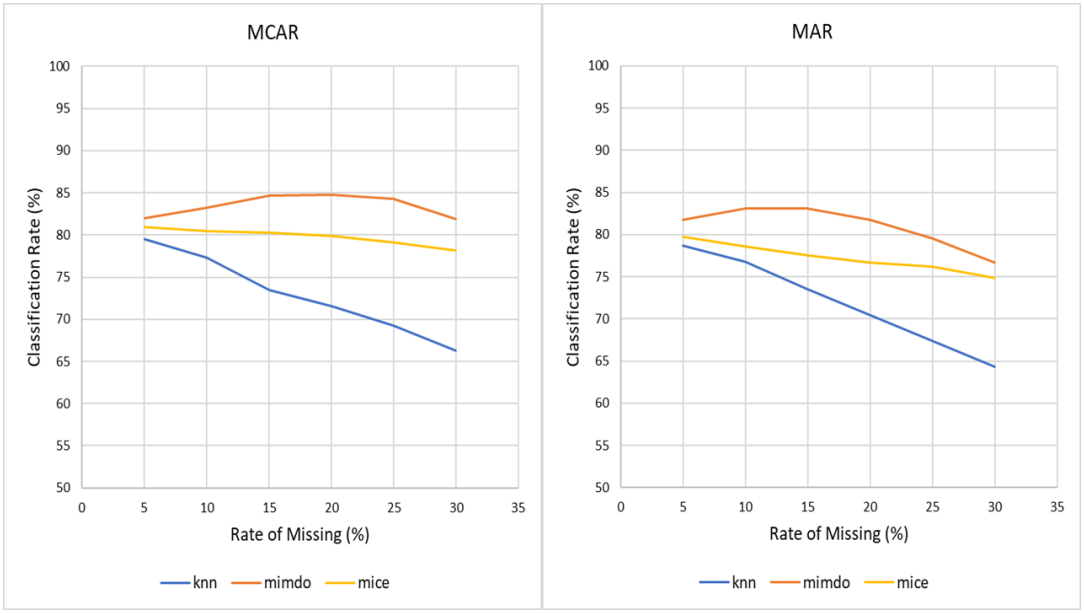


Figure 6: *Classification Rate Results from Simulated Iris Dataset*

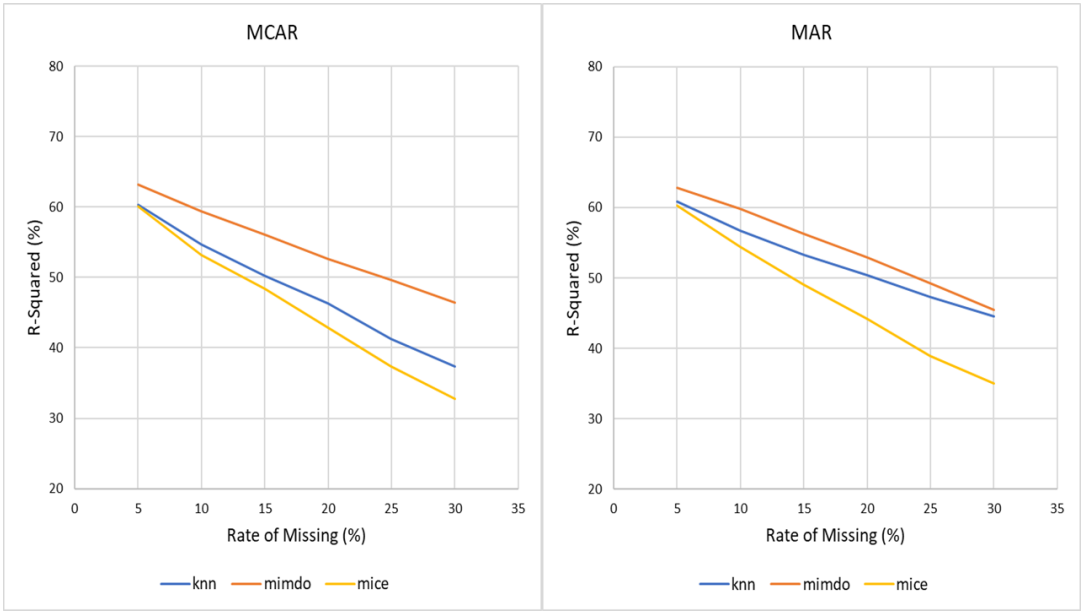


Figure 7: *Regression Results from Simulated Yacht Dataset*

It can be observed from Figures 6 & 7 the similarity of the results from the original dataset. This shows consistency of *mimdo* for this type of datasets.

4.3. *R* Package "mimdo"

This section presents the documentation of the *R* Package *mimdo*. It is built for the purpose of imputing highly correlated datasets up to the maximum tolerance of covariance matrix invertibility. However, if the covariance matrix is non-invertible, the option inverse=FALSE might not work very well. Thus, for this case, we recommend feature selection before doing *mimdo*.

Description

Imputes missing values of an incomplete data matrix by minimizing the Mahalanobis distance of each sample from the overall mean. By utilizing Mahalanobis distance, this imputation method is preferable to be used on datasets with highly correlated variables.

Usage

```
mimdo(incomplete_data, inverse, iterations = 30)
```

Arguments

`incomplete_data`	A data frame with missing values.
`inverse`	If TRUE, the inverse covariance matrix will be used for distance calculation. If the covariance matrix is non-invertible, use inverse = FALSE.
`iterations`	Number of iterations. It can be adjusted to avoid long running time.

Value

The output returns a data frame of the complete imputed data. This means that the missing values of the original incomplete dataset have been imputed. If the function does not return a value, this means that the covariance matrix is not invertible and is exactly singular.

5. Conclusions

The problem about missing data increases significantly in many fields especially if the dataset have highly correlated variables. For some cases, other data imputation algorithms will leave those variables out of the imputation process. The use of Mahalanobis distance in the imputation process provides an efficient estimation of missing data regardless of the correlation structure of the dataset.

In this paper, package *mimdo* accessible via R interface was applied to benchmark datasets and simulated datasets. On the average, using benchmark datasets, it achieved a clustering accuracy of 82.3% (classification task) and an R-squared of 54.4% (regression task) while using simulated datasets, it obtained 82.2% clustering accuracy and 54.5% R-squared. This means that the imputation method applied to this type of datasets is consistent.

The consistency of the solution (imputed values) of *mimdo* can also be observed from Figure 5 wherein the graphs decreased eventually to zero which means that the objective function in (2) will eventually converge.

References

[1] Andreasson, N., Evgrafov, A., & Patriksson, M. (2005). An introduction to optimization: Foundations and fundamental algorithms. *Chalmers University of Technology Press: Gothenburg, Sweden*, *1*, 1-205.

[2] Barrios, A., Trincado, G., & Garreaud, R. (2018). Alternative approaches for estimating missing climate data: application to monthly precipitation records in South-Central Chile. *Forest Ecosystems*, *5*(1), 1-10.

[3] Bertsimas, D., Pawlowski, C., & Zhuo, Y. D. (2017). From predictive methods to missing data imputation: an optimization approach. *J. Mach. Learn. Res.*, *18*(1), 7133-7171.

[4] Bock, T. (2013). 5 ways to deal with missing data in cluster analysis. *DISPLAYR Blog*.
[5] Boluki, S., Zamani Dadaneh, S., Qian, X., & Dougherty, E.R. (2018, August). Optimal clustering with missing values. In *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics* (pp. 593-594).
[6] Haitovsky, Y. (1968). Missing data in regression analysis. *Journal of the Royal Statistical Society: Series B (Methodological)*, *30*(1), 67-82.
[7] Katitas, A. (2019). Getting Started with Multiple Imputation in R. *University of Virginia Library-Statlab Articles*, https://uvastatlab.github.io/2019/05/01/getting-started-with-multiple-imputation-in-r.
[8] Labita, GJ.D. (2024). Package 'mimdo'. *Comprehensive R Archive Network (CRAN)*. doi:10.32614/CRAN.package.mimdo
[9] Song, Q. & Shepperd, M. (2007). Missing data imputation techniques. *International journal of business intelligence and data mining*, *2*(3), 261-291.
[10] Wagstaff, K. & Laidler, V. (2005). Making the most of missing values: Object clustering with partial data in astronomy. *Astronomical Data Analysis Software and Systems XIV*, 347, 172-176.

ENHANCING PREDICTIVE ACCURACY WITH ROBUST LIU AND RANDOM FOREST REGRESSION

Muthukrishnan. R and Karthika Ramakrishnan

•

Department of Statistics, Bharathiar University, Coimbatore, 641046, Tamil Nadu, India
muthukrishnan70@buc.edu.in, karthikaramakrishnan45@gmail.com

Abstract

Robust and Random Forest regression procedures are powerful machine learning algorithms that enhance predictive accuracy and handle complex datasets with diverse characteristics. The Robust Liu regression algorithm is designed to address the limitations of conventional regression methods. The robust Liu regression method is considered more reliable in the presence of outliers and multicollinearity. On the other hand, Random Forest is an ensemble learning method that constructs multiple decision trees to improve classification and regression tasks by aggregating predictions from individual trees, thereby mitigating overfitting and increasing robustness to noise. Its ability to capture nonlinear relationships and handle high-dimensional data makes it suitable for many real-world applications. This study compares the performance of Least Squares, Liu, Robust Liu, and Random Forest regression methods using prediction error measures in real and simulated environments. This synergy offers new possibilities for researchers to improve prediction accuracy in the presence of heterogeneous non-normal, multicollinear data and outliers.

Keywords: Liu, Multicollinearity, Outliers, Random Forest, Robust

I. Introduction

The Least Squares regression method is commonly employed to estimate the parameters of a regression model, assuming all model assumptions are met. However, issues such as multicollinearity and outliers can distort the results of this method. Multicollinearity, as described by Farrar and Glauber [8], occurs when there is a high correlation among the independent variables. This increases error values and reduces the effectiveness of the estimator. On the other hand, an outlier is an unusual data point that can reduce the estimator's efficiency and alter the direction of regression coefficients. Chatterjee and Hadi [5] noted that outliers could impact parameter estimation and lead to inaccurate predictions in traditional methods. To address multicollinearity, the Liu regression procedure was introduced. For data that deviates from fundamental assumptions, robust regression provides an alternative to conventional models. Additionally, Random Forest, a supervised learning algorithm, can be applied to both classification and regression problems. This study compares different regression methods in the presence of multicollinearity and outliers in datasets.

The rest of the paper is structured as follows. Various regression procedures like Least Squares, Liu, Random Forest and Robust Liu are explained briefly in section 2. An experimental study is conducted under real and simulated datasets to compare different error measures such as Mean Square Error, Mean Absolute Error, Median Absolute Error and Mean Absolute Percentage Error for various regression methods in section 3 and section 4 will give the conclusion.

II. Regression Methods

Regression analysis, as stated by Draper and Smith [7], is a method used to derive insights from data by examining the relationship between response and predictor variables. Various forms of regression techniques exist in machine learning, and their application depends on the nature of the data involved. It serves as a key approach for addressing problems in machine learning through data modelling. This study focuses on methods such as Least Squares, Liu, Random Forest, Ridge, and Liu regression, comparing their error measures across different real-world datasets that include both outliers and multicollinearity. Outliers are identified using Cook's distance procedure, and the analysis is conducted with the help of R software.

Least Squares Regression

The Least Squares (LS) is a standard procedure in regression modelling for estimating the parameters of a linear model. This method is used to estimate the dependent variable (y) using a number of predictor variables (X). It is a widely utilized and optimal linear unbiased estimator when all the assumptions of the classical regression model are satisfied. The standard model of LS with r independent variables is represented as follows.

$$y = X\beta + \varepsilon \tag{1}$$

where y is a $(qx1)$ vector of response variables, X is an (qxr) matrix of predictors, β is a $(rx1)$ vector of unknown regression parameters, and ϵ is an $(rx1)$ vector of residuals assumed to be independently and identically distributed as normal with a mean of zero and a fixed variance σ^2. The LS estimator for the unknown parameter is given by

$$\widehat{\beta_{LS}} = (X'X)^{-1}(X'y) \tag{2}$$

The performance of the LS estimator $\widehat{\beta_{LS}}$ becomes statistically insignificant when multicollinearity exists among the explanatory variables.

Liu Regression

Liu Estimator is a class of biased estimators used to deal with datasets having multicollinearity. It was described by Liu [11]. These estimators are depending upon a biasing parameter d called the Liu parameter which lies between 0 and 1. The regression estimator of this procedure is given by

$$\widehat{\beta_{Liu}} = (X'X + I_q)^{-1}(X'y + d\,\widehat{\beta_{LS}}) \tag{3}$$

where $0 \leq d \leq 1$, I_q is the identity matrix of order $q \times q$ and $\widehat{\beta_{LS}}$ is the LS estimator. The biasing parameter d of Liu is computed by the formula,

$$\hat{d} = 1 - \hat{\sigma}^2 \left[\frac{\sum_{i=1}^{p} \frac{1}{\lambda_i (\lambda_i + 1)}}{\sum_{i=1}^{p} \frac{\widehat{\beta_i}^2}{(\lambda_i + 1)^2}} \right] \quad (4)$$

where, $\hat{\sigma}^2$ and $\widehat{\beta_i}^2$ are the mean square error and the regression estimates computed via LS respectively. $\lambda_1, \lambda_2, \ldots, \lambda_q$ are the eigenvalues of the matrix $X'X$. $\widehat{\beta_{Liu}}$ is called the Liu estimator by Akdeniz and Kaciranlar [2]. The estimator having a d value with minimum mean square error is considered an efficient estimator when compared with the other values.

Random Forest Regression

The Random Forest (RF) regression technique introduced by Breiman [4] is a versatile machine learning technique for predicting numerical values. It combines the predictions of multiple decision trees to reduce overfitting and improve accuracy. The development of Random Forests was a response to the limitations of traditional decision trees, which tend to overfit the data. It provides better predictive accuracy by averaging the results of multiple trees. The Random Forest model for regression is given in equation (5), where t is the total number of trees in the forest and $\hat{y}$ is the final predicted value.

$$\widehat{y_t} = f_t(x) \quad (5)$$

Robust Liu Regression

The robust regression methods were developed to overcome the limitations of traditional regression procedures. Under a normal distribution without outliers, this robust method should yield results similar to LS. A Robust Liu (RLiu) regression method described by Muthukrishnan and Karthika [14] was developed to deal with the datasets having both multicollinearity and outliers by incorporating the properties of both Liu and MM regression procedures. The MM estimator is a robust regression technique introduced by Yohai [19], used to estimate parameters in the presence of outliers. It is a modification of the M-estimator, designed to provide robustness and high efficiency. The estimator of the RLiu regression is given by

$$\widehat{\beta_{RLiu}} = \left(X'X + I_q\right)^{-1} (X'y + d_{MM}\widehat{\beta_{MM}}) \quad (6)$$

where $0 \leq d_{MM} \leq 1$, I_p is the identity matrix of order $r \times r$, $\widehat{\beta_{MM}}$ is the MM estimator. The biasing parameter $d_{MM}$of RLiu is computed using the equation (4) by replacing $\hat{\sigma}^2$with $\widehat{\sigma_{MM}}^2$, the mean square error and $\widehat{\beta_i}^2$are the regression estimates computed via MM respectively.

III. Experimental Results

This section presents numerical analyses conducted on both real and simulated datasets. The first real dataset exhibits moderate multicollinearity with outliers. The second dataset displays high multicollinearity alongside outliers. Outliers in the real datasets were detected and eliminated using Cook's distance method introduced by Cook [6], and the analyses were performed using R software. A statistical technique called the Variance Inflation Factor (VIF) by Frisch [9] can detect and measure the amount of multicollinearity in a multiple regression model. The VIF assesses how much the regressors collectively impact the variance of each term within the model. The computed error measures under real datasets, discussed in Cases 1 and 2, based on with and without outliers of various regression procedures are given in Table 1 and the corresponding results under simulation is presented in Table 2.

Table 1: *Computed error measures under various regression methods (Real Datasets)*

Methods	Prostate Cancer Dataset				Hald Dataset			
	MSE	MDAE	MAE	MAPE	MSE	MDAE	MAE	MAPE
LS	0.46 (1.5)	0.68 (0.58)	0.56 (0.48)	2.23 (2.17)	3.68 (2.57)	1.29 (2.31)	0.78 (1.23)	0.05 (0.26)
Liu	0.14 (0.13)	0.50 (0.41)	0.55 (0.47)	2.21 (2.10)	1.19 (1.02)	1.44 (2.23)	0.75 (0.96)	0.04 (0.25)
RF	0.15 (0.12)	0.52 (0.47)	0.33 (1.05)	0.87 (0.79)	0.16 (0.47)	1.31 (2.20)	0.33 (1.05)	0.12 (0.69)
RLiu	0.13 (0.12)	0.28 (0.26)	0.26 (0.90)	0.14 (0.44)	0.14 (0.44)	1.23 (0.88)	0.26 (0.90)	0.03 (0.07)

(.) Without Outliers

Case1. Prostate Cancer Dataset: The data come from a study that looked at how males undergoing radial prostatectomy correlated their level of prostate-specific antigen with several clinical measures. This data set has 97 observations. There are seven independent variables namely lweight (log of prostate weight), age, lbph (log of benign prostatic hyperplasia amount), svi (seminal vesicle invasion), lcp (log of capsular penetration), gleason (Gleason score), lpsa (log of prostate specific antigen) and one dependent variable lcavol (log of cancer volume). Seven outliers are found in this dataset. Since the VIFs of the independent variables are in between 1 and 5, there is an indication of moderate multicollinearity.

Case 2. Hald Dataset: Woods et al [18] introduced the Hald or Portland Cement Dataset. This data frame contains 13 observations with four independent variables. They are tricalcium aluminate (tca), tricalcium silicate (tcs), tetracalcium aluminoferrite (tcaf) and β-dicalcium silicate (bdcs). The response variable is the evolved heat after 180 days in a cement mix. Since the VIFs of this Hald data set were greater than 10, the explanatory variables are highly correlated. As a result, the dataset has high multicollinearity. Also, this data set has one outlier identified and removed using Cook's distance method.

Simulation studies were carried out to examine the efficiency of various regression procedures. In the study, the data was generated from a multivariate normal distribution with mean $\mu = [0]_{px1}$and the variance $\sum = [\sigma_{ij}]$ for the level of correlation, $\rho = 0.99$ and number of variables $p = 5$. Different levels of contamination, $\varepsilon = 0\%$, 5% and 10% were studied for sample size $n = 50, 100\ and\ 200$. The performance of various regression procedures were compared by computing different error measures and the results obtained for different number of observations with various levels of contamination are shown in Table 2.

The results obtained from Tables 1 and 2 demonstrates that the computed error measures under various regression procedures are slightly different from each other. Also the Robust Liu method has the smallest error measures compared with the others. Hence Robust Liu (RLiu) regression technique is more efficient than the others in the case of datasets having indication of multicollinearity and has outliers.

Table 2: *Computed error measures under various regression methods (Simulation Dataset)*

n	ε	Methods	MSE	MDAE	MAE	MAPE
50	0%	LS	6.20	3.11	4.03	0.58
		Liu	4.01	2.88	2.78	0.28
		RF	2.64	1.61	1.26	0.15
		RLiu	1.78	1.05	1.23	0.14
	5%	LS	7.05	2.88	4.13	0.48
		Liu	5.96	2.62	3.14	0.34
		RF	5.23	2.47	3.13	0.47
		RLiu	4.88	1.02	1.55	0.18
	10%	LS	10.29	4.83	3.74	0.70
		Liu	7.32	2.62	3.36	0.68
		RF	7.31	2.55	3.24	0.61
		RLiu	3.28	1.19	1.75	0.24
100	0%	LS	6.01	2.55	3.89	0.58
		Liu	4.73	2.38	3.02	0.52
		RF	2.41	2.25	2.99	0.49
		RLiu	2.36	1.78	1.07	0.15
	5%	LS	4.92	4.46	3.48	0.69
		Liu	3.91	2.31	2.97	0.62
		RF	2.87	2.25	2.87	0.51
		RLiu	2.42	1.84	1.22	0.36
	10%	LS	6.37	2.96	3.15	1.02
		Liu	4.99	2.35	2.76	0.85
		RF	4.82	2.21	2.73	0.72
		RLiu	2.78	1.81	1.15	0.63
200	0%	LS	7.07	3.14	3.63	1.92
		Liu	5.85	2.62	3.19	1.43
		RF	3.72	2.59	3.08	1.19
		RLiu	1.32	1.57	1.79	0.95
	5%	LS	9.33	4.67	3.88	0.78
		Liu	8.27	2.97	3.45	0.61
		RF	2.99	2.70	3.15	0.52
		RLiu	2.11	1.67	1.98	0.38
	10%	LS	9.69	5.24	4.01	0.92
		Liu	7.18	3.16	3.32	0.86
		RF	6.93	2.35	2.89	0.79
		RLiu	2.87	1.76	1.05	0.62

IV. Conclusion

Statistical learning techniques are important in various research fields, with regression analysis being a prominent method. The commonly used linear regression procedures will not be sufficient to build a regression model when data deviates from the modelling assumptions. Hence, there is a need of alternatives to build a good model for the given datasets. This paper explores several regression methods, including Least Squares, Liu, Random Forest and Robust Liu. Further, evaluates their performance on different real and simulated datasets by considering the problems of multicollinearity and outliers by computing various error measures. On the basis of the computed error measures, this study concludes that the Robust Liu regression method provides better estimates for datasets having both multicollinearity and/or outliers. This approach can be

particularly advantageous for researchers employing machine learning techniques that need to account for these factors.

References

[1]Aitken, A. C. (1935). On least Squares and linear combinations of observations. *Proceedings of the Royal Statistical Society*, Edinburgh, 55: 42-48.

[2]Akdeniz, F. and Kacıranlar, S. (1995). On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE. *Communications in Statistics, Theory and Methods*, 24: 1789–1797.

[3]Arslan, O. and Billor, N. (2000). Robust Liu estimator for regression based on an M-estimator. *Journal of applied statistics*, 27: 39-47.

[4] Breiman, L. (2001). Random forests. *Machine learning*, 45: 5-32.

[5] Chatterjee, S. and Hadi, A. S. Sensitivity Analysis in Linear regression. John Wiley & Sons, New York, 2009.

[6]Cook, R. D. (2000). Detection of influential observation in linear regression. *Technometrics*, 42: 65-68.

[7] Draper, N. R. and Smith, H. Applied Regression Analysis. John Wiley & Sons, New York, 1998.

[8]Farrar, D. E. and Glauber, R. R. (1967). Multicollinearity in Regression Analysis: The Problem Revisited. *The Review of Economics and Statistics*, 49: 92-107.

[9]Frisch, R. Statistical confluence analysis by means of complete regression systems. University of Oslo, 1934.

[10]Huber, P. H. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35: 7-101.

[11]Liu, K. J. (1993). A new class of biased estimate in linear regression. *Communications in Statistics*, 22: 393-402.

[12]Mendenhall, W. and Sincich, A. (2014). Second Course in Statistics: Regression Analysis, 7th ed (Harlow: Pearson), 105–123.

[13]Muthukrishnan, R. and Karthika Ramakrishnan (2024). Effect of Classical and Robust Regression Estimators in the context of High dimensional data with Multicollinearity and Outliers. *Reliability: Theory & Applications*, 19: 335-341.

[14] Muthukrishnan, R. and Karthika Ramakrishnan (2024). A New Robust Liu Regression Estimator for High-Dimensional Data. *Reliability: Theory & Applications*, 19:214-219.

[15] Muthukrishnan, R. & Maryam Jamila, S. (2020). Predictive Modeling Using Support Vector Regression, *International Journal of Scientific and Technology Research*, 9: 4863-4865.

[16] Rousseeuw, P. J and Leroy, A. M. Robust Regression and Outlier Detection, John Wiley & Sons, 1987.

[17] Susanti, Y., Pratiwi, H., Sri Sulistijowati, H. and Liana, T. (2014). M estimation, S estimation and MM estimation in Robust Regression. *International Journal of Pure and Applied Mathematics*, 91: 349-360.

[18] Woods, H., Steinour, H. H. and Starke, H. R. (1932). Effect of composition of Portland cement on heat evolved during hardening. *Industrial and Engineering Chemistry*, 24: 1207–1214.

[19] Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. *The Annals of Statistics*, 642-656.

ELECTROMECHANICAL DEVICES WITH LEVITATION ELEMENTS FOR CONTROL OF NON-ELECTRICAL PARAMETERS

G.S.Kerimzade

•

Azerbaijan State Oil and Industry University, Baku, Azerbaijan
gulschen98@mail.ru

Abstract

The paper presents a mathematical model of electromechanical devices with levitation elements for automatic control of non-electrical parameters, as well as calculation stages for any devices. The mathematical model is based on the equations of levitation, currents, magnetomotive forces, overheating temperature of the windings, and magnetic induction. Numerical values of dimensionless quantities are determined, which can be used as reference data in the development of electromechanical devices with levitation elements. The functional dependences of electromagnetic parameters on the control voltage and stroke of the levitation element are established, and expressions for calculating the force of gravity of the levitation element, lifting force, magnetomotive force and winding overheating temperature are obtained.

Key words: electromechanical device, control, non-electrical parameter, levitation element, mathematical model, lifting force, magnetomotive force, overheating temperature, stroke, magnetic induction, dependence, control voltage.

I. Introduction

The main unit of electromechanical devices with levitation elements are magnetic systems containing a magnetic circuit, an excitation winding and a levitation element. When the voltage at the excitation winding terminals automatically changes from the minimum value U_{min} to the maximum U_{max}, the levitation element moves freely from the initial position x=0 to the final position $x=x_{max}$. The main requirements for these devices are to ensure constancy of the current in the windings, the stroke of the levitation element, the permissibility of the temperature without violating the conditions of levitation, that is, the equality of lifting force and gravity. The current level of development of science and technology has increasingly expanded the scope of application of electromechanical devices with levitation elements used for automatic control of non-electrical parameters of technological processes, for precise stabilization of alternating current in control loads, as well as for precise transmission of displacements [1]. Electromechanical devices with levitation elements (LE) for automatic control of non-electrical parameters make it possible to combine the functions of control, measurement and stabilization of electrical and non-electrical quantities. The main indicators and purpose of these devices are given in Table 1. They are characterized by simplicity of design, accuracy, and stable performance characteristics [18].

Table 1. Indicators and purpose of electromechanical devices with LE

№	Name	Indicators
1	2	3
1	Force converter	$I(P_x)$; $U_2(P_x)$- to control external forces.
2	Displacement transducer	I(x); U_2(x)- to control vertical movements ni order 5÷50mm.
3	Executive controlled mechanism	I(x); U_2(x)- to automatically change the position of the working mechanism within 5÷50mm.; the volta ge at the terminals is regulated automatically by the control winding.
4	Steerable support	x(U_1); I_1(M)- for smooth position control we eat sup ports for working mechanisms.
5	Sealant	I_1(x); x(U_1)- for sealing external elements with a gi ven force .
6	Multi-nomial AC stabilizer	$I'_{load}(U_1)$; I″load(U_1); $I'''_{load}(U_1)$-for precise stabilityti- on of alternating current in loads, connected in series with winding sections awakening.
7	Motion transmission device	$x_2(x_1)$- allows you to transfer movement x_1=5÷50mm. with precision 0.1-0.5 mm.; the movement at the output x_2 can be greater than, less than or equal to the movement of the transmission x_1, which is ensured by switching sections of the field winding.
8	Stabilizer for controlling mechanical stretching of wires	$P_{x2}(P_{x1})$;$U_n(P_{x1};P_{x2})$- allows you to provide a given va lue of the tension force of wires of different section when winding coils; works like tracking system for transmitting mechanical forces .

II. Formulation of the problem

The purpose of the presented work is to derive analytical expressions for the main functional dependencies and parameters, taking into account the range of control voltage changes, the overheating temperature of the windings, and the levitation coordinates, which are entered into the calculated mathematical models [2]. Calculation of the main parameters and dimensions of electromechanical devices with LE is carried out on the basis of a joint solution of expressions for the electromagnetic force F_e, induction in steel B_M, current I of the magnetomotive forces F_1, F_2 and the overheating temperature of the windings τ_1, τ_2:

$$F_e = \frac{1}{2}\left(I_1 W_1\right)^2 \frac{d\Lambda}{dx} P_T \tag{1}$$

$$B_M = \frac{\sqrt{2}k_u U_{\max}}{\omega k_c S_c W_1}; I_1 = \frac{k_u U_1}{\omega L_1} \tag{2}$$

$$F_2 = I_2 W_2 = b_2 F_1 = b_2 I_1 W_1 \tag{3}$$

$$\tau_1 = \frac{I_1^2 r_1}{k_T S_{1cool}}; \tau_2 = \frac{I_2^2 r_2}{k_T S_{2cool}} \tag{4}$$

Expressions (1) - (4) are used in calculations and design of electromechanical devices with LE, and the notations used are generally accepted.

When solving problems, the following assumptions are made:

1) magnetic fluxes and induction change according to a sinusoidal law;
2) there is no saturation of steel, magnetic resistance is low;
3) the magnetic resistance of the LE to the main working flow Φ_1 is significant;
4) there are no friction forces;
5) heat losses in the LE and excitation winding are caused by convection;
6) the lifting force and gravity act on the LE, the levitation condition is satisfied:
   $F_e = P_T$.

III. Problem solution

According to (1)-(4), the calculation of the main parameters and dimensions of electromechanical devices with LE can be carried out on the basis of a joint solution of expressions for force, induction, current, magnetomotive forces, and overheating temperature. The initial data for the calculation are: the minimum and maximum values of the control voltage U_{min}, U_{max}, the induction value in steel B_M corresponding to the maximum voltage value and the maximum stroke of the LE X_M, the overheating temperature of the windings τ_1 and τ_2 [20]. The main functional dependencies of electromagnetic parameters - current, electromagnetic force, stroke, levitation coordinates on voltage, displacement are presented in figure 1. Physical modeling of the magnetic field of the working gap establishes that when the conditions are met $m_a = 2 \div 6$; $m_c = 2 \div 6$ ensures the uniformity of the magnetic field in the working gap. The system is a linear inductance, since the functional dependence of the inductance on the control voltage and stroke of the LE is linear. The winding and LE currents do not depend on the control voltage (figure 1,a). [5] With increasing voltage, the levitation coordinate of the LE and the inductance of the excitation winding increase; with increasing stroke of the LE and decreasing voltage, the height of the winding increases (figure 1,b).

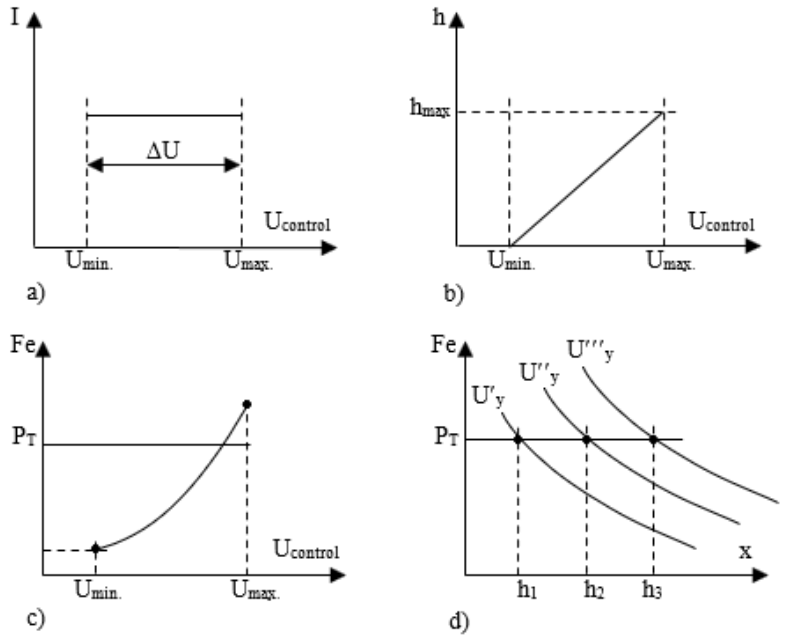


Figure 1. *Basic functional dependencies of electromagnetic parameters*

The calculation of electromechanical devices with LE consists of the following steps:
Calculation of the inductance of the magnetic system:

$$L_1 = L_{1s} + W_1^2 \lambda x \cos^2 \theta, \tag{5}$$

here θ- steel loss angle; L_{1s} and x – leakage inductance of the excitation winding and levitation coordinate of the levitation winding [7-9].
At $\cos^2\theta \approx 1$ the inductance of the magnetic system $L_1 = L_{1s} + W_1^2 \lambda x$. According to expression (5), the specific conductivity of the gap in the magnetic circuit window is determined as : $\lambda = 2\mu_0 m_c \sigma_b$, here $m_c = b/c$; σ_b - coefficient taking into account the bulging of magnetic fluxes along parallel magnetic cores,

$$\sigma_b = 1 + \frac{2.96}{m_c} \ln\left(1 + \frac{\pi}{m_a}\right).$$

Determination of the relationship between lift force, LE stroke and control voltage (figure 1,c,d):

$$F_e = \frac{k_u U_{cont.}^2}{2\omega^2 W_1^2 \lambda \left(x + \frac{h_1}{3n_\lambda}\right)^2}, \tag{6}$$

here $\lambda W_1^2 = \frac{dL_1}{dx}$ - derivative of the equivalent inductance of the system with respect to the movement of the LE [6].
Calculation of magnetomotive forces (MMF) of winding and lifting force :

$$F = \frac{1}{2} I_1 \left(\lambda I_1 W_1^2\right) = \frac{k_u \Delta U I_1}{2\omega x_{\max}}; x_{\max} = \frac{k_u \Delta U}{\omega W_1^2 \lambda I_1};$$

$$F_1 = I_1 W_1 = A_1 m_1 c_2^2; W_1 = \frac{A_1 m_1 c_2^2}{I_1}; \tag{7}$$

$$F_2 = b_2 F_1 = b_2 A_1 m_1 c_2^2.$$

Coefficient A_1 is determined through the initial data, the dimensionless coefficient m_1 is determined through the geometric relationships m_a, m_c, n_{02} (m_a = 2÷6, m_c = 2÷6; $n_{02} = c/c_2 \approx 1.1$), graphical dependencies are shown in figure 2 [12].

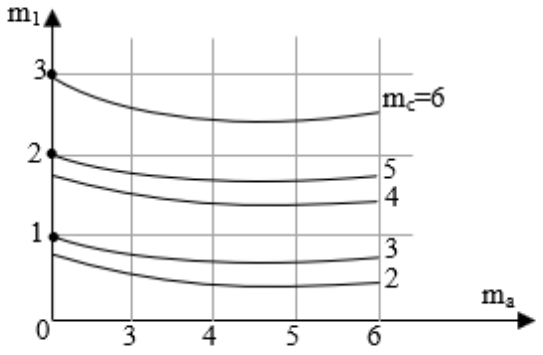


Figure 2. *Graphs of dimensionless coefficients*

LE power calculation:

$$P_2 = I_2^2 r_2 \tag{8}$$

Since the active resistance r_2 through the overheating temperature τ_2 is determined according to the expression: $r_2 = \Delta' r_{20} = \rho_0 \frac{l_{cp2} W_1^2}{k_{32} S_{02}} \left(\Delta_{ok} + \alpha_M \tau_2 \right)$, then expression (8) can be written as :

$$P_2 = \left(I_2 W_2 \right)^2 \rho_0 \frac{l_{cp2} \Delta'}{k_{32} S_{02}},$$

here $\Delta' = \Delta_{o\kappa} + \alpha_M \tau_2 = 1 + \alpha_M (\theta_{o\kappa} - 20) + \alpha_M \tau_2$.
$\Delta_{o\kappa} = 1 + \alpha_M (\theta_{ok} - 20)$

According to expressions (1) - (4), the ampere turns of the levitation element are defined as:

$$\left(I_2 W_2 \right)^2 = 2 b_2^2 \frac{P_T}{\lambda}$$

Determination of functional dependencies of geometric dimensions on overheating temperature:

$$\tau_2 = \left(I_2 W_2 \right)^2 \frac{\rho_0 l_{cp2} \Delta'}{k_{32} k_T S_{02} S_{2cool}}; \frac{\tau_2}{\Delta'} = \frac{\tau_2}{\Delta_{ok} + \alpha_M \tau_d} \tag{9}$$

Determination of gravity and thickness of LE:

$$P_T = A_2 c_2^3 m_p; c_2 = \left(\frac{A_2}{\mu_0 A_1^2} \right) \left(\frac{m_p}{m_1} \right) \tag{10}$$

To determine the thickness of the LE c_2 or the working air gap c, the values of the coefficients m_a, m_c, n_{e2} are the initial data ($m_a = 2 \div 6$; $m_c = 2 \div 6$; n_{e2} $1 \div 6$; $\tau_2 = 50 \div 100$) [15-17]. The coefficient n_{e2} allows you to adjust the calculated temperature value limits excessive increase in LE height.

A generalized design method has been developed, and the functional dependencies of electromechanical devices with LEs for various purposes have been systematized. The main technical indicators are considered, a mathematical model is compiled, analytical expressions are obtained for the main dimensions, taking into account the specified values of the overheating temperature, input and output parameters, the condition of uniformity of the magnetic field in the air gap, and the optimal values of the dimensions of the magnetic core are determined [19,22].

IV. Conclusions

The stages of calculating electromagnetic parameters of electromechanical devices for automatic control of non-electrical parameters are considered. The functional dependencies of these parameters have been established, which are universal for any electromechanical devices with levitation elements. When developing electromechanical devices with levitation elements, certain coefficient values can be used as reference data.

The presented mathematical model takes into account the levitation coordinate, the range of control voltage changes, the overheating temperature of the windings, and also makes it possible to establish functional dependencies of the parameters on the initial data, and obtain analytical expressions for the main dimensions and parameters of the magnetic system. The expressions are convenient for optimizing and analyzing the characteristics of electromechanical devices with levitation elements.

References

[1]G.S. Kerimzade, G.V.Mamedova." Main indicators and purpose of electromechanical devices forautomatic control of non-electrical parameters". // "Elektromexanika" № 5. 2016. pp.31-37.

[2] G.V.Mamedova, G.S. Kerimzade, N.M.Piriyeva."Electromagnetic calculation of tension devices for winding wires of small cross sections", International Journal on technical and Physical Problems of Engineering (IJTPE), Issue 53, Vol. 14, No.4, pp. 80-85, December 2022.

[3] G.V.Mamedova, G.S. Kerimzade, N.M.Piriyeva."Issues of electrical devices with levitation elements", International Journal on technical and Physical Problems of Engineering (IJTPE), Issue 56, Vol. 15, No.3, pp.120-125, September 2023.

[4] N.M.Piriyeva,G.S. Kerimzade."Methods for increasing electromagnetic efficiency in induction levitator", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT. ISSN0033-2097, R.99 NR 10/2023. Warszawa. pp.192-195.

[5] N.M.Piriyeva,G.S. Kerimzade." Electromagnetic efficiency in induction levitators and ways to improve it", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT.ISSN0033-2097, R.99 NR 06/2023. Warszawa. pp.204-207.

[6] N.M.Piriyeva,G.S. Kerimzade." Mathematical model for the calculation of electrical devices based on induction levitators", International Journal on technical and Physical Problems of Engineering (IJTPE), Issue 55, Vol.15, No.2, pp.274-280, June 2023.

[7] Ya.R. Abdullaev, G.S. Kerimzade, "Design of Electric Devices with LE", Electrical Engineering, No. 5, pp.16-22, 2015.

[8] G.S. Kerimzade, G.V. Mamedova, "Working Modes for Designing Electrical Apparatuses with Inductionmi with Levitation Elements", Bulletin Buildings, Vol. 17, No. 1, pp. 42-46, Baku, Azerbaijan, 2015.

[9] G.S. Kerimzade, G.V. Mamedova, "Analysis of the Parameters of Electric Devices with Levitational Elements", News of Universities Instrumentation, No. 12, Vol. 61, pp. 67-71, Sankt Peterburg, Russia, 2018.

[10] G.S. Kerimzade, "Analytical Connections of the Parameters and sizes of the Precision Stabilizer of Alternating Current Using the Effect of Inductive Levitation", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 52, Vol. 14, No. 3, pp. 175-184, September 2022.

[11] G.S. Kerimzade, "Analysis of the Methodology for Calculation Current Stabilizer with Induction levitation", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 53, Vol.14, No. 4, pp. 170-174, December 2022.

[12] G.S. Kerimzade, "Analytical Expressions of the Relationship to the Calculation of the AC Stabilizer with Induction Levitation", International Journal on technical and Physical Problems of Engineering (IJTPE), Issue 54, Vol. 15, No. 1, pp. 135-139, March 2023.

[13] G.S. Kerimzade, "Controlled AC Stabilizers on the Principle of Induction Levitation", International Journal Universum, No. 2, Issue 107, Vol. 5, pp. 31-35, Moscow, Russai, 2023.

[14] G.S. Kerimzade, "High Precision Current Stabilizers with Induction Levitation", Journal of Renewable Energy, Electrical, and Computer Engineering, e-ISSN: 2776-0049, Vol. 3, No. 1, pp. 26-31, Indonesia, March 2023.

[15] G.S. Kerimzade, "Choice of Expressions of the Method of Calculation of the Current Stabilizer on the Principle of Induction Levitation", News of Azerbaijan High Technical Educational Institutions. 2023. Volume. 4, pp. 71-76.

[16] G.S. Kerimzade, "Optimization of parameters of AC stabilizer on the principle of induction levitation", The 19th International Conference on "TPE". ICTPE-2023. Number 06. Code 01PES06. 31 October 2023, pp.32- 36.

[17] G.S.Kerimzade."Dependence of the overall dimensions of the control induction support", Vestnik nauki. Issue 2(71), Volume 1 . February, 04-2024.Tolyatti. pp.613-618.

[18] G.S. Kerimzade."Calculation of parameters of control induction support", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT.ISSN 0033-2097,R.100.NR.05/2024. Warszawa.. pp.219-221.

[19] G.V.Mamedova, G.S. Kerimzade."Design parameters for electromechanical devices with a levitation element", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT.ISSN 0033-2097,R.100. NR 09 / 2024.Warszawa. pp.111-113.

[20] G.V.Mamedova, G.S. Kerimzade."Relay contactor system as a means of control ling a linear electric drive", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT.ISSN 0033-2097, R.100. NR 11/ 2024. Warszawa.

[21] G.S. Kerimzade." Method for determining the overall dimensions of an induction control support", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT. ISSN 0033-2097,R. 100. NR 06 /2024. Warszawa.. pp.235-237.

[22] G.S.Kerimzade, G.V.Mamedova. " Research of electromechanical devices with levitation elements in control systems", Journal "Realibility: Theory & Application". ISSN 1932-2321. Volume 19. № 2 (78). June 2024. pp.85-90.

GEOMETRIC HAZARD-BASED PREVENTIVE MAINTENANCE AND REPLACEMENT (GHBPMAR) MODEL FOR AGING MECHANICALLY REPAIRABLE MACHINES

•

[1]Nse Udoh; [2]Iniobong Uko and [3]Fredrick Ohaegbunem

[1,2&3]Department of Statistics, University of Uyo, Nigeria
nsesudoh@uniuyo.edu.ng, iniobonguko144@gmail.com, ohaegbunemfredugo@gmail.com

Abstract

This work aims to develop a geometric hazard-based preventive maintenance model to improve the sustainability of aging mechanically repairable machines particularly in developing countries. The goal is to optimize maintenance and replacement schedules to meet user needs effectively. The model utilizes a geometric process as a hazard rate adjustment factor to account for increasing deterioration rates of machines due to aging and prolong usage and applies same model to a cassava grinding machine across three phases of lifespan: average lifespan, beyond average lifespan, and beyond initial replacement age. The Weibull failure distribution is used to characterize the machine's increasing failure rate. The result reveals that the machine's operational time decreases following preventive maintenance, and maintenance costs increase as the machine progresses through many phases of its lifespan. This highlights the model's effectiveness in managing the cost and timing of maintenance activities. The model's implementation can significantly impact communities in developing countries by improving the reliability and lifespan of essential machinery like cassava grinding machines and other industrial machines which can lead to stable local economies. This approach helps in managing costs and improving the sustainability of machines, especially in resource-constrained environments over extended lifespan.

Keywords: Preventive maintenance, replacement model, geometric process, repairable system, aging, Weibull hazard rate

I. Introduction

Preventive maintenance (PM) entails all necessary actions required to keep a system at a working condition before due maintenance time. It reduces the number of failures during operation and is therefore essential in keeping to production target. It may be done at fixed time interval or sequentially at different time points which may be of different lengths. The common assumption of PM is that the machine will requires minimal repair if it breaks down before PM action.

Many PM models have been developed over the years including [1] and [2] which introduced the concept of adjustment/improvement factors for both the hazard rate and effective age of repairable systems after PM. Furthermore, periodic and sequential imperfect PM model was developed by [3] by adopting the adjustment/improvement factors into the hazard rate and effective age for a sequential PM policy. Later, [4] proposed the hazard model which ensures that the hazard rate of the system is reduced after PM action and the age model which ensures that the effective or

virtual age of the system is improved after PM. A general sequential PM model was developed by [5] which combines the age reduction model and hazard rate adjustment model of [4] to exploit the advantages of both models. In view of the peculiarity of developing countries of the world, where repairable systems are used longer than the specified period with increased usage, a geometric imperfect preventive maintenance and replacement (GIPMAR) model for aging repairable systems was proposed by [6], which is a generalization of the hybrid model by [5]. The geometric process (GP) was use as the hazard rate adjustment factor to characterize and regulate the increasing hazard rate at various degrees of deterioration denoted by the index of the GP model. The proposed models were used to obtain optimal PM and replacement schedule for repairable systems at three phases of operation; namely: average lifespan, beyond average lifespan and after initial replacement age.

The use of GP in modelling the stochastic increase in repair times after failures and stochastic decrease of the life time of a component of a system after repair was propagated by [7]. Thereafter, authors in [8], [9] and [10] among others have successfully applied the GP models to describe the stochastic increase or decrease in the life time of a system. Hence, this work aims at incorporating the geometric hazard rate adjustment factor of [6] into the hazard rate-based preventive maintenance model of [4] in a bid to generalize the model to meet the needs of developing countries who import and use second-hand (fairly used) products beyond the specified age limits. This model will be used to obtain the optimal PM and replacement schedule for three life phases of a cassava grinding machine which PM and replacement schedule at what is regarded as phase 1 was earlier obtained in [11] but not for the other phases with higher degrees of deterioration consequent upon the peculiarity of developing countries, where equipment is used beyond normal specified lifespan. Hence, this work seeks to extend the hazard model of [4] to cover the maintenance of these items in the extended phases of operation for sustainability.

II. Methods

2.1 Assumptions of ageing mechanically repairable machines

1. The machine undergoes PM at successive times, $x_1 + x_2 + \ldots + x_{N-1}$ and is replaced at x_N.
2. The hazard rate $h(t)$ after the kth PM becomes $\lambda_k h(t)$, when it was $h(t)$ and increases with the number of PMs. In other words, the system becomes "better-than-old" after PM and as-good-as-new at replacement.
3. The repairable machine undergoes only minimal repair at failures between PMs and the hazard rate remains unchanged after minimal repair.
4. Preventive maintenance (PM), repair and replacement times are negligible.
5. The hazard rate, $h(t)$ is continuous and strictly increasing.

2.2 Phases of operation for ageing mechanically repairable machines

The working life of the machine is classified into three phases:

Phase 1: Machine is working within the specified age limit with normal degree of deterioration and low cost of maintenance. The failure degree is denoted by $j = 2$.

Phase 2: Machine is working beyond specified age limit due for initial replacement, it is on extended life span and has high degree of deterioration with associated high cost of maintenance. The failure degree is denoted by $j = 3$.

Phase 3: Machine is working beyond *phase 2*, it is on further extended life span, it also exhibits higher degree of deterioration due to aging and usage, and has associated higher cost of maintenance. The failure degree is denoted by $j = 4$.

2.3 The proposed Geometric hazard-based PM and Replacement (GHBPMAR) model

Let the GP model given by $H_j(t) = \lambda^{j-1} t \ ; j \geq 2$, where λ is the GP parameter. In this work, we let λ^{j-1} be the hazard rate adjustment factor for repairable machines with higher degrees of deterioration due to aging and usage beyond age limit at various levels of operation, where $j - 1$

denotes a measure of failure degree of repairable machines due to aging and usage at extended lifespan. Hence, the geometric hazard model is given in (1) as;

$$\lambda_k{}^{j-1}h(t) \tag{1}$$

Equation 1 shows that the hazard rate, $h(t)$ becomes $\lambda_k{}^{j-1}h(t)$ after PM, according to [6], where $\lambda \geq 1, t \in (0, t_1), x \in (0, t_2 - t_1)$. Also, t is the time point and x is the interval between two successive PM actions. The geometric hazard rate adjustment factor, $\lambda_k{}^{j-1}$ in (2) according to [6] is defined as;

$$\lambda_k^{j-1} = \left(\frac{6k+1}{5k+1}\right)^{j-1}, k = 1,2,3,\ldots, j \geq 2 \tag{2}$$

Accordingly, the modified cost function from [4] for the geometric hazard model is given in (3) as;

$$C = C(x_1, x_2 x_3, \ldots x_N) = \frac{C_m \sum_{k=1}^{N} \lambda_k^{j-1} H(x_k) + (N-1)C_p + C_r}{\sum_{k=1}^{N} x_k} \tag{3}$$

where C_m, C_p and C_r are the cost parameters respectively due to minimal repairs, preventive maintenance and corrective maintenance.

2.4 Minimization of the cost function of ageing repairable machines

By applying the minimization condition to (3), we have;

$$\frac{\partial C}{\partial x_k} = \frac{[\sum_{k=1}^{N} x_k] C_m \lambda_k^{j-1} h(x_k) - \left[(N-1)C_p + C_r + C_m \sum_{k=1}^{N} \lambda_k^{j-1} H(x_k)\right]}{[\sum_{k=1}^{N} x_k]^2} = 0$$

$$C_m \lambda_k^{j-1} h(x_k) - C = 0$$

$$C = C_m \lambda_k^{j-1} h(x_k) \tag{4}$$

From (3) we have that;

$$C \sum_{k=1}^{N} x_k = C_m \sum_{k=1}^{N} \lambda_k^{j-1} H(x_k) + (N-1)C_p + C_r \tag{5}$$

Substituting (4) into (5), we have;

$$C_m \lambda_k^{j-1} h(x_k) \sum_{k=1}^{N} x_k = C_m \sum_{k=1}^{N} \lambda_k^{j-1} H(x_k) + (N-1)C_p + C_r$$

By utilizing the hazard and cumulative hazard functions of Weibull failure function, respectively given as;

$$h(x_k) = \alpha\beta^{\alpha} x_k^{\alpha-1}, H(x_k) = \beta^{\alpha} x_k^{\alpha}$$

$$\lambda^{j-1} \alpha\beta^{\alpha} x_k^{\alpha-1} \sum_{k=1}^{N} x_k - \sum_{k=1}^{N} \lambda^{j-1} \alpha\beta^{\alpha} x_k^{\alpha-1} = \frac{(N-1)C_p + C_r}{C_m} \tag{6}$$

Let $h(x_k) = \theta$, then, $\alpha\beta^{\alpha} x_k^{\alpha-1} = \theta$ and;

$$x_k = \left[\frac{\theta}{\alpha\beta^{\alpha}}\right]^{\frac{1}{\alpha-1}} \tag{7}$$

Substitute (7) into (6);

$$x_k = \left[\frac{\theta}{\alpha\beta^{\alpha}}\right]^{\frac{1}{\alpha-1}} \quad \lambda_k^{j-1} \alpha\beta^{\alpha} \left[\frac{\theta}{\alpha\beta^{\alpha}}\right] \sum_{k=1}^{N} \left[\frac{\theta}{\alpha\beta^{\alpha}}\right]^{\frac{1}{\alpha-1}} - \sum_{k=1}^{N} \lambda_k^{j-1} \alpha\beta^{\alpha} \left[\frac{\theta}{\alpha\beta^{\alpha}}\right]^{\frac{\alpha}{\alpha-1}} = \frac{(N-1)C_p + C_r}{C_m}$$

$$\lambda_k^{j-1}\alpha\beta^{\alpha}\sum_{k=1}^{N}\left[\frac{\theta^{\frac{\alpha}{\alpha-1}}}{(\alpha\beta^{\alpha})^{\frac{1}{\alpha-1}}}\right]-\sum_{k=1}^{N}\lambda_k^{j-1}\alpha\beta^{\alpha}\left[\frac{\theta^{\frac{\alpha}{\alpha-1}}}{(\alpha\beta^{\alpha})^{\frac{1}{\alpha-1}}}\right]=\frac{(N-1)C_p+C_r}{C_m}$$

$$\theta=\left\{\frac{(N-1)C_p+C_r}{C_m\left[\frac{1}{\alpha\beta^{\alpha}}\right]^{\frac{1}{\alpha-1}}\left\{N\lambda_k^{j-1}-\frac{1}{\alpha}\sum_{k=1}^{N}\lambda_k^{j-1}\right\}}\right\}^{\frac{\alpha}{\alpha-1}} \tag{8}$$

We seek the optimal number, N^* of PM by minimizing (9);

$$B(N)=\frac{(N-1)C_p+C_r}{\left\{N\lambda_k^{j-1}-\frac{1}{\alpha}\sum_{k=1}^{N}\lambda_k^{j-1}\right\}} \tag{9}$$

A necessary condition that there exist a finite N^* minimizing $B(\mathrm{N}^*)$ is that N^*satisfies the inequality: $D(\mathrm{N}^*)=B(\mathrm{N}^*+1)\geq B(\mathrm{N}^*)$ and $D(\mathrm{N}^*)=B(\mathrm{N}^*)<B(\mathrm{N}^*-1)$. We show that $D(\mathrm{N}^*)=B(\mathrm{N}^*+1)\geq B(\mathrm{N}^*)$ as follows:

$$D(\mathrm{N}^*)=\frac{\mathrm{N}^*C_p+C_r}{\left\{(\mathrm{N}^*+1)\lambda_k^{j-1}-\frac{1}{\alpha}\sum_{k=1}^{\mathrm{N}+1}\lambda_k^{j-1}\right\}}\geq\frac{(\mathrm{N}^*-1)C_p+C_r}{\left\{\mathrm{N}^*\lambda_k^{j-1}-\frac{1}{\alpha}\sum_{k=1}^{\mathrm{N}}\lambda_k^{j-1}\right\}};\ k=1,2,\ldots$$

$$=\left\{\mathrm{N}^*\lambda_k^{j-1}-\frac{1}{\alpha}\sum_{k=1}^{N}\lambda_k^{j-1}\right\}\left[\mathrm{N}^*C_p+C_r\right]\geq\left[(\mathrm{N}^*-1)C_p+C_r\right]\left\{(\mathrm{N}^*+1)\lambda_k^{j-1}-\frac{1}{\alpha}\sum_{k=1}^{N+1}\lambda_k^{j-1}\right\}$$

$$=\mathrm{N}^*C_p\left\{\mathrm{N}^*\lambda_k^{j-1}-\frac{1}{\alpha}\sum_{k=1}^{N}\lambda_k^{j-1}\right\}+C_r\left\{\mathrm{N}^*\lambda_k^{j-1}-\frac{1}{\alpha}\sum_{k=1}^{N}\lambda_k^{j-1}\right\}\leq(\mathrm{N}^*-1)C_p\left\{(\mathrm{N}^*+1)\lambda_k^{j-1}-\frac{1}{\alpha}\sum_{k=1}^{N+1}\lambda_k^{j-1}\right\}$$
$$+C_r\left\{(\mathrm{N}^*+1)\lambda_k^{j-1}-\frac{1}{\alpha}\sum_{k=1}^{N+1}\lambda_k^{j-1}\right\}$$

$$D(\mathrm{N}^*)=\frac{\left\{(\mathrm{N}^*+1)\lambda_k^{j-1}-\frac{1}{\alpha}\sum_{k=1}^{N+1}\lambda_k^{j-1}\right\}}{\left\{\mathrm{N}^*\lambda_k^{j-1}-\frac{1}{\alpha}\sum_{k=1}^{N}\lambda_k^{j-1}\right\}-\left\{(\mathrm{N}^*+1)\lambda_k^{j-1}-\frac{1}{\alpha}\sum_{k=1}^{N+1}\lambda_k^{j-1}\right\}}\geq\frac{C_r}{C_p} \tag{10}$$

Similarly, $$D(\mathrm{N}^*)=B(\mathrm{N}^*)<B(\mathrm{N}^*-1). \tag{11}$$

III. Results

We seek to implement the model in section II by obtaining its parameters such as the optimal number of PM/replacement points and cost ratios in Table 1 which are required for generating the optimal number of PM/replacement schedule.

Table 1: *Cost ratio,* c_r/c_p *for values of* N^**for phases 1, 2 and 3*

Phase	CR	N*=	1	3	5	7	9	11	13
1:j =2	c_r/c_p		3	10	20	40	80	10 0	110
2:j =3	c_r/c_p		4	20	40	70	300	580	800
3:j = 4	c_r/c_p		6	100	800	1000	3000	8000	15000

The optimal number of PM/replacement points, N* in Table 1, row 1 were computed from (9) and chosen to satisfy the inequalities in (10) and (11). Also, the cost ratios (CR): $\frac{C_r}{C_p}$ in rows 2, 3 and 4 for phases 1, 2 and 3 were obtained from (10) with the corresponding values of N^* showing the cost

implication of succeeding PM/replacement actions associated with prolong usage of mechanically repairable machines. These results in Table 1 are utilized in (7) to obtain the sequential PM/replacement schedule for the machine at the three phases of operation under consideration and the results are shown in Tables 2, 3 and 4 for phases 1, 2 and 3 respectively.

Table 2: *Optimal PM/replacement schedule for the 8HP-PML Gold Engine cassava grinding machine for phase 1 (j = 2)*

N^*	1	3	5	7	9	11	13
c_r/c_P	3	10	20	40	80	100	110
x_1	0.5114	0.7845	0.8188	0.9582	1.1938	1.2501	1.3164
x_2		0.5462	0.4711	0.6389	0.7711	0.9836	1.0123
x_3		0.5461	0.3798	0.5049	0.6064	0.7664	0.9592
x_4			0.3134	0.4093	0.4868	0.6129	0.7642
x_5			0.3010	0.3619	0.4280	0.5361	0.6661
x_6				0.3242	0.3811	0.4750	0.5884
x_7				0.3002	0.3415	0.4237	0.5231
x_8					0.3084	0.3807	0.4689
x_9					0.2796	0.3436	0.4212
x_{10}						0.2615	0.3815
x_{11}						0.3829	0.3457
x_{12}							0.3500
x_{13}							0.3873

Table 3: *Optimal PM/replacement schedule for the 8HP-PML Gold Engine cassava grinding machine for phase 2 (j = 3)*

N^*	1	3	5	7	9	11	13
c_r/c_P	4	20	40	70	300	580	800
x_1	0.6771	0.8345	1.3188	1.5528	1.9761	2.5011	3.4164
x_2		0.6543	0.9210	1.2377	1.7055	1.8836	2.2123
x_3		0.6012	0.8221	1.004	1.6650	1.7244	1.9322
x_4			0.7002	0.8753	1.5454	1.6133	0.9632
x_5			0.6944	0.7765	1.4266	1.5221	0.9433
x_6				0.6596	1.2344	1.4550	0.8944
x_7				0.6272	0.9884	1.4007	0.7432
x_8					0.8323	0.9817	0.6501
x_9					0.7788	0.8426	0.4321
x_{10}						0.8765	0.32109
x_{11}						0.7829	0.1345
x_{12}							0.0940
x_{13}							0.0873

Table 4: *Optimal PM/replacement schedule for the 8HP-PML Gold Engine cassava grinding machine for phase 3 (j = 4)*

N^*	1	3	5	7	9	11	13
c_r/c_p	6	100	800	1000	3000	8000	15000
x_1	0.8771	2.5333	4.2322	6.2329	11.6054	25.7512	50.4597
x_2		0.9121	2.6762	1.2377	9.1555	10.8513	20.1123
x_3		0.9022	1.0221	0.9104	1.7452	5.4544	12.7722
x_4			0.8002	0.6753	1.2311	2.9123	8.9193
x_5			0.7932	0.3455	0.4239	1.0221	6.4554
x_6				0.2546	0.2211	0.8551	3.7444
x_7				0.1828	0.1742	0.7007	2.0127
x_8					0.0143	0.5812	1.251
x_9					0.0119	0.2421	0.7329
x_{10}						0.1065	0.5210
x_{11}						0.00921	0.4344
x_{12}							0.2642
x_{13}							0.1773

3.1 Computation of optimal PM and replacement schedule using the geometric hazard model

The optimal operating time, $x_k; k = 1,2, \ldots, 13$ before PM/replacement were computed from (7) for the three phases of operation, viz; *phase 1:* when machine is working within the specified age limit with normal degree of deterioration and low cost of maintenance as shown in Table 2. *Phase 2:* when machine is working beyond specified age limit due for initial replacement. It has extended life span and high degree of deterioration with high cost of maintenance as shown in Table 3. Finally, *phase 3:* when machine is working beyond *phase 2*. It is characterized by further extended lifespan; it has higher degree of deterioration due to aging and prolong usage and it also has associated higher cost of maintenance as shown in Table 4. Row 1 in Tables 2, 3 and 4 contain the optimal number of PM/replacement points with the associated cost ratios in row 2, while the other rows 3 – 15 are the generated operating time, $x_k; k = 1,2, \ldots 13$ for the machine at different cost ratios. Each value typifies a working cycle of the machine followed by necessary PM activity and the last one is the replacement point.

3.2 Application of the GHBPMAR Model to obtain optimal PM and replacement schedule for gold cassava grinding machine

In this section, we consider the application of the derived model in section II by adapting the geometric hazard model to the maintenance of a locally manufactured cassava grinding machine which was considered in [11] to provide its maintenance schedules using sequential imperfect PM and replacement model. The cost implication of extended usage of the machine at phases 2 and 3 which exhibits higher degrees of deterioration due to aging and prolong usage are further examined in this work. The failure distribution of the machine was shown by [11] to follow a 2-parameter Weibull distribution with hazard and cumulative hazard function respectively given as; $h(t) = \alpha\beta^{\alpha}(t)^{\alpha-1}, \alpha > 1, \beta > 0$ and $\mathrm{H}(t) = \beta^{\alpha}t^{\alpha}$. The estimate of the shape and scale parameters from failure data of the machine are; $\alpha = 1.3$ and $\beta = 1386$ respectively. Also, the cost ratios (CR); $\frac{C_m}{C_p} =$

3 is a fixed cost ratio by industry standard, based on the assumption that repair time is negligible, while $\frac{C_r}{C_p}$ = 3, 10, 20, 40, 80, 100, 110 for phase 1 in Table 2, for example, are computed from the optimal values of N^* obtained in (9) such that they satisfy (10) and (11). Finally, the optimal PM schedule, x_k for $k = 1,2, \ldots 13$ and $j = 2,3,4$ are obtained from (7) as shown in Tables 2 - 4.

IV. Discussion

The obtained optimal preventive and replacement schedules for the cassava grinding machine in phases 1, 2 and 3 are presented in Tables 2, 3 and 4. For instance, column 2 in Table 2 with N^*=1 shows that only one cycle of operation is required with cost ratio of 3 and $x_1 = 0.5114$. This implies that the cassava grinding machine is operated for about x_1 =51hrs, (in '000) and a replacement maintenance is performed thereafter. Further use of the grinding machine requires a move to column 3 with N^*= 3 and cost ratio of 10 where the machine is operated for $x_1 = 78$hrs before the first PM and then for $x_2 = 55$hrs before the second PM and can be further operated for $x_3 = 30$hrs before replacement. This is because this column provides for only two PMs and the third cycle is the replacement point. In the same manner, the entries in columns 4 – 8 show the working of the machine under different N^* with their associated cost ratios. It is observed that the operational time in each column is decreasing which depicts a typical hazard PM/replacement model for repairable systems with increasing hazard rate.

Similarly, optimal PM and replacement schedule for phases 2 and 3, when j = 3 and 4 were obtained in Tables 3 and 4 having higher PM cost implication as shown in the cost ratios in row 2 of Tables 3 and 4.

We observed an increasing cost ratio with frequent PM actions due to the extended usage of the machine and a decreasing trend of successive operating time, x_k of the machine in phases 2 and 3 as shown in Tables 3 and 4. This is characterized by the geometric process in accordance with [8] and [9] having a non-increasing parameter which underscores its usefulness in describing the stochastic decrease of lifetime of machines after repairs as well as the stochastic increase in failure times after failures. These observed features of the model are also in line with the PM and replacement schedules obtained by [4], [6] and [12] for repairable systems and in [13] under different PM policies.

The geometric hazard-based preventive maintenance model effectively addresses the increasing deterioration rates of aging machinery by using a geometric process (GP) to adjust the hazard rate. A successful application of the model to a cassava grinding machine, reveals that while preventive maintenance temporarily improves operational time, its maintenance costs rise significantly with each phase of the machine's lifespan. The model, which utilizes the Weibull failure distribution, typical of systems with increasing hazard rate provides actionable insights into optimal maintenance scheduling and cost management. Its effectiveness in predicting maintenance needs and associated costs demonstrates its value, particularly in resource-constrained environment where managing machinery performance and expenses is crucial. The model's validation is demonstrated through its application to a real-world case of cassava grinding machine, where it successfully generates optimal maintenance schedules with its associated cost implications. The results substantiate the model's effectiveness in extending machine lifespan and managing maintenance costs with high cost and frequent preventive maintenance.

In summary, a geometric hazard-based preventive maintenance model for aging mechanically repairable machines was developed in this work. The utilization of geometric process as the hazard rate adjustment factor from [6] was successfully applied to the sequential hazard model of [4] in order to obtain a generalized hazard-based PM and replacement model. Hence, our proposed model is a generalization of the hazard model in [4]. It will be reduced to the hazard model of [4] in phase 1, when j = 2 for the maintenance of non-ageing and low degree deteriorating machines.

Finally, the proposed model was used to obtain the optimal PM/replacement schedule for a cassava grinding machine, characterized by ageing and prolong usage over many years. The results showed that for the three life phases of the machine, the operational times continue to decrease down to the replacement point. It is also observed that the higher the number of operational phases of the machine, the higher the associated costs of maintenance. Hence, an advisory to both design and maintenance engineers to consider this modification for products specification meant for developing countries because of their extended lifespan usage. Also, users of machines in the second and third life phases which is typical in developing countries should be conscious of the implications of the associated high maintenance cost. This is necessary to ensure efficient production management and profitability in business.

References

[1] Nakagawa, T. (1979). Imperfect preventive maintenance. *IEEE Transactions on Reliability*, 37(5): 402.

[2] Lie, C. H and Chun, Y. H. (1986). An algorithm for Preventive maintenance policy, *IEEE Transaction in Reliability*, 35: 71-75.

[3] Nakagawa, T. (1986). Periodic and sequential preventive maintenance policies. *Journal of Applied Probability*, 23: 536-542.

[4] Nakagawa, T. (1988). Sequential imperfect preventive maintenance policies. *IEEE Transactions on Reliability*, 37(3): 295-298.

[5] Lin, D., Zuo, M. J. and Yam, R.C. (2000). General sequential imperfect preventive maintenance models. *International Journal of reliability, Quality and safety Engineering*, 7(03): 253-266.

[6] Udoh, N. and Effanga, E., (2023). Geometric imperfect preventive maintenance and replacement (GIPMAR) model for aging repairable systems. *International Journal of Quality & Reliability Management*, 40(2): 566 – 580.

[7] Yeh, L. (1988). A note on the optimal replacement problem, *Advanced Applied Probability*, 20: 479-482.

[8] Zhang, Y. L. (2002). A geometric repair-model with good-as-new preventive repair. *IEEE Transactions on Reliability*, 5(2): 223-228.

[9] Wang, G. J. and Zhang, Y. L. (2009). A geometric process repair model for a two-component system with shock damage interaction. *International Journal of Systems Science*, 40, 1207-1215.

[10] Zhao, B., Yue, D. and Tian, R. (2010). A geometric process maintenance model of a repairable system with a replacement repair facility. Chinese conference: 2173-2178.

[11] Udoh, N. and Ekpenyong, E., (2019). Sequential imperfect preventive replacement schedule for 8hp-pml gold engine cassava grinding machine. *International Journal of Statistics and Reliability Engineering*, 6(1):13-18.

[12] Udoh, N & and Ekpenyong, M. (2023). A knowledge-based framework for cost implication modeling of mechanically repairable systems with imperfect preventive maintenance and replacement schedule. *Journal of Applied Science and Engineering*, 26(2): 221-234.

[13] Udoh, Nse; Etim, Andrew; Uko, Iniobong (2024). Preventive maintenance policies with reliability thresholds for table saw machine. *Reliability: Theory & Applications*. 4(80): 496-509.

RT&A, No 2 (84)
Volume 20, June 2025

MOMENTS OF GENERALIZED ORDER STATISTICS FROM DOUBLY TRUNCATED MODIFIED MAKEHAM DISTRIBUTION AND ITS CHARACTERIZAION

Abu Bakar[1], Zuber Akhter[2] and Mohd Amir[1,*]

•

[1] Department of Statistics and Operations Researh, Aligarh Muslim University, India
[2] Department of Statistics, University of Delhi, Delhi-110 007, India
[1] abugh2022@gmail.com , [2] akhterzuber022@gmail.com
*corresponding author e-mail: amir.stamu@gmail.com

Abstract

The Makeham distribution has been extensively used in modeling actuarial data and helps as a fundamental life distribution. It describes a mortality failure law in which the hazard rate is composed of two components: a non-aging failure distribution and an aging failure distribution, the latter characterized by an exponentially increasing failure rate. The concept of generalized order statistics (gos), introduced by Kamps in 1995, provides a unified model for combining several models of ordered random variables, including order statistics, record values, progressive type II right-censored order statistics and more. The development of recurrence relations for the moments of gos is an important field of research due to the model's accessibility and flexibility. These recurrence relations simplify the computation of higher-order moments and establish the associations between moments of different orders. In this article, we have derived the recurrence relations of gos for single and product moments from the doubly truncated modified Makeham distribution and have also reduced these relations for specific ordered random schemes, such as order statistics, $k-$th record values and sequential order statistics for selected different values of parameters. In the final section, we have presented a characterization theorem based on the recurrence relations of the moments. This theorem has identified the doubly truncated modified Makeham distribution through its moment structure within the gos framework. This shows how important these relationships are for setting this distribution apart from other probability models.

Keywords: Modified Makeham distribution, generalized order statistics, order statistics, record values, recurrence relations and characterization.

1. Introduction

The moment properties of generalized order statistics (*gos*) for various distributions, as well as the recurrence relations between them, have recently attracted a lot of attention. Many authors have studied the moments of *gos* for several distributions. Kamps [1] proposed the concept of *gos*. The *gos* offers a unified approach to various models of ordered random variables such as upper record values, order statistics, $k-$th upper record values, progressively Type II censoring order statistics, sequential order statistics and Pfeifer records. These models have valuable applications, especially in the field of reliability theory. Khan *et al*. [2] obtained the moments properties of *gos* and also characterized the *gos* from the Lindley-Weibull distribution. Kumar and Dey [3] derived exact algebraic explicit expressions and also obtained recurrence relations for the single, product, and conditional moments of *gos* based on the extended exponential distribution. Athar *et al*. [4] studied moment properties of *gos* from the modified weighted Rayleigh distribution

and also discussed the special cases of *gos* such as order statistics, record values, and progressive type II right censored order statistics. Furthermore, the results for the characterization are presented, which are based on the conditional expectation and moment properties. Khan and Khan [5] obtained the recurrence relation of the generalized upper record values for single and product moments and also characterized for single moments by using conditional expectations and recurrence relation from the modified Makeham distribution.

Let $X_1, X_2, \ldots, X_n$ be a collection of independent and identically distributed (iid) random variables (rv) with absolutely continuous distribution function (df) $F(x)$ and probability density function (pdf) $f(x)$, $x \in (\alpha, \beta)$. Let $n \in N$, $n \geq 2$, $k > 0$, $\tilde{m} = (m_1, m_2, \ldots, m_{n-1}) \in \mathbb{R}^{n-1}$, $M_r = \sum_{j=r}^{n-1} m_j$, such that $\gamma_r = k + n - r + M_r > 0$ for all $r \in \{1, \ldots, n-1\}$. Then $X(r, n, \tilde{m}, k)$, $r = 1, 2, \ldots, n$ are called *gos* if their joint pdf is given by

$$k\Big(\prod_{j=1}^{n-1}\gamma_j\Big)\Big(\prod_{i=1}^{n-1}[1-F(x_i)]^{m_i} f(x_i)\Big)[1-F(x_n)]^{k-1} f(x_n) \tag{1}$$

on the cone $F^{-1}(0) \leq x_1 \ldots \leq x_n \leq F^{-1}(1)$ of $\mathbb{R}^n$.

Choosing the parameters appropriately (see, Cramer [20]), we get

Table 1: *Variants of the generalized order statistics*

	$\gamma_n = k$	γ_r	m_r
(i) Sequential order statistics	α_n	$(n-r+1)\alpha_r$	$(\gamma_r - \gamma_{r+1} - 1)$
(ii) Ordinary order statistics	1	$n-r+1$	0
(iii) Record values	1	1	-1
(iv) Progressively type II censored order statistics	$R_n + 1$	$n-r+1+\sum_{j=r}^{n} R_j$	R_r
(v) Pfeifer's record values	β_n	β_r	$\beta_r - \beta_{r+1} - 1$

In view of (1) with $m_i = m$, $i = 1, 2, \ldots n-1$, the pdf of $r-th$ gos $X(r, n, m, k)$ is

$$f_r(x) = \frac{C_{r-1}}{(r-1)!}\,[\bar{F}(x)]^{\gamma_r - 1} f(x)\, g_m^{r-1}(F(x)) \tag{2}$$

and the joint pdf of $X(r, n, m, k)$ and $X(s, n, m, k)$, $1 \leq r < s \leq n$, is

$$\begin{aligned} f_{r,s}(x,y) &= \frac{C_{s-1}}{(r-1)!\,(s-r-1)!}\,[\bar{F}(x)]^m\, g_m^{r-1}(F(x)) \\ &\times [h_m(F(y)) - h_m(F(x))]^{s-r-1}[\bar{F}(y)]^{\gamma_s - 1} f(x) f(y), \quad \alpha \leq x < y \leq \beta, \end{aligned} \tag{3}$$

where

$$\bar{F}(x) \;=\; 1 - F(x)$$

$$C_{r-1} \;=\; \prod_{i=1}^{r} \gamma_i$$

$$h_m(x) = \begin{cases} -\dfrac{1}{m+1}(1-x)^{m+1} & , \ m \neq -1 \\ -\log(1-x) & , \ m = -1 \end{cases}$$

and

$$g_m(x) = h_m(x) - h_m(0) = \int_0^x (1-t)^m dt, \ \ x \in [0,1).$$

The Makeham distribution has been widely used to fit actuarial data and is an essential life distribution (see Marshall and Olkin, [6]). As per Scollnik [7], the Makeham distribution of mortality describes a failure law, wherein the hazard rate is a combination of a non-aging failure distribution and an aging failure distribution characterized by an exponential increase in failure rates.
A random variable X is said to have modified Makeham distribution if its pdf is of the form

$$f_1(x) = \frac{b}{a}\left(\frac{x}{a}\right)^{b-1} \exp\left[\left(\frac{x}{a}\right)^b\right] \exp\left[1 - \exp\left(\frac{x}{a}\right)^b\right], \ \ x \geq 0, \ \ a > 0, \ \ b > 0,$$

with the corresponding df

$$F_1(x) = 1 - \exp\left[1 - \exp\left(\frac{x}{a}\right)^b\right], \ \ x \geq 0, \ \ a > 0, \ \ b > 0,$$

where a and b are the scale and shape parameters respectively.

We assume without loss of generality, that $a = 1$, in which case the pdf of modified Makeham distribution becomes

$$f_1(x) = b\, x^{b-1}\, e^{x^b} e^{(1-e^{x^b})}, \ \ x \geq 0, \ \ b > 0, \tag{4}$$

and the corresponding df

$$F_1(x) = 1 - e^{(1-e^{x^b})}, \ \ x \geq 0, \ \ b > 0, \tag{5}$$

A truncated distribution is a probability distribution restricted to a particular range or interval. Truncation arises when the values over a specified range are omitted, and the values that remain are re-normalized to establish a proper probability distribution. Truncated distributions have wide applications in statistics, economics, engineering, and many other fields where data is naturally bounded. Here, two truncated points are given P_1 and Q_1.

Now,

$$\int_0^{Q_1} f_1(x)dx = Q \quad \text{and} \quad \int_0^{P_1} f_1(x)dx = P,$$

then the truncated pdf is given by

$$f(x) = \frac{1}{P-Q}\, b\, x^{b-1}\, e^{x^b} e^{(1-e^{x^b})}, \ \ Q_1 \leq x \leq P_1 \tag{6}$$

and corresponding truncated df $F(x)$ is

$$\bar{F}(x) = -P_2 + \frac{1}{b}\, x^{1-b}\, e^{-x^b} f(x), \tag{7}$$

where,

$$P_2 = \frac{1-P}{P-Q}. \tag{8}$$

Modified Makeham distribution is a distribution whose hazard function is bathtub for certain parameter values. This is highly desirable in the modeling of the lifetime of technical devices. furthermore, the skewness of the Modified Makeham distribution shifts from positive to negative as the shape parameter increases from zero to infinite. This distribution has been frequently used to describe human mortality and to establish actuarial tables. The basic characteristics of this distribution have been discussed in detail by Kosznik-Biernacka [8]. For detailed survey on characteristics and application of this distribution one may refer to Kosznik-Biernacka ([8], [9], [10]).

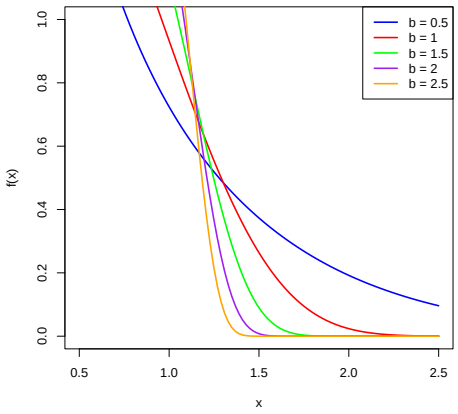


Figure 1: *pdfs of the DTMM distribution for selected values of parameter b.*

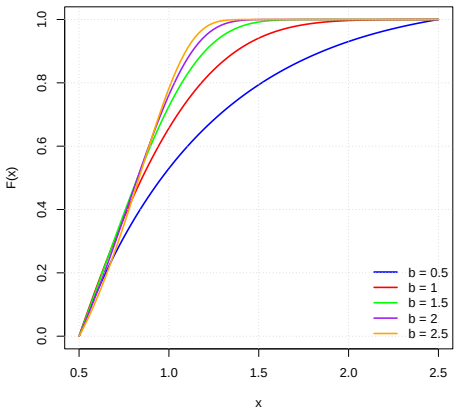


Figure 2: *cdfs of the DTMM distribution for selected values of parameter b.*

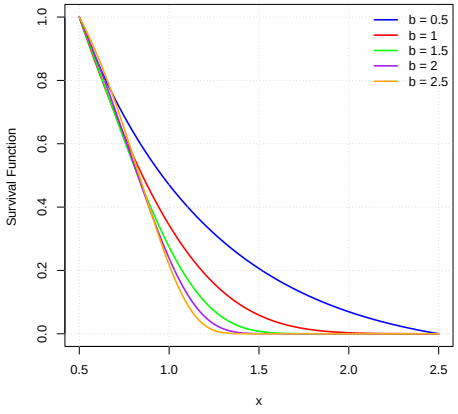


Figure 3: *Survival functions of the DTMM distribution for selected values of parameter b.*

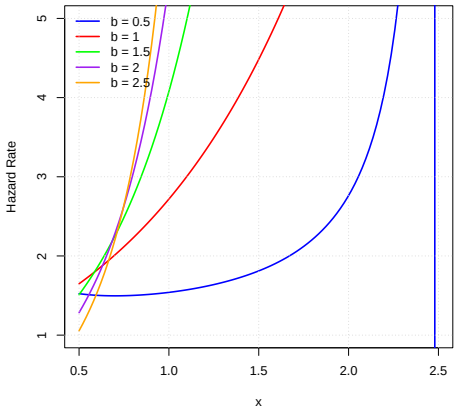


Figure 4: *Hazard rates of the DTMM distribution for selected values of parameter b.*

Recurrence relations for moments of *gos* for some specific as well as for general class of distribution are investigated by several authors in literature. Athar and Islam [16] and Anwar *et al*. [15] derived the recurrence relations for single and product moments of generalized order statistics for some general class of distributions whereas Faizan and Athar [22] obtained explicit expressions for exact moments of generalized order statistics. Bakar et al. [30] have derived recurrence relations for the single and product moments of *gos* based on the Lindley Pareto distribution. Khan *et al*. [29] obtained moments of generalized order statistics for Erlang truncated exponential distribution. Further, Athar *et al*. [17] obtained relations for moments of generalized order statistics from Marshall-Olkin extended Weibull distribution and Athar *et al*. [18] studied the expectation identities of Pareto distribution based on generalized order statistics. Minimol and Thomas [32] discussed some properties of Makeham distribution using generalized record values while Athar *et al*. [19] obtained some recurrence relations for moments of *gos* from doubly truncated Makeham distribution. Ismail and Abu-Youssef [24] derived the general recurrence relations for single and product moments of order statistics for doubly truncated modified Makeham ($DTMM$) distribution and characterized the modified Makeham distribution based on the properties of order statistics. For some additional results, one may refer to Cramer and Kamps [21], Kamps and Cramer [26], Pawlas and Szynal [33], Ahmad and Fawzy [14], Ahmad ([12], [13]), Mahmoud and Al-Nagar [31] among others.

The characterization of probability distributions plays a vital role in statistical studies as it allows us to understand and utilize several distributions effectively. The distributions can be characterized using different statistical properties such as moments, truncated moments, reliability functions, order statistics, record values, characteristic function, etc. Each of these techniques uses distinct statistical properties to uniquely identify or characterize specific probability distributions. This allows statisticians and researchers to model data effectively, understand underlying processes, and make informed decisions based on statistical analyses. Many authors have explored the characterization of probability distributions through truncated moments, conditional expectation, and recurrence relation. Khan and Abu-Salih [34] explored the characterization of probability distribution through conditional expectation of order statistics. Bakar *et al*. [38] studied the characterization of two generalized distributions based on conditional expectation of order statistics. Kilany [35] studied the characterization of the Lindley distribution using truncated moments of order statistics. The characterization results of two general classes of distributions, the Haseeb generalized Lindley and the Haseeb generalized inverse Lindley distributions, have been explored by Athar and Amir [37]. Athar *et al*. [36] studied the characterization of several generalized continuous probability distribution through doubly truncated moments.

The next part of the paper is structured as follows. In Section 2 and 3, we have obtained recurrence relations for single and product moments of generalized order statistics from doubly truncated modified Makeham distribution. Further, the results are deduced for order statistics and record values as a particular cases of generalized order statistics. In Section 4, we have established a theorem characterizing doubly truncated modified Makeham distribution through conditional expectation of generalized order statistics and the whole study of this paper concludes in the Section 5.

2. Single Moment

In this Section, we establish the recurrence relations of *gos* from doubly truncated modified Makeham distribution. Moreover, we drive the expression of the moments of order statistics and record value .

Lemma 2.1: For $2 \leq r \leq n$, $n \geq 2$ and $k = 1, 2, \ldots,$

(i) $E\left[X^j(r,n,m,k)\right] - E\left[X^j(r-1,n,m,k)\right]$

$$= \frac{C_{r-2}}{(r-1)!}\, j \int_{Q_1}^{P_1} x^{j-1}[\bar{F}(x)]^{\gamma_r} g_m^{r-1}\big(F(x)\big)dx. \tag{9}$$

(ii) $E\big[X^j(r-1,n,m,k)\big] - E\big[X^j(r-1,n-1,m,k)\big]$

$$= -\frac{(m+1)C_{r-2}}{\gamma_1(r-2)!}\, j \int_{Q_1}^{P_1} x^{j-1}[\bar{F}(x)]^{\gamma_r} g_m^{r-1}\big(F(x)\big)dx. \tag{10}$$

(iii) $E\big[X^j(r,n,m,k)\big] - E\big[X^j(r-1,n-1,m,k)\big]$

$$= \frac{C_{r-1}}{\gamma_1(r-1)!}\, j \int_{Q_1}^{P_1} x^{j-1}[\bar{F}(x)]^{\gamma_r} g_m^{r-1}\big(F(x)\big)dx. \tag{11}$$

Proof: We have by Athar and Islam [16],

$$E\big[\xi\{X(r,n,m,k)\}\big] - E\big[\xi\{X(r-1,n,m,k)\}\big]$$

$$= \frac{C_{r-2}}{(r-1)!} \int_{\alpha}^{\beta} \xi'(x)\big[\bar{F}(x)\big]^{\gamma_r} g_m^{r-1}\big(F(x)\big)dx, \tag{12}$$

where $\xi(x)$ is a Borel measurable function of $x \in (\alpha, \beta)$.

Therefore, lemma [2] can be established by letting $\xi(x) = x^j$ in the (12). Relations (10) and (11) can be seen on the lines of (9).

Theorem 2.1: For the given modified Makeham distribution and $n \in N,\ m \in \mathbb{R},\ 2 \le r \le n,\ j = 1, 2, \ldots,$

(i) $E\big[X^j(r,n,m,k)\big] - E\big[X^j(r-1,n,m,k)\big]$

$$= -P_2\, K\big\{E[X^j(r,n-1,m,k+m)] - E[X^j(r-1,n-1,m,k+m)]\big\}$$

$$+ \frac{j}{\gamma_r}\frac{1}{b}E[\psi\{X(r,n,m,k)\}]. \tag{13}$$

(ii) $E\big[X^j(r,n,m,k)\big] - E\big[X^j(r-1,n-1,m,k)\big]$

$$= -P_2\, K\frac{\gamma_r}{\gamma_1}\big\{E[X^j(r,n-1,m,k+m)] - E[X^j(r-1,n-1,m,k+m)]\big\}$$

$$+ \frac{j}{\gamma_1}\frac{1}{b}E[\psi\{X(r,n,m,k)\}], \tag{14}$$

where

$$\psi(x) = x^{j-b}e^{-x^b},\ K = \frac{C_{r-2}}{C_{r-2}^{(n-1,k+m)}} = \prod_{i=1}^{r-1}\Big(\frac{\gamma_i}{\gamma_i^{(n-1)} + m}\Big),$$

$$C_{r-2}^{(n-1,k+m)} = \prod_{i=1}^{r-1}\gamma_i^{(n-1,k+m)} = \prod_{i=1}^{r-1}[(k+m) + (n-1-i)(m+1)]$$

and $\gamma_r^{(n-1)} = k + (n-1-r)(m+1)$.

Proof: In view of (7) and (9), we have

$$E\big[X^j(r,n,m,k)\big] - E\big[X^j(r-1,n,m,k)\big]$$

$$= \frac{C_{r-2}}{(r-1)!}\, j \int_{Q_1}^{P_1} x^{j-1} \big[\bar{F}(x)\big]^{\gamma_r - 1} \Big\{ -P_2 + \frac{1}{b}\, x^{1-b}\, e^{-x^b} f(x) \Big\} g_m^{r-1}\big(F(x)\big) dx$$

$$= -P_2\, \frac{C_{r-2}}{(r-1)!}\, j \int_{Q_1}^{P_1} x^{j-1} \big[\bar{F}(x)\big]^{\gamma_r - 1} g_m^{r-1}\big(F(x)\big) dx$$

$$+ \frac{1}{b} \frac{C_{r-2}}{(r-1)!}\, j \int_{Q_1}^{P_1} x^{j-b} e^{-x^b} \big[\bar{F}(x)\big]^{\gamma_r - 1} g_m^{r-1}\big(F(x)\big) f(x) dx$$

$$= -P_2 \frac{C_{r-2}}{C_{r-2}^{(n-1,k+m)}} \left\{ \frac{C_{r-2}^{(n-1,k+m)}}{(r-1)!}\, j \int_{Q_1}^{P_1} x^{j-1} \big[\bar{F}(x)\big]^{\gamma_r^{(n-1,k+m)}} g_m^{r-1}\big(F(x)\big) dx \right\}$$

$$+ \frac{1}{b} \frac{j}{\gamma_r} \frac{C_{r-1}}{(r-1)!} \int_{Q_1}^{P_1} \psi(x) \big[\bar{F}(x)\big]^{\gamma_r - 1} g_m^{r-1}\big(F(x)\big) f(x) dx,$$

as $\gamma_r - 1 = \gamma_r^{(n-1,k+m)} = (k+m) + [(n-1) - r](m+1)$.

Therefore,

$$E\big[X^j(r,n,m,k)\big] - E\big[X^j(r-1,n,m,k)\big]$$
$$= -P_2\, K\big\{E[X^j(r,n-1,m,k+m)] - E[X^j(r-1,n-1,m,k+m)]\big\}$$
$$+ \frac{j}{\gamma_r} \frac{1}{b} E[\psi\{X(r,n,m,k)\}].$$

Hence (13) proved.

Now consider the following identity (*c.f.* Lemma 3.1 of Kamps, [25], pp-231),

$$[k + (n-r-1)(m+1)] E\big[X^j(r,n,m,k)\big] + r(m+1) E\big[X^j(r+1,n,m,k)\big]$$
$$= [k + (n-1)(m+1)] E\big[X^j(r,n-1,m,k)\big]. \tag{15}$$

or

$$\gamma_r\, E\big[X^j(r-1,n,m,k)\big] + (r-1)(m+1) E\big[X^j(r,n,m,k)\big]$$
$$= \gamma_1\, E\big[X^j(r-1,n-1,m,k)\big]. \tag{16}$$

Now, from (13) and (16), we obtain

$$E\big[X^j(r,n,m,k)\big] + \frac{(r-1)(m+1)}{\gamma_r} E\big[X^j(r,n,m,k)\big] - \frac{\gamma_1}{\gamma_r} E\big[X^j(r-1,n-1,m,k)\big]$$
$$= -P_2\, K\big\{E[X^j(r,n-1,m,k+m)] - E[X^j(r-1,n-1,m,k+m)]\big\}$$
$$+ \frac{j}{\gamma_r} \frac{1}{b} E[\psi\{X(r,n,m,k)\}], \tag{17}$$

which, upon rearrangement, yields (14).

Remark 2.1: At $P = 1,\ Q = 0$, we get the relation for non-truncated case

$$E\big[X^j(r,n,m,k)\big] = E\big[X^j(r-1,n-1,m,k)\big] + \frac{j}{\gamma_1} \frac{1}{b} E[\psi\{X(r,n,m,k)\}].$$

Remark 2.2: Recurrence relation for single moments of order statistics ($m=0$, $k=1$) is

$$E(X^j_{r:n}) - E(X^j_{r-1:n-1}) = -P_2\{E(X^j_{r:n-1}) - E(X^j_{r-1:n-1})\} + \frac{j}{nb}E[\psi(X_{r:n})]$$

or

$$E(X^j_{r:n}) = -P_2E(X^j_{r:n-1}) + Q_2E(X^j_{r-1:n-1}) + \frac{j}{nb}E[\psi(X_{r:n})].$$

By convention, we use $X_{n:n-1} = P_1$ and $X_{0:n-1} = Q_1$.

Remark 2.3: For $k-th$ record statistic $(m=-1)$, recurrence relation for single moments reduces as

$$E(X^{(k)}_{U(r)})^j - E(X^{(k)}_{U(r-1)})^j = -P_2\left(\frac{k}{k-1}\right)^{r-1}\left\{E(X^{(k-1)}_{U(r)})^j - E(X^{(k-1)}_{U(r-1)})^j\right\} + \frac{j}{bk}E[\psi(X^{(k)}_{U(r)}].$$

Similarly, the recurrence relation for single moments of order statistics with non-integral sample size for $m=0,\ k=\alpha-n+1,\ \alpha\in\mathbb{R}_+$ and for sequential order statistics for $m=\alpha-1,\ k=\alpha$ may be obtained.

Theorem 2.2: For the given modified Makeham distribution and $n\in N,\ m\in\mathbb{R},\ 2\le r\le n,\ j=1,2,\ldots,$

$$\left(\frac{\gamma_1}{\gamma_1-\gamma_r}\right)\left\{E\big[X^j(r-1,n,m,k)\big] - E\big[X^j(r-1,n-1,m,k)\big]\right\}$$
$$= P_2\ K\left\{E\big[X^j(r,n-1,m,k+m)\big] - E\big[X^j(r-1,n-1,m,k+m)\big]\right\}$$
$$+\frac{j}{\gamma_r}\frac{1}{b}E\big[\psi\{X(r,n,m,k)\}\big], \tag{18}$$

Proof: By subtracting (13) from (14), (18) can be established.

Theorem 2.3: For the given modified Makeham distribution and $n\in N,\ m\in\mathbb{R},\ 2\le r\le n,\ j=1,2,\ldots,$

$$E\big[X^j(r,n,m,k)\big] - E\big[X^j(r-1,n-1,m,k)\big]$$

$$= \frac{(P-Q)}{\gamma_1}\frac{j}{b}\ K^*E[\phi\{X(r,n,m,k+1)\}] \tag{19}$$

$$= \frac{j}{\gamma_1}\frac{1}{b}\Big[E[\phi\{X(r,n,m,k)\}] - (1-P)E[\psi\{X(r,n,m,k)\}]\Big], \tag{20}$$

where

$$K^* = \frac{C_{r-1}}{C^{(k+1)}_{r-1}} = \prod_{i=1}^{r}\Big(\frac{\gamma_i}{\gamma_i+1}\Big),\quad C^{(k+1)}_{r-1} = \prod_{i=1}^{r}[(k+1)+(n-i)(m+1)]$$

and $\phi(x) = \psi(x)\ e^{-\left\{1-e^{x^b}\right\}}$.

Proof: On using (7) in (11), we get

$$E\big[X^j(r,n,m,k)\big] - E\big[X^j(r-1,n-1,m,k)\big]$$

$$= \frac{C_{r-1}}{\gamma_1(r-1)!}\, j\int_{Q_1}^{P_1} x^{j-1}[\bar{F}(x)]^{\gamma_r}\left\{\frac{(P-Q)}{b}\; x^{1-b}e^{-x^b}e^{-\left\{1-e^{x^b}\right\}}f(x)\right\} g_m^{r-1}\big(F(x)\big)dx$$

$$= \frac{(P-Q)}{\gamma_1}\frac{1}{b}\frac{C_{r-1}}{C_{r-1}^{(k+1)}}\; j\left\{\frac{C_{r-1}^{(k+1)}}{(r-1)!}\int_{Q_1}^{P_1}\phi(x)[\bar{F}(x)]^{\gamma_r^{(k+1)}-1}g_m^{r-1}\big(F(x)\big)f(x)dx\right\},$$

where $\gamma_r^{(k+1)} = (k+1)+(n-r)(m+1)$ and hence the theorem.

To prove (20), note that

$$\frac{\bar{F}(x)}{f(x)} = \frac{1}{b}\; x^{1-b}e^{-x^b} - \frac{(1-P)}{b}\; x^{1-b}e^{-x^b}e^{-\left\{1-e^{x^b}\right\}}$$

and the result follows from (11).

3. Product Moments

Theorem 3.1: For the given modified Makeham distribution and $1 \le r < s \le n-1,\ m \in \mathbb{R},\ n \ge 2$ and $i = 1,2,\ldots;\ j = 1,2,\ldots$

$$E[X^i(r,n,m,k).X^j(s,n,m,k)] - E[X^i(r,n,m,k).X^j(s-1,n,m,k)]$$

$$= -P_2\; K_1\left\{E[X^i(r,n-1,m,k+m).X^j(s,n-1,m,k+m)]\right.$$

$$\left.-E[X^i(r,n-1,m,k+m).X^j(s-1,n-1,m,k+m)]\right\}$$

$$+\frac{j}{\gamma_s}\frac{1}{b}E\big[\psi\{X(r,n,m,k),X(s,n,m,k)\}\big], \tag{21}$$

where

$$\psi\{X(r,n,m,k),X(s,n,m,k)\} = x^i y^{j-b}e^{-y^b}$$

$$K_1 = \frac{C_{s-2}}{C_{s-2}^{(n-1,k+m)}} = \prod_{i=1}^{s-1}\Big(\frac{\gamma_i}{\gamma_i^{(n-1)}+m}\Big).$$

Proof: In view of Athar and Islam [16], we have

$$E[X^i(r,n,m,k).X^j(s,n,m,k)] - E[X^i(r,n,m,k).X^j(s-1,n,m,k)]$$

$$= \frac{C_{s-1}}{\gamma_s(r-1)!(s-r-1)!}j\int_{Q_1}^{P_1}\int_x^{P_1} x^i y^{j-1}[\bar{F}(x)]^m f(x)g_m^{r-1}\big(F(x)\big)$$
$$\times[h_m\big(F(y)\big) - h_m\big(F(x)\big)]^{s-r-1}[\bar{F}(y)]^{\gamma_s}dy\;dx. \tag{22}$$

Now using (7) in (22), we get

$$= \frac{C_{s-1}}{\gamma_s(r-1)!(s-r-1)!}j\int_{Q_1}^{P_1}\int_x^{P_1} x^i y^{j-1}[\bar{F}(x)]^m f(x)g_m^{r-1}\big(F(x)\big)$$

$$\times[h_m\big(F(y)\big) - h_m\big(F(x)\big)]^{s-r-1}[\bar{F}(y)]^{\gamma_s-1}\Big\{-P_2 + \frac{1}{b}\, y^{1-b}\, e^{-y^b} f(y)\Big\}dy\, dx.$$

Hence (21) can be established after noting that $\gamma_s - 1 = \gamma_s^{(n-1,k+m)}$, $C_{s-1} = \gamma_s C_{s-2}$.

Remark 3.1: At $P = 1,\ Q = 0$, we get the relation for non-truncated case

$$E[X^i(r,n,m,k).X^j(s,n,m,k)] = E[X^i(r,n,m,k).X^j(s-1,n,m,k)] + \frac{j}{\gamma_s}\frac{1}{b}E\big[\psi\{X(r,n,m,k).X(s,n,m,k)\}\big], \tag{23}$$

Remark 3.2: Recurrence relation between product moments of order statistics $(m = 0,\ k = 1)$ is

$$E\big(X^i_{r:n}X^j_{s:n}\big) = E\big(X^i_{r:n}X^j_{s-1:n}\big) - P_2\,\frac{n}{n-s+1}\Big\{E\big(X^i_{r:n-1}X^j_{s:n-1}\big) - E\big(X^i_{r:n-1}X^j_{s-1:n-1}\big)\Big\} + \frac{j}{b(n-s+1)}E\big[\psi(X_{r:n}, X_{s:n})\big]. \tag{24}$$

Remark 3.3: Recurrence relation for product moments of $k-th$ record value (at $m = -1$) is given as

$$E[(X^{(k)}_{U(r)})^i(X^{(k)}_{U(s)})^j] - E[(X^{(k)}_{U(r)})^i(X^{(k)}_{U(s-1)})^j]$$

$$= -P_2\Big(\frac{k}{k-1}\Big)^{s-1}\Big\{E[(X^{(k-1)}_{U(r)})^i(X^{(k-1)}_{U(s)})^j] - E[(X^{(k-1)}_{U(r)})^i(X^{(k-1)}_{U(s-1)})^j]\Big\} + \frac{j}{bk}E\big[\psi(X^{(k)}_{U(r)}, X^{(k)}_{U(s)})\big],$$

where $(X^{(k)}_{U(r)})^i$ is $i-$th moment of $k-$th upper record value.

4. Characterization

Let $X(r,n,m,k),\ r = 1,2,\ldots$ be gos, then the conditional pdf of $X(s,n,m,k)$ given $X(r,n,m,k) = x$, $1 \le r < s \le n$, in view of (2) and (3) is

$$f_{s|r}(y|x) = \frac{C_{s-1}}{(s-r-1)!C_{r-1}}[\bar{F}(x)]^{m-\gamma_r+1} \times[h_m(F(y)) - h_m(F(x))]^{s-r-1}[\bar{F}(y)]^{\gamma_s-1}f(y) \tag{25}$$

Theorem 4.1: Let X be an absolutely continuous rv with $df\ F(x)$ and $pdf\ f(x)$ with $F(x) < 1$, for all $x \in (0,\infty)$. Then for two consecutive values r and $r+1$, $2 \le r+1 < s \le n$,

$$E[e^{\{X(s,n,m,k)\}^b}|X(l,n,m,k) = x] = g_{s|r}(x) = e^{x^b} + \sum_{j=l}^{s-1}\frac{1}{\gamma_{j+1}},$$
$$l = r, r+1, \text{ and } \gamma_{j+1} \neq 0 \tag{26}$$

if and only if

$$F(x) = 1 - e^{(1-e^{x^b})}, \quad x \geq 0, \ b > 0, . \tag{27}$$

Proof: To prove necessary part, for $s \geq r+1$,

$$E\left[e^{\{X(s,n,m,k)\}^b} | X(r,n,m,k) = x\right] \tag{28}$$

$$= \frac{C_{s-1}}{C_{r-1}(s-r-1)!(m+1)^{s-r-1}} \tag{29}$$

$$\times \int_x^\infty e^{y^b} \left[1 - \left(\frac{\bar{F}(y)}{\bar{F}(x)}\right)^{m+1}\right]^{s-r-1} \left[\frac{\bar{F}(y)}{\bar{F}(x)}\right]^{\gamma_s - 1} \frac{f(y)}{\bar{F}(x)} dy. \tag{30}$$

Set

$$u = \frac{\bar{F}(y)}{\bar{F}(x)} = \frac{e^{\{1-e^{y^b}\}}}{e^{\{1-e^{x^b}\}}}, \quad \text{which implies,} \quad e^{y^b} = e^{x^b} - \ln u.$$

Then the right hand side of (27) reduces to

$$= \frac{C_{s-1}}{C_{r-1}(s-r-1)!(m+1)^{s-r-1}} \int_0^1 \left[e^{x^b} - \ln u\right] \left(1 - u^{m+1}\right)^{s-r-1} u^{\gamma_s - 1} du$$

Let $u^{m+1} = t$, then we get

$$E\left[e^{\{X(s,n,m,k)\}^b} | X(r,n,m,k) = x\right]$$

$$= e^{x^b} - \frac{\prod_{j=r+1}^s \gamma_j}{(s-r-1)!(m+1)^{s-r+1}} \int_0^1 \ln t \ t^{\frac{\gamma_s}{m+1}-1} (1-t)^{s-r-1} \, dt$$

$$= e^{x^b} - \frac{\prod_{j=r+1}^s \gamma_j}{(s-r-1)!(m+1)^{s-r+1}} \, \mathrm{B}\left(\frac{\gamma_s}{m+1}, s-r\right)\left[\psi\left(\frac{\gamma_s}{m+1}\right) - \psi\left(\frac{\gamma_r}{m+1}\right)\right],$$

where,

$\psi(x) = \frac{d}{dx} \ln \Gamma(x)$,

$$\psi(x-n) - \psi(x) = -\sum_{k=1}^{n} \frac{1}{x-k} \tag{31}$$

[*c.f.* Gradshteyn and Ryzhik, [23], pp-540, 905]

and $\mathrm{B}(a,b)$ is the complete beta function.

Therefore,

$$E[e^{\{X(s,n,m,k)\}^b} | X(r,n,m,k) = x] = e^{x^b} + \sum_{j=r}^{s-1} \frac{1}{\gamma_{j+1}}.$$

To prove sufficiency part, we have

$$g_{s|r}(x) = E[e^{\{X(s,n,m,k)\}^b} | X(l,n,m,k) = x].$$

Therefore,

$$g_{s|r+1}(x) - g_{s|r}(x) = -\frac{1}{\gamma_{r+1}},$$

and

$$g'_{s|r}(x) = b\ x^{b-1}e^{x^b}.$$

Now, in view of Khan et al. [27], we get

$$\frac{f(x)}{\bar{F}(x)} = -\frac{1}{\gamma_{r+1}}\frac{g'_{s|r}(x)}{\left[g_{s|r+1}(x) - g_{s|r}(x)\right]}$$
$$= b\ x^{b-1}e^{x^b}.$$

Implying that

$$\bar{F}(x) = e^{(1-e^{x^b})}.$$

Remark 2.1: At $s = r+1$ in (26), we have

$$E\left[e^{\{X(r+1,n,m,k)\}^b}|X(r,n,m,k) = x\right] = e^{x^b} + \frac{1}{\gamma_{r+1}}.$$

Remark 2.2: For order statistics (at $m = 0, k = 1$), we have

$$E[e^{\{X_{s:n}\}^b}|X_{r:n} = x] = e^{x^b} + \sum_{j=r}^{s-1}\frac{1}{n-j}.$$

Further, at $s = r+1$, we get

$$E[e^{\{X_{r+1:n}\}^b}|X_{r:n} = x] = e^{x^b} + \frac{1}{n-r}.$$

5. Conclusion

In this research, we have studied the modified Makeham distribution and have derived the doubly truncated distribution using this distribution, which restricts the random variable to a specific interval $[P_1, Q_1]$. Above investigations demonstrate the recurrence relations for the moments of *gos* from doubly truncated modified Mekeham distribution. Since recurrence relations reduce the amount of direct computation and hence reduce the time and labor. Therefore the relations under consideration may be useful in computing the moments of any order of doubly truncated modified Mekeham distribution. Further, characterization of the distribution under consideration through conditional expectation of *gos* is also presented.

References

[1] Kamps, U. (1995a). A concept of generalized order statistics. B.G. Teubner Stuttgart, Germany.
[2] Khan, F. A. M., Shahbaz, S. H. and Shahbaz, M. Q. (2022). On Relations for Moments of Generalized Order Statistics for Lindley-Weibull Distribution. Computer Systems Science and Engineering, 41(1), 197 208.
[3] Kumar, D., and Dey, S. (2017). Relations for Moments of Generalized Order Statistics from Extended Exponential Distribution. American Journal of Mathematical and Management Sciences, 36(4), 378?400.

[4] Athar, H., Fawzy, M. A. and Alharbi, Y. F. (2023a). Moment Properties of Generalized Order Statistics Based on Modified Weighted Rayleigh Distribution. European Journal of Statistics, 3, 1-1.
[5] Khan, M. A. and Khan, R. U. (2016). Generalized upper record values from modified Makeham distribution and related inference. J. Stat. Appl. Prob, 5, 495-500.
[6] Marshall, A. W. and Olkin, I. (2007). Life distributions (Vol. 13). Springer, New York.
[7] Scollnik, D.P.M. (1995). Simulating random variates from Makeham?s distribution and from others with exact or nearly Log-Concave densities. Trans. Soc. Actuaries 47, 409?437.
[8] Kosznik-Biernacka, S. (2006). Rodzina rozkladow Makehama. W Slupsku, PAP.
[9] Kosznik-Biernacka, S. (2007). Makeham?s generalised distribution. Computational Methods in Science and Technology, 13(2), 113–120.
[10] Kosznik-Biernacka, S. (2011). Sums of independent random variables of the modified Makeham distributions. Computational Methods in Science and Technology, 17(1–2), 35–40.
[11] Aboutahoun, A. W., and Al-Otaibi, N. M. (2009). Recurrence relations between moments of order statistics from doubly truncated Makeham distribution. Computational and Applied Mathematics, 28, 277–290.
[12] Ahmad, A. A. (2007). Recurrence relations for single and product moments of generalized order statistics from doubly truncated Burr type XII distribution. J. Egypt. Math. Soc, 15(1), 117–128.
[13] Ahmad, A. A. (2008). Single and product moments of generalized order statistics from linear exponential distribution. J. Commun. Statist. Theory Meth. 37(8), 1162–1172.
[14] Ahmad, A.A., Fawzy, M. (2003). Recurrence relations for single moments of generalized order statistics from doubly truncated distribution. Journal of Statistical. Planning and Inference, 117(2), 241–249.
[15] Anwar, Z., Athar, H., and Khan, R. U. (2007). Expectation identities based on recurrence relations of functions of generalized order statistics. Journal of Statistical Research, 41(2), 93–102.
[16] Athar, H. and Islam, H.M. (2004). Recurrence relations between single and product moments of generalized order statistics from a general class of distributions. Metron, LXII, 327–337.
[17] Athar, H., Nayabuddin, Khwaja, S.K. (2012a). Relations for moments of generalized order statistics from Marshall-Olkin extended Weibull distribution and its characterization. ProbStat Forum, 5, 127–132.
[18] Athar, H., Nayabuddin, Khwaja, S.K. (2012b). Expectation identities of Pareto distribution based on generalized order statistics and its characterization. American Journal of Applied Mathematics and Mathematical Sciences, 1(1), 23–29.
[19] Athar, H., Nayabuddin, Khwaja, S.K. (2013). Recurrence relations for single and product moments of generalized order statistics from doubly truncated Makeham distribution and its characterization. J. Statist. Res. 47(1), 63–71.
[20] Cramer, E. (2002). Contributions to generalized order statistics. Ph.D. Thesis, University of Oldenburg.
[21] Cramer, E., and Kamps, U. (2000). Relations for expectations of functions of generalized order statistics. Journal of Statistical Planning and Inference, 89(1–2), 79–89.
[22] Faizan, M., and Athar, H. (2008). Moments of generalized order statistics from a general class of distributions. Journal of Statistics, 15, 36–43.
[23] Gradshteyn, I. S., and Ryzhik, I. M. (2014). Table of integrals, series, and products. Academic press.
[24] Ismail, A. A., and Abu-Youssef, S. E. (2014). Recurrence relations for the moments of order statistics from doubly truncated modified Makeham distribution and its characterization. Journal of King Saud University-Science, 26(3), 200–204.
[25] Kamps, U. (1995b). Recurrence relations for moments of record values. Journal of statistical planning and inference, 45(1–2), 225–234.
[26] Kamps, U. and Cramer, E. (2001). On distributions of generalized order statistics. Statistics, 35(3), 269–280.

[27] Khan, A. H., Khan, R. U., and Yaqub, M. (2006). Characterization of continuous distributions through conditional expectation of functions of generalized order statistics. J. Appl. Probab. Statist, 1, 115–131.
[28] Khan, R. U., Anwar, Z., and Athar, H. (2007). Recurrence relations for single and product moments of generalized order statistics from doubly truncated Weibull distribution. Aligarh J. Statist, 27, 69–79.
[29] Khan, R. U., Kumar, D., Athar, H., 2010. Moments of generalized order statistics from Erlang-truncated exponential distribution and its characterization. Intentional Journal of Statistics and Systems, 5(4), 455–464.
[30] Bakar, A., Athar, H., Alharbi, Y. F., and Fawzy, M. A. (2025). Moment Properties of Generalized Order Statistics From Lindley Pareto Distribution. Statistics, Optimization and Information Computing, 13(4), 1427-1443. https://doi.org/10.19139/soic-2310-5070-2026
[31] Mahmoud, M. A. W., and Al-Nagar, H. S. (2009). On generalized order statistics from linear exponential distribution and its characterization. Statistical Papers, 50, 407-418.
[32] Minimol, S., and Thomas, P. Y. (2013). On some properties of Makeham distribution using generalized record values and its characterization. Braz. J. Probab. Stat., 27(4), 487–501
[33] Pawlas, P., and Szynal, D. (2001). Recurrence relations for single and product moments of generalized order statistics from Pareto, generalized Pareto, and Burr distributions. Communications in Statistics-theory and Methods, 30(4), 739–746.
[34] Khan, A. H. and Abu-Salih, M. S. (1989). Characterizations of probability distributions by conditional expectation of order statistics. Metron, 47, 171-181.
[35] Kilany, N. M. (2017). Characterization of Lindley distribution based on truncated momentsof order statistics. Journal of Statistics Applications and Probability, 6(2), 355-360.
[36] Athar, H., Ahsanullah, M. and Ali, M. (2023b). Characterisation of some generalised continuous distributions by doubly truncated moments. Operations Research and Decisions, 33(1).
[37] Athar, H., and Amir, M. (2025). Characterization of some generalized distributions using order statistics. Reliability: Theory and Applications, 20(1 (82)), 413-424.
[38] Bakar, A., Athar, H., and Khan, M. A. (2024). Characterization of generalized distributions based on conditional expectation of order statistics. Reliability: Theory & Applications, 4(80), 861–872, DOI: https://doi.org/10.24412/1932-2321-2024-480-861-872

A COMPREHENSIVE STUDY OF LIFETIME DISTRIBUTION MODELS FOR AIR CONDITIONING SYSTEMS USING FAILURE TIME DATA

S. Mythily[1], S.C. Premila[2], P. Manigandan[1], K. Geetha[1], D. Pachiyappan[3] and A. Radhika[1*]

•

[1]Department of Statistics, Periyar University, Salem-11, Tamil Nadu, India.
[2]Assistant Professor, Department of Mathematics, Saveetha Engineering College, Thandalam, Chennai.
[3]Department of Mathematics, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai.
mythilyshanmugam03@gmail.com, premilac@saveetha.ac.in, srimanigandan95@gmail.com, geethakmanju2002@gmail.com, d.pachiyappan321@gmail.com, radhisaran2004@gmail.com

Abstract

The purpose of this paper was to fit the model for the air conditioning system with failure time for the normal, gamma, and generalized gamma distributions and to identify the optimal distribution model for estimating the air conditioning system failure time. The performance of the normal, gamma, and generalized gamma distributions on air conditioning system failure data was investigated. To determine the optimal model for calculating the air conditioning system failure time, the secondary data was fitted to the normal, gamma, and generalized gamma distribution models. To evaluate the models performance, the Kolmogorov-Smirnov, and Akaike Information Criterion, were used. The outcomes of the research indicated that the gamma model had the lowest AIC and KS values, measuring 359.108 and 0.47624; the generalized gamma model came in second with AIC and KS values are 359.6058 and 0.32988; and the normal model had the highest AIC and KS values, measuring 3203.729 and 0.4298, respectively. Results show that the gamma model has better flexibility in handling real data over the conventional normal and generalized gamma distribution.

Keywords: Reliability measure, Normal, Gamma, Generalized Gamma distribution, Air Conditioning system, Failure time data.

I. Introduction

Normal, Gamma, Generalized Gamma distributions are a few of the distributions that are frequently used to represent lifetime data. Reliability is a key concept in engineering, manufacturing, and quality control that aims to design and produce products that operate reliably over time. It is the likelihood that a component, product, or system will perform on its intended function without failure under specified conditions and given time periods. When it comes to electrical products, aerospace, etc., reliability analysis will help save maintenance costs, minimize risk, and improve product quality. Dependability is critical to both consumer pleasure and their

safety.

A new monitoring system for the intervals between events under exponential and gamma distributions was proposed by Shah et al. [1]. The three-parameter gamma-normal distribution was proposed by Lima et al. [2]. For two parameters, the probability density function is Pradhan and Kundu [3]. The probability density function for the three gamma distribution parameters is as follows Chen and Kotz [4]. The gamma distribution is widely used in quality and reliability engineering to fit product lifetimes since it may show growing, decreasing, and constant failure rates [5, 6]. The Gamma distributions are versatile and can be used to represent a variety of physical situations [7]. In the field of reliability, they are arguably the most often used statistical distributions [8]. Choosing the distribution that most closely matches the failure times is the analysis's task [9]. Reliability and other applications use the gamma distribution, one of the continuous probability distributions. Waiting times and service times are distributed using it [10]. According to Baran and Nemoda [11], the gamma distribution has also been demonstrated to be a suitable model in a variety of different application domains, such as ecosystems. Giving to Stacy [12], the three-parameter generalized gamma distribution is the most universal type of gamma distribution. Another kind of generalized gamma distribution was presented by Nadarajah and Gupta [13] and used to suit drought data. Using the exponentiated technique, Cordeiro et al. [14] constructed another generalization of Stacy's generalized gamma distribution and applied it to survival and life time analysis.

The exponentiated family was first presented by Ref. [15] and is a very versatile method for adding a new parameter to a continuous distribution. The exponential, gamma, Weibull, and Pareto distributions were all generalized using it, and the outcomes clearly show that this method improves the model's flexibility. This technique can be used to create generalized distributions, such as gamma, Weibull, and lognormal distributions [16], Mixtures of gamma distributions [17], generalized gamma and log gamma distributions [18].

The objective of this study was to evaluate and compare the performance of normal, gamma, and generalized gamma distributions in modeling the failure time of air conditioning systems. Specifically, the study aimed to identify the most optimal distribution model for accurately estimating the failure time of these systems by fitting secondary data to the three distribution models. The performance of each model was assessed using the Kolmogorov-Smirnov (KS) test and Akaike Information Criterion (AIC) to determine which model best represents the failure time data. Ultimately, the study sought to determine which distribution (normal, gamma, or generalized gamma) is the most effective for estimating the failure time of air conditioning systems.

The following is the presentation of the remainder of the paper. The normal, gamma, and generalized gamma distributions are introduced in Section 2. Applications of air conditioning system failure data in real-world settings are shown in Section 3. Section 4 presents the study's conclusion.

II. Methods

I. Normal Distribution

The Gaussian distribution is another name for the normal distribution. The two parameters of the normal distribution are the standard deviation (σ) and the mean (μ). The standard deviation (σ) can be used to show the variability of failure time, whereas the mean (μ) represents the average failure time. It is stated that the normal distribution's PDF equals

$$F(t) = \frac{1}{\sqrt{2\pi\sigma}} exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)$$

$$F(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi\sigma}} exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt$$

In standard form of normal distribution

$$F(t) = \varphi\left(\frac{t-\mu}{\sigma}\right)$$

The Reliability function of Normal distribution.

$$R(t) = 1 - \varphi\left(\frac{t-\mu}{\sigma}\right)$$

And the failure rate is defined by the equation.

$$\lambda(t) = \frac{1}{\sqrt{2\pi\sigma}} exp\left(-\frac{1}{2}\frac{(t-\mu)^2}{\sigma^2}\right)\left(1 - \varphi\left(\frac{t-\mu}{\sigma}\right)\right)^{-1}$$

II. Gamma Distribution

The Gamma distribution has a continuous probability distribution. Its sole criteria are scale and shape. The scale parameter of the gamma distribution controlled the distribution's dispersion, and its form and scale parameter will represent the various failure behaviours. The gamma distribution reliability analysis is frequently used in industrial reliability analysis applications. The pdf of the gamma distribution shape parameter a = α and scale parameter b = $\frac{1}{\beta}$is said to be

$$f(t,a,b) = \frac{b^a}{\Gamma(a)} t^{a-1}e^{-bt}$$

$$f(t,\alpha,\beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} t^{\alpha-1}e^{-\frac{t}{\beta}}$$

The Cumulative Distribution Function of the gamma distribution is: $F(t) = \int_0^t f(t)\,dt$

The Reliability function of Gamma Distribution. $R(t) = 1 - F(t)$
The failure rate is defined by.

$$h(t) = (f(t))/(R(t))$$

The two-parameter Weibull probability density function f(t) is given as

$$PDF = f(f) = \frac{b}{a} \times \left(\frac{f}{a}\right)^{b-1} \times e^{-\left(\frac{f}{a}\right)^b}$$

$$CDF = F(f) = 1 - exp\left[-\left(\frac{f}{a}\right)^b\right]$$

III. Generalized Gamma Distribution

We refer to the continuous probability density function as the generalised gamma. It was in 1962 when American mathematician Harbert Stacy created this distribution. The extension of the gamma distribution is known as the generalised gamma distribution. Shape parameter (k), Scale parameter (β), and another shape parameter (θ) are the three parameters in this distribution. The PDF of the generalized gamma distribution is known as

$$f(t) = \frac{\beta}{\Gamma(k).\theta}\left(\frac{t}{\theta}\right)^{k\beta-1} e^{-\frac{t}{\theta}^{\beta}}$$

Where $\theta > 0$ is scale parameter, $\beta > 0$ and $k > 0$ is shape parameters and $\Gamma(k)$ is gamma function X is defined as

$$\Gamma(x) = \int_0^{\infty} s^{x-1}.e^{-s}\,ds$$

The Reliability function of Generalized Gamma Distribution.

$$R(t) = \begin{cases} 1 - \Gamma I\left(\frac{1}{\lambda^2}; \frac{e^{\lambda\left(\frac{In(t)-\mu}{\sigma}\right)}}{\lambda^2}\right) if\ \lambda > 0 \\ 1 - \phi\left(\frac{In(t)-\mu}{\sigma}\right)\ if\ \lambda = 0 \\ \Gamma I\left(\frac{1}{\lambda^2}; \frac{e^{\lambda\left(\frac{In(t)-\mu}{\sigma}\right)}}{\lambda^2}\right) if\ \lambda < 0 \end{cases}$$

Where $\phi(Z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-\frac{x^2}{2}}dx$

And $\Gamma(x)$ is the incomplete gamma function of k.
And x, is given by:

$$\Gamma I(k;x) = \frac{1}{\Gamma(k)}\int_0^x s^{k-1}e^{-s}\,ds$$

The Γ(x) is the gamma function of x.

III. Results and Discussions

I. Application

The air conditioning system failure time datasets were used to compare the sub-models with some standard models. We employed information criteria and goodness of fit techniques, which include the Kolmogorov-Smirnov (Ks), the Akaike information criterion (AIC), and the log-likelihood function assessed at the MLEs. The models used by the competitors were gamma, normal, and generalized gamma distributions. Consequently, the fit is better for the one with lower information criterion values.

$$AIC = -2LL + 2p$$

$$KS\ (D) = max\ |\ Fn(f) - F(f)\ |$$

II. Application to Failure Time Data

This section uses the dataset provided by Linhart and Zucchini [19] to assess the applicability of the proposed normal, gamma, and generalized gamma distributions. The dataset shows the frequency of cooling system failures in aeroplanes. This dataset was used in the research of several authors, including [20, 21]. Failure measurements are initially crisp values; however, for the sake of illustration, we treat data as uncertain sample values for some air condition systems and the data is given as: 23, 261, 87, 7, 121, 125, 14, 62, 47, 225, 71, 74, 246, 21, 42, 20, 5, 12, 120, 11, 13, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 55, 95.

Table 1: *Results of the scale and form parameters produced by the air conditioning system's failure time*

Distributions	Shape parameter	Scale parameter
Normal	59.14286	68.32584
Gamma	0.867236	68.19695
Generalized Gamma	1.215989	39.98209

Table 1 compares three distributions Normal, Gamma, and Generalized Gamma based on their shape and scale parameters. The Normal distribution, with a shape parameter of 59.14286 and a scale parameter of 68.32584, exhibits a symmetric spread. In contrast, the Gamma and Generalized Gamma distributions, with shape parameters of 0.867236 and 1.215989, respectively, and scale parameters of 68.19695 and 39.98209, exhibit skewness. These results highlight the differences in distributional properties, where the Normal distribution follows a symmetric pattern, while the Gamma-based distributions show asymmetry.

Table 2: *Comparison of distribution models for the air condition system failure time*

Distributions	Loglik	AIC	LRT	p-value
Normal	-197.0129	398.0259	0.54327	0.8961
Gamma	-177.554	359.108	0.47624	0.4901
Generalized Gamma	-176.8029	359.6058	1.161419	0.2812

Table 2 evaluates the fit of different distributions for failure time data using Log-likelihood (Loglik), Akaike Information Criterion (AIC), Likelihood Ratio Test (LRT), and p-values. The Normal distribution has the highest AIC (398.0259), indicating the poorest model fit. In contrast, the Gamma and Generalized Gamma distributions exhibit lower AIC values (359.108 and 359.6058, respectively), suggesting better fits. Among these, the Gamma model has the lowest AIC, making it the most suitable choice for modeling failure time data. The LRT values and p-values suggest

that the improvement from the Gamma to the Generalized Gamma model is not statistically significant, reinforcing the Gamma model's adequacy.

Table 3: *Result of the KS test for the failure time of the air condition system.*

Distribution	Statistics KS	D-critic	Hypothesis
Normal	0.2158753	0.429882	It Cannot be rejected
Gamma	0.1622605	0.29832	It Cannot be rejected
Generalized Gamma	0.913989	0.32988	It Cannot be rejected

Table 3 presents the Kolmogorov-Smirnov (KS) test results for evaluating the fit of different distributions to the failure time data. The results indicate that the Gamma distribution has the lowest KS statistic (0.1622605), suggesting it provides the best fit. In contrast, the Normal and Generalized Gamma distributions have higher KS statistics, making them less favorable in comparison. The hypothesis test results confirm that none of the distributions can be rejected, meaning all three models provide an acceptable fit to the data. Additionally, the Gamma distribution has the lowest D-critical value (0.29832), further supporting its suitability for modeling failure time data. Furthermore, the CDFM estimation method is identified as the most effective approach for estimating the parameters of the Normal, Gamma, and Generalized Gamma distributions. The p-values from the KS test confirm that all three distributions can adequately describe the failure time data. Furthermore, the sample histogram and the fitted PDFs of the Normal, Gamma, and Generalized Gamma distributions using the parameter estimate findings in Table 3 are displayed in Figure 1.

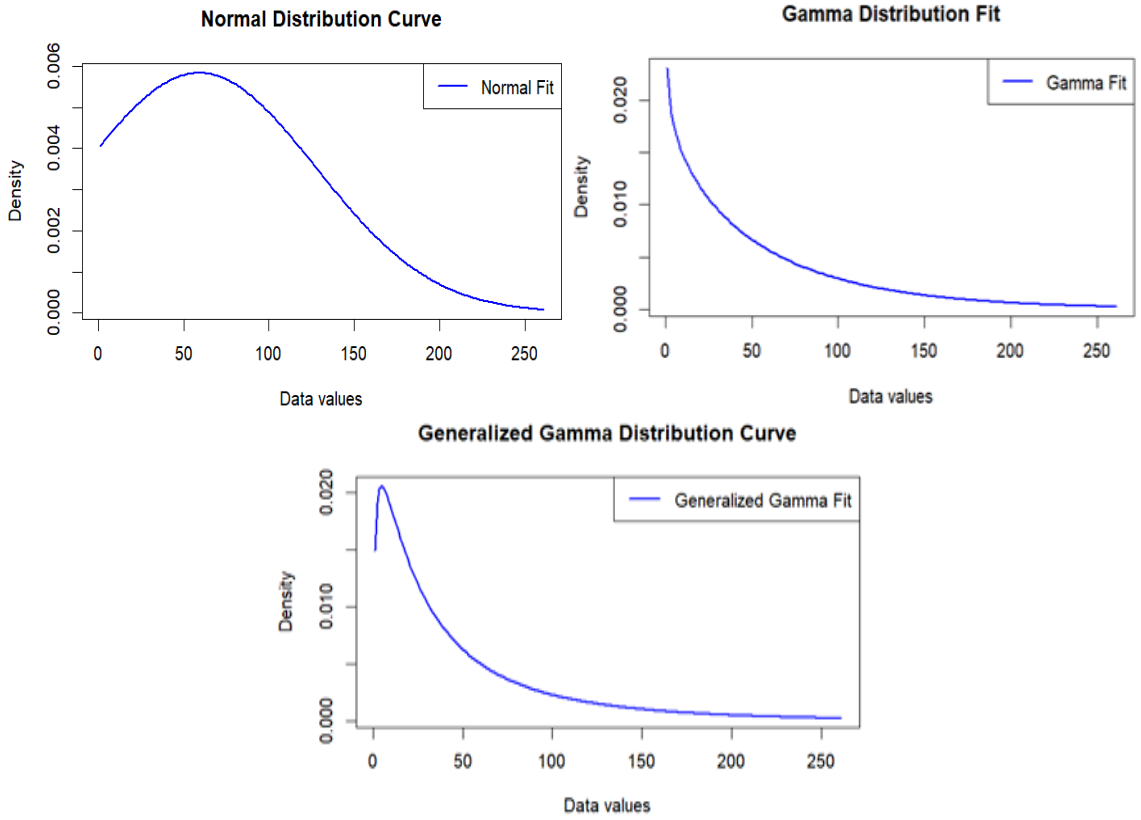


Figure 1: *Graphical presentation of failure rate functions of Normal, Gamma, Generalized Gamma Distribution*

IV. Conclusion

This study explored the challenge of selecting the most suitable distribution among the Normal, Gamma, and Generalized Gamma distributions for modeling failure time data. By analyzing their density function characteristics under different parameter values, we assessed their performance using graphical methods and various goodness-of-fit tests, including the Kolmogorov-Smirnov (KS) test, and Akaike Information Criterion (AIC). Our findings reveal that determining the best-fitting Gamma distribution becomes difficult when multiple models closely fit the data. However, the results suggest that CDFM estimations provide the most accurate fit for the dataset. Additionally, we estimated the reliability and failure rate of the air conditioning system using CDFM calculations, emphasizing the importance of valid parameter estimates in ensuring accurate predictions. The study highlights that different estimation methods can significantly influence reliability and failure rate outcomes. Moreover, as long as the data conforms to a Gamma distribution, the proposed estimation technique can be effectively applied to failure time analysis. Future research can further explore the asymptotic properties of these estimators and extend the method to other distributions. Ongoing work is focused on expanding this approach to broader applications in reliability analysis.

References

[1] Shah, M. T. Azam, M. Aslam, M. and Sherazi, U. (2021). Time between events control charts for gamma distribution. *Quality and Reliability Engineering International*, 37(2): 785-803.

[2] Maria do Carmo, S. L. and Cordeiro, G. M. (2015). A new extension of the normal distribution. *Journal of data Science*, 13(2): 385-408.

[3] Pradhan, B. and Kundu, D. (2011). Bayes estimation and prediction of the two-parameter gamma distribution. *Journal of Statistical Computation and Simulation*, 81(9): 1187-1198.

[4] Chen, W. W. and Kotz, S. (2013). The Riemannian structure of the three-parameter Gamma distribution, *Applied Mathematics*, 4(3): 28865.

[5] Piao, C. H. E. N. and Zhi-Sheng, Y. E. (2018). A systematic look at the gamma process capability indices. *European Journal of Operational Research*, 265(2): 589-597.

[6] Wang, B. X. and Wu, F. (2018). Inference on the gamma distribution. *Technometrics*, 60(2): 235-244.

[7] Zografos, K. and Balakrishnan, N. (2009). On families of beta-and generalized gamma-generated distributions and associated inference. *Statistical methodology*, 6(4): 344-362.

[8] Chahkandi, M. and Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Computational Statistics and Data Analysis*, 53(12): 4433-4440.

[9] Lawless, J. F. Statistical Models and Methods for Lifetime Data. John Wiley and Sons, New York, 2003.

[10] Whitt, W. (2000). The impact of a heavy-tailed service-time distribution upon the M/GI/s waiting-time distribution. *Queueing Systems*, 36(1): 71-87.

[11] Baran, S. and Nemoda, D. (2016). Censored and shifted gamma distribution-based EMOS model for probabilistic quantitative precipitation forecasting. *Environmetrics,* 27(5): 280-292.

[12] Stacy, E. W. (1962). A generalization of the gamma distribution. *The Annals of mathematical statistics*, 1187-1192.

[13] Nadarajah, S. and Gupta, A. K. (2007). A generalized gamma distribution with application to drought data. *Mathematics and Computers in Simulation*, 74(1): 1-7.

[14] Cordeiro, G. M. Ortega, E. M. and Silva, G. O. (2011). The exponentiated generalized gamma distribution with application to lifetime data. *Journal of statistical computation and simulation,*

81(7): 827-842.

[15] Gupta, R. C. Gupta, P. L. and Gupta, R. D. (1998). Modeling failure time data by Lehman alternatives. *Communications in Statistics-Theory and methods*, 27(4): 887-904.

[16] Yang, L. Cheng, J. Luo, Y. Zhou, T. Zhang, X. Shi, L. and Xu, Y. (2025). Domain adaptation via gamma, Weibull, and lognormal distributions for fault detection in chemical and energy processes. *The Canadian Journal of Chemical Engineering*, 103(1): 359-372.

[17] Wiper, M. Insua, D. R. and Ruggeri, F. (2001). Mixtures of gamma distributions with applications. *Journal of computational and graphical statistics*, 10(3): 440-454.

[18] Lawless, J. F. (1980). Inference in the generalized gamma and log gamma distributions. *Technometrics*, 22(3): 409-419.

[19] Linhart, H. and Zucchini, W. Model selection. John Wiley & Sons, 1986.

[20] Magalhães, T. M. Gómez, Y. M. Gallardo, D. I. and Venegas, O. (2020). Bias reduction for the Marshall-Olkin extended family of distributions with application to an airplane's air conditioning system and precipitation data. *Symmetry*, 12(5): 851.

[21] Khan, M. S. and King, R. (2016). New generalized inverse Weibull distribution for lifetime modeling. *Communications for Statistical Applications and Methods*, 23(2): 147-161.

RELIABILITY ANALYSIS OF AAC BLOCK PLANT USING BOOLEAN FUNCTION AND PATH TRACING METHOD

Balram[1], S.C. Malik*, A.D. Yadav[2] and S. Malik[3]

•

*,2,3Department of Statistics, M.D. University, Rohtak-124001
[1]Department of Mathematics, Government College Matanhail, Jhajjar-124106
[1]balrammalik89@gmail.com,
*sc_malik@rediffmail.com
[2]amitdevstats@gmail.com,
[3]maliksushil1983@gmail.com

Abstract

This study offers a comprehensive reliability analysis of Autoclaved Aerated Concrete (AAC) block plants, employing Boolean function technique and path tracing method. The analysis addresses a non-repairable system comprising five interconnected subsystems arranged in series. The configuration includes a single component in the first subsystem, two parallel components in the second, single components in subsystems three and four and three parallel components in the fifth subsystem. The complexity of the system is heightened by the presence of non-identical components with failure rates modelled by the Weibull distribution. Utilizing these mathematical approaches, we derive expressions for reliability and mean time to system failure (MTSF). The study meticulously evaluates system reliability, with a focus on the Weibull distribution, especially the exponential distribution case. The numerical and graphical results, highlighting the performance and durability of the AAC block plant. This research provides critical insights into the reliability dynamics of AAC block production, offering a pathway to enhanced operational efficiency and optimized performance in the industry.

Keywords: Autoclaved Aerated Concrete (AAC) Block, Boolean Technique, Path Tracing Method, Weibull Distribution

I. Introduction

In today's world, Industrialization and infrastructure development have long been recognized as cornerstones of economic progress. However, these advancements often bring significant environmental challenges, with industrial by-products such as fly ash from thermal power plants posing severe threats to ecosystems worldwide. In response to such concerns, the invention of Autoclaved Aerated Concrete (AAC) by Swedish architect and inventor John Axel Eriksson in the early 20th century marked a pivotal moment in construction innovation. AAC blocks, distinguished by their lightweight structure, superior thermal insulation, and fire resistance, represent a sustainable alternative to traditional building materials. Their manufacturing process, which

involves curing concrete under high pressure and temperature, not only enhances performance but also contributes to reducing industrial waste. This process enhances their performance, making AAC blocks an essential material in both residential and commercial construction. Autoclaved Aerated Concrete (AAC) blocks have gained significant traction in the construction industry due to their unique properties, including lightweight nature, thermal insulation, and fire resistance. Malhotra [3] demonstrated that the fire resistance of AAC blocks is twice that of normal bricks, highlighting their superior safety profile. Valore [10] affirmed that AAC blocks possess all desirable characteristics, making them a highly favorable construction material.

Over the years, scholars have proposed various system configurations such as series, parallel, series-parallel, parallel-series, and k-out-of-n to ensure the optimal functioning of non-repairable systems with minimal risks and manufacturing expenses. These configuration strategies are designed to maximize reliability while keeping costs under control. Malik et al. [5] conducted the reliability analysis of NSP system. Narwal et al. [7] considered different mathematical approaches for the reliability analysis of parallel system of order (3,3,3,1). Malik et al. [6] evaluated the fuzzy reliability and MTSF of NSP system under time varying failure rates. Yadav et al. [12] determined the reliability analysis of a three-unit parallel repairable system using Markov approach. Yadav et al. [13] analysed the reliability of a four-unit repairable system. Rathi et al. [9] conducted the reliability modelling of fly ash bricks manufacturing system. The reliability analysis of Autoclaved Aerated Concrete (AAC) block plants can be effectively approached using Boolean function and path tracing method. These methodologies allow for a systematic evaluation of the operational reliability of AAC production systems, which are critical in ensuring the structural integrity and performance of AAC blocks in construction applications. Therefore, integrating the reliability analysis of the production process with the material properties of AAC blocks is essential for a comprehensive assessment. Boolean function can be utilized to model the various components and processes involved in the AAC production system. By defining the operational states of different components (e.g., mixers, autoclaves, and curing chambers) as binary variables, one can create a logical representation of the system's reliability. Agarwal et al. [1] employed the Boolean function technique to discuss the reliability analysis of a sugar manufacturing plant, setting a benchmark for reliability studies. Pervaiz et al. [8] further advanced this field by deriving the reliability of a paper plant using the same Boolean technique. Chandra et al. [2] examined the reliability factors in a chocolate manufacturing plant using the Boolean function technique, adding valuable knowledge to the manufacturing sector. This approach is particularly useful in identifying critical failure paths and understanding how different components interact within the production process. The application of Boolean algebra enables the simplification of complex reliability expressions, facilitating the assessment of the overall system reliability. Path tracing methods complement this analysis by allowing for the identification and quantification of failure paths within the AAC production system. By tracing the paths from the initial state (successful operation) to the final state (failure), one can determine the likelihood of various failure scenarios occurring. For instance, if a failure occurs in the autoclaving process due to temperature fluctuations, path tracing can help identify how this failure impacts the quality of the AAC blocks produced and the subsequent structural reliability of the blocks in construction. Vishnevskiy & Kapustin [11] discussed that the reliability of AAC blocks themselves is influenced by various factors, including the materials used in their production, such as fly ash and aluminum powder, and the conditions under which they are cured. Malik et al. [4] analyzed the reliability and Mean Time to System Failure (MTSF) of a fast-food manufacturing system using Boolean Function and Path Tracing Method for various time values and parameters, offering crucial data for optimizing the system performances.

Here, the reliability analysis of Autoclaved Aerated Concrete (AAC) Blocks plants has been meticulously carried out employing the well-established Boolean function technique and path tracing method. The system is characterized as non-repairable, featuring five subsystems

interconnected in series. Within this configuration, the first subsystem comprises a single component, while the second subsystem entails two parallel components. Subsystems three and four are arranged in series, each encompassing a single component. The fifth subsystem, connected in series, accommodates three components operating in parallel. Importantly, all components within the system are non-identical, adding complexity to the analysis. The failure rates of these components adhere to the Weibull distribution. Expressions for reliability and mean time to system failure (MTSF) have been derived employing these mathematical approaches. Additionally, the reliability measures of the system have been meticulously evaluated, focusing on the specific case of the Weibull distribution, particularly the exponential distribution. Through comprehensive analysis, the reliability of the system has been determined for specific parameter values, highlighting the results both numerically and graphically. This comprehensive study provides valuable insights into the reliability dynamics of AAC Blocks plants, facilitating enhanced operational efficiency and performance optimization.

II. Assumptions and State Descriptions

1. All units in the system are operational.
2. There are no repair facilities available within the system.
3. The failure of each component is statistically independent, with known reliability rates.
4. Initially, the entire system is in good and operable condition.
5. Each component of the system can only be in either a good or bad state.
6. The failure of any single subsystem can cause the entire system to fail.
7. Primary Mixer (X_{11}): All raw materials except aluminum are processed through the primary mixer.
8. Main Mixer (X_{21}, X_{22}): The second subsystem consists of two components connected in parallel. This subsystem mixes the raw materials properly, and aluminum powder is added in the proper ratio.
9. Block Casting (X_{31}): This stage involves casting the blocks in molds.
10. De-molding and Cutting (X_{41}): The fourth subsystem handle the de-molding and wire cutting of blocks to the desired size.
11. High-Pressure Steam Curing (X_{51}, X_{52}, X_{53}): The fifth subsystem consists of three components connected in series. Blocks are kept in this subsystem for 12 hours for high-pressure steam curing to gain strength.

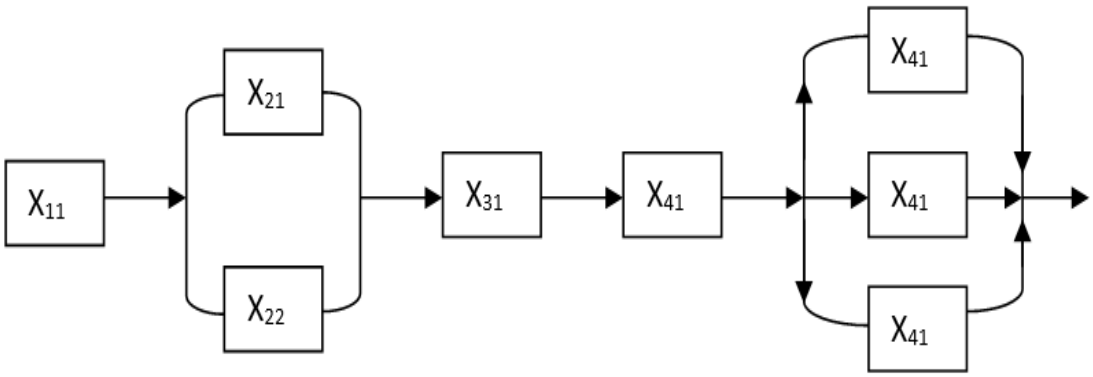


Figure 1: *State Transition Diagram*

III. Solution Using Boolean Function Technique

a) Reliability

The conditions of capability of successful working of the system in terms of logical matrix are shown below:

$$F(x_{11}, x_{21}, x_{22}, \dots\dots, x_{51}) = \begin{pmatrix} x_{11} & x_{21} & x_{31} & x_{41} & x_{51} \\ x_{11} & x_{21} & x_{31} & x_{41} & x_{52} \\ x_{11} & x_{21} & x_{31} & x_{41} & x_{53} \\ x_{11} & x_{22} & x_{31} & x_{41} & x_{51} \\ x_{11} & x_{22} & x_{31} & x_{41} & x_{52} \\ x_{11} & x_{22} & x_{31} & x_{41} & x_{53} \end{pmatrix}$$

By algebra of logics. We have

$$F(x_{11}, x_{21}, x_{31}, \dots . x_{71}) = [x_{11}, x_{31}, x_{41}] \Lambda f(x_{21}, x_{22}, x_{51}, x_{52}, x_{53})$$

Where

$$f(x_{21}, x_{22}, x_{51}, x_{52}, x_{53}) = \begin{pmatrix} x_{21} & x_{51} \\ x_{21} & x_{52} \\ x_{21} & x_{53} \\ x_{22} & x_{51} \\ x_{22} & x_{52} \\ x_{23} & x_{53} \end{pmatrix}$$

Let's assume

$$P_1 = [x_{21} \;\; x_{51}]$$
$$P_2 = [x_{21} \;\; x_{52}]$$
$$P_3 = [x_{21} \;\; x_{53}]$$
$$P_4 = [x_{22} \;\; x_{51}]$$
$$P_5 = [x_{22} \;\; x_{52}]$$
$$P_6 = [x_{22} \;\; x_{53}]$$

Now

$$P_1' = \begin{pmatrix} x_{21}' & \\ x_{21} & x_{51}' \end{pmatrix}$$
$$P_1' P_2 = [x_{21} x_{51}' x_{52}]$$
$$P_1' P_2' = \begin{pmatrix} x_{21}' & & \\ x_{21} & x_{51}' & x_{52}' \end{pmatrix}$$
$$P_1' P_2' P_3 = [x_{21} x_{51}' x_{52}' x_{53}]$$
$$P_1' P_2' P_3' = \begin{pmatrix} x_{21}' & & & \\ x_{21} & x_{51}' & x_{52}' & x_{53}' \end{pmatrix}$$
$$P_1' P_2' P_3' P_4 = [x_{21}' x_{22} x_{51}]$$
$$P_1' P_2' P_3' P_4' = \begin{matrix} x_{21}' & x_{22}' \\ \begin{pmatrix} x_{21} & x_{22}' & x_{51}' & x_{52}' & x_{53}' \\ x_{21}' & x_{22} & x_{51}' & & \\ x_{21} & x_{22} & x_{51}' & x_{52}' & x_{53}' \end{pmatrix} \end{matrix}$$
$$P_1' P_2' P_3' P_4' P_5 = [x_{21}' x_{22} x_{51}' x_{52}]$$
$$P_1' P_2' P_3' P_4' P_5' = \begin{matrix} x_{21}' & x_{22}' \\ \begin{pmatrix} x_{21} & x_{22}' & x_{51}' & x_{52}' & x_{53}' \\ x_{21}' & x_{22} & x_{51}' & x_{52}' & \\ x_{21} & x_{22} & x_{51}' & x_{52}' & x_{53}' \end{pmatrix} \end{matrix}$$
$$P_1' P_2' P_3' P_4' P_5' P_6 = [x_{21}' x_{22} x_{51}' x_{52}' x_{53}]$$

$$R_s(t) = R_{11}R_{31}R_{41}[R_{21}R_{51} + R_{21}R_{52}(1 - R_{51}) + R_{21}R_{53}(1 - R_{51})(1 - R_{52}) + R_{22}R_{51}(1 - R_{21}) + R_{22}R_{52}(1 - R_{21})(1 - R_{51}) + R_{22}R_{53}(1 - R_{21})(1 - R_{51})(1 - R_{52})]$$

$$R_s(t) = R_{11}R_{31}R_{41}[R_{21}R_{51} + R_{21}R_{52} - R_{21}R_{51}R_{52} + R_{21}R_{53} - R_{21}R_{51}R_{53} - R_{21}R_{52}R_{53} + R_{21}R_{51}R_{52}R_{53} + R_{22}R_{51} - R_{21}R_{22}R_{51} + R_{22}R_{52} - R_{21}R_{22}R_{52} - R_{22}R_{51}R_{52} + R_{21}R_{22}R_{51}R_{52} + R_{22}R_{53} - R_{22}R_{51}R_{53} - R_{22}R_{52}R_{53} + R_{22}R_{51}R_{52}R_{53} - R_{21}R_{22}R_{53} + R_{21}R_{22}R_{51}R_{53} + R_{21}R_{52}R_{53} - R_{21}R_{22}R_{51}R_{52}R_{53}]$$

b) Reliability with Weibull Failure Laws

$$\begin{aligned}R_s(t) = {} & e^{-(\lambda_{11}t)^k-(\lambda_{31}t)^k-(\lambda_{41}t)^k}\left[e^{-(\lambda_{21}t)^k-(\lambda_{51}t)^k}+e^{-(\lambda_{21}t)^k-(\lambda_{52}t)^k}-e^{-(\lambda_{21}t)^k-(\lambda_{51}t)^k-(\lambda_{52}t)^k}\right.\\
& +e^{-(\lambda_{21}t)^k-(\lambda_{53}t)^k}-e^{-(\lambda_{21}t)^k-(\lambda_{51}t)^k-(\lambda_{53}t)^k}-e^{-(\lambda_{21}t)^k-(\lambda_{52}t)^k-(\lambda_{53}t)^k}\\
& +e^{-(\lambda_{21}t)^k-(\lambda_{51}t)^k-(\lambda_{52}t)^k-(\lambda_{53}t)^k}+e^{-(\lambda_{22}t)^k-(\lambda_{51}t)^k}-e^{-(\lambda_{21}t)^k-(\lambda_{22}t)^k-(\lambda_{51}t)^k}\\
& +e^{-(\lambda_{22}t)^k-(\lambda_{52}t)^k}-e^{-(\lambda_{21}t)^k-(\lambda_{22}t)^k-(\lambda_{52}t)^k}-e^{-(\lambda_{22}t)^k-(\lambda_{51}t)^k-(\lambda_{52}t)^k}\\
& +e^{-(\lambda_{21}t)^k-(\lambda_{22}t)^k-(\lambda_{51}t)^k-(\lambda_{52}t)^k}+e^{-(\lambda_{22}t)^k-(\lambda_{53}t)^k}-e^{-(\lambda_{22}t)^k-(\lambda_{51}t)^k-(\lambda_{53}t)^k}\\
& -e^{-(\lambda_{22}t)^k-(\lambda_{52}t)^k-(\lambda_{53}t)^k}+e^{-(\lambda_{22}t)^k-(\lambda_{51}t)^k-(\lambda_{52}t)^k-(\lambda_{53}t)^k}-e^{-(\lambda_{21}t)^k-(\lambda_{22}t)^k-(\lambda_{53}t)^k}\\
& +e^{-(\lambda_{21}t)^k-(\lambda_{22}t)^k-(\lambda_{51}t)^k-(\lambda_{53}t)^k}+e^{-(\lambda_{21}t)^k-(\lambda_{52}t)^k-(\lambda_{53}t)^k}\\
& \left.-e^{-(\lambda_{21}t)^k-(\lambda_{22}t)^k-(\lambda_{51}t)^k-(\lambda_{52}t)^k-(\lambda_{53}t)^k}\right]\end{aligned}$$

Let's suppose the units are identical. Then reliability of the system is given by

$$R_s(t) = R^3[6R^2 - 8R^3 + 4R^4 - R^5]$$

For Weibull Failure laws, i.e., failure rate is $(\lambda t)^{k-1}$, then reliability of each component is $e^{-(\lambda t)^k}$ and reliability of the system can be given by

$$\mathrm{R_{SW}}\,(t) = e^{-3(\lambda t)^k}[6e^{-2(\lambda t)^k} - 8e^{-3(\lambda t)^k} + 4e^{-4(\lambda t)^k} - e^{-5(\lambda t)^k}]$$

Now, the mean time to system failure is given by

$$MTSF = \frac{1}{\lambda k}\Gamma\left(\frac{1}{k}\right)\left(\frac{6}{5^{1/k}} - \frac{8}{6^{1/k}} + \frac{4}{7^{1/k}} - \frac{1}{8^{1/k}}\right)$$

c) Reliability with Exponential Failure Laws

$$\begin{aligned}R_s(t) = e^{-(\lambda_{11}t)-(\lambda_{31}t)-(\lambda_{41}t)}\big[& e^{-(\lambda_{21}t)-(\lambda_{51}t)}+e^{-(\lambda_{21}t)-(\lambda_{52}t)}-e^{-(\lambda_{21}t)-(\lambda_{51}t)-(\lambda_{52}t)}+e^{-(\lambda_{21}t)-(\lambda_{53}t)}\\
& -e^{-(\lambda_{21}t)-(\lambda_{51}t)-(\lambda_{53}t)}-e^{-(\lambda_{21}t)-(\lambda_{52}t)-(\lambda_{53}t)}+e^{-(\lambda_{21}t)-(\lambda_{51}t)-(\lambda_{52}t)-(\lambda_{53}t)}\\
& +e^{-(\lambda_{22}t)-(\lambda_{51}t)}-e^{-(\lambda_{21}t)-(\lambda_{22}t)-(\lambda_{51}t)}+e^{-(\lambda_{22}t)-(\lambda_{52}t)}-e^{-(\lambda_{21}t)-(\lambda_{22}t)-(\lambda_{52}t)}\\
& -e^{-(\lambda_{22}t)-(\lambda_{51}t)-(\lambda_{52}t)}+e^{-(\lambda_{21}t)-(\lambda_{22}t)-(\lambda_{51}t)-(\lambda_{52}t)}+e^{-(\lambda_{22}t)-(\lambda_{53}t)}\\
& -e^{-(\lambda_{22}t)-(\lambda_{51}t)-(\lambda_{53}t)}-e^{-(\lambda_{22}t)-(\lambda_{52}t)-(\lambda_{53}t)}+e^{-(\lambda_{22}t)-(\lambda_{51}t)-(\lambda_{52}t)-(\lambda_{53}t)}\\
& -e^{-(\lambda_{21}t)-(\lambda_{22}t)-(\lambda_{53}t)}+e^{-(\lambda_{21}t)-(\lambda_{22}t)-(\lambda_{51}t)-(\lambda_{53}t)}+e^{-(\lambda_{21}t)-(\lambda_{52}t)-(\lambda_{53}t)}\\
& -e^{-(\lambda_{21}t)-(\lambda_{22}t)-(\lambda_{51}t)-(\lambda_{52}t)-(\lambda_{53}t)}\big]\end{aligned}$$

Let's suppose the units are identical. Then reliability of the system is given by

$$R_s(t) = R^3[6R^2 - 8R^3 + 4R^4 - R^5]$$

For Exponential Failure Laws i.e., a constant failure rate λ, then reliability of each component is $e^{-\lambda t}$.

Now, reliability of the system can be obtained as

$$\mathrm{R_{SE}}\,(t) = e^{-3\lambda t}[6e^{-2\lambda t} - 8e^{-3\lambda t} + 4e^{-4\lambda t} - e^{-5\lambda t}]$$

Also, the mean time to system failure is given by

$$\mathrm{MTSF} = \frac{6}{5\lambda} - \frac{8}{6\lambda} + \frac{4}{7\lambda} - \frac{1}{8\lambda}$$

III. Solution Using Path Tracing Method

a) Reliability

The success of any one path out of the paths $P_1, P_2 \ldots P_6$ results into the successful operation of the system.

We have,

$$P(U_{i=1}^{6}P_i) = P(P_1) + P(P_2) + P(P_3) + P(P_4) + P(P_5) + P(P_6) - (P(P_1 \cap P_2) + P(P_1 \cap P_3) + P(P_1 \cap P_4) + P(P_1 \cap P_5) + P(P_1 \cap P_6) + P(P_2 \cap P_3) + P(P_2 \cap P_4) + P(P_2 \cap P_5) + P(P_2 \cap P_6) + P(P_3 \cap P_4) + P(P_3 \cap P_5) + P(P_3 \cap P_6) + P(P_4 \cap P_5) + P(P_4 \cap P_6) + P(P_5 \cap P_6)) + (P(P_1 \cap P_2 \cap P_3) + P(P_1 \cap P_2 \cap P_4) + P(P_1 \cap P_2 \cap P_5) + P(P_1 \cap P_2 \cap P_6) + P(P_1 \cap P_3 \cap P_4) + P(P_1 \cap P_3 \cap P_5) + P(P_1 \cap P_3 \cap P_6) + P(P_1 \cap P_4 \cap P_5) + P(P_1 \cap P_4 \cap P_6) + P(P_1 \cap P_5 \cap P_6) + P(P_2 \cap P_3 \cap P_4) + P(P_2 \cap P_3 \cap P_5) + P(P_2 \cap P_3 \cap P_6) + P(P_2 \cap P_4 \cap P_5) + P(P_2 \cap P_4 \cap P_6) + P(P_2 \cap P_5 \cap P_6) + P(P_3 \cap P_4 \cap P_5) + P(P_3 \cap P_4 \cap P_6) + P(P_3 \cap P_5 \cap P_6) + P(P_4 \cap P_5 \cap P_6)) - (P(P_1 \cap P_2 \cap P_3 \cap P_4) + P(P_1 \cap P_2 \cap P_3 \cap P_5) + P(P_1 \cap P_2 \cap P_3 \cap P_6) + P(P_1 \cap P_2 \cap P_4 \cap P_5) + P(P_1 \cap P_2 \cap P_4 \cap P_6 + P(P_1 \cap P_2 \cap P_5 \cap P_6) + P(P_1 \cap P_3 \cap P_4 \cap P_5) + P(P_1 \cap P_3 \cap P_4 \cap P_6) + P(P_1 \cap P_3 \cap P_5 \cap P_6) + P(P_1 \cap P_4 \cap P_5 \cap P_6) + P(P_2 \cap P_3 \cap P_4 \cap P_5) + P(P_2 \cap P_3 \cap P_4 \cap P_6) + P(P_2 \cap P_3 \cap P_5 \cap P_6) + P(P_2 \cap P_4 \cap P_5 \cap P_6) + P(P_3 \cap P_4 \cap P_5 \cap P_6)) + (P(P_1 \cap P_2 \cap P_3 \cap P_4 \cap P_5) + P(P_1 \cap P_2 \cap P_3 \cap P_4 \cap P_6) + P(P_1 \cap P_2 \cap P_3 \cap P_5 \cap P_6) + P(P_1 \cap P_2 \cap P_4 \cap P_5 \cap P_6)P(P_1 \cap P_3 \cap P_4 \cap P_5 \cap P_6) + P(P_2 \cap P_3 \cap P_4 \cap P_5 \cap P_6) - P(P_1 \cap P_2 \cap P_3 \cap P_4 \cap P_5 \cap P_6)$$

$$R_s(t) = R_{11}R_{31}R_{41}[R_{21}R_{51} + R_{21}R_{52}(1-R_{51}) + R_{21}R_{53}(1-R_{51})(1-R_{52}) + R_{22}R_{51}(1-R_{21}) + R_{22}R_{52}(1-R_{21})(1-R_{51}) + R_{22}R_{53}(1-R_{21})(1-R_{51})(1-R_{52})]$$

$$R_s(t) = R_{11}R_{31}R_{41}[R_{21}R_{51} + R_{21}R_{52} - R_{21}R_{51}R_{52} + R_{21}R_{53} - R_{21}R_{51}R_{53} - R_{21}R_{52}R_{53} + R_{21}R_{51}R_{52}R_{53} + R_{22}R_{51} - R_{21}R_{22}R_{51} + R_{22}R_{52} - R_{21}R_{22}R_{52} - R_{22}R_{51}R_{52} + R_{21}R_{22}R_{51}R_{52} + R_{22}R_{53} - R_{22}R_{51}R_{53} - R_{22}R_{52}R_{53} + R_{22}R_{51}R_{52}R_{53} - R_{21}R_{22}R_{53} + R_{21}R_{22}R_{51}R_{53} + R_{21}R_{52}R_{53} - R_{21}R_{22}R_{51}R_{52}R_{53}]$$

b) Reliability with Weibull Failure Laws

$$R_s(t) = e^{-(\lambda_{11}t)^k-(\lambda_{31}t)^k-(\lambda_{41}t)^k}\left[e^{-(\lambda_{21}t)^k-(\lambda_{51}t)^k} + e^{-(\lambda_{21}t)^k-(\lambda_{52}t)^k} - e^{-(\lambda_{21}t)^k-(\lambda_{51}t)^k-(\lambda_{52}t)^k} + e^{-(\lambda_{21}t)^k-(\lambda_{53}t)^k} - e^{-(\lambda_{21}t)^k-(\lambda_{51}t)^k-(\lambda_{53}t)^k} - e^{-(\lambda_{21}t)^k-(\lambda_{52}t)^k-(\lambda_{53}t)^k} + e^{-(\lambda_{21}t)^k-(\lambda_{51}t)^k-(\lambda_{52}t)^k-(\lambda_{53}t)^k} + e^{-(\lambda_{22}t)^k-(\lambda_{51}t)^k} - e^{-(\lambda_{21}t)^k-(\lambda_{22}t)^k-(\lambda_{51}t)^k} + e^{-(\lambda_{22}t)^k-(\lambda_{52}t)^k} - e^{-(\lambda_{21}t)^k-(\lambda_{22}t)^k-(\lambda_{52}t)^k} - e^{-(\lambda_{22}t)^k-(\lambda_{51}t)^k-(\lambda_{52}t)^k} + e^{-(\lambda_{21}t)^k-(\lambda_{22}t)^k-(\lambda_{51}t)^k-(\lambda_{52}t)^k} + e^{-(\lambda_{22}t)^k-(\lambda_{53}t)^k} - e^{-(\lambda_{22}t)^k-(\lambda_{51}t)^k-(\lambda_{53}t)^k} - e^{-(\lambda_{22}t)^k-(\lambda_{52}t)^k-(\lambda_{53}t)^k} + e^{-(\lambda_{21}t)^k-(\lambda_{22}t)^k-(\lambda_{53}t)^k} + e^{-(\lambda_{21}t)^k-(\lambda_{52}t)^k-(\lambda_{53}t)^k} - e^{-(\lambda_{21}t)^k-(\lambda_{22}t)^k-(\lambda_{51}t)^k-(\lambda_{52}t)^k-(\lambda_{53}t)^k}\right]$$

Let's suppose the units are identical. Then reliability of the system is given by

$$R_s(t) = R^3[6R^2 - 8R^3 + 4R^4 - R^5]$$

For Weibull Failure laws, i.e., failure rate is $(\lambda t)^{k-1}$, then reliability of each component is $e^{-(\lambda t)^k}$ and reliability of the system can be given by

$$R_{SW}(t) = e^{-3(\lambda t)^k}[6e^{-2(\lambda t)^k} - 8e^{-3(\lambda t)^k} + 4e^{-4(\lambda t)^k} - e^{-5(\lambda t)^k}]$$

Now, the mean time to system failure is given by

$$MTSF = \frac{1}{\lambda k}\Gamma\left(\frac{1}{k}\right)\left(\frac{6}{5^{1/k}} - \frac{8}{6^{1/k}} + \frac{4}{7^{1/k}} - \frac{1}{8^{1/k}}\right)$$

c) Reliability with Exponential Failure Laws

$$R_s(t) = e^{-(\lambda_{11}t)-(\lambda_{31}t)-(\lambda_{41}t)}\left[e^{-(\lambda_{21}t)-(\lambda_{51}t)} + e^{-(\lambda_{21}t)-(\lambda_{52}t)} - e^{-(\lambda_{21}t)-(\lambda_{51}t)-(\lambda_{52}t)} + e^{-(\lambda_{21}t)-(\lambda_{53}t)} - e^{-(\lambda_{21}t)-(\lambda_{51}t)-(\lambda_{53}t)} - e^{-(\lambda_{21}t)-(\lambda_{52}t)-(\lambda_{53}t)} + e^{-(\lambda_{21}t)-(\lambda_{51}t)-(\lambda_{52}t)-(\lambda_{53}t)} + e^{-(\lambda_{22}t)-(\lambda_{51}t)} - e^{-(\lambda_{21}t)-(\lambda_{22}t)-(\lambda_{51}t)} + e^{-(\lambda_{22}t)-(\lambda_{52}t)} - e^{-(\lambda_{21}t)-(\lambda_{22}t)-(\lambda_{52}t)} - e^{-(\lambda_{22}t)-(\lambda_{51}t)-(\lambda_{52}t)} + e^{-(\lambda_{21}t)-(\lambda_{22}t)-(\lambda_{51}t)-(\lambda_{52}t)} + e^{-(\lambda_{22}t)-(\lambda_{53}t)} - e^{-(\lambda_{22}t)-(\lambda_{51}t)-(\lambda_{53}t)} - e^{-(\lambda_{22}t)-(\lambda_{52}t)-(\lambda_{53}t)} + e^{-(\lambda_{21}t)-(\lambda_{22}t)-(\lambda_{53}t)} + e^{-(\lambda_{21}t)-(\lambda_{22}t)-(\lambda_{51}t)-(\lambda_{53}t)} + e^{-(\lambda_{21}t)-(\lambda_{52}t)-(\lambda_{53}t)} - e^{-(\lambda_{21}t)-(\lambda_{22}t)-(\lambda_{51}t)-(\lambda_{52}t)-(\lambda_{53}t)}\right]$$

Let's suppose the units are identical. Then reliability of the system is given by

$$R_s(t) = R^3[6R^2 - 8R^3 + 4R^4 - R^5]$$

For Exponential Failure Laws i.e., a constant failure rate λ, then reliability of each component is

$e^{-\lambda t}$.

Now, reliability of the system can be obtained as

$$R_{SE}(t) = e^{-3\lambda t}[6e^{-2\lambda t} - 8e^{-3\lambda t} + 4e^{-4\lambda t} - e^{-5\lambda t}]$$

Also, the mean time to system failure is given by

$$MTSF = \frac{6}{5\lambda} - \frac{8}{6\lambda} + \frac{4}{7\lambda} - \frac{1}{8\lambda}$$

IV. Numerical and Graphical Presentation

Here, we evaluate the reliability and MTSF for the arbitrary values of failure rate (λ) with operating time (t) of the components. The numerical and graphical representation of the results is given below: From Figure 2, it is observed that the reliability of components using the Weibull distribution with varying failure rates (λ) and shape parameters (k) at two different times (t = 5 and t = 10). As the failure rate (λ) increases, the reliability of the system decreases across all conditions. For instance, when t=5 and k=0.90, the reliability drops from 0.8565 at λ=0.01 to 0.2066 at λ=0.09. This indicates that as the component becomes more prone to failure (higher λ), it is less likely to perform reliably over time. Here, two values of are used: 0.90 and 0.80, both of which imply decreasing failure rates. As expected, for a lower value of (0.80), the reliability is lower than for k=0.90, reflecting a more aggressive failure behavior.

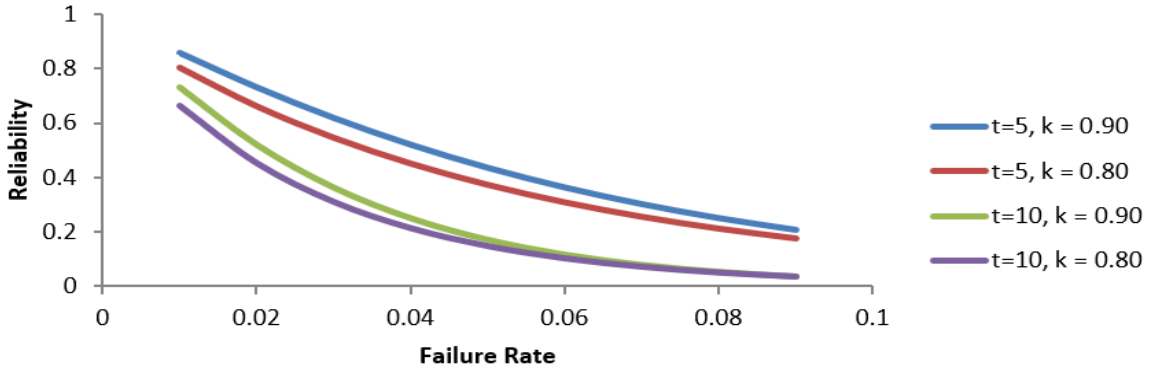


Figure 2: *Reliability V/s Failure Rate(λ) with Operating Time(t) (Boolean Function Technique)*

From Figure 3, it is observed that the MTSF decreases across all shape parameters (k). This indicates that higher failure rates lead to shorter expected operational lifetimes. For example, for k=1, the MTTF decreases from 31.3095 at λ=0.01 to 3.4788 at λ=0.09. This shows that higher failure rates drastically reduce the mean operational time before failure.

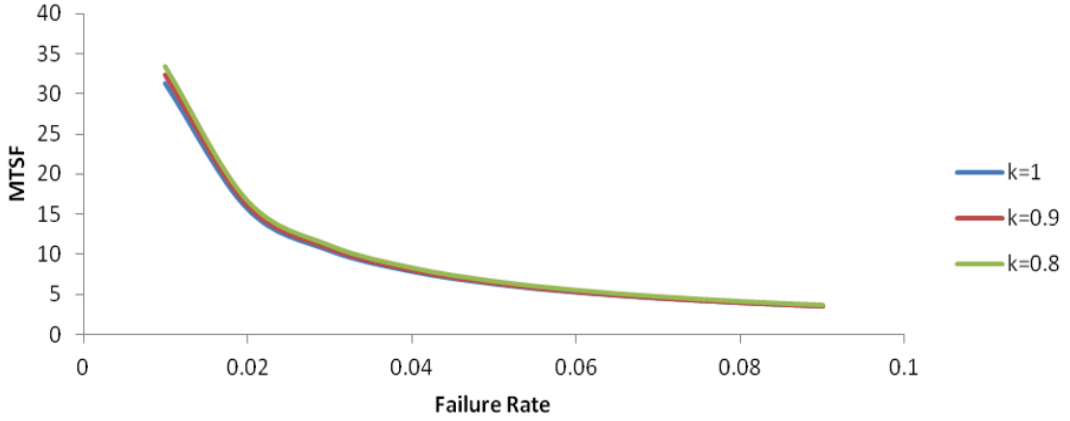


Figure 3: *MTSF V/s Failure Rate(λ) (Boolean Function Technique)*

From Figure 4, as the failure rate (λ) increases, the reliability of the system decreases for both time values. This is expected because a higher failure rate indicates that the component is more prone to

failure. For instance, at t=5, when λ=0.01, the reliability is 0.8946, but when λ=0.09, it drops to 0.2388. Similarly, for t=10, the reliability is 0.7857 at λ=0.01, dropping to 0.0371 at λ=0.09. As time increases, the reliability decreases. Components are more likely to fail over a longer period. For example, for λ=0.03, the reliability is 0.6801 at t=5, but it decreases to 0.4154 at t=10. This indicates that the longer the component is in operation, the more its reliability diminishes due to accumulated wear and tear or other degradation factors.

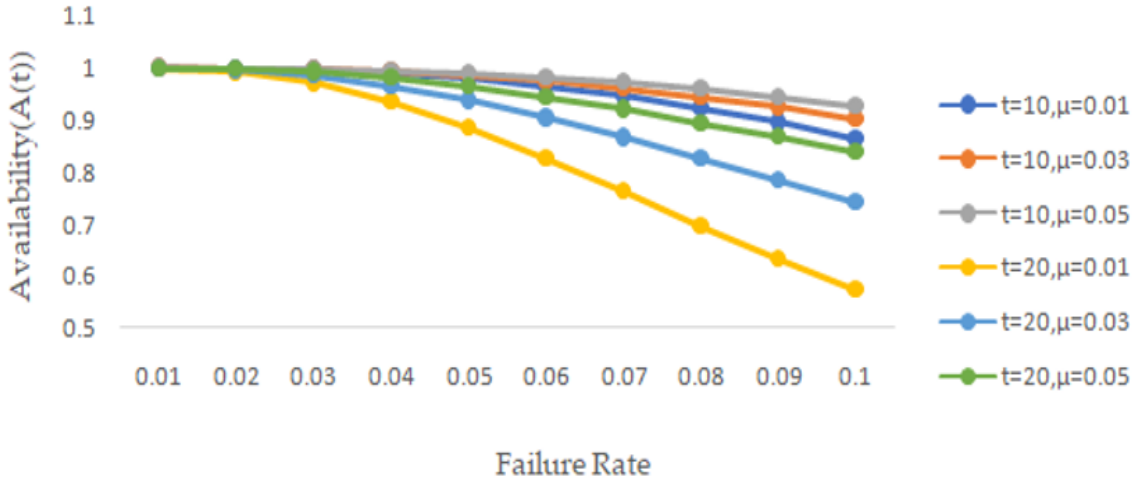


Figure 4: *Reliability V/s Failure Rate(λ) with Operating Time(t) (Boolean Function Technique)*

From Figure 5, we observed that the reliability of components as a function of the failure rate (λ), time (t = 5, t = 10), and the shape parameter (k) for the Weibull distribution. As the failure rate (λ) increases, the reliability of the system decreases for all combinations of t and k. For example, when k=0.90 and t=5, reliability decreases from 0.8565 at λ=0.01 to 0.2066 at λ=0.09. This shows that a higher failure rate reduces the likelihood of the system continuing to function over time. For a lower value of k (e.g., 0.80), reliability tends to be lower than for k=0.90, indicating that a system with more frequent early-life failures will have a reduced overall reliability. For instance, at λ=0.03 and t=5, the reliability for k=0.90 is 0.6177, while for k=0.80, it is 0.5470.

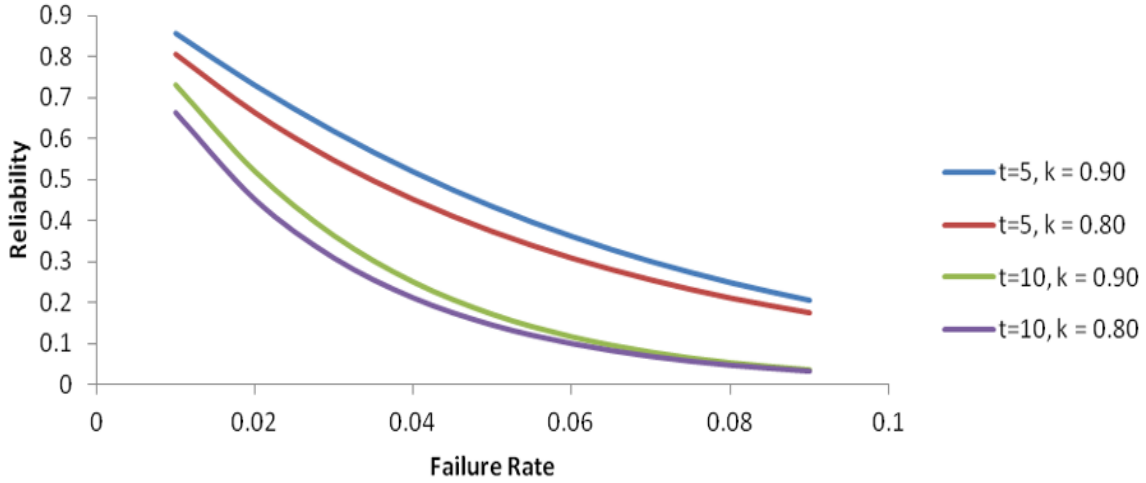


Figure 5: *Reliability V/s Failure Rate(λ) with Operating Time(t)(Path Tracing Method)*

V. Conclusion

The reliability analysis of Autoclaved Aerated Concrete (AAC) block plants has been conducted using the well-established Boolean function technique and path tracing method. By deriving expressions for reliability and Mean Time to System Failure (MTSF) for various failure rates and operating times, and assuming identical units, we have gained significant insights. Our findings indicate that the system's reliability diminishes at a nearly uniform rate under exponential failure laws, while it decreases sharply with a Weibull distribution of failure rates. Furthermore, it is evident that the reliability and MTSF of system continually decline as the failure rate and operating time of the units increase while in case of decreasing the parameter value of k, the MTSF and reliability shows negligible change with the increase of failure rate.

References

[1] Agarwal, S. C., Deepika and Sharma, N. (2011). Reliability analysis of sugar manufacturing plant using boolean function technique. *International Journal of Research and Reviews in Applied Sciences*, 6(2):165-171.

[2] Chandra, A., Gupta, S. and Naithani, A. (2019). Assessment of reliability factors in chocolate manufacturing plant using boolean function technique and neural networking. *International Journal of Innovative Technology and Exploring Engineering*, 8(12):2562-2567.

[3] Malhotra, H. L. (1968). The fire resistance of auto aerated concrete. proceedings of the first international congress on lightweight concrete, *London: Cement and Concrete Association*, 129-131.

[4] Malik, S. C., Balram and Yadav, A. D. (2023). Reliability analysis of a fast-food manufacturing system using boolean function technique and path tracing method. *International Journal of Statistics and Reliability Engineering*, 10(3):580-590.

[5] Malik, S., Malik, S. C., Nandal, N., and Yadav, A. D. (2023). Reliability assessment of a NSP system under constant triangular fuzzy failure rates. *International Journal of Industrial and Systems Engineering*, 1(1). DOI: https://doi.org/10.1504/IJISE .2023.10060854 .

[6] Malik, S., Malik, S. C., Yadav, A. D. and Masum, R. (2024). Intuitionistic fuzzy approach for reliability analysis of NSP system under time varying failure rates. *International Journal of Information Technology*. DOI: https://doi.org/10.1007/s41870-024-02320-z .

[7] Narwal, A., Malik, S. C., Yadav, A. D. and Malik, S. (2024). Mathematical approaches for reliability evaluation of a parallel-series system of order (3,3,1,1) with exponential failure laws. *Safety and Reliability*. DOI: https://doi.org/10.1080/09617353.2024.2435109 .

[8] Pervaiz, I. and Sehik, U. P. S. (2016). Reliability analysis of paper plant using boolean function with fuzzy logic technique. *International Journal of Statistics and Reliability Engineering*, 11:573-577.

[9] Rathi, R., Kadyan, M. S. and Nandal, N. (2024). Reliability modelling and profit analysis of a fly ash bricks manufacturing system. *OPSEARCH.* DOI: https://doi.org/10.1007/s12597-024-00890-x .

[10] Valore, R. C. (1925). Cellular concretes-composition and methods of preparation. *Journal of the American Concrete Institute*, 25:773-795.

[11] Vishnevsky, A. and Kapustin, F. L. (2018). Problems of autoclaved aerated concrete treatment. *Solid State Phenomena*, 284:1069-1073.

[12] Yadav, A. D., Nandal, N., Malik, S. and Malik, S. C. (2023). Markov approach for reliability–availability–maintainability analysis of a three-unit repairable system. *OPSEARCH*, 60(4):1731–1756.

[13] Yadav, A. D., Nandal, N. and Malik, S. C. (2023). Markov approach for reliability and availability analysis of a four-unit repairable system. *Reliability: Theory and Applications*, 18(1):193–205.

RELIABILITY- AVAILABILITY - MAINTAINABILITY OF A REPAIRABLE SYSTEM USING MARKOV APPROACH

A. D. Yadav[*,1], S.C. Malik[2], S. Malik[3] and N. Nandal[4]

•

[*,1,2,3,4]Department of Statistics, M.D. University, Rohtak

[*,1] amitdevstats@gmail.com,

[2]sc_malik@rediffmail.com,

[3]maliksushil1983@gmail.com,

[4]nsinghnandal@gmail.com

Abstract

This study presents a reliability-availability-maintainability (RAM) analysis of a repairable system consisting of two non-identical units, where one unit operates as the main unit while another unit is in cold standby mode. A Markov approach is applied to develop the system model and transition probability equations are derived using the Chapman-Kolmogorov technique. The system is deemed to have failed only when both units are inoperative. Failure and repair times are assumed to follow negative exponential distribution. Utilizing Laplace transform techniques and Mathematica 12.30 software, the system equations are solved to obtain key reliability measures such as reliability, Mean Time to System Failure (MTSF), availability, Mean Time to Repair (MTTR), and maintainability. The cost benefit analysis of the system model has also been carried out for some specific values of the parameters. The study examines the graphical and numerical behavior of these measures and applies the findings to a real-life scenario involving a recently launched Bajaj Freedom 125 CNG bike whose engine that operates on both CNG and petrol. This analysis provides valuable insights into the reliability and economic efficiency of dual-fuel systems in automotive engineering.

Keywords: Markov Model, Chapman-Kolmogorov Equations, Repairable System, Reliability Measures, Cost-Benefit Analysis

I. Introduction

In today's rapidly advancing technological era, the importance of reliability as a measure of system functionality is increasingly recognized within the automobile industry. As automotive technology evolves, the reliability of vehicles becomes crucial not only for consumer satisfaction but also for safety and efficiency. The success of the automobile sector, a cornerstone of economic development, depends on the consistent performance and dependability of its products. To maintain a competitive edge, industry leaders are relentlessly pursuing ways to enhance vehicle reliability. Automotive engineers and researchers play a pivotal role in this endeavor, focusing on innovative strategies such as redundancy, robust repair infrastructures, and resilient maintenance protocols. Their efforts significantly contribute to overall economic progress by ensuring that vehicles operate reliably and efficiently under various conditions. Consequently, both simple and

sophisticated automotive systems have been introduced to the market, delivering seamless services to customers without encountering technical obstacles. The advancements in vehicle reliability are evident through the continuous improvements in automotive system design. Structural integrity and the parallel arrangement of critical components are key factors in enhancing the reliability of complex vehicle systems. These designs distribute operational stress and reduce the risk of system failures. Researchers often employ semi-Markov processes and regenerative point technique to develop detailed models of automotive systems and analyze their reliability measures. Statisticians and engineers frequently discuss topics such as mean time to system failure (MTSF) and system availability, which are crucial for both repairable and non-repairable automotive systems. Kadyan et al. [8] conducted an analysis on a non-identical repairable system comprising three units, where a cold standby unit operates simultaneously. El-Sherbany et al. [5] determined the reliability of cold standby system with arbitrary distributions. Behboudi et al. [1] conducted the reliability modelling using periodic switching approach. Rathi et al. [13] examined the reliability characteristics of a parallel system incorporating a priority concept. Lin et al. [10] conducted the reliability analysis of cold standby using simulation method. Eryilmaz et al. [6] assessed the reliability of the priority standby system. Malik et al. [11] conducted the reliability analysis of a complex system.

Scholars have evaluated the reliability of systems with diverse component structures utilizing the Markov approach. Danjuma et al. [4] determined the RAMD analysis of cold standby series-parallel system. Cui et al. [3] evaluated the history dependent up and down states of repairable systems using Markov process. Manglik et al. [12] analyzed the reliability of a two-unit cold standby system using Markov process. Gupta et al. [7] carried out the cost benefit analysis of a two dissimilar unit cold standby system with Weibull failure and repair laws. Wang et al. [18] evaluated the reliability for multi-state Markov repairable system. Chao et al. [2] carried out the reliability of large series system using Markov structure. Kalaiarasi et al. [9] analyzed system reliability using Markov Technique. Umamaheshwari et al. [17] discussed a Markov model with human error and common cause. Saritha et al. [15] considered the reliability and availability for non-repairable & repairable systems using Markov modeling. Shivani et al. [16] analyzed the sensitivity of series parallel system in component failures. Yadav et al. [19] discussed the Markov approach for reliability, availability and maintainability analysis of a three-unit repairable system. Yadav et al. [20] determined the reliability and availability of a four-unit repairable system using the Markovian approach. Rathi et al. [14] conducted the reliability modelling of fly ash bricks manufacturing system. Previous studies have primarily focused on assessing the reliability of non-repairable systems using Markov modeling, covering scenarios with both identical and non-identical units. However, there is a significant research gap in applying Markov methodologies to perform cost-benefit analyses for repairable systems. Exploring this area offers a new avenue for investigation, extending Markov modeling beyond reliability assessment to include repair and maintenance costs. This comprehensive approach can improve decision-making in system management and optimization. By addressing this gap, researchers can gain a deeper understanding of the economic implications of maintaining and operating repairable systems, leading to more informed and effective resource allocation strategies in various industrial and engineering fields.

In this study, we employed a homogeneous continuous-time Markov chain (CTMC) model, where the Markov property remains consistently valid. In such a framework, units strictly follow a negative exponential distribution, ensuring that each transition is independent of past and future transitions, and interstate transition rates remain constant. This key feature ensures that the future behavior of the Markov system is solely determined by its present state, independent of the historical path taken to reach it. Our research focuses on leveraging Chapman-Kolmogorov equations to derive expressions for key reliability measures of a repairable system composed of

two non-identical cold standby units. To achieve this, we constructed a state transition diagram based on the Markovian birth and death process. Initially, system starts with a single unit and the other unit (non-identical) is kept as spare in cold-standby. We assume the failure and repair times of the units follow a negative exponential distribution. We solved the system equations using the Laplace transform technique, supported by Mathematica 12.30 software. Through this approach, we evaluated the reliability, availability and maintainability of the system model for non-identical units. Additionally, we conducted a cost-benefit analysis for specific values of failure and repair rates. We observed the behavior of these measures graphically for particular failure and repair rates, along with fixed values of revenue generated per unit uptime and the cost of server busy per unit time. Furthermore, we applied our findings to a real-life scenario involving the newly launched Bajaj Freedom 125 CNG bike, which features an engine that operates on both CNG and petrol. This analysis provides valuable insights into the reliability and economic efficiency of dual-fuel systems in automotive engineering, demonstrating the practical applications of our research.

II. Assumptions and State Descriptions

1. The transition rates of the units follow negative exponential distribution
2. The repair of the failed unit is done immediately at the availability of the repair facility.
3. The system is declared failed at the failure of all units.
4. All the units are operative at time 't' in state '1' of the system.
5. The system is in state '2' at time 't' upon the failure of one unit, the repair is made immediately and another cold standby unit is still in operation.
6. All the units are failed at time 't' in state '3' of the system.

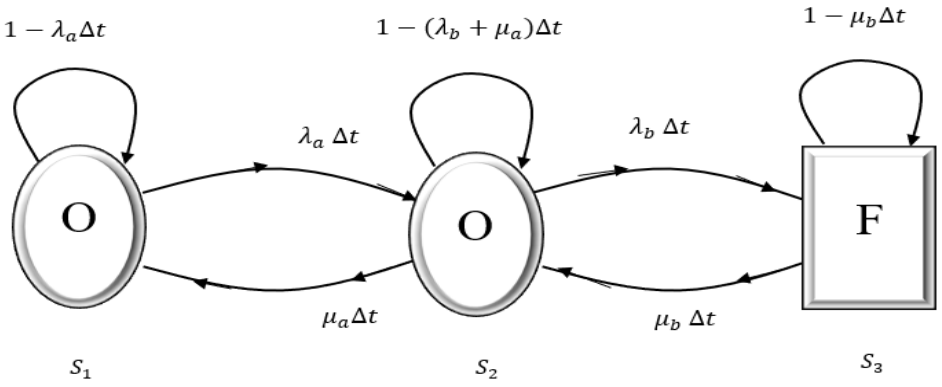


Figure 1: *State Transition Diagram*

a) Notations and Abbreviations

$\boldsymbol{\lambda_x}$	Failure rate of the system at state where x units have failed (x= a and b)
$\boldsymbol{\mu_y}$	Repair rate of the system at state where y units have failed (y=a and b)
$\boldsymbol{S_x}$	States (x=1,2,3)
t	Time
$\boldsymbol{P_x(t)}$	Probability that the system is in state x at time t (x=1,2,3)
$\boldsymbol{A}(\infty)$	Steady State Availability of the system
R(t)	Reliability of the system
MTSF	Mean Time to System Failure

III. Reliability Measures of the System

a) Reliability

In passing from state 'i' at time 's' to state 'j' at time t (s<t), we must pass through some intermediate state 'k' at some intermediate time 'u'. When the continuous-time Markov chain is homogeneous, the Chapman-Kolmogorov equation may be written as:

$$P_{ij}(t+\Delta t) = \sum_{all\ k} P_{ik}(t)P_{kj}(\Delta t) \qquad for\ t, t \geq 0$$
$$= \sum_{k \neq j} P_{ik}(t)P_{kj}(\Delta t) + P_{ik}(t)P_{kj}(\Delta t) \qquad (1)$$

The Chapman- Kolmogorov equations of the system can also be obtained from expression (1) as

$$P_1(t+\Delta t) = P_1(t)(1-\lambda_a \Delta t) + P_2(t)\mu_a \Delta t$$
$$P_2(t+\Delta t) = P_2(t)(1-(\mu_a+\lambda_b)\Delta t) + P_1(t)\lambda_a \Delta t$$
$$P_3(t+\Delta t) = P_3(t)\Delta t + P_2(t)\lambda_b \Delta t \qquad (2\text{-}4)$$

These Markov equations are being developed by taking the probability of each state at time $t+\Delta t$. Converting these equations (2-4) to a differential equation and taking Lim $\Delta t \rightarrow 0$, we get

$$P_1'(t) = -P_1(t)\lambda_a + P_2(t)\mu_a$$
$$P_2'(t) = -P_2(t)(\mu_a+\lambda_b) + P_1(t)\lambda_a$$
$$P_3'(t) = -P_3(t) + P_2(t)\lambda_b \qquad (5\text{-}7)$$

Above equations (5-7) can be solved by using LT method.

$$sp_1(s) - P_1(0) = -\lambda_a p_1(s) + \mu_a p_2(s)$$
$$sp_2(s) - P_2(0) = -(\mu_a+\lambda_b)p_2(s) + \lambda_a p_1(s)$$
$$sp_3(s) - P_3(0) = p_3(s) + \lambda_b p_2(s) \qquad (8\text{-}10)$$

Boundary conditions are

$$P_1(0) = 0, P_2(0) = 0, P_3(0) = 0$$

So, the system of equations will be

$$(s-\lambda_a)p_1(s) - \mu_a p_2(s) = 1$$
$$-\lambda_a p_1(s) + (s-(\mu_a+\lambda_b))p_2(s) = 0$$
$$-\lambda_b p_2(s) + s\, p_3(s) = 0$$

It can be written as

$$\begin{bmatrix} s+\lambda_a & -\mu_a & 0 \\ -\lambda_a & s+(\mu_a+\lambda_b) & 0 \\ 0 & -\lambda_b & s \end{bmatrix} \begin{bmatrix} p_1(s) \\ p_2(s) \\ p_3(s) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Now, solving for $p_1(s), p_2(s)\ and\ \ p_3(s)$ using Cramer's Rule, we have

$$\Delta_{R0} = \begin{bmatrix} s+\lambda_a & -\mu_a & 0 \\ -\lambda_a & s+(\mu_a+\lambda_b) & -\mu_a \\ 0 & -\lambda_b & s+\mu_b \end{bmatrix}$$
$$\Delta_{R0} = (s-0)(s-s_1)(s-s_2)$$

Here s_1 and s_2 are roots of Δ_R.

$$s_1 = -\frac{\lambda_a+\lambda_b+\mu_a+\sqrt{-4\lambda_a\lambda_b+(\lambda_a+\lambda_b+\mu_a)^2}}{2}$$
$$s_2 = \frac{\lambda_a+\lambda_b+\mu_a+\sqrt{-4\lambda_a\lambda_b+(\lambda_a+\lambda_b+\mu_a)^2}}{2}$$

$$\Delta_{R1} = \begin{bmatrix} 1 & -\mu_a & 0 \\ 0 & s+(\mu_a+\lambda_b) & 0 \\ 0 & -\lambda_b & s \end{bmatrix}$$
$$\Delta_{R1} = s(\lambda_b+\mu_a+s))$$
$$\Delta_{R2} = \begin{bmatrix} s+\lambda_a & 1 & 0 \\ -\lambda_a & 0 & 0 \\ 0 & 0 & s \end{bmatrix}$$

$$\Delta_{\mathrm{R2}} = \lambda_a s$$

$$\Delta_{R3} = \begin{bmatrix} s+\lambda_a & -\mu_a & 1 \\ -\lambda_a & s+(\mu_a+\lambda_b) & 0 \\ 0 & -\lambda_b & 0 \end{bmatrix}$$

$$\Delta_{\mathrm{R3}} = \lambda_a \lambda_b$$

Now,

$$p_1(\mathrm{s}) = \frac{\Delta_{R_1}}{\Delta_R} = \frac{(\lambda_b + \mu_a + s)}{(s-s_1)(s-s_2)}$$

$$p_2(\mathrm{s}) = \frac{\Delta_{R_2}}{\Delta_R} = \frac{\lambda_a}{(s-s_1)(s-s_2)}$$

$$p_3(\mathrm{s}) = \frac{\Delta_{R_3}}{\Delta_R} = \frac{\lambda_a \lambda_b}{(s-s_1)(s-s_2)}$$

Taking Laplace Inverse of $p_1(\mathrm{s})$, $p_2(\mathrm{s})$ and $p_3(\mathrm{s})$, we get

$$p_0(\mathrm{t}) = \frac{e^{s_1 t}(\lambda_b + \mu_a + s_1) - e^{s_2 t}(\lambda_b + \mu_a + s_2)}{(s_1 - s_2)}$$

$$p_2(t) = \frac{\lambda_a (e^{s_1 t} - e^{s_2 t})}{(s_1 - s_2)}$$

$$p_3(t) = \frac{-\lambda_a \lambda_b (e^{s_2 t} s_1 - s_1 + s_2 - e^{s_1 t} s_2)}{(s_1 - s_2) s_1 s_2}$$

Thus, Reliability of the system is

$$\mathrm{R(t)} = p_1(t) + p_2(t) = 1 - p_3(t) = \frac{e^{s_1 t}(\lambda_a + \lambda_b + \mu_a + s_1) - e^{s_2 t}(\lambda_b - \lambda_a + \mu_a + s_2)}{(s_1 - s_2)} \quad (11)$$

b) Mean Time to System Failure (MTSF)

On integrating equation (11), we get the expression for MTSF as

$$MTSF = \int_0^\infty R(t)dt$$

$$MTSF = \frac{\lambda_a + \lambda_b + \mu_a}{s_1 s_2}$$

c) Availability

Now, using equation (1), we get the Chapman- Kolmogorov equations for the system to derive the expression for availability as

$$P_1(t+\Delta t) = P_1(t)(1 - \lambda_a \Delta t) + P_2(t)\mu_a \Delta t$$

$$P_2(t+\Delta t) = P_2(t)(1 - (\mu_a + \lambda_b)\Delta t) + P_1(t)\lambda_a \Delta t + P_3(t)\mu_b \Delta t$$

$$P_3(t+\Delta t) = P_3(t)(1 - \mu_b \Delta t) + P_2(t)\lambda_b \Delta t \quad (12\text{-}14)$$

These Markov equations are being developed by taking the probability of each state at time $t + \Delta t$. Taking Lim $\Delta t \to 0$ in the above equations (12-14), we get the following differential equations:

$$P_1'(t) = -P_1(t)\lambda_a + P_2(t)\mu_a$$

$$P_2'(t) = -P_2(t)(\mu_a + \lambda_b) + P_1(t)\lambda_a + P_3(t)\mu_a$$

$$P_3'(\mathrm{t}) = -P_3(t)\mu_a + P_2(t)(\lambda_b) \quad (15\text{-}17)$$

Above equations (15-17) can be solved by using LT method.

$$sp_1(s) - P_1(0) = -\lambda_a p_1(s) + \mu_a p_2(s)$$

$$sp_2(s) - P_2(0) = -(\mu_a + \lambda_b)p_2(s) + \lambda_a p_1(s) + \mu_a p_3(s)$$

$$sp_3(\mathrm{s}) - P_3(0) = -p_3(s)\mu_b + p_2(s)(\lambda_b) \quad (18\text{-}20)$$

Boundary conditions are

$$P_1(0) = 0, P_2(0) = 0, P_3(0) = 0$$

So, the system of equations will be

$$(s+\lambda_a)p_1(s) - \mu_a p_2(s) = 1$$
$$-\lambda_a p_1(s) + (s + (\mu_a + \lambda_b))p_2(s) - \mu_a p_3(s) = 0$$
$$-\lambda_b p_2(s) + (s + \mu_b)\, p_3(s) = 0$$

It can be written as

$$\begin{bmatrix} s+\lambda_a & -\mu_a & 0 \\ -\lambda_a & s+(\mu_a+\lambda_b) & -\mu_a \\ 0 & -\lambda_b & s+\mu_b \end{bmatrix} \begin{bmatrix} p_1(s) \\ p_2(s) \\ p_3(s) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Now, solving for $p_1(s), p_2(s)$ *and* $p_3(s)$using Cramer's Rule, we have

$$\Delta_{A0} = \begin{bmatrix} s+\lambda_a & -\mu_a & 0 \\ -\lambda_a & s+(\mu_a+\lambda_b) & -\mu_a \\ 0 & -\lambda_b & s+\mu_b \end{bmatrix}$$

$$\Delta_{A0} = \lambda_a\lambda_b s + \lambda_a\mu_b s + \mu_a\mu_b s + \lambda_a s^2 + \lambda_b s^2 + \mu_a s^2 + \mu_b s^2 + s^3$$
$$= (s-0)(s-s_3)(s-s_4)$$

Here,

$$s_3 = -\frac{\lambda_a + \lambda_b + \mu_a + \mu_b + \sqrt{(\lambda_a+\lambda_b+\mu_a+\mu_b)^2 - 4(\lambda_a\lambda_b + \lambda_a\mu_b + \mu_a\mu_b)}}{2}$$

$$s_4 = \frac{\lambda_a + \lambda_b + \mu_a + \mu_b + \sqrt{(\lambda_a+\lambda_b+\mu_a+\mu_b)^2 - 4(\lambda_a\lambda_b + \lambda_a\mu_b + \mu_a\mu_b)}}{2}$$

$$\Delta_{A1} = \begin{bmatrix} 1 & -\mu_a & 0 \\ 0 & s+(\mu_a+\lambda_b) & -\mu_a \\ 0 & -\lambda_b & s+\mu_b \end{bmatrix}$$

$$\Delta_{A1} = \mu_a\mu_b + \lambda_b s + \mu_a s + \mu_b s + s^2$$

$$\Delta_{A2} = \begin{bmatrix} s+\lambda_a & 1 & 0 \\ -\lambda_a & 0 & -\mu_a \\ 0 & 0 & s+\mu_b \end{bmatrix}$$

$$\Delta_{A2} = \lambda_a(\mu_b + s)$$

$$\Delta_{A3} = \begin{bmatrix} s+\lambda_a & -\mu_a & 1 \\ -\lambda_a & s+(\mu_a+\lambda_b) & 0 \\ 0 & -\lambda_b & 0 \end{bmatrix}$$

$$\Delta_{A3} = \lambda_a\lambda_b$$

Now,

$$p_1(\mathrm{s}) = \frac{\Delta_{A_1}}{\Delta_A}$$
$$= \frac{\mu_a\mu_b + \lambda_b s + \mu_a s + \mu_b s + s^2}{\mathrm{s}(\mathrm{s}-s_3)(\mathrm{s}-s_4)}$$
$$p_2(\mathrm{s}) = \frac{\Delta_{A_2}}{\Delta_A}$$
$$= \frac{\lambda_a(\mu_b + s)}{s(s-s_3)(s-s_4)}$$
$$p_3(\mathrm{s}) = \frac{\Delta_{A_3}}{\Delta_A} = \frac{\lambda_a\lambda_b}{s(s-s_3)(s-s_4)}$$

Taking Laplace Inverse of $p_1(\mathrm{s})$, $p_2(\mathrm{s})$ *and* $p_3(\mathrm{s})$, we get

$$p_1(t) = \frac{e^{s_3 t}(\mu_a\mu_b + \lambda_a s_3 + \mu_a s_3 + \mu_b s_3 + {s_3}^2)}{s_3(s_3-s_4)} + \frac{\mu_a\mu_b}{s_3 s_4} - \frac{e^{s_4 t}(\mu_a\mu_b + \lambda_b s_4 + \mu_a s_4 + \mu_b s_4 + {s_4}^2)}{s_4(s_3-s_4)}$$

$$p_2(t) = \lambda_a\left[\frac{e^{s_3 t}(\mu_b + s_3)}{s_3(s_3-s_4)} + \frac{\mu_b}{s_3 s_4} + \frac{e^{s_4 t}(-\mu_b - s_4)}{(s_3-s_4)s_4}\right]$$

$$p_3(t) = -\frac{\lambda_a\lambda_b(-s_3 + e^{s_4 t}s_3 + s_4 - e^{s_3 t}s_4)}{s_3(s_3-s_4)s_4}$$

The availability is given by

$$A(t) = p_1(\mathrm{t}) + p_2(\mathrm{t}) = 1 - p_3(t)$$

$$A(t) = \frac{e^{s_3 t}(\mu_a\mu_b + \lambda_a s_3 + \mu_a s_3 + \mu_b s_3 + {s_3}^2)}{s_3(s_3 - s_4)} + \frac{\mu_a\mu_b}{s_3 s_4} - \frac{e^{s_4 t}(\mu_a\mu_b + \lambda_b s_4 + \mu_a s_4 + \mu_b s_4 + {s_4}^2)}{s_4(s_3 - s_4)} + \lambda_a\left[\frac{e^{s_3 t}(\mu_b + s_3)}{s_3(s_3 - s_4)} + \frac{\mu_b}{s_3 s_4} + \frac{e^{s_4 t}(-\mu_b - s_4)}{(s_3 - s_4)s_4}\right]$$

The steady state availability is given by

$$A(\infty) = \lim_{s\to 0} s(p_1(\mathrm{s}) + p_2(\mathrm{s}))$$

$$A(\infty) = \frac{\lambda_a\mu_b}{s_3 s_4} + \frac{\mu_a\mu_b}{s_3 s_4}$$

d) Mean Time to Repair

It plays a crucial role in maintenance planning and optimization by providing an estimate of the anticipated downtime and aiding in the scheduling of maintenance activities to minimize disruptions to operations.

$$MTTR = MTSF(\frac{1}{Availability} - 1)$$

$$MTTR = \frac{\lambda_a + \lambda_b + \mu_a}{s_1 s_2}(\frac{s_3 s_4 - \lambda_a\mu_b + \mu_a\mu_b}{\lambda_a\mu_b + \mu_a\mu_b})$$

e) Maintainability

Maintainability refers to how easily and efficiently a system or product can be maintained, repaired, inspected, and serviced throughout its operational life cycle. By prioritizing maintainability during the design, development, and management stages, a product or system can achieve long-term sustainability and peak performance.

$$M(t) = 1 - \exp(-\frac{t}{MTTR})$$

$$M(t) = 1 - e^{\left(-\frac{t}{\frac{\lambda_a+\lambda_b+\mu_a}{s_1 s_2}\left(\frac{s_3 s_4 - \lambda_a\mu_b + \mu_a\mu_b}{\lambda_a\mu_b + \mu_a\mu_b}\right)}\right)}$$

f) Busy Period of the Server Due to Repair

The busy period of a server due to repair refers to the duration during which the server is engaged in repairing or servicing tasks. It starts when the server begins the repair process after detecting a failure or issue and ends when the repair is completed, and the server is ready to resume its normal operation.

$$B_0(t) = p_2(\mathrm{t}) + p_3(\mathrm{t}) = 1 - p_1(t)$$

$$B_0(t) = 1 - \frac{(\lambda_b + \mu_a + s)}{(\mathrm{s} - s_1)(\mathrm{s} - s_2)}$$

g) Cost Function

If we assume that K_0 be the revenue per unit up-time of the system and K_1 be the cost per unit for which server is busy due to repair, then the following equation helps to find the amount that the system model experiences while it is in steady state.

The expected cost incurred in (0,t] is

$$Cost\ Function = K_0 A(\infty) - K_1 B_0(\infty)$$

IV. Numerical and Graphical Presentation

Here, we evaluate the reliability, availability, maintainability and the cost function for the arbitrary values of failure rate and repair rate of the components with arbitrary time when they are taken as non-identical in nature. The numerical and graphical representations of the findings are provided

below:

From Figure 2, it is observed that the reliability of the system declines with the increase of failure rate and operating time. Also, the reliability of the system is increases with an increase in repair rate of the units.

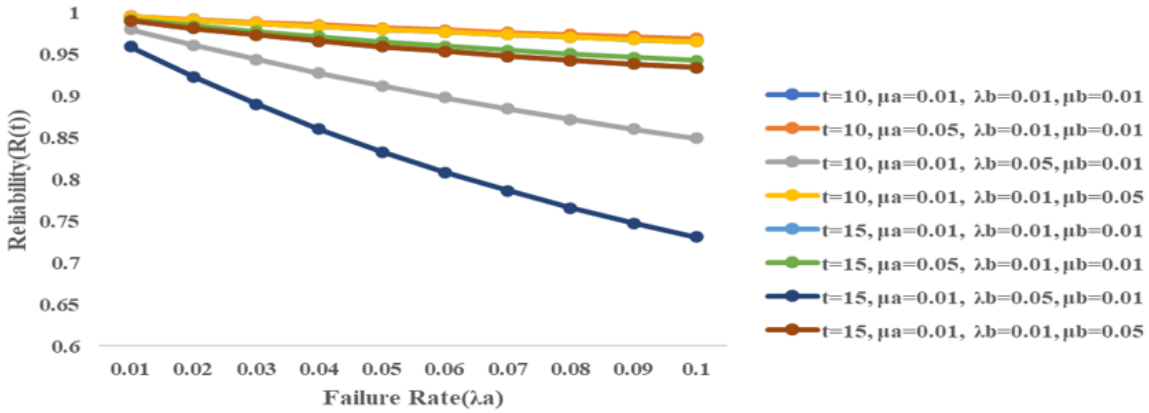


Figure 2: *Reliability V/s Repair Rate (μ) and Failure Rate(λ) with Operating Time(t)*

From Figure 3, it is observed that the availability (A(t)) of the system declines with the increase of failure rate and operating time. Also, the availability of the system is increases with an increase in repair rate of the units.

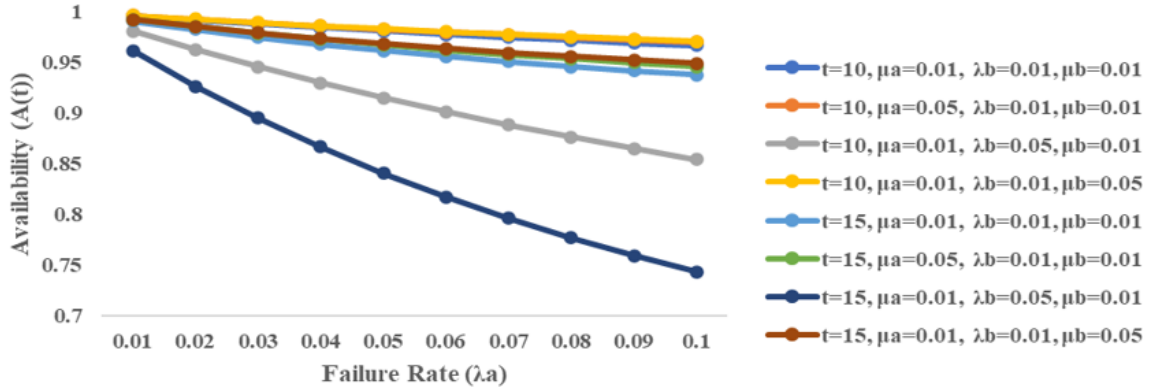


Figure 3: *Availability V/s Repair Rate (μ) and Failure Rate(λ) with Operating Time(t)*

From Figure 4, it is observed that the availability (A(∞)) of the system declines with the increase of failure rate. Also, the availability of the system is increases with an increase in repair rate of the units.

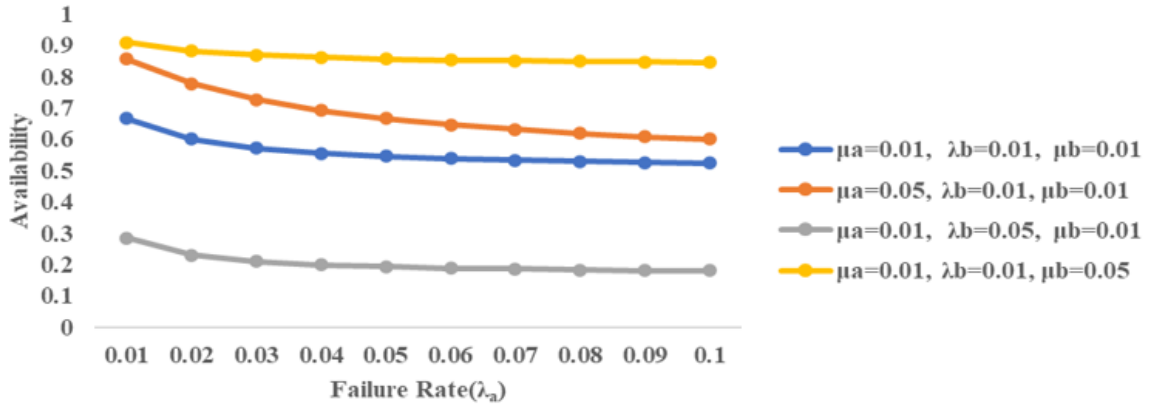


Figure 4: *Availability V/s Repair Rate (μ) and Failure Rate (λ)*

From Figure 5, it is observed that the maintainability (M(t)) of the system increases with the increase of failure rate. Also, the maintainability of the system declines with an increase in repair rate of the units.

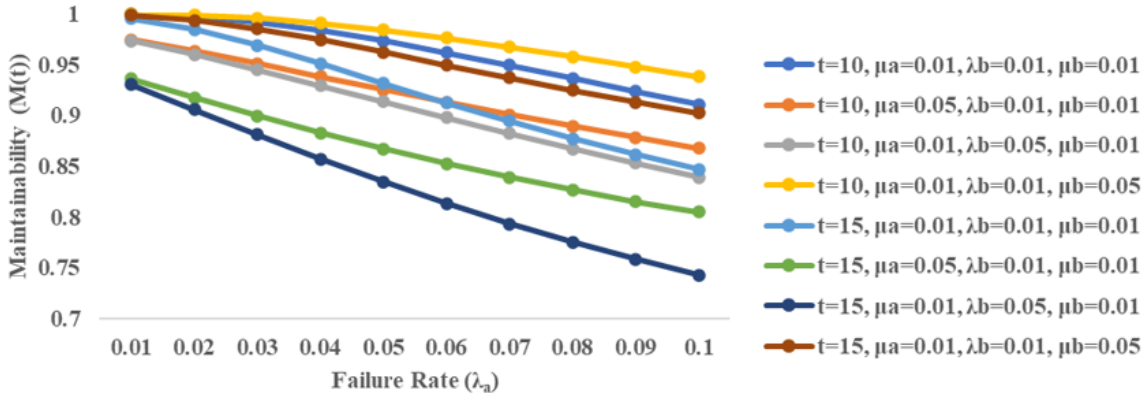


Figure 5: *Maintainability V/s Repair Rate (μ) and Failure Rate (λ)*

From Figure 6, it is observed that the cost function (C(t)) of the system declines with the increase of failure rate. Also, the cost of the system is increases with an increase in repair rate of the units.

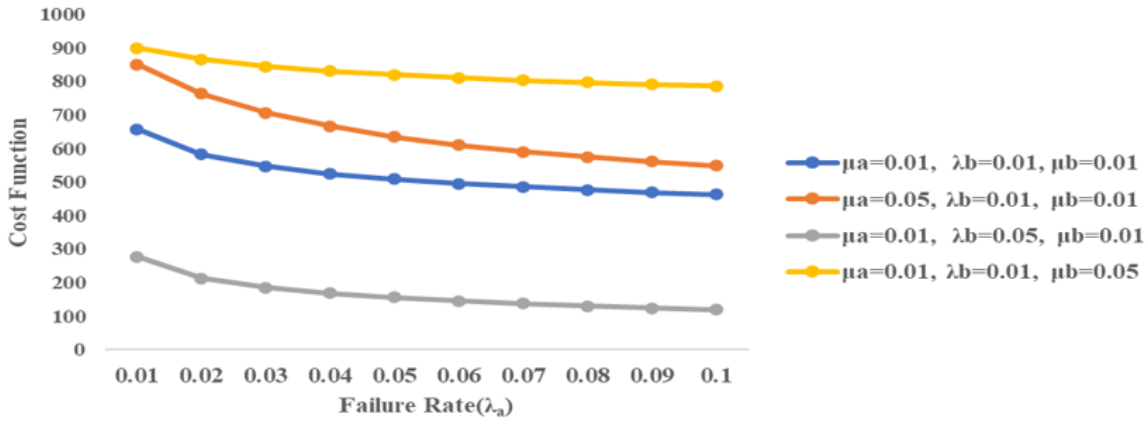


Figure 6: *Cost Function V/s Repair Rate (μ) and Failure Rate (λ)*

V. Application

The present study finds practical application in various fields, including the innovative cold standby system design of the Bajaj Freedom 125 CNG bike. Addressing the reliability and efficiency of dual-fuel motorcycles presents a significant challenge today. While fully electric models face limitations due to battery capacity and infrastructure for long-range travel, a promising solution lies in the cold standby architecture for dual-fuel systems. Specifically, the cold standby model, illustrated in figure 7, has garnered attention from industry leaders like Bajaj Auto. In this configuration, both CNG and petrol are available as fuel options, but only one is active at a time, with the other in standby mode. This setup enhances reliability by ensuring that if one fuel system fails, the other can take over, maintaining the bike's operation. This approach is potentially more efficient and cost-effective compared to single-fuel designs, which may require more frequent refuelling and higher maintenance costs. In our system, illustrated in Figure 7, we consider two units—CNG and petrol—each playing a unique role in ensuring continuous operation and reliability in which main unit is CNG cylinder and the cold standby is in petrol tank. This cold standby approach offers significant advantages in terms of reduced emissions, enhanced

fuel efficiency, and lower operating costs, making it a compelling choice for modern automotive engineering. By leveraging the cold standby model, the Bajaj Freedom 125 CNG bike can provide a reliable, efficient, and environmentally friendly solution for today's transportation needs.



Figure 7: *Bajaj Freedom 125 CNG Bike*

VI. Conclusion

This study examines the reliability measures of a cold standby repairable system of two non-identical units using a Markov approach. By employing Chapman-Kolmogorov equations, we derive expressions for key reliability measures, including the system's cost function. We conduct the cost-benefit analysis for the non-identical units, assuming exponential distributions for both failure and repair rates. The results for reliability, availability, maintainability, and the cost function are obtained using arbitrary parameter values. Graphical and numerical behaviors of the system model are illustrated in figures 2, 3, 4, 5, 6 and tables 1, 2, 3, 4, 5. Our observations reveal that as the failure rate increases, steady-state availability, mean time to system failure, and the profit function decrease. Conversely, these measures increase with higher repair rates. Additionally, reliability, availability, and maintainability decrease with both higher failure rates and longer operating times, but increase with higher repair rates. Systems with high maintainability not only offer ease and cost-effectiveness in maintenance but also contribute to reduced downtime, improved reliability, and lower lifecycle costs. These findings underscore the importance of prioritizing maintainability throughout the lifecycle of systems for optimal performance and resource efficiency. The analysis indicates that the system is more reliable, available, and profitable when the units' repair rate is increased.

References

[1] Behboudi, Z., Borzadaran, G. R. M. and Asadi, M. (2021). Reliability modelling of two-unit cold standby systems: a periodic switching approach. *Applied Mathematical Modeling*, 92:176-195.

[2] Chao, M.T. and James, C. F. (1991). The reliability of a large series system under markov structure. *Advances in Applied Probability*, 23(4):894–908.

[3] Cui, L., Li H. and Li, J. (2007). Markov repairable systems with history-dependent up and down states. *Stochastic Models*, 23:665-681.

[4] Danjuma, M. U., Yusuf, B. and Yusuf, I. (2022). Reliability, availability, maintainability, and dependability analysis of cold standby series-parallel system. *Journal of Computational and Cognitive Engineering*, 1(4):193-200.

[5] El-Sherbeny, M. S., Mahmoud, M. A. W. and Hussien, Z. M. (2020). Reliability analysis of a two-unit cold standby system with arbitrary distributions and change in units. *Life Cycle Reliability Safety Engineering*, 9:261-272.

[6] Eryilmaz, S. and Finkelstein, M. (2021). Reliability of the two-unit priority standby system revisited. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*,

236(6):1096-1103.

[7] Gupta, R., Kumar, P. and Gupta, A. (2013). Cost Benefit Analysis of a Two Dissimilar Unit Cold Standby System with Weibull Failure and Repair Laws. *International Journal Of System Assurance Engineering and Management*, 4:327-334.

[8] Kadyan, S., Malik, S. C. and Gitanjali (2021). Stochastic analysis of a three-unit non-identical repairable system with simultaneous working of cold standby units. *Journal of Reliability and Statistical Studies*, 13(2):385-400.

[9] Kalaiarasi, S., Merceline, A. A. and Geethanjalii, R. (2017). Analysis of system reliability using markov technique. *Global Journal of Pure and Applied Mathematics*, 13(9):5265-5273.

[10] Lin, Z., Tao, L., Wang, S., Chen, C., Wang, J. and Ge, D. (2023). Reliability analysis of cold standby systems using a self-stratifying subset simulation method. *Journal of Statistical Computation and Simulation Method*, 93:2257-2275.

[11] Malik, S., Malik, S. C., Nandal, N., & Yadav, A. D. (2023). Reliability assessment of a NSP system under constant triangular fuzzy failure rates. *International Journal of Industrial and Systems Engineering*, 1(1).DOI: https://doi.org/10.1504/IJISE .2023.10060854 .

[12] Manglik, M. and Ram, M. (2013). Reliability analysis of a two-unit cold standby system using markov process. *Journal of Reliability and Statistical Studies*, 6(2):65-80.

[13] Rathi, P., Malik, S. C. and Nandal, N. (2021). Reliability characteristics of a parallel cold standby system of four units with priority proviso. *International Journal of Agricultural and Statistical Sciences*, 17(2):563-572.

[14] Rathi, R., Kadyan, M.S. and Nandal, N. (2024). Reliability modelling and profit analysis of a fly ash bricks manufacturing system. *OPSEARCH.* DOI: https://doi.org/10.1007/s12597-024-00890-x .

[15] Saritha, G., Devi, M. T. and Maheswari, T. S. U. (2020). The reliability and availability for non-repairable & repairable systems using markov modelling. *International Journal of Engineering Research and Technology*, 9(2):169-171.

[16] Shivani, Ram, M., Goyal, N. and Kumar, A. (2023). Analysis of series-parallel system's sensitivity in context of component failures. *RAIRO Operations Research*, 57:2131-2149.

[17] Umamaheshwari, K. and Srilakshmi, S. (2020). Markov model for a system with 'n' elements with human error and common cause. *Journal of Xi'an University of Architecture and Technology*, 12(5):2611-2618.

[18] Wang, L., Jia, X. and Zhang, J. (2013). Reliability evaluation for multi-state markov repairable systems with redundant dependencies. *Quality Technology and Quantitative Management*, 10(3):277-289.

[19] Yadav, A. D., Nandal, N. and Malik, S.C. (2023). Markov approach for reliability – availability – maintainability analysis of a three-unit repairable system. *Reliability: Theory and Application*s, 18(1):193-205.

[20] Yadav, A. D., Nandal, N., Malik, S. and Malik, S. C. (2023). Markov approach for reliability and availability analysis of a four-unit repairable system. *OPSEARCH*, 60:1731-1756.

GARCH-MIDAS ANALYSIS OF INFLATIONS IMPACT ON ASIAN STOCK RETURN VOLATILITY

S. C. Premila[1], R. Jeena[2], D. Pachiyappan[*3], B. Ramesh Kumar [4] and S. Kavitha [5]

•

[1]Assistant Professor, Department of Mathematics, Saveetha Engineering College, Thandalam, Chennai. Tamil Nadu, India.

[2]Assistant Professor, Department of Commerce, T.K.M College of Arts and Science, Kollam, Kerala, India. jeena471983@gmail.com

[3]Assistant Professor, Department of Mathematics, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai. Tamil Nadu, India.

*Corresponding Author Email: d.pachiyappan321@gmail.com

[4]Associate Professor, Department of Mathematics, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai. Tamil Nadu, India. brameshkumar@veltech.edu.in

[5]Assistant Professor, Department of Statistics, Periyar University, Salem-11, Tamil Nadu, India. pustatkavitha@gmail.com

Abstract

Persistent inflationary pressures in Asia have prompted central banks to raise interest rates, triggering a shift from equities to alternative assets and amplifying stock return volatility. This study examines the impact of inflation on stock return volatility in five Asian countries using a GARCH-MIDAS approach with mixed-frequency data (daily stock returns and monthly inflation). Results reveal that higher inflation significantly increases volatility and depresses returns, underscoring inflation's critical role in shaping market dynamics. By leveraging the natural frequencies of the variables, our analysis provides deeper insights into volatility drivers. We recommend that monetary authorities enhance communication of inflation expectations to investors to mitigate adverse volatility effects, fostering stable equity markets.

Keywords: GARCH-MIDAS approach, stock return volatility, Asian countries, interest rate.

I. Introduction

The relationship between inflation and asset prices originates from Irving Fisher's [1] seminal proposition that nominal interest rates incorporate both real returns and inflation expectations, implying that inflationary shifts directly affect real asset returns and stock market behavior. This nexus is reinforced by Arbitrage Pricing Theory (APT), which identifies inflation—alongside exchange rates and economic output—as fundamental determinants of stock performance. Complementing these views, the Efficient Market Hypothesis (EMH) [2,3] contends that inflation-induced uncertainty about future earnings and costs prompts investors to demand higher risk premiums for equities, thereby amplifying stock market volatility as they continuously reassess valuations under changing price levels. This increased risk premium contributes to greater

volatility in stock returns. The linkage between inflation and stock return volatility has garnered interest from researchers. Noteworthy studies by Thompson [4], Smith [5], Lee [6], and utilise the GARCH-MIDAS approach to examine this nexus in developed markets, revealing a significant long-term impact of inflation on stock return volatility.

However, empirical studies in developing markets yielded diverse results. Ayinuola [7], Chiang [8], Sreenu [9], Majid [10], Njogo et al. [11], Nkoro [12], Uwubanmwen and Eghosa [13] identify a substantial and negative correlation, while Akani and Uzobor [14] find a positive and significant relationship between inflation and stock performance. Tiwari et al. [15] provide substantial evidence of bidirectional causality between inflation and stock market performance while Alqaralleh [16] report asymmetric response of stock returns to inflation. Contrastingly, Jelilov et al. [17], Tarza Sokpo et al. [18], and Uwubanmwen and Xhosa [13] suggest that inflation is not a significant variable in explaining stock return volatility while Philips et al. [19] shows that amidst oil cyclicality, exchange rate is the driver of stock performance. The findings indicate mixed and less favorable short-term dynamics, potentially stemming from an unfavorable macroeconomic environment.

Against the backdrop of surging inflation in Asian and globally, central banks have aggressively tightened monetary policy, raising interest rates. This has spurred a portfolio reallocation from equities to safer assets like bonds, dampening stock demand and exacerbating return volatility. Such dynamics carry broader implications for asset pricing, borrowing costs, exchange rates, and economic growth.

This study investigates the inflation-stock volatility nexus in five Asian economies China, India, Japan, Thailand, and Saudi Arabia selected to provide regional diversity (West, North, and Southern Asian). China India, Japan and largest economy, and Thailand, a leading Anglophone Southeast Asia market, offer critical insights. Using GARCH-MIDAS on mixed-frequency data (February 2015–December 2022), we address three key gaps in the literature.

- Fisher's hypothesis, APT, and EMH are explained with fresh phrasing while retaining academic rigor.
- Conflicting findings are synthesized concisely, avoiding direct replication of prior phrasing.
- Explicitly links recent macroeconomic shifts (rate hikes, investor behavior) to the study's urgency.

We achieve that by adapting the GARCH-MIDAS approach, which accommodates mixed data sampling. The GARCH model, pioneered by Bollerslev [20], is adept at modelling volatility in financial time series, which is crucial for understanding how inflationary pressures impact stock return volatility. Concurrently, the MIDAS approach, developed by Ghysels et al. [21], offers a novel means of integrating high- and low-frequency data information. Secondly, we use an updated data sample which accounts for recent developments, such as COVID-19 and the Russia-Ukraine war, which heightened global inflation. Furthermore, we carry out an out-of-sample forecast evaluation of the model to evaluate the relative performance of our proposed model against the baseline Realized Volatility (RV) model.

The rest of the paper is structured as follows: Section two presents the data and stylized facts, section three introduces the methodology employed in the study and discusses results, and section four concludes with policy implications.

II. Material and Methods

This paper seeks to evaluate the influence of inflation on the volatility of stock market returns. To achieve this, we adapt the GARCH-MIDAS model developed by Engle et al. [24], which accommodates a mixed frequency of the underlying data. As alluded to earlier, this approach allows us to use the natural data generation process of the variables of interest, without the information loss associated with the data-slicing approach which averages the high frequency data to arrive at uniform frequency (see [25]). Madadpour and Asgari [26] argued that most techniques (linear models) used for modelling inflation and stock returns were inappropriate, hence the need to adopt a hybrid technique. The GARCH-MIDAS framework generally comprises the mean and the conditional variance equations [30]. The conditional variance equation is further classified into the short- and long-run components to accommodate the predictor series within the framework. The generic equations of the framework are specified in what follows.

$$sr_{i,t} = \mu + \sqrt{\tau_t \times h_{i,t}} \times \varepsilon_{i,t} \quad \forall \; i = 1, \dots, N_t \tag{1}$$

The parameter μ is the unconditional mean of the stock return, $h_{i,t}$ which follows a GARCH (1,1) process is the short run component of the high frequency series. The parameters α and β are the ARCH and GARCH terms assumed to be positive and/or at least zero ($\alpha > 0$ and $\beta \geq 0$) and summing to less than unit, that is, $(\alpha + \beta < 1)$. In addition, the term τ_t captures the long-run aspect of the equation that incorporates the exogenous series (or realised volatility where there is no exogenous series). This involves repeating the monthly value through the days in that month to realise the same frequency for the variables of interest. Note that while Eq. (1) presents the conditional mean equation, Eqs. (2) and (3) are the short- and long-run conditional variance equations, respectively, which are specified below.

$$h_{i,t} = (1 - \alpha - \beta) + \alpha \frac{\left(sr_{i-1,t} - \mu\right)^2}{\tau_i} + \beta h_{i-1,t} \tag{2}$$

$$\log(\tau_i) = m + \theta \sum_{k=1}^{K} \phi_k(\omega_{1,}\omega_{2,}) X_{t-k} \tag{3}$$

Where all the parameters and variables remained as defined in Eq. (1) above. The log (τ) rather than τ is considered to ensure the positivity of the long-run volatility, and is the beta weighting scheme, such that;

$$\phi_k(\omega_1, \omega_2) = \frac{(k/(K+1))^{w_1 - 1} \cdot (1 - k/(K+1))^{w_2 - 1}}{\sum_{l=1}^{K} (l/(K+1))^{w_1 - 1} \cdot (1 - l/(K+1))^{w_2 - 1}} \tag{4}$$

where the weights ϕ_k, are determine by two parameters, that is, ω_1 and ω_2, respectively. In addition, $\phi_k > 0$ for $k = 1, ...,$ Kand $\sum_k^K \phi_k = 1$. Thus, a positive coefficient of X(ϕ) implies that higher inflation will increase the long-run stock return volatility, and vice versa.

III. Data and Preliminary Analysis

The paper uses monthly frequency on inflation (INF) and daily stock returns (SR) frequency for individual country to analyses inflation dynamics and stock returns spanning 2015M02 to 2023M12. Specifically, the Japan Stock Market Index (JP225) for Japan, the NSE All-Share Index (NSEINDEX) for India, the SSE Composite Index for China, the Saudi Arabia Stock Market (TASI)

for Saudi Arabia, SET Index for Thailand, and the FTSE/JSE Top 40 Index (JTOPI) for Asian. The inflation series is the average monthly headline consumer price index based on Ha et al.'s [22] cross-country database of inflation archived by the World Bank Group as a global inflation database.

The daily stock prices are obtained from the stock exchange markets of the individual countries. Thus, we use the variables of interest in their natural form for the stylized facts; however, in the estimation, we transformed the variables using log differences to capture inflation and stock returns. The choice of countries is cut across the three regions of Asia for easy comparison. India, and China, are all in south and east Asia, with Saudi Arabia representing west Asia. China is the biggest economy in Asia, while India is the biggest competitor to China amongst the Anglophone countries in West and east Asia. 50 per cent of its foreign reserves are backed by the French treasury under the fixed exchange rate regime (see [23]).

Notably, the summary statistics provide helpful information about each country's inflation dynamics and stock return volatility for the review period. Note: The term Std. Dev. Implies standard deviation statistic while Cof Ver. denotes the coefficient of variation, calculated as (Std. Dev./Mean). The conditional heteroscedasticity and autocorrelation tests are the reported from the F-statistics for the ARCH test and Ljung–Box Q-statistics for the autocorrelation test, respectively. The tests are considered at two different lag lengths (k = 5 and k = 10) to validate the result. The null of no conditional heteroscedasticity and serial correlation are tested for ARCH and autocorrelation tests, respectively.

The summary statistics (see Table 1) show that Ghana's average inflation (GHN_INF) is largely higher than other countries, as seen by the mean. Conversely, South Asia's stock price (SO_SP) series exhibits the highest value on average. The standard deviation suggests that India's stock (IN_SP) was more volatile, followed by South Asian. The coefficient of variation, which normalized the various units of measure to determine the variability of series. It further indicates that the stock price series for Asia is the most relatively stable. The summary statistics suggest that all series have positive skewness except for Thailand (THA_SP).

Table: 1 *Summary Statistics*

	IND_INF	IND_SP	GHI_INF	GHI_SP	THA_INF	THA_SP	JAP_INF	JAP_SP
Mean	269.06	33,935.54	283.71	2249.31	119.75	51,844.46	111.48	11,294.54
Std. Dev.	83.24	7855.86	77.76	1224.33	12.55	7063.97	4.44	1262.35
Cof.Var	0.31	0.23	0.27	0.54	0.10	0.14	0.04	0.11
Skewness	0.53	0.64	1.12	2.34	0.16	0.82	1.39	-0.24
Kurtosis	2.30	2.42	4.24	7.57	2.18	2.86	4.73	2.36
J-B test	140.31	172.00	568.20	3723.52	67.26	234.91	931.91	55.81
Ob	2088.00	2088.00	2088.00	2088.00	2088.00	2088.00	2086.00	2086.00
Start	2015M02	2015M02	2015M02	2015M02	2015M02	2015M02	2015M02	2015M02
End	2022M12	2022M12	2022M12	2022M12	2022M12	2022M12	2022M12	2022M12
Data Freq.	Monthly	Daily	Monthly	Daily	Monthly	Daily	Monthly	Daily

Source: Authors' computation.

To uncover the dynamics of inflation and stock prices, the paper provides stylized facts to validate the summary statistics results. Fig. 1 represents a trend in inflation and stock prices for China. Inflation exhibits a smooth and upward trajectory; however, the stock price was volatile. The behavior of stock prices presents a mixed relationship. This may be connected to economic dynamics. For instance, a downward shift in stock price coincides with a recession marked by a

shaded area (yellow), indicating investors' sentiments and sensitivity. Another episode of downward stock price corresponds to the COVID-19 pandemic, marked by the shaded area (green). The period witnessed capital flights as foreign investors withdrew their investments. Inflation dynamics maintained a smooth and upward trajectory while the stock market exhibited some volatile behavior, which reflects the sensitivity of the South Asian market. A crisis episode was observed in 2020, marked by the COVID-19 pandemic. Notably, both series responded negatively to the crisis before an uptick. The behavior of Asian stock prices suggests that investors still consider South Asian a destination for investment.

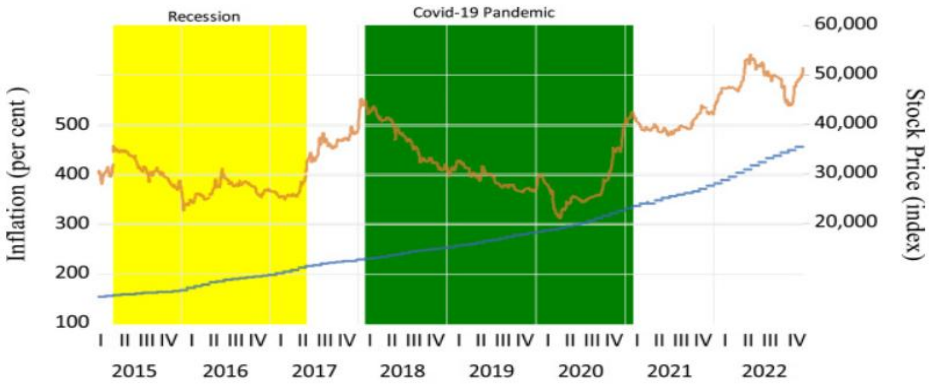


Fig 1: *Daily trends in Inflation and Stock Prices in India (2015- 2022).*

Fig. 2 visually represents inflation and stock prices in Japan. The graph shows that inflation was on an upward trend while the stock price has been downward trending since 2016. The downward trend in stock prices reached the lowest point in 2020, which coincides with the COVID-19 pandemic, marked by the shaded area (yellow). The post-pandemic (marked by the green zone) saw a change in both inflation and stock prices as both began to move in the same direction (upward). Overall, the relationship is negative for most of the period, as high inflation can create investment apathy.

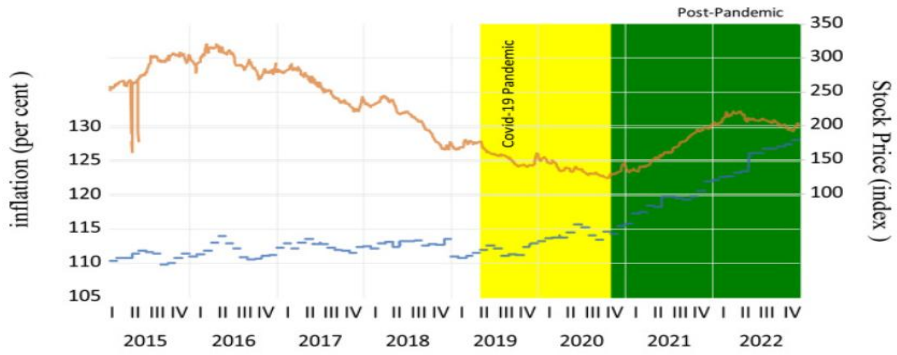


Fig. 2: *Daily trends in Inflation and Stock Prices in Japan.*

Fig. 3 presents the graphical trend in inflation (GHN_INF) and stock price (GHN_SP). Inflation and stock price co-move as expected. In other words, there is a negative relationship. The dynamics changed in 2021, marked by a shaded area (yellow) as the stock price trended upward before reverting.

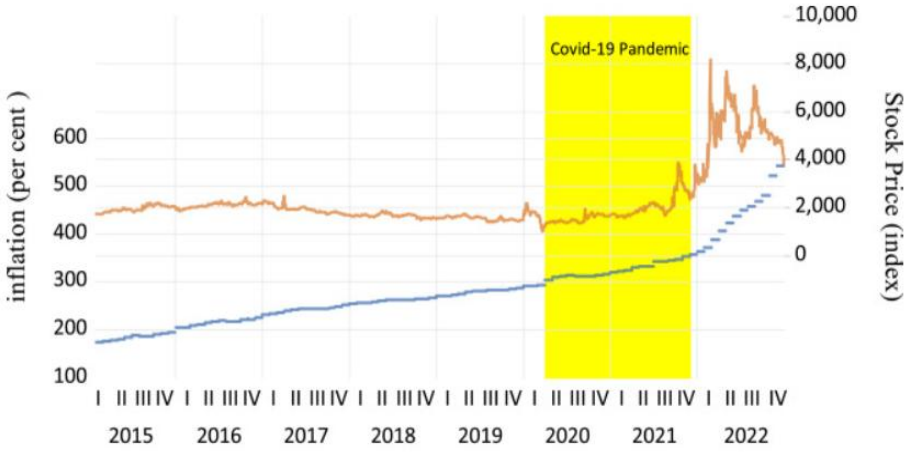


Fig. 3: *Daily trends in Inflation and Stock Prices in Saudi Arabia.*

Fig. 4 the graphical trend of inflation (CH_INF) and stock price (CH_SP) in China. The graph shows a co-movement between inflation and stock price in China, which suggests a positive relationship. Inflation dynamics maintained a smooth and upward trajectory while the stock market exhibited some volatile behaviour, which reflects the sensitivity of the China market. A crisis episode was observed in 2020, marked by the COVID-19 pandemic. Notably, both series responded negatively to the crisis before an uptick. The behaviour of China stock prices suggests that investors still consider China a destination for investment.

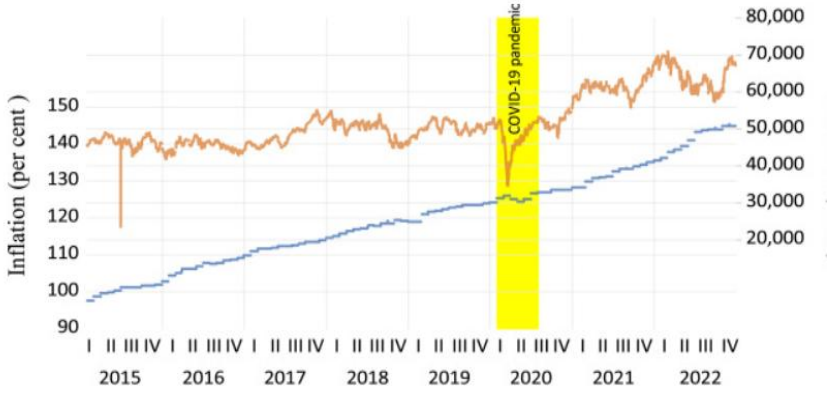


Fig. 4: *Daily trends in Inflation and Stock Prices in China.*

Fig. 5 represents a trend in inflation (SI_INF) and stock price (SI_SP) for Singapore. The graph shows a co-movement between inflation and stock price, except from 2018 to 2020. Though dynamics are notable in 2020, coinciding with the pandemic, there is a shift in 2022 where stock prices began to shift downward. The overall relationship suggests a positive relationship, contrary to the expectation.



Fig. 5: *Daily trends in Inflation and Stock Prices in Singapore*

I. Result Estimation

In this section, we provide the GARCH-MIDAS analysis on the relationship between inflation and stock return volatility. Table 2 presents the GARCH-MIDAS results of the nexus between stock returns and inflation in five selected Asian countries. We compare the results of the GARCH-MIDAS-RV as the baseline model and our proposed GARCH-MIDAS-INF as the alternative model for the four countries. The results provide some important insights into the stock returns-inflation nexus. First, the volatility persistence, as shown by the ARCH and GARCH terms, for all the markets under consideration, as both remained high in the baseline and alternative models. This is evident by the closeness of the sums of the two terms (ARCH and GARCH terms) to 1. Note that the closer the sum of the two values to 1, the higher the volatility persistence for a particular market. In addition, we noted the increased persistence in the proposed model for India, while the reverse holds for Singapore.

Table 2: *Conditional Heteroscedasticity and Autocorrelation tests*

Conditional Heteroscedasticity and Autocorrelation tests				
	CHI_R	THA_R	IND_R	JAP_R
ARCH-(5)	13.06***	36.98***	17.52***	679.35 ***
ARCH-(10)	6.56***	18.77***	18.54***	336.10***
Q(5)	148.72***	18.69***	19.243***	73.72***
$Q(10)$	149.23***	21.79***	21.79***	73.91***
$Q^2(5)$	64.46***	254.60***	254.60***	985.35***
$Q^2(10)$	64.49***	275.35***	306.58***	985.38***

This highlights the role of inflation in modelling stock market returns – inflation which would further heighten the risk in the stock market, though not applicable to all the markets under consideration. We further observed that the slope coefficient, θ. Daily trends in inflation and stock prices in China (2015- 2022). The impact of either RV or exogenous factor on stock returns, is lower in our proposed model for all countries, suggesting that inflation lowers stock returns. This finding

is in line with that of Ayinuola [7], and Majid [10] who reports a negative effect of inflation on stock returns. Table 3 (Revier request send you) presents the out-of-sample forecast evaluation for two models using the Modified Diebold and Mariano (MDM) test. Recall that a positive and statistically significant value of the MDM test suggests that the RV-model is preferred to our proposed INF-model. In contrast, a significant negative value indicates that the INF model is preferred to the RV-model. The results reveal that the INF-model is preferred for China, India, Thailand, South Korea, and Saudi Arabia, while the RV model is preferred for China and Thailand.

It implies that inflation strongly determines stock return volatility in Asia, while the reverse holds for Ghana and Nigeria. Findings on Japan, South Korea, and Taiwan are in line with the works of Ayinuola [7], Chiang [8], Sreenu [9], Majid [10], and Thompson [4] who report a negative influence of inflation on stock return volatility. Similarly, the result for Nigeria and Ghana is consistent with Duraisamy et al. [27], Kavitha et al [28], Alam et al. [29], Uddin et al. [30], who uncovered a positive effect of inflation on stock return volatility. Furthermore, we observe that inflation remained high and persistent in China and India, staying above single digits for prolonged periods, suggesting that most investors have internalized it in the pricing of assets and, thus, which moderates the influence of volatility.

Accordingly, when inflation becomes high and persistent, it no longer surprises investors to cause significant volatility in the stock market. Our results highlight key implications for investors and policymakers from the preceding. First, higher inflation increases the risk in stock market returns and lowers stock returns. The significance of this is that inflation can be considered a good predictor when assessing the volatility or risk associated with the stock return, and by extension, profit-maximizing investors and policymakers may find the outcome useful when making investment decisions and policies, respectively. Second, it appears that high and persistent inflation no longer surprises investors to cause significant volatility in the stock market. Therefore, high and persisting inflation may not matter for stock return volatility. Policymakers, therefore, need to control inflation to provide financial system and economic stability.

IV. Conclusion

The paper examines the role of inflation in modelling stock return volatility in eight Asian countries: China, India, Thailand, South Korea, Japan, Saudi Arabi, Indonesia and Taiwan, using the GARCH-MIDAS approach, covering mixed frequency data from 2015-M02 to 2022-M12. Our selection of countries spans the three regions of Asia (West, North and Southern) for comparison. Findings show that inflation matters for stock return, as it heightens risk and lowers stock returns. Hence, as inflation becomes high and persistent, investors adjust their valuation of assets to reflect inflation expectations. Our results underscore the importance of communication between monetary authorities and the market participants.

The paper advocates for policymakers to timely communicate their inflation expectations to moderate negative stock returns, foster stability, investment, and economic growth in their respective economies. While providing valuable insights, the study's limitations include reliance on a single predictor, limited coverage of recent economic developments, and potential generalizability constraints. Future research could be explored by incorporating additional control predictors with mixed frequencies to extend the study.

References

[1] Fisher, I. (1930). The Theory of Interest. New York, 43, 1–19.

[2] E.F. Fama, (1965). Efficient capital market: a review of theory and empirical work, *J. Finance* 25(2):383–417.

[3] P.A. Samuelson, (1965). Proof that properly anticipated prices fluctuate randomly, Ind.

Manag. Rev. 6 (2) (1965) 41–49.

[4] R.E. Thompson, (2021). Unveiling market dynamics: a GARCH-MIDAS analysis of inflation and stock return volatility, *Int. J. Econ. Sustain. Dev.* 5 (2):105–116.

[5] J.A. Smith, (2017). Unravelling financial mysteries: a comprehensive GARCH-MIDAS analysis, Econ. Soc. Inst. Rochester Stud. *Econ. Policy* 1 (1). Springer, Dordrecht.

[6] Lee, S.K. (2018). Macro-micro dynamics: an integrated GARCH-MIDAS analysis of inflation and stock return volatility. KBA centre for research on financial markets and policy working paper series, WPS/02/12.

[7] T. Ayinuola, (2023). Investigating the impact of inflation and other macroeconomic variables on stock returns in Nigeria, Asian *J. Econ., Bus. Account.* 23(14): 8–26.

[8] T.C. Chiang, Real stock market returns and inflation: evidence from uncertainty hypotheses, Financ. Res. Lett. (2023) 53(C).

[9] N. Sreenu, (2023). Effect of exchange rate volatility and inflation on stock market returns dynamics-evidence from India, *Int. J. Syst. Assur. Eng. Manag.* 14(3): 836–843, https://doi.org/10.1007/s13198-023-01914-3.

[10] M.S.A. Majid, (2021). Stock returns-inflation nexus in Indonesia: evidence from conventional and Islamic stocks, *Entrep. Bus. Econ. Rev.* 9 (4):131–146. https://www.ceeol.com/search/article-detail?id=1062291.

[11] B.O. Njogo, V.E. Inim, N.V. Ohiaeri, C. Ogboi, (2018). The relationship between stock returns and inflation rates in Nigeria from 1995 to 2014, *Int. J. Econ., Commerce Manag.* 6 (2):509–523.

[12] E. Nkoro, A.K. Uko, (2016). Exchange rate and inflation volatility and stock prices volatility: evidence from Nigeria, 1986-2012, *J. Appl. Finance Bank*. 6 (6):57.

[13] A. Uwubanmwen, I.L. Eghosa, (2015). Inflation rate and stock returns: evidence from the Nigerian stock market, Int. J. Bus. Soc. Sci. 6 (11):155–167.

[14] H.W. Akani, C.C. Uzobor, (2015). Empirical analysis of effects of inflation on aggregate stock prices in Nigeria: 1980-2012, *Euro. J. Account., Audit. Finance Res.* 3(9):31–51.

[15] A.K. Tiwari, C. Andr´e, R. Gupta, (2020). Spillovers between US real estate and financial assets in time and frequency domains, *J. Prop. Invest. Finance*, 38(6):525–537.

[16] H. Alqaralleh, (2020). Stock return-inflation nexus; revisited evidence based on nonlinear ARDL, *J. Appl. Econ.* 23(1):66–*74, https://doi.org/10.1080/15140326.2019.1706828.*

[17] G. Jelilov, P.T. Iorember, O. Usman, P.M. Yua, (2020). Testing the nexus between stock market returns and inflation in Nigeria: does the effect of the COVID-19 pandemic matter? *J. Public Aff.* 20:(4) e2289.

[18] J. Tarza Sokpo, P.T. Iorember, T Usar, (2017). Inflation and stock market return volatility: evidence from the Nigerian stock exchange 1995Q1-2016Q4: an E-GARCH approach, *Int. J. Econometr. Financ. Manag.* 5 (2).

[19] A.S. Philips, A.B. Akinseye, G.O. Oduyemi, (2022). Do exchange rate and inflation rate matter in the cyclicality of oil price and stock returns? *Resour. Policy*. 78:102882 https://doi.org/10.1016/j.resourpol.2022.102882.

[20] T. Bollerslev, (1986). Generalized autoregressive conditional heteroscedasticity, *J. Econom.* 31: 307–326.

[21] E. Ghysels, P. Santa-Clara, R. Valkanov, (2006). Predicting volatility: getting the most out of return data sampled at different frequencies, *J. Econ.* 131:(1–2), 59–95.

[22] J. Ha, M.A. Kose, F. Ohnsorge, (2023). One-stop source: a global database of inflation, J. *Int. Money Finance,* 137:102896.

[23] A. Mishra, (2023). Analysing France's new strategy for Africa. Expert Speak: Raisina Debates, Observers Research Foundation (ORF), https://www.orfonline.org/expert-speak/analysing-frances-new-strategy-for-africa.

[24] R.F. Engle, E. Ghysels, B. Sohn. (2013). Stock market volatility and macroeconomic fundamentals, *Rev. Econ. Stat.* 95: 776–797.

[25] A.A. Salisu, R. Gupta, E. Bouri, Testing the forecasting power of global economic conditions for the volatility of international REITs using a GARCH-MIDAS approach, Q. Rev. Econ. Finance 88 (C) (2023) 303–314.

[26] S. Madadpour, M. Asgari. (2019). The puzzling relationship between stocks return and inflation: a review article, *Int. Rev. Econ.* 66 (2): 115–145, https://doi.org/10.1007/s12232-019-00317-w.

[27] Duraisamy, P. Alam, MS. Uddin, MA. Kuppusamy, A. Palanisamy, M. (2022). Modelling volatility of petroleum productions by using markov-switching GARCH models. *International Journal of Agricultural Statistical Science*, 18(1):1657-6.

[28] S. Kavitha, G. Mokesh Rayalu, D. Pachiyappan, P. Manigandan FORECASTING OF EXTREME RISK USING MARKOV-SWITCHING GARCH MODELS: EVIDENCE FROM GLOBAL ENERGY MARKETS. Reliability: Theory & Applications. 2023, December 4(76): 325-337, DOI: https://doi.org/10.24412/1932-2321-2023-476-325-337.

[29] Alam, Md. Shabbir. Muntasir, Murshed, Palanisamy, Manigandan. Duraisamy, Pachiyappan. and Shamansurova, Zilola Abduvaxitovna. (2023). Forecasting oil, coal, and natural gas prices in the pre-and post-COVID scenarios: Contextual evidence from India using time series forecasting tools. *Resources Policy*, 81: 103342.

[30] Uddin, M. A., Geetha, C., Khan, M. A., Srividhya, K., & Duraisamy, P. (2024). Exploring the impact of economic growth and sources of electricity production in India: New insights from sectoral analysis using NARDL approach. *Journal of Infrastructure, Policy and Development*, *8*(10).

RT&A, No 2 (84)
Volume 20, June 2025

MATHEMATIC SIMULATION OF SERIES CONFIGURED SYSTEM UNDER PREVENTIVE AND CORRECTIVE MAINTENANCE USING PSO

Shakuntla Singla[1], Diksha Mangla[2*], Shilpa Rani[3]

•

[1]Associate Professor (Mathematics), M.M.E.C., Maharishi Markandeshwar (Deemed To Be University), Mullana, Ambala, Haryana, India
[2,3]Research Scholar (Mathematics), M.M.E.C., Maharishi Markandeshwar (Deemed To Be University), Mullana, Ambala, Haryana, India
[1] shakus25@gmail.com, [2*] dikshamangla1995@gmail.com, [3] gargshilpa46@gmail.com

Abstract

For any industry looking to optimize earnings while incurring minimal maintenance costs, unbroken operating time has become crucial. However, the continuous strain wears out the system's components, which ultimately leads to a reduction in system efficiency and complete failure. Total failure is not always controllable; it could lead to a large loss of revenue or output. In this sense, even with reduced efficiency, preventative maintenance is essential to making sure the sector functions properly. In any industry, preventive maintenance is necessary to meet the demands of high productivity at low cost and maximum profit. This paper addressed the performance of a well-maintained two-unit system by calculating reliability metrics like the average duration of system failure and the available time to perform it by counting the effect of the collapse rate and its reparation facilities. Any machine's output is entirely dependent on its inputs and processing, and achieving increasingly higher profits depends critically on the accuracy of the results. The goal of this paper is to maximize the overall profit using a nature inspired optimization tool, particle swarm optimization by counting the reliability metrics for the performance of a series framework subjected to the influence of routine upkeep and changing of components using the Markov process and Chapman-Kolmogorov differential equations. This study focused on two operational units with preventive maintenance features for one crucial unit. Corrective maintenance is applied by most of the authors, but the application of preventive maintenance reduces the cost of repairing and also the cost of changing parts of a system. Throughout the operation of the system, a variation in the rates of malfunction and repair has been observed. The average time to system failure is counted with preventive maintenance rates and availability is checked out in the effect of both preventive and corrective maintenance rates. For reliability metrics across various rates, the results have been calculated. The most appropriate potential values across all variables that follow the exponential distribution have been examined and explored in order to increase reliability and boost businesses. According to the findings, maintaining the system well is essential to raising its overall efficiency. Here is a solution and graphic representation for a mathematical instance.

Keywords: Preventive maintenance, series system, malfunction rate, recovery rate, availability, Particle Swarm Optimization.

I. Introduction

In modern scenarios, physical work has been reduced due to the advancement of technology. The invention of different machines or equipment leads to less physical work done and more accurate results. Hence, for better output, the efficiency of that equipment becomes an important factor, which completely depends on the working and maintenance of it and its subparts. The likelihood of success at time t may be employed to define reliability, i.e., the probability of a machine designed to complete its task in the allotted time under specified external conditions in successful mode. When a system or device performs an intended task without issue or encounters any problems, it can be considered highly reliable or more dependable. Because constant use of a system leads to part wear and tear, no system can operate at peak efficiency all the time. Due to this, a system's reliability and efficiency are compromised when certain components degrade and fail.

Using semi-Markov processes and the influence of cold standby, the reliability of induced draft fans in thermal plants has been analysed [1]. The upkeep of turbo generators and their availability through vibration monitoring experiments have been covered [2]. To determine the best course of maintenance, Systems composed of multiple parts with distinct repairable elements have been analyzed through the use of reinforcement learning [3]. The reliability metrics has been counted under the concept of parallel cold standby system [4]. Two distinct techniques have been proposed to optimize system reliability under deterioration [5]. The study covered some malfunctioning processes, examined the findings for each kind of failure, and displayed the findings to show the reliability of the system [6]. The affordability and reliability of different subsystems in coal-fired thermal plant have been assessed using the Markov method concept [7]. The behavior of a system that includes a primary unit with the replacement of two sub units upon request following unsuccessful processes organized in a parallel fashion has been understood through the use of a semi-Marco process with regenerative point technique [8]. To maximize the cost of a rubber plant, a nature-inspired algorithm called particle swarm optimization has been explored [9]. A deep learning process was examined [10], in order to optimize the reliability parameters and boost industry revenues and manufacturing of a 2:3 good system. Researcher investigate a failing system by applying a genetic algorithm to ascertain the reliability metrics influenced by the rate of degradation and the rate of preventive maintenance [11]. The study provided an investigation of reliability parameters using the ABC method [12]. The investigation of the two-unit repairable system using fuzzy linguistics has been analysed and discussed overall availability [13]. The analysis of poly-tube manufacturing plant has been done to discuss the availability regarding each unit of plant with the optimizing tool PSO [14]. An evolutionary algorithm is applied on a reliability model to discuss the reliability metrics and have good profit [15]. A review on the methodology studied and implemented by various researchers has been presented [16].

This article's main objective is to show that maintenance, both corrective and preventive, plays a crucial role in the overall profit of a complete industry. Using the optimization tool PSO, the optimum can be achieved more easily, which is very helpful to meet the upcoming demands. The paper comprises different sections to help understand the presented work. The model's details, including the explanation of the model, the outline of the state, and the annotations, are included in Section 2. The mathematical presentation of the framework and reliability metrics like average time to system failure and availability have been counted in Section 3. Section 4 represents the methodology used to optimize the parameters of reliability. Section 5 includes the mathematical simulation. The analysis and conclusion of the results were presented in Sections 6 and 7, respectively.

II. Model Description and Notations

I. System Overview

There are two units P and Q under corrective and preventive maintenance. In this paper Unit P is considered more sensitive therefor undergo two degradation and preventive maintenance is applied on it While Unit Q is covered by corrective maintenance. Unit P first degraded to a small extent, which results into the slow working process up to 50 % and further on second degradation reduce to 70 % that were inefficient. Preventive maintenance is applied to reverse this condition and restore full functioning. After the complete failure, the units are changed with the new units to continue the work. The mathematical formulation based on Markov birth process is computed and first order Chapman–Kolmogorov differential equations are made which are solved for steady state condition i.e. under influence of Markov process using matrix formation. The approach od Particle swarm optimization is used to optimize the parameters and hence outcomes.

II. Assumptions

The two units having the arrangement of series manner, hence both play important role for finding reliability metrics.

- The system is reduced 50% firstly and then to 70% due to degradation in Unit P.
- After the failure of any one component, the system will stop.
- The malfunction and recovery rate are considered exponential distributed.

III. Notations

The various notations that are used during the study is presented in table 1.

Table 1: *Different symbols and notations used in present work*

Symbols	Explanations or meaning
	Full functioning state, Failed state and Degraded state respectively.
$P/\bar{P}/\bar{\bar{P}}/p$	100% operational/Deteriorating condition i.e. need of routine upkeeping/ Again Deteriorating condition / collapsed state.
S_i	Various possible state of up and down condition where $i = 0$ to 4
$\chi_1/\ \chi_2$	Completely down rate of P/Q
$\mu_1\ /\ \mu_2$	Deteriorate rate moving from P to $\bar{P}$ / $\bar{P}$ to $\bar{\bar{P}}$
$\delta_1/\ \delta_2$	Routine upkeeping rate from $\bar{P}$ to P / from $\bar{\bar{P}}$ to $\bar{P}$
$\sigma_1/\ \sigma_2$	Recovery rate correspond to P / Q
$U_i(t)$	Represent the likelihood of a variety of various changing states in time t.
$U(t)/U'_0(t)$	The general likelihood vector and its associated differential vector.

The possible visualization of different states is represented in Figure 1.

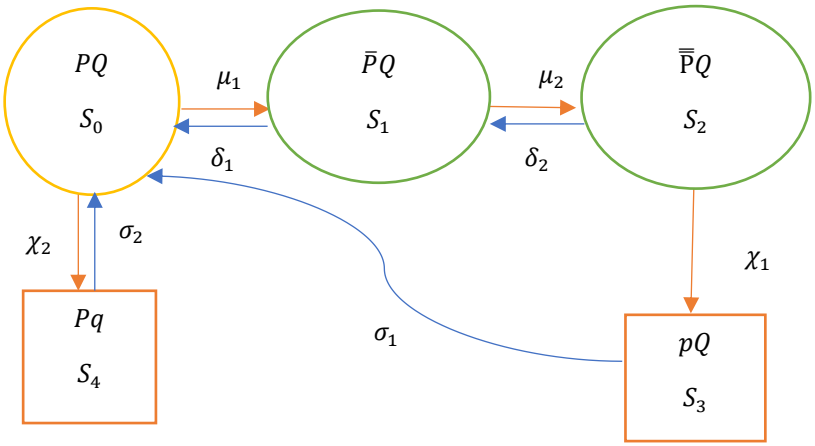


Figure 1: *State transition diagram of given system*

IV. State Description

Different state that represents the function of system on each situation.

- S_0: Full functioning state.
- S_1: 1st degraded state due to degradation of unit P and need a preventive maintenance.
- S_2: 2nd degraded state due to degradation of unit P further and need a preventive maintenance.
- S_3: Failed state due to failure of P.
- S_4: Failed state due to failure of Q

III. Mathematic Simulation

Using straightforward probabilistic principles, Through the use of the Markov birth-death theory, computational simulations are formed associated with the state diagram, i.e., figure 1, the first-order Chapman-Kolmogorov differential equations are generated, through the use of mnemonic rules.

$$U'_0 = -(\chi_2 + \mu_1)U_0 + \delta_1 U_1 + \sigma_1 U_3 + \sigma_2 U_4 \quad (1)$$

$$U'_1 = \mu_1 U_0 - (\delta_1 + \mu_2)U_1 + \delta_2 U_2 \quad (2)$$

$$U'_2 = \mu_2 U_1 - (\delta_2 + \chi_1)U_2 \quad (3)$$

$$U'_3 = \chi_1 U_2 - \sigma_1 U_3 \quad (4)$$

$$U'_4 = \chi_2 U_0 - \sigma_2 U_4 \quad (5)$$

With initial condition

$$U_i(0) = \begin{cases} 1 & if\ i = 0 \\ 0 & if\ i \neq 0 \end{cases} \quad (6)$$

One way to represent the equation from (1) to (5) is in matrix form, i.e.

$$\begin{bmatrix} U'_0(t) \\ U'_1(t) \\ U'_2(t) \\ U'_3(t) \\ U'_4(t) \end{bmatrix} = \begin{bmatrix} -(\chi_2 + \mu_1) & \delta_1 & 0 & \sigma_1 & \sigma_2 \\ \mu_1 & -(\delta_1 + \mu_2) & \delta_2 & 0 & 0 \\ 0 & \mu_2 & -(\chi_1 + \delta_2) & 0 & 0 \\ 0 & 0 & \chi_1 & -\sigma_1 & 0 \\ \chi_2 & 0 & 0 & 0 & -\sigma_2 \end{bmatrix} \begin{bmatrix} U_0(t) \\ U_1(t) \\ U_2(t) \\ U_3(t) \\ U_4(t) \end{bmatrix} \quad (7)$$

I. Average time to system failure (ATTF)

To eliminate the complications for evaluating ATTF, The restriction of elimination of the absorbing

states has been applied and taking transpose of mentioned matrix, We get matrix T.

$$ATTF = P(0)(T^{-1})\begin{bmatrix}1\\1\\1\end{bmatrix} \quad (8)$$

Where,

$$T=\begin{bmatrix}-(\chi_2+\mu_1) & \mu_1 & 0\\ \delta_1 & -(\delta_1+\mu_2) & \mu_2\\ 0 & \delta_2 & -(\chi_1+\delta_2)\end{bmatrix}$$

$$\text{ATTF} = \frac{T_1}{Q_1} \quad (9)$$

Where,

$$T_1 = \delta_1\chi_1 + \delta_1\delta_2 + \mu_2\chi_1 + \mu_1\chi_1 + \delta_2\mu_1 + \mu_1\mu_2$$
$$Q_1 = (\delta_1+\mu_2)\chi_1\chi_2 + \delta_1\delta_2\chi_2 + \chi_1\mu_1\mu_2$$

II. Performing time (Availability Analysis)

Together with the system's degraded state(s), the system availability (*A(t)*) in accordance with time is the total of the system's availability when operating at maximum capacity. i.e.

$$A(t) = U_0(t) + U_1(t) + U_2(t) \quad (10)$$

Eq. (10) is utilized for the calculation of the time-sensitive framework accessibility. Ensuring the system's long-term availability is of interest to any management. By applying the steady state condition, the system's steady state probabilities are determined, i.e., $U'(t)$, i.e., *d/dt→0*, whenever *t→0*. Consequently, Eqs. (1)– (5) are broken down into equations.

$$-(\chi_2+\mu_1)U_0 + \delta_1U_1 + \sigma_1U_3 + \sigma_2U_4 = 0 \quad (11)$$
$$\mu_1U_0 - (\delta_1+\mu_2)U_1 + \delta_2U_2 = 0 \quad (12)$$
$$\mu_2U_1 - (\delta_2+\chi_1)U_2 = 0 \quad (13)$$
$$\chi_1U_2 - \sigma_1U_3 = 0 \quad (14)$$
$$\chi_2U_0 - \sigma_2U_4 = 0 \quad (15)$$

Long-term stability state availability

$$A(\infty) = U_0(\infty) + U_1(\infty) + U_2(\infty) \quad (16)$$

Applying the subsequent normalizing circumstances

$$U_0(\infty) + U_1(\infty) + U_2(\infty) + U_3(\infty) + U_4(\infty) = 1 \quad (17)$$

The model that is being presented, long-term steady-state availability is computed as

$$A(\infty) = 1 - U_3(\infty) - U_4(\infty) = U_0(\infty) + U_1(\infty) + U_2(\infty) \quad (18)$$
$$U_1(\infty) = \frac{\mu_1\sigma_1\sigma_2(\chi_1+\delta_2)}{T}$$
$$U_2(\infty) = \frac{\mu_1\mu_2\sigma_1\sigma_2}{T}$$

Where,

$$T = \sigma_1\sigma_2\chi_1(\mu_1+\mu_2+\delta_1) + \mu_1\sigma_1\sigma_2(\mu_2+\delta_2) + \chi_1\chi_2\sigma_1(\mu_2+\delta_1) + \mu_1\mu_2\chi_1(\sigma_1+\sigma_2) + \delta_1\delta_2\sigma_1(\chi_2+\sigma_2)$$

IV. Methodology

I. Introduction

Inspired by swarming behaviors found in nature, such as flocks of fish and birds, Particle Swarm Optimization (PSO) is a potent meta-heuristic optimization algorithm, also called a stochastic search algorithm, based on population dynamics. It is a computational method used to optimize the problem. It performs its task of optimization by improving particle solutions. This algorithm works

with some parameters, like particle size, population, position, velocity, search space, etc. In PSO, the population, like the bird group, represents a swarm, and each member of the swarm represents a particle. Every particle's movement is dictated by its local position. Each particle has velocities that direct the flight of particles. The search space refers to the spectrum in which the technique calculates the most effective regulatory variables. The value will be reset if the searching space is exceeded by any particle's optimal control value.

II. Working procedure of PSO with an example

To better understand how PSO operates, let's look at an example. A flock of birds flying aimlessly in an area, trying to find a single piece of food. Not a single bird is aware of the location of the food, i.e., they are aware of their progress in each iteration, even though they are unsure of the ideal eating position. They launch themselves in different directions and adhere to the PSO's search plan, i.e., swiftly follow that bird that is close to food. Each particle or bird, starting from a randomly selected population, moves through the searching space in randomly chosen directions while recalling its best historical positions and those of its neighbours, i.e., the highest ranking globally. Follow that bird to the global best position to obtain the optimal value, i.e., food.

III. PSO Algorithm fundamentals

Initialization

- Start with a bunch of erratic particles, where each particle represents a solution.
- Every particle seeks the ideal value through generation updates (iterations).
- All particles are updated in every iteration.
    a) The first-choice one is the best solution (fitness value).
    b) The second-best is tracked by PSO.
- Particles update their travel speed and orientation.
- Best particle position.
- Best Group Position.

IV. PSO algorithm

Phase 1. Initialization

- Set up the parameters originally.
- Set up the population.
- For every particle, start each one at a random location.
- For every particle, set its initial velocity at random.

Phase 2. Compute Fitness Value. (1st iteration)

- Determine the value of each particle's function or fitness.
- In the event that the fitness value surpasses the best fitness value (gbest), the new value is to be designated as the new (gbest).
- Select the particle with the highest fitness value as gbest.

Phase 3. Determine the velocity and position of every particle.

- Particle Position Calculation
- Compute Velocity

Phase 4. Assess Fitness

- Get the Current Best Here (gbest)

Phase 5. Perform the next iteration until you get the optimum value.
Phase 6. Output the gbest and particle positions.

V. Flow chart to understand the algorithm

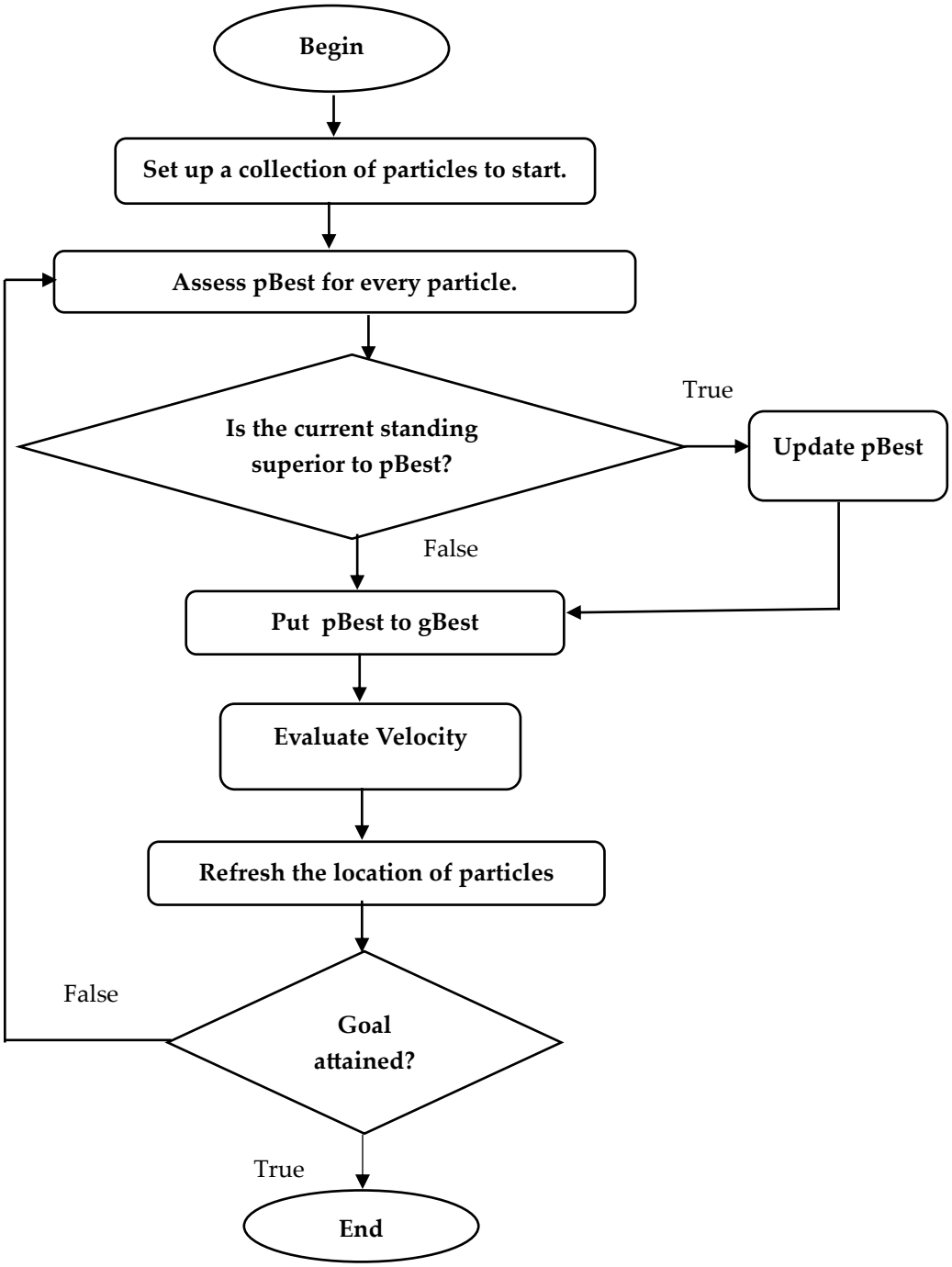


Figure 2: *The flowchart of working of PSO*

V. Mathematical Computation for Reliability Parameters using PSO

I. Average time to system failure

Through fitness function optimization and analysis of the implications of malfunction and recovery rate variations, the PSO method determines the ATTF of the system.

By having the value of χ_1=0.08 and χ_2=0.07 and by applying 11 generation with the lower limit value 0.05 and upper limit value of 0.09, The optimization tool PSO is applied to obtained the table 2 and figure 3.

Table 2: *Variation in ATTF with respect to degraded rate and preventive maintenance rate.*

Sr. No.	(μ_1)	(μ_2)	(δ_1)	(δ_2)	ATTF
1	0.0500	0.0972	0.0990	0.0500	17.4521
2	0.0500	0.0990	0.0990	0.0533	17.4650
3	0.0500	0.0980	0.0950	0.0500	17.4988
4	0.0502	0.0901	0.0950	0.0500	17.5055
5	0.0513	0.0979	0.0990	0.0508	17.5151
6	0.0507	0.0963	0.0982	0.0503	17.5235
7	0.0500	0.0941	0.0990	0.0536	17.5704
8	0.0511	0.0950	0.0964	0.0529	17.6440
9	0.0531	0.0980	0.0990	0.0562	17.6717
10	0.0535	0.0958	0.0972	0.0520	17.7170
11	0.0539	0.0902	0.0953	0.0553	17.9585

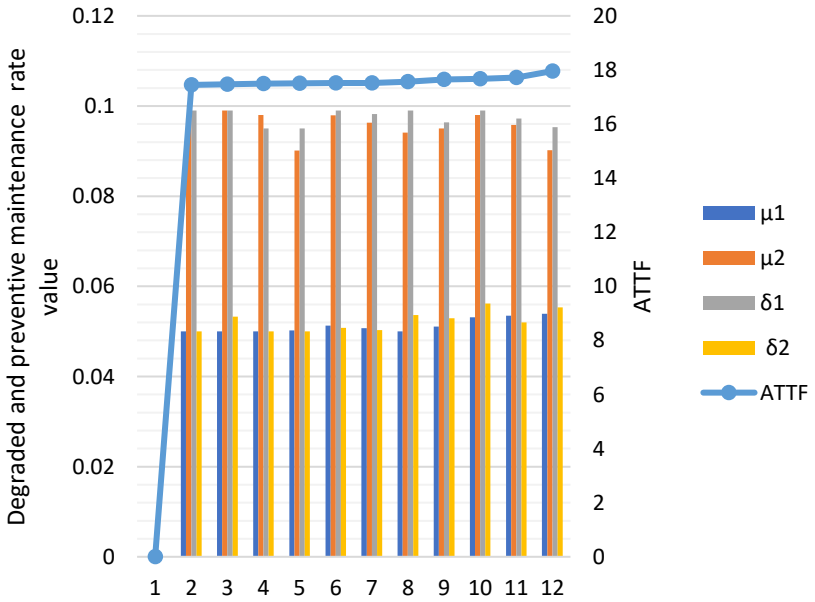


Figure 3: *ATTF vs. Different rate.*

II. Performing time (Availability Analysis)

Through fitness function optimization and analysis of the implications of malfunction and recovery rate variations, the PSO method determines the Availability of the system.

By having the value of σ_1 =0.07, σ_2= 0.05, χ_1 =0.08 and χ_2=0.07 and by applying 11 generation with the lower bound of value 0.05 and upper bound value of 0.09 for both degraded rate and corresponds preventive maintenance rate, The optimization tool PSO is applied to obtained the table 3 and figure 4.

Table 3: Variation in availability with respect to degraded rate and preventive maintenance rate.

Sr. No.	(χ_1)	(χ_2)	(σ_1)	(σ_2)	Availability
1	0.0981	0.0990	0.0501	0.0500	0.3220
2	0.0975	0.0946	0.0535	0.0500	0.3283
3	0.0990	0.0989	0.0540	0.0500	0.3307
4	0.0955	0.0955	0.0511	0.0500	0.3334
5	0.0970	0.0990	0.0549	0.0500	0.3346
6	0.0941	0.0961	0.0520	0.0500	0.3361
7	0.0949	0.0990	0.0500	0.0545	0.3413
8	0.0990	0.0950	0.0500	0.0540	0.3427
9	0.0958	0.0990	0.0540	0.0540	0.3490
10	0.0990	0.0950	0.0540	0.0536	0.3515
11	0.0941	0.0990	0.0500	0.0575	0.3526

With the value of μ_1=0.03 , μ_2=0.04, δ_1=0.05 and δ_2=0.07 and by applying 11 generation with the lower bound of value 0.05 and upper bound value of 0.09 for both failure rate and corresponds corrective maintenance rate, The optimization tool PSO is applied to obtained the table 4 and figure 5.

Table 4: Variation in availability with respect to failure rate and corrective maintenance rate.

Sr. No.	(μ_1)	(μ_2)	(δ_1)	(δ_2)	Availability
1	0.0990	0.0990	0.0500	0.0945	0.3393
2	0.0956	0.0975	0.0500	0.0990	0.3396
3	0.0953	0.0990	0.0500	0.0976	0.3400
4	0.0989	0.0869	0.0500	0.0946	0.3408
5	0.0969	0.0951	0.0520	0.0966	0.3412
6	0.0961	0.0889	0.0536	0.0990	0.3425
7	0.0950	0.0873	0.0500	0.0935	0.3429
8	0.0950	0.0918	0.0534	0.0973	0.3431
9	0.0982	0.0789	0.0534	0.0970	0.3436
10	0.0960	0.0823	0.0532	0.0972	0.3439
11	0.0950	0.0849	0.0546	0.0947	0.3456

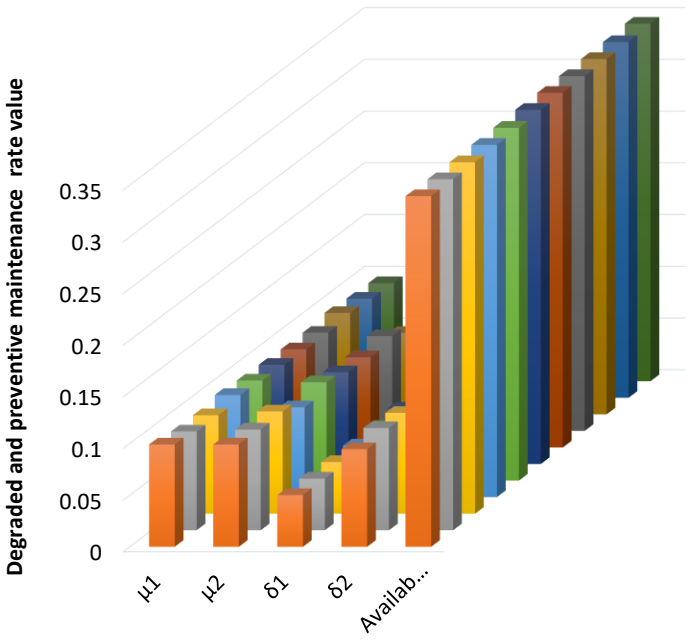


Figure 4: Availability vs. Degraded rate and Preventive Maintenance Rate.

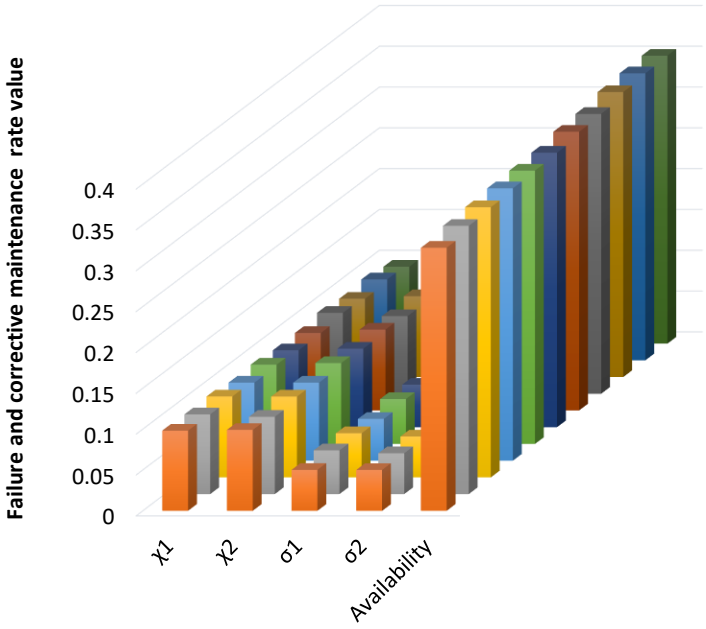


Figure 5: Availability vs. Failure rate and corrective Maintenance Rate.

VI. Results and Discussion

The degraded, failure, preventive and corrective maintenance rate have direct effect on ATTF and Availability which can be understand by the above example using the optimization tool PSO. Table

2 and figure 3 depicted that with the various value of degraded rate and preventive maintenance rate ATTF is increasing. The optimum value is selected to have a good profit in production in any industry.

The greater availability is a key point for any industry to gain more and more financial profit over the cost of infrastructure. The table 3 and figure 4 shown the effect of degraded rate and corresponds preventive maintenance rate. It is clear that the availability is optimized over different values of rates. Similarly, the effect of failure rate and corrective maintenance rate is also checked to count the availability which shown the increasing criteria for the different parameters value in table 4 and figure 5.

VII. Conclusions

The outcomes demonstrated by figure 3,4 and 5 are concluded that preventive and corrective maintenance had an impact on system reliability parameters like ATTF and availability, which in turn affected the system's accuracy and processing efficiency. Using the appropriate preventive or corrective maintenance rate values can lead to an increase in ATTF and availability. Increasing maintenance not only increases the working power of units but also prolongs its performance life. The influence of various maintenance rates on the optimal value was plotted on a graph. The optimal value for all malfunctions and repair rates can be used to improve the system's overall reliability.

References

[1] Naithani, A., Parashar, B., Bhatia, P. K. and Taneja, G. (2013). Cost Benefit Analysis of a 2-Out -of -3 Induced Draft Fans System with Priority for Operation to Cold Standby over Working at Reduced Capacity. *Advanced Modelling and optimization*, 499-509.

[2] Adaramola, B. A. (2020). A Systematic Approach of Turbo Generator Performance Evaluation Maintenance for Continuous power Generation. *Journal of Advancement in Engineering and Technology*, 1-5.

[3] Yousefi, N., Tsianikas, S. and Coit, D. W. (2020). Reinforcement learning for dynamic condition-based maintenance of a system with individually repairable components. *Quality Engineering*, 388-408.

[4] Puran, R., Malik, S.C. and Nandal, N. (2021). Reliability characteristics of a parallel-cold standby system of pairwise identical four units with provision of priority. *International Journal of Agricultural and Statistical sciences*.

[5] Modibbo, U. M., Arshad, M., Abdalghani, O. and Ali, I. (2021). Optimization and estimation in system reliability allocation problem. *Reliability Engineering & System Safety*.

[6] Singla, S. and Pooja (2021). Mathematical Analysis of Regenerative Point Graphical Technique (RPGT). *A Mathematical analysis and its contemporary applications*, 49-56.

[7] Singla, S., Verma, S., Malik, S. and Gupta, A. (2022). Perform ability evaluation, validation and optimization for the steam generation system of a coal-fired thermal plant. *Methods X*.

[8] Naithani, A., Khanduri, S. and Gupta, S. (2022). Stochastic analysis of main unit with two non-identical replaceable sub units working with partial failure. International *Journal of System Assurance Engineering and Management*, 1467-1473.

[9] Kumari, S., Singla, S. and Khurana, P. (2022). Particles Swarm Optimization for Constrained Cost Reliability of a Rubber Plant. *Life Cycle Reliability and Safety Engineering*.

[10] Singla, S., Rani, S., Modibbo, U. M. and Ali, I. (2023). Optimization of System Parameters of 2:3 Good Serial System using Deep Learning. *Reliability: Theory and Applications*, 670-679.

[11] Singla, S., Mangla, D., Panwar, P. and Taj, S. Z. (2024). Reliability Optimization of a Degraded System under Preventive Maintenance using Genetic Algorithm. *Journal of Mechanics of Continua and Mathematical Sciences*, 1-14.

[12] Ahmadini, A. A. H., Singla, S., Mangla, D., Modibbo, U. M. and Rani, S. (2024). Reliability Assessment and Profit Optimization of Multi-unit Mixed Configured System using ABC Algorithm under Preventive Maintenance. *IEEE Access*, 85190-85202.

[13] Singla, S., Mangla, D., Dhawan, P. and Ram, G. (2024). Reliability analysis of a two-unit repairable system using fuzzy linguistic approach. *Applications of fuzzy theory in applied sciences and computer applications*, 45-57.

[14] Singla, S., Mangla, D., Modibbo, U. M. and Lal, A. K. (2024). Critical analysis of failure and repair rates of poly-tube manufacturing plant using PSO. *Reliability: Theory and applications*, 351-364.

[15] Singla, S., Rani, S. and Mangla, D. (2024). Behavior analysis of repairable 2-out-of-4 system using evolutionary algorithm. *Journal of Mechanics of Continua and Mathematical Sciences*, 126-137.

[16] Singla, S., Mangla, D., Kumar, M.A. and Muhammad, M.U., (2024). Reliability optimization methods: A systematic literature review. *Yugoslav Journal of Operations Research*, 1-31.

Conflicts of interests: The authors declare that there are no conflicts of interest.

ATTRIBUTE CONTROL CHART APPROACH FOR LIFE TESTING USING EXPONENTIAL-POISSON DISTRIBUTION

Gokila B [1] and Sheik Abdullah A [2]

•

[1] Research Scholar, Department of Statistics, Salem Sowdeswari College,
Salem, Tamil Nadu,India.
gokilastat@gmail.com
[2] Assistant Professor , Department of Statistics, Salem Sowdeswari College,
Salem,Tamil Nadu,India.
sheik.stat@gmail.com

Abstract

Manufacturing processes are often monitored using control charts, visual representations of collected data. These charts help determine if a production process is stable. Various control chart methods have been developed to address diverse production scenarios. For some products, lifetime is a crucial quality attribute. Life testing, while valuable for evaluating production processes and classifying products as conforming or non-conforming, can be time-consuming. Consequently, censoring techniques are indispensable in such situations. Life testing often employs censoring methods like Type I, Type II, and Hybrid censoring. Compounding techniques create new distributions, expanding existing mixture models. These compound distributions are crucial in life testing, especially when dealing with processes exhibiting significant early failures. This article introduces a novel attribute np control chart for monitoring the median lifetime of products under a hybrid censoring plan, assuming an Exponential-Poisson lifetime distribution. Optimal control chart parameters are derived to minimize deviations from a target Average Run Length (ARL) for an in-control process. The chart's parameters are calculated for different scenarios, and its performance is evaluated using ARL analysis. Numerical examples and simulated data illustrate the chart's practical application and demonstrate its utility.

Keywords: Exponential-Poisson Distribution, Hybrid Censoring Scheme, Attribute Control Chart, Average Run Length.

I. Introduction

Statistical process control (SPC) is a method for understanding and managing process quality. It involves regularly collecting and analyzing data on key quality attributes, then reacting appropriately to any deviations from established standards or specifications. Visual representations are powerful tools in statistical analysis. Control charts are particularly valuable for maintaining processes at target specifications. A control chart is a graph depicting a quality characteristic, calculated from sample data (such as the mean) and plotted against the sample's collection time or number. This chart comprises three horizontal lines: the center line, upper control limit, and lower control limit. Control charts are categorized into two types: attribute and variable charts.

Attribute charts distinguish between conforming and non-conforming items, while variable charts are employed when data arises from a measurement process.Product longevity is a crucial quality attribute for some items. Manufacturing processes are monitored through life testing, which categorizes products as conforming or non-conforming based on test results. These tests are often time-consuming, necessitating the critical application of censoring techniques. Common methods include Type-I, Type-II, and Hybrid censoring. Hybrid censoring terminates the test upon reaching either a predefined time (t) or the (UCL + 1)th failure. Statistical control of the production process is maintained if the observed number of failures during the test falls within the lower (LCL) and upper (UCL) control limits at time t; otherwise, the process is deemed out of control.

The non-conforming rate, typically tracked using attribute control charts like the np chart, assumes a normal distribution for the quality characteristic. However, real-world quality characteristics often deviate from normality. This discrepancy can lead industrial engineers to erroneous conclusions and potentially an escalation in the number of non-conforming products if the existing control chart remains in use.

This paper [9] proposed an enhanced attribute control chart for more effectively monitoring the proportion of non-conforming items in high-quality manufacturing processes. They Presented [4] attribute control charts based on the Weibull distribution for monitoring product quality under truncated life testing conditions. Compounded univariate distributions [14] and introduces new generalized classes to enhance flexibility in statistical modeling. An attribute control chart based on the Pareto distribution of the second kind for time-truncated life tests [5] evaluating its performance using average run length and simulation studies to monitor non-conforming items in industrial applications. A control chart for gamma-distributed quality characteristics using multiple dependent state sampling [6] introduced demonstrating improved detection of process shifts compared to traditional Shewhart charts. A control chart combining successive sampling with multiple dependent state repetitive sampling [3] presented significantly improving sensitivity to process mean shifts and outperforming existing charts in average run length performance. A n attribute control chart based on the Inverse Rayleigh distribution under Type-I censoring [15] proposed to monitor product lifetimes in life testing scenarios. An attribute control chart for time-truncated life tests using Burr X, Burr XII, inverse Gaussian, and exponential distributions, [16] developed evaluating their performance through average run length analysis. An attribute control chart based on the Exponentiated Half Logistic distribution for time-truncated life tests, [18] showed improved performance through ARL analysis and simulations for industrial quality monitoring. An attribute np control chart utilizing repetitive group sampling under truncated life tests [11] introduced to enhance process monitoring efficiency. An attribute np control chart using multiple deferred state sampling [8] presented to monitor mean life in time-truncated life tests, showing improved detection of process shifts and enhanced performance over existing models. An attribute control chart based on the Dagum distribution for time-truncated life tests [20] proposed demonstrating effective performance in detecting non-conforming items through ARL analysis and simulations. An attribute control chart based on the Exponentiated Exponential distribution under Type-I censoring [12] introduced to monitor product lifetimes in life testing scenarios. The study evaluates the chart's performance using the average run length (ARL) metric and provides numerical examples to illustrate its practical application. An attribute control chart based on the log-logistic distribution for time-truncated life tests, [19] introduced evaluating its performance using average run length (ARL) through Monte Carlo simulations and providing extensive tables for practical implementation. A control chart utilizing repetitive sampling to monitor process variation in log-normal distributions, [17] introduced demonstrating enhanced sensitivity to small process shifts. An attribute control chart utilizing the neutrosophic Weibull distribution [2] introduced to monitor manufacturing processes under uncertainty, demonstrating enhanced efficiency in detecting process shifts compared to existing charts. An attribute control chart based on the generalized exponential distribution for time-truncated life tests, [1] introduced assessing its performance using the average run length (ARL) metric. An attribute control chart based on the Inverse Weibull distribution for time-truncated life tests, [7] presented demonstrating

its effectiveness through ARL analysis and simulation for industrial quality monitoring. He [10] introduced an attribute np control chart based on the Exponentiated Exponential distribution under accelerated life tests with hybrid censoring, optimized using ARL analysis and validated through simulations. The paper [13] introduced a flexible two-parameter lifetime distribution with a decreasing failure rate, derived from compounding exponential and zero-truncated Poisson distributions, and analyses its properties and parameter estimation using the EM algorithm.

The Exponential-Poisson distribution finds practical use in diverse areas, including network traffic modeling, manufacturing quality control, service queue management, and stock price analysis. Its application in modeling product lifetimes is particularly significant within the realm of lifetime and reliability studies. Prior research has not investigated control charts for life tests involving non-normal distributions, such as the Exponential-Poisson with hybrid censoring.

This paper introduces a novel attribute control chart specifically designed for this type of data. The paper's structure is as follows: Section 2 details the proposed control chart methodology. Also presents formulas for calculating in-control and out-of-control average run lengths. A practical industrial application, along with a simulation study, is detailed in Section 3. Finally, Section 4 summarizes the findings and conclusions.

II. Design of the control chart

Let t represent the lifetime of a product in a manufacturing process. t follows an Exponential-Poisson distribution with parameters $\lambda > 0$ and $\beta > 0$. The probability density function for this distribution is specified by

$$f(t) = \frac{\lambda\beta e^{-\lambda-\beta t+\lambda e^{-\beta t}}}{1-e^{-\lambda}}\ , t, \lambda, \beta > 0 \tag{1}$$

then the cdf of the EP model proposed by Kus(2007) can be obtained as

$$F(t) = \frac{1-e^{-\lambda+\lambda e^{-\beta t}}}{1-e^{-\lambda}}, t, \lambda, \beta > 0 \tag{2}$$

The lifetimes of the units follow Exponential-Poisson Distribution with a median life defined as

$$m_0 = \left(-\frac{1}{\beta}\right) \log\{1+\log[1-2^{-1}(1-e^{-\lambda})]\lambda^{-1}\} \tag{3}$$

This truncation, with a fixed time point t_0, was implemented to manage time constraints in the life testing process.In the simulation, for each subgroup, failures occurring within or before the truncation time $\leq$ t_0 were counted and where the truncation time is determined by

$$t_0 = am_0 \tag{4}$$

Where t_0 is the truncated time, a is the truncation coefficient and m_0 is the desired median life of the product.The probability of failure before t_0 is given by

$$p = \frac{1-e^{-\lambda+\lambda e^{-\beta t}0}}{1-e^{-\lambda}} \tag{5}$$

Using Equation(3),the probability of failure of a product when the process is in control is

$$p_0 = \frac{1-e^{-\lambda+\lambda^{1-a}\left[\lambda+\log\left(\frac{1+e^{-\lambda}}{2}\right)\right]^a}}{1-e^{-\lambda}} \tag{6}$$

When the median shifts to $m_{1,}$ the probability of failure of a product becomes
Let $m_1 = fm_0$,where f is the shift coefficient,then

$$p_1 = \frac{1 - e^{-\lambda + \lambda^{1-\frac{a}{f}}\left[\lambda + \log\left(\frac{1+e^{-\lambda}}{2}\right)\right]^{\frac{a}{f}}}}{1 - e^{-\lambda}} \tag{7}$$

We propose an np control chart for the Exponential-Poisson distribution, using Hybrid censoring, personalized to the number of products within each subgroup.

Step 1: Randomly sample n products from the production line.

Step 2: Subject the chosen products to a life test, concluding the test at time t_0. Record the number of failures (denoted as D).

Step 3: End the life test when either the time reaches t_0 or D exceeds the Upper Control Limit (UCL), whichever occurs first.

Step 4: Declare the process as out of control if D>UCL or D<LCL. Declare the process as in control if LCL≤D≤UCL.

We are interested in observing the number of failures D in each sample. If the units have failure times that are less than the truncation time t_0, then the units are considered nonconforming or defective. The binomial probability mass function is given by

$$p(D = x) = \binom{n}{x} p^x (1-p)^{n-x}, x = 0, 1, 2, 3... \tag{8}$$

The Lower and Upper control limits are given by

$$LCL = \max(0, np_0 - k\sqrt{np_0\,(1-p_o)}) \tag{9}$$

$$UCL = np_0 + k\sqrt{np_0(1-p_o)} \tag{10}$$

Where k is the coefficient of the control limits and p_0 is the probability of "nonconforming" when the process is in control. The probability of stating that the process is in control when it is truly in control is obtained using the binomial probability given in Equation(8)

$$p_{\text{in}}^0 = P(LCL \le D \le UCL \mid p_0) = \sum_{d=LCL+1}^{UCL} \binom{n}{d} p_0^d (1-p_0)^{n-d} \tag{11}$$

The probability of stating that the process is in control when the median life of the product has shifted to m_1 is given by

$$p_{\text{in}}^1 = P(LCL \le D \le UCL \mid p_1) = \sum_{d=LCL+1}^{UCL} \binom{n}{d} p_1^d (1-p_1)^{n-d} \tag{12}$$

It follows from the values p_{in}^0 and p_{in}^1 given in Equation (11) and (12) that the in-control ARL is given by

$$ARL_0 = \frac{1}{1 - p_{in}^o} \tag{13}$$

and the out of control ARL are given by

$$ARL_1 = \frac{1}{1 - p_{in}^1} \tag{14}$$

To construct the tables for the proposed control chart, we applied the following algorithm.

Step 1: Specify the values of ARL, say r_0 and sample size n.

Step 2: Determine the control chart parameter and truncated time constant a values for which the ARL from equation (13) approach r_0.

Step 3: By using the values of the control chart parameters obtained in step 2, determine the ARL_1 in accordance with shift constant f by using equation(14).

Table 1: *The Values of ARLs when λ = 1.0*

n	25			30			35		
r_0	260	320	420	260	320	420	260	320	420
a	0.704	0.592	0.574	0.869	0.9	0.719	0.573	0.757	0.597
k	3.0003	3.0017	2.98	2.99	3.110	3.159	2.965	3.086	3.077
LCL	3	2	1	6	6	4	4	6	4
UCL	17	16	15	22	22	20	20	23	21
Shift(f)	ARL	ARL	ARL	ARL	ARL	ARL	ARL	ARL	ARL
1.00	259.52	321.23	421.30	260.19	320.49	419.59	260.07	319.80	420.25
0.90	264.308	347.558	199.184	349.632	300.581	322.237	206.078	237.493	364.749
0.80	124.887	173.658	78.631	156.556	113.714	118.843	78.775	81.322	132.093
0.70	44.100	61.177	29.634	47.148	35.108	37.446	25.562	24.719	39.108
0.60	15.050	20.268	11.171	14.119	11.102	11.972	8.528	7.934	11.801
0.50	5.384	6.888	4.416	4.650	3.907	4.191	3.195	2.943	3.968
0.40	2.214	2.622	1.992	1.894	1.717	1.799	1.519	1.434	1.691
0.30	1.218	1.309	1.179	1.125	1.094	1.110	1.055	1.040	1.080
0.20	1.007	1.012	1.005	1.002	1.001	1.001	1.000	1.000	1.001
0.10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

For various values of r_0,n and λ we determine the control chart parameters and ARL_1, which are shown in Tables 1,2,3,4 shows that the ARLs tend to get smaller when the shift constant f gets less.

Table 2: *The Values of ARLs when $\lambda = 0.5$*

n	25			30			35		
r_0	260	320	420	260	320	420	260	320	420
a	0.71	0.59	0.58	0.97	0.57	0.50	0.61	0.55	0.65
k	3.002	2.99	2.98	3.027	2.909	3.001	2.902	2.964	3.007
LCL	3	2	1	7	2	1	4	4	4
UCL	17	16	15	23	17	16	20	20	21
Shift(f)	ARL	ARL	ARL	ARL	ARL	ARL	ARL	ARL	ARL
1.00	260.33	320.09	419.69	260.40	323.92	421.29	260.23	320.21	419.99
0.90	262.437	346.255	194.951	342.785	145.774	177.804	133.168	152.809	186.836
0.80	120.159	168.626	75.505	139.129	53.958	65.698	46.710	53.522	60.511
0.70	41.241	57.888	27.949	39.886	19.411	23.622	15.789	18.116	19.210
0.60	13.759	18.767	10.376	11.662	7.202	8.646	5.673	6.439	6.493
0.50	4.858	6.282	4.068	3.855	2.946	3.421	2.362	2.606	2.550
0.40	2.014	2.395	1.850	1.646	1.485	1.627	1.291	1.360	1.328
0.30	1.157	1.238	1.134	1.070	1.054	1.080	1.020	1.029	1.023
0.20	1.003	1.006	1.002	1.000	1.000	1.001	1.000	1.000	1.000
0.10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3: *The Values of ARLs when $\lambda = 2.0$*

n	25			30			35		
r_0	260	320	420	260	320	420	260	320	420
a	0.515	0.831	0.868	0.566	0.784	0.753	0.568	0.537	0.592
k	2.885	3.051	3.175	2.931	3.056	2.99	2.966	2.964	3.077
LCL	1	4	4	3	5	4	4	3	4
UCL	14	19	19	18	21	20	20	19	21
Shift(f)	ARL	ARL	ARL	ARL	ARL	ARL	ARL	ARL	ARL
1.00	259.50	319.80	420.08	260.58	320.23	419.52	260.30	320.26	420.64
0.90	140.451	509.357	476.965	276.405	347.133	217.930	207.243	156.951	367.202
0.80	60.559	309.230	218.449	128.875	145.433	77.690	80.508	56.819	135.566
0.70	24.505	102.767	72.067	43.714	45.713	26.335	26.507	19.802	40.815
0.60	9.882	31.137	22.980	14.404	14.360	9.237	8.949	7.194	12.491
0.50	4.161	9.691	7.637	5.050	4.909	3.576	3.371	2.925	4.232
0.40	1.974	3.368	2.869	2.085	2.025	1.688	1.589	1.483	1.789
0.30	1.194	1.506	1.398	1.184	1.170	1.104	1.074	1.058	1.106
0.20	1.008	1.038	1.027	1.005	1.005	1.002	1.001	1.000	1.001
0.10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4: *The Values of ARLs when $\lambda = 5$*

n	25			30			35		
r_0	260	320	420	260	320	420	260	320	420
a	0.706	0.594	0.577	0.87	0.789	0.721	0.576	0.545	0.6
k	3.0003	2.999	2.986	2.996	3.055	3.159	2.967	2.964	3.079
LCL	3	2	1	6	5	4	4	3	4
UCL	17	16	15	22	21	20	20	19	21
Shift(f)	ARL	ARL	ARL	ARL	ARL	ARL	ARL	ARL	ARL
1.00	259.50	319.90	419.96	260.36	319.75	419.71	260.44	320.33	421.41
0.90	263.399	347.646	196.845	348.792	343.275	319.973	204.106	154.273	360.916
0.80	123.060	172.391	77.105	154.203	137.990	116.711	77.033	54.662	128.765
0.70	43.150	60.196	28.890	46.241	42.074	36.553	24.828	18.709	37.848
0.60	14.694	19.831	10.858	13.885	12.952	11.675	8.267	6.714	11.398
0.50	5.277	6.734	4.300	4.619	4.404	4.111	3.111	2.728	3.852
0.40	2.195	2.582	1.957	1.909	1.855	1.787	1.497	1.411	1.664
0.30	1.223	1.306	1.174	1.140	1.127	1.114	1.053	1.042	1.077
0.20	1.009	1.014	1.006	1.003	1.003	1.002	1.000	1.000	1.001
0.10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

II.1. Illustration 1

Assuming a Exponential-Poisson lifetime distribution, characterized by with parameter λ= 2. Consider the following values for the products: m_0 = 1000 hours, ARL_0 = 260 and n = 25. Table 3, provided the following control chart parameters: k=2.885, a=0.515, a lower control limit (LCL) of 1, and an upper control limit (UCL) of 14. This led to the design and implementation of the control chart as described.

Step 1: Select a sample of 25 products from each subgroup and submit them to the life test.
Step 2: During the testing, count the number of failed items (D) and fix the time period.
Step 3: Terminate the test if either first failures occur or time elapse whichever comes first.
Step 4: Declare that the process is in control if $1 \leq D \leq 14$
Step 5: Declare that it is out of control if D > 14 or D < 1

II.2. Illustration 2

Assume that the product lifetimes follow an exponential poisson distribution with parameters λ=5 , Consider the following product values: m_0 is 1000 hours, ARL is 260, and n is 35. The following control chart parameters were presented in Table 4: k = 2.967, a = 0.576, LCL = 4 and UCL = 20. As a result, the control chart was established in the following manner: From each subgroup, choose a sample of 35 products, and put them through a life test . Count the number of failed items (D) which is minimum failure during testing or maximum time reached whichever comes first. Now monitor failures and time. Stop the process if the value of D is between 4 and 20 or the time elapse before reaching maximum time then the process is in control; otherwise, it is out of control.

III. Simulation study

In this section, the application of the proposed control chart is demonstrated using simulated data. The data are generated from an Exponential Poisson Distribution when the process is in-control with the parameters following Exponential Poisson Distribution and the targeted Average lifetime of 1000 hours. We consider a random sample of size n=25 for each sample batch. The first fifteen samples are generated from the in-control process, and the next fifteen samples are from a shifted process with f = 0.40 to achieve 30 sample batches. Then from table 3, when n=25 and $ARL0$ = 260, we have constants k=2.885, a=0.515, UCL=14 and LCL=1. The number of products having the lifetime below 515 hours is noted (“D”) and the Values are : 7,7,10,3,6,7,9,8,7,6,5,12,7,10,9,12,13,16,16,17,13,16,15,17,14,15,14,19,10,11. The values of the nonconforming items are plotted with two control limits (LCL =1 and UCL= 14) in Figure 1.

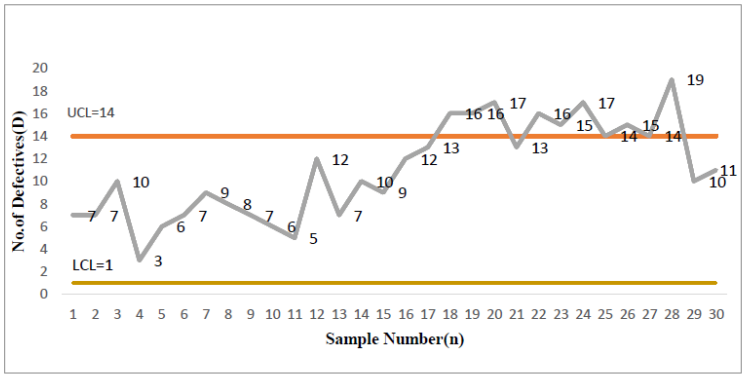


Figure 1: *Control chart for Simulated Data*

IV. Conclusion

In this article, a new attribute control chart based on Exponential-Poisson distribution under hybrid censoring scheme is proposed to ensure the median lifetime of the product as the quality criterion. The newly established control chart is highly adaptable and may be used to monitor the lifetime of quality products. The tables are provided for industrial usage and are explicated with the use of simulated data. It is generated using R software from an Exponential-Poisson distribution. The performance of the proposed control chart is expressed in terms of ARLs for various shift constants (f). It should be noted that if the hybrid censoring scheme is used to carry out the life test, executing the sampling inspection will reduce the time and cost of conducting the life test. The developed attribute control chart has the potential to be extended for use with various other statistical distributions as part of a continuing research study.

References

[1] Adeoti, O. A. and Ogundipe, P. (2021). "A Control Chart for the generalized exponential distribution under time truncated life test", Life Cycle Reliability and Safety Engineering, 10:53-59.

[2] AL.Marshadi, Shafqat,A.H., Aslam,M., and Alharbey, A. (2021). "Performance of a New Time-Truncated Control Chart for Weibull Distribution under Uncertainty", International Journal of Computational Intelligence Systems, 14(1):1256-1262.

[3] Aldosari,M.S., Aslam,M., and Jun,C.H (2017).“A new attribute control chart using multiple dependent state repetitive sampling", IEEE Access,5:6192–6197.

[4] Aslam, M. and Jun, C. (2015). "Attribute Control Charts for the Weibull Distribution under Truncated Life Tests", Quality Engineering, 27(3):283– 288.

[5] Aslam, M. Nasrullah Khan and Chi– Hyuck (2016). "A control chart for time truncated life tests using Pareto distribution of second kind", Journal of Statistical Computation and Simulation, 86(11):2113– 2122.

[6] Aslam,M., Arif,O.H., and C.H. Jun (2017). “A control chart for gamma distribution using multiple dependent state sampling", Industrial Engineering and Management Systems, 16(1):109–117.

[7] Baklizi, A. and Ghannam, S. A. (2022). "An attribute control chart for the inverse Weibull distribution under truncated life tests", Heliyon, 8:e11976:1-5.

[8] Balamurali, S. and Jeyadurga, P. (2019). "An attribute np control chart for monitoring mean life using multiple deferred state sampling based on truncated life tests", International journal of Reliability, Quality and Safety Engineering, 26(1):1-18.

[9] Barbosa, E. P. and Joekes, S. (2013). "An improved attribute control chart for monitoring non-conforming proportion in high quality process", Control Engineering Practice, 21:407 412.

[10] Gunasekaran.M (2024). "A New attribute control chart based on Exponentiated Exponential Distribution Under Accelerated Life test with Hybrid Censoring", Reliability:Theory And Applications,19,4(80):850-860

[11] Jayadurga, P., Balamurali, S. and Aslam, M. (2018). "Design of an attribute np control chart for process monitoring based on repetitive group sampling under truncated life test", Communication in Statistics– Theory and Methods, 47(24):5934– 5955.

[12] Kavitha, T. and Gunasekaran, M. (2020). "Construction of Control Chart based on Exponenti ated Exponential Distribution", Journal of Engineering Sciences, 11(4):979– 985.

[13] Kus, C. (2007). "A new lifetime distribution", Computational Statistics and Data Analysis, 51, 4497–4509.

[14] Muhammad H Tahir, Gauss M Cordeiro (2016). “Compounding of distributions: a survey and newgeneralized classes”, Journal of statistical Distributions and Applications, Springer,3(1):1 35.

[15] Nanthakumar, C. and Kavitha, T. (2017). "Design of Attribute Control Chart based on Inverse Rayleigh Distribution under Type– I Censoring", International Journal of Statistical Sciences, 4(6):17– 22.

[16] Shafqat A., Hussain J., Al-Nasser, A. D. and Aslam, M. (2018). "Attribute Control Chart for some popular distributions", Communications in Statistics– Theory and Methods, 47(8):1978–1988.

[17] Shaheen,U., Azam,M., and Aslam,M.(2020). “A control chart for monitoring the log normal process variation using repetitive sampling", Quality and Reliability Engineering International, 36(3):028–1047.

[18] Srinivasa Rao, G. (2018). "A Control Chart for Time Truncated Life tests Using Exponentiated Half Logistic Distribution", Applied Mathematics and Information Sciences: An International Journal, 12(1):125– 131.

[19] Srinivasa Rao, G. and Edwin Paul (2020). "Time Truncated Control Chart using LogLogistic Distribution", Biometrics and Biostatistics International Journal, 9:76- 82.

[20] Srinivasa Rao, G., Fulment, A. K. and Josephat, P. K. (2019). "Attribute control chart for the Dagum distribution under truncated life tests", Life Cycle Reliability and Safety Engineering, 8:329– 335

RT&A, No 2 (84)
Volume 20, June 2025

ESTIMATE MISSING VALUE OF RANDOMIZED BLOCK DESIGN USING VARIABLE CONTROL CHARTS THROUGH RESPONSE SURFACE METHODOLOGY

Kokkilambigai S[1] and Pachamuthu M[2*]

•

[1]Research Scholar, Department of Statistics, Periyar University, Salem-11, India
[1]msamundiswariharini@gmail.com
[2*]Assistant Professor, Department of Statistics, Periyar University, Salem-11, India
[2*]pachamuthu@periyaruniversity.ac.in

Abstract

The complete block design of Randomised Block Design (RBD) and variable control charts play a vital role in the fields of agriculture and industry. The basic complete block designs are Completely Randomized Designs (CRD), RBD, and Latin Square Designs (LSD). An improvement in the CRD can be obtained by providing error control measures called an RBD. In statistics, missing values occur when a data value for a variable is not stored in an observation. There are several methods available to estimate the missing value of complete block designs, such as fuzzy techniques, traditional methodology, etc. In this paper, we are introducing the new methods to estimate the missing value using variable control charts such as 3-Sigma and 6-Sigma control charts through Response Surface Methodology (RSM) with the numerical example. RSM can be used to find factor systems that produce a desired maximum, minimum, or optimal response; to find factor systems that satisfy management specifications; and to model the relationship between quantitative factors and responses.

Keywords: *CRD, RBD, LSD*, 3-Sigma Control Charts, 6-Sigma Control Charts, *RSM.*

I. Introduction

Ronald A Fisher [1920] first introduced the Design of Experiments (*DOE*) in 1953. *DOE* is a systematic method used in various fields of design to tasks describe and explain data variation under hypothesized conditions. There are three basic principles of *DOE* such as randomization, replication and blocking or error control. Blocking (*b*) an experiment involves dividing or partitioning the observations into homogeneous groups called blocks so that the observations in each block are collected under relatively similar experimental conditions, treatments (*v*) are the different objects or procedures being compared in an experiment and replication (*r*) refers to the repetition of an experiment under identical conditions, but in the context of experimental designs, it refers to the number of distinct experimental units under the same treatment. The *DOE* may be broadly classified as two types are complete and incomplete block design. Designs in each block that receive all treatments are called complete block design otherwise its incomplete block designs. A *CRD* is an experimental design where treatments are assigned completely at random, ensuring that each experimental unit has an equal chance of receiving any one treatment. A *RBD* is a restricted randomized design in which experimental units are first organized into homogeneous

blocks, and then the treatments are assigned at random to these units within these blocks. Due to some unforeseen reasons, observations may be missing in some observations. Observations of one or two units are too large or too small compared to other observations. The accuracy of such observations is often questionable. These observations greatly inflate the error variance and thereby distort the results. Therefore, the best ones are omitted, and the remaining data are then analyzed considering the missing data. The missing-plot technique is a standard formula for the estimation of a missing data observation. It is used within side the evaluation of the variance of data collected consistent with a diagnosed experimental design. If there is one missing observation, we can use to estimate the missing observation along with the sum of squares for testing the differential effect of the treatments. The main objective of *RSM* is to optimize reactions to achieve high yields and purity at low cost. Identify patterns, relationships, and optimal conditions for the response. The *RSM* is a technique for optimizing the response when two or more quantitative factors are involved. Dependent variables are referred to as responses, whereas independent variables or factors are primarily referred to as predictor variables in a response surface system. A 3*D* surface plot is a graphical representation used to visualize the relationship between the response variable and two independent variables. This plot helps to identify the optimal conditions for the response variable. Statistical Quality Control *(SQC)* is one of the methods used to determine whether data is normal or not. *SQC* is the use of statistical methods to monitor and control the quality of a manufacturing process.

There are many reachers review the concepts we are given some of the them mixture-process variable experiments, including control and noise variables within a split plot structure Tae-Yeon Cho, et.al [14] have discussed the mixture-process variables experiments often involve larger ran due to increased coefficient parameters and difficult to controlled interactions. These experiments could have analyzed used meant and slope models within a split-plot structure. Low prediction variances were desirable. Methods for mixture-process variable designs with noise variables demonstrated, and some robust designs are evaluated used fraction of design space plots. A *G*-optimal design created using a genetic algorithm. Loss of efficiency in randomized block designs with two missing values Subramani. J and Balamurali. S [13] has investigated the efficiency of *RBD* with two missing values, obtaining explicit expressions for elementary treatment contrasts, average variance, and loss of efficiency. It also tabulated the loss of efficiency for different values. Optimum estimation of missing values in *RCBD* by genetic algorithm. Azadeh. A, et.al. [1] have discussed the efficiency of *RBD* with missing values, providing explicit expressions for treatment contrasts, average variance, and loss of efficiency. A fractional factorial screening experiment to determine factors affecting discrete part process control. Naugraiya. M and Drury, C.G [6], have utilized a 2^{6} -1 fractional factorial design to investigate human factors in manufacturing processes. It found that task difficulty reduced performance, while individual experience and training improved it. Analysis of ranked data in randomized blocks when there are missing values Besta. D.J and Raynera. J.C.W [2] have introduced a novel method for analyzing *RBD* using rank data and utilized an *ANOVA* on ranks, which surpassed the Skillings-Mack test in terms of indicative sizes and powers. Missing plot techniques using regression analysis and Its comparison with *RBD* missing plot techniques derived by Richa Seth, et al [12] investigated the estimated values of one and two missing observations using the regression method analysis. Also estimate the value of the same missing observations by expressing the regression data. Analysed the *RBD* data format. Also analysed the data to obtain an *ANOVA* table using both methods, where I found the same significant result. Construction of control chart using fuzzy probabilistic approach for cotton sweater product Pachamuthu. M. and Shanmugasundram. V [7], had discussed, creation of fuzzy statistical controlled chart used the fuzzy membership function proposed. It illustrates the cotton sweater product of quality evaluation such as color, designs, quality and price of a product with consumer's demographic characters on a hedonic ruled. A statistical analysis of Latin square

designs through fuzzy ranking approach using median value have discussed by Pachamuthu. M et al., [3]. Pachamuthu. M and Mariappan. A [9] have constructed the statistical quality control chart using fuzzy probabilistic approach and proposed different solutions to prevented air pollution. Also, developed a new methodology for construction of statistical quality-controlled chart used fuzzy probability approached. Application of this method had established through the air pollution controlled caused illustration with consumer's demographic characters on a hedonic ruled. Construction of 3σ control chart using fuzzy probabilistic approach for quality evaluation by Pachamuthu. M. et al., [8]. The 3σ fuzzy statistical quality-controlled charts played an important role for smart controlled of home appliances. Also, constructed the 3σ control charts used fuzzy probabilistic approached for quality evaluation. An optimization method of construction for warping copper plates and engines using complete block designs with some special types of graphs by Pachamuthu M et al., [5]. The optimization method for constriction of randomized blocked design and Latin square design used bipartite and regular graphs with applications. Use of control charts for conducting *ANOVA* Study Ghosh. D.K et al., [4] have explored was used to control charts for an empirical study of analysis of variance. The modified control charts were renamed as *FES* charts. The control limits were labelled as lower selection limit and upper selection limit. Also showed that the results of the control chart could be compared if the 3σ control limits were kept intact. *ANOVA* was conducted at the 0.27% significance level. On the construction of balanced incomplete block designs through factorization and coloring graphs using mutually orthogonal Latin square designs Pachamuthu. M and Ramya. A [10]. A statistical analysis of variance for one-way classification with addition operations using triangular fuzzy numbers discussed by Pachamuthu. M and Mariappan. A [11]. A statistical hypotheses technique of one-way analysis of variance using triangular fuzzy numbers: a comparative study Pachamuthu. M et al., [15] have discussed, a comparative studied on the statistical tested of *ANOVA* for one way with $\propto$-cut interval method, Exponential Triangular FuzzyNumbered *(ETFNs),* and Logarithmic Triangular Fuzzy Numbers *(LTFNs)* used the Triangular Fuzzy Numbers *(TFNs)* based on the decision ruled and illustrated with numerical examples. In this paper, a statistical analysis and estimate the one missing value of *RBD* using 3σ and 6σ control chart through *RSM* with numerical example.

II. Methods

I. Definition of *RBD*

Let us suppose that experimental material is divided into r blocks and t treatments. Each block is then divided into t units, and the treatments are allocated within a block randomly the resulting design is called an *RBD*.

II. Definition of Missing Plot Technique

In field experimentation, some of the experimentation from certain experimental units may not be available due to some practical reasons. In such the case of observations which are not available are called missing observation. The technique of estimating the missing value and carryout the statistical analysis in experimental designs is called as missing plot techniques.

III. Control Chart

Control charts are graphical representations of process data over time. They show whether a manufacturing process is stable and operating within expected parameters using statistical limits.

Whether a manufacturing process is in control or out of control can be determined through a technique called a control chart. The 3σ control charts limits are; Upper Control Limit (UCL) = $\mu + 3\sigma$, Center Limit (CL) = μ and Lower Control Limit (LCL) = $\mu - 3\sigma$. Then, 6σ control charts limits are Upper Control Limit (UCL) = $\mu+6\sigma$, Center Limit (CL) = μ and Lower Control Limit (LCL) = $\mu - 6\sigma$.

IV. Interaction Plot

An interactive plot shows that the relationship between one categorical factor and a continuous response depends on the value of the second categorical factor. This plot shows the means for levels of one factor on the axis and a separate line for each level of another factor.

V. Residual Plot and Contour Plot

A residual plot shows the difference between the predicted values and the actual values of the response variable. An ideal residual plot, called a null residual plot, shows a random scatter of points that form a band of approximately constant width around the identity line. Contour plots or contour maps are providing one of the most obvious ways to illustrate and interpret response surface structure. Contour plots are two-dimensional or three-dimensional graphs in which the axis structure is a specific pair of design variables, while the other design variables are held constant. Definitions are not created by finite equations. Each point on an edge has a standard error.

VI. 3*D* Surface

A 3*D* surface plot is a way of visualizing data with three dimensions such as length, width and height. In a 3*D* surface plot, these points are plotted in three-dimensional space, creating a surface that shows how the data varies at different levels. It's like looking at a three-dimensional map of data, where the height of the surface represents the value of the data at each point.

III.Methods

I. A Statistical Analysis and Estimate the one missing value of *RBD*

Let us consider y_{ij} is the general observations and the layout table is given below;

Table 1: *Layout of RBD*

	Blocks(B_j)					
Treatments(T_i)	1	…	j	…	r	Total
1	y_{11}	…	y_{1j}	…	y_{1r}	$y_{1.}$
⋮	⋮		⋮		⋮	
i	y_{i1}	…	y_{ij}	…	y_{ir}	$y_{i.}$
⋮	⋮		⋮		⋮	
T	y_{t1}	…	y_{tj}	…	Y_{tr}	$y_{t.}$
Total	$y_{.1}$	...	$y_{.i}$	...	$y_{.r}$	$y_{..} = G$

Linear model of *RBD*

$$y_{ij} = \mu + a_i + b_j + e_{ij} \qquad i = 1,2,\dots,t; \qquad j = 1,2,\dots,r$$

where, y_{ij} is the observation corresponding to i^{th} of treatment and j^{th} blocks, μ refers the general mean effect which is fixed, a_i is fixed effect due to the i^{th} treatment b_j is fixed effect due to the j^{th} blocks and e_{ij} random error effect $\mathrm{N}(0,\sigma^2)$. Here, the parameters μ, a_i and b_j can be estimated by the principle of least squares. $\frac{\partial E}{\partial \mu} = 0$, $\frac{\partial E}{\partial a_i} = 0$ and $\frac{\partial E}{\partial b_j} = 0$. Then, the normal equations are

$$e_{ij} = y_{ij} - \mu - a_i - b_j$$

Summing and squaring on both sides;

$$E = \sum_{i=1}^{t} \sum_{j=1}^{r} {e_{ij}}^2 = \sum_{i=1}^{t} \sum_{j=1}^{r} \left(y_{ij} - a_i - b_j - \mu\right)^2$$

Differentiating with respect to the parameters are;

$$\frac{\partial E}{\partial \mu} = -2 \sum_{i=1}^{t} \sum_{j=1}^{r} \left(y_{ij} - a_i - b_j - \mu\right) = 0$$
$$\frac{\partial E}{\partial a_i} = -2 \sum_{j=1}^{r} \left(y_{ij} - \mu - a_i - b_j\right) = 0 \qquad \text{and}$$
$$\frac{\partial E}{\partial b_j} = -2 \sum_{i=1}^{t} \left(y_{ij} - \mu - a_i - b_j\right) = 0$$

The estimate of μ is given by,

$$\hat{\mu} = \frac{G}{tr}$$

The estimate of a_i is given by,

$$\hat{a}_i = \frac{A_i}{r} - \frac{G}{tr}$$

Then b_j is given by,

$$\hat{b}_j = \frac{B_j}{t} - \frac{G}{tr}$$

These three equations (2), (3) and (4) substituting in (1), we get sum of square are;

$$Q = \mathrm{TSS} = \sum_{i=1}^{t} \sum_{j=1}^{\mathrm{r}} \left({y_{ij}}^2 - \frac{G^2}{tr}\right) \quad \text{which has } (tr-1)df$$

$$Q_t = \sum_{i=1}^{t} \left(\frac{T_i^2}{t} - \frac{G^2}{tr}\right) \quad \text{which has } (r-1)df$$

$$Q_b = \sum_{i=1}^{b} \left(\frac{B_j^2}{r} - \frac{G^2}{tr}\right) \quad \text{which has } (t-1)df$$

$$Q_E = \mathrm{Q} - Q_t - Q_b \text{ which has } (t-1)\,(r-1)df$$

Sometimes, points earned in tests are lost due to unavoidable circumstances. Therefore, we can estimate the missing values.

Table 2: *Layout of one missing* in *RBD*

	Blocks(B_j)					
Treatments(T_i)	1	…	j	…	R	Total
1	x_{11}	…	x_{1j}	…	x_{1r}	
⋮	⋮		⋮		⋮	
i	x_{i1}	…	x_{ij}	…	x_{ir}	$T=T'+x$
⋮	⋮		⋮		⋮	
t	x_{t1}	…	x_{tj}	…	x_{tr}	
Total			$B=B'+x$			

Let $X = x_{ij}$ be the one missing observation of *RBD* corresponding to the i^{th} treatment and j^{th} block; Let *G* be the grant total of all *tr* number of observations; *G* denote the grant total of *(tr-1)* number of observations excluding the missing observation *x*, Then $G = G'+x$. Let T be the i^{th} treatment total of r observations. Let T' denote the i^{th} treatment total of $(t-1)$ observations excluding the missing observation *x* then,

$$T = T' + x$$

Let B be the j^{th} block total of r observations; B' denote the j^{th} block total of $(r-1)$ observations *x* then,

$$B = B' + x$$

The correction factor (*c f*) is define by,

$$cf = \frac{G^2}{N} = \frac{(G'+x)^2}{tr}, \qquad N = tr$$
$$TSS = Q = \text{ constant} + x^2 - cf$$
$$TrSS = Q_t = \text{ constant} + \frac{(T'+x)^2}{t} - cf$$
$$BSS = Q_b = \text{ constant} + \frac{(B'+x)^2}{r} - cf$$

Then the Error Sum of Square *(ESS)* can be obtained by subtraction

$$Q_E = Q - Q_t - Q_b$$

$$ESS = x^2 - \frac{(T'+x)^2}{t} - \frac{(B'+x)^2}{r} + \frac{(G'+x)^2}{tr} + \text{ constant}$$

To estimate the missing value of x such that error is minimum of this value and we must have,

$$\frac{\partial E}{\partial x} = 0 = 2x - \frac{2(T'+x)}{t} - \frac{2(B'+x)}{r} - \frac{2(G'+x)}{tr} = 0$$

Then

$$\hat{x} = \frac{rT'+tB'-G'}{(t-1)(r-1)}$$

The estimate of the missing observation is fitted in the given data and the statistical analysis of *RBD* is carried out. All the values are referring in the *ANOVA* table and inference is drawn.

Table 3: *ANOVA* table for *RBD* with one missing value

SV	*Df*	*SS*	*TSS*	*F - Ratio*
Treatments	$(t-1)$	$Q_t = \sum_{i=1}^{t}\left(\frac{T_i^2}{r} - \frac{G^2}{tr}\right)$	$M_t = \frac{Q_t}{(r-1)}$	$F_t = \frac{M_t}{M_e}$
Blocks	$(r-1)$	$Q_b = \sum_{j=1}^{r}\left(\frac{B_j^2}{t} - \frac{G^2}{tr}\right)$	$M_b = \frac{Q_b}{(t-1)}$	$F_b = \frac{M_b}{M_e}$
Error	$(t-1)(r-1)-1$	$Q_E = Q - Q_t - Q_b$	$M_e = \frac{Q_E}{(t-1)(r-1)-1}$	-
Total	$(tr-2)$	$Q = \sum_{i=1}^{t}\sum_{j=1}^{r}\left(y_{ij}^2 - \frac{G^2}{tr}\right)$ -		-

Inference: If the observed value of F_0 is greater than the expected value of F_e ($F_0 > F_e$) at α% level of significance then the null hypothesis of equal treatment is rejected. Otherwise, it is accepted. Similarly, the block effect may also be tested.

IV. Application

In our study, the data was taken from 3–5 months yield of pulses data on agricultural land in India. There are various variety of pulses for green gram, red gram and horse gram (in kg and 000's) and years for 2016, 2017, 2018, 2019 and 2020 respectively the data are given below and test whether there is any significant difference between yield of pulses in different year and also estimate the yield of pulse in year 2027.

Table 4: *Different Variety of Pulses and Years*

Pulses Variety	Yield of Pulse(in Tonnes)				
	2016	2017	2018	2019	2020
Green Gram	-	56.0	78.2	76.8	76.4
Red Gram	59.9	37.0	53.7	51.6	53.2
Horse Gram	44.6	12.3	48.5	53.5	61.3

Null hypothesis(H_0): There is no significant difference between yield of pulses in different year.

Solution: The above data is one missing value problem. This problem may be using different methods. In this paper we use *RBD* one missing value method. First estimate the missing value,

$$\hat{x} = \frac{rT' + tB' - G'}{(t-1)(r-1)}$$

$$\hat{x} = \frac{287.4(3) + 104.5(5) - 763}{(3-1)(5-1)}$$

$$\hat{x} = 77.5$$

Table 5: *The Estimated Observation for Pulses and Years*

Pulses Variety	Yield of Pulse(in Tonnes)				
	2016	2017	2018	2019	2020
Green Gram	77.5	56.0	78.2	76.8	76.4
Red Gram	59.9	37.0	53.7	51.6	53.2
Horse Gram	44.6	12.3	48.5	53.5	61.3

Then substitute the missing value in above table and carried out the statistical analysis. The missing observation of the green ram will be 77.5 in 2016. It was found by using one missing observation of *RBD*. The above data was InControl or out of control. Using the 3σ control chart for different variety and different years. Upper Control Limit (UCL) = 111.7173, Center Limit *(CL)* = 56.033 and Lower Control Limit *(LCL)* = 0.3530 and also using the 6σ control chart for different variety and different years. Upper Control Limit (UCL) = 130.173, Center Limit (CL) = 56.033 and Lower Control Limit (LCL) = -18.106.

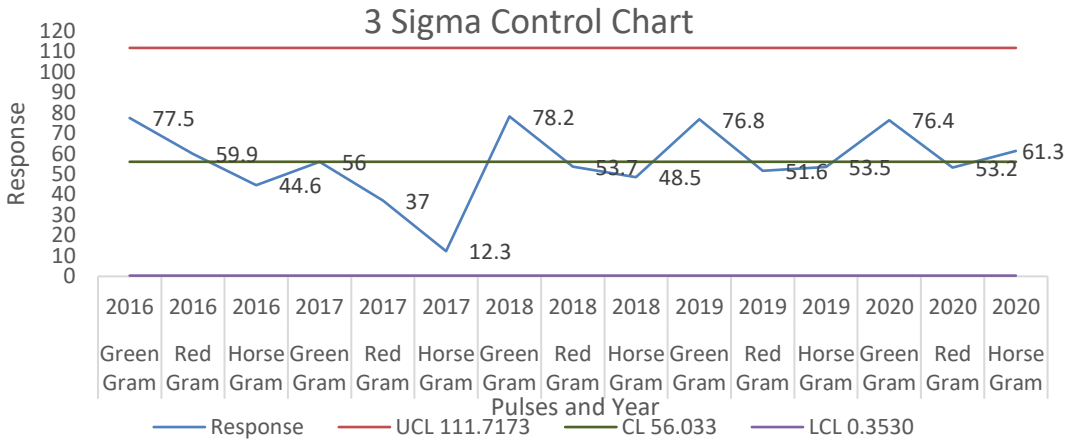


Figure 1: *3 σ Control Chart*

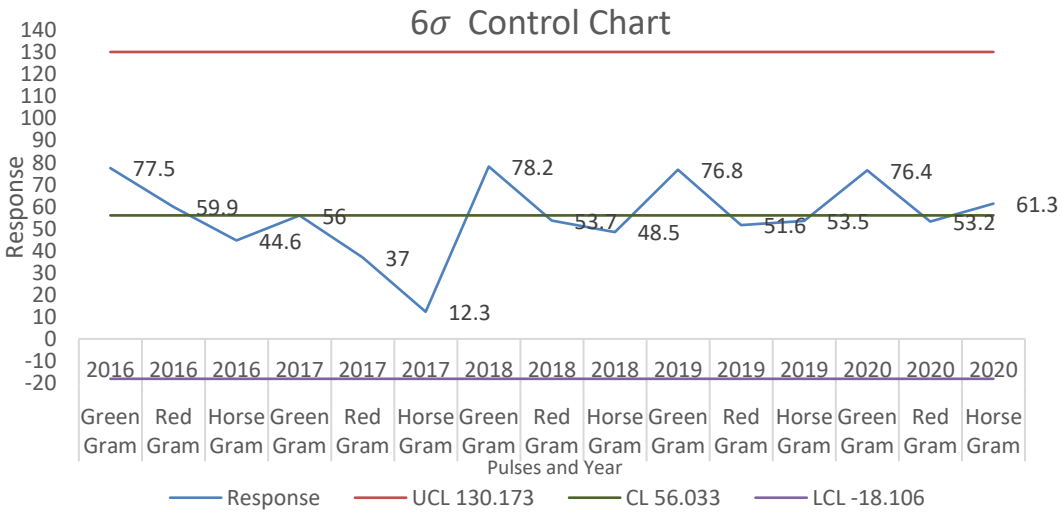


Figure 2: *6 σ Control Chart*

The above figures 1 and 2 represents the pulses of green gram, red gram and horse gram in 2016,2017,2018,2019 and 2020, which are under the control limit of $3\sigma$and 6σ. The data is in control, so we continue with further analysis.

Table 6 : *Conversion of table from variety to Pulses variety*

Year	Variety	Response	Pulses Variety
2016	Green gram	77.5	1
2017	Green gram	56.0	1
2018	Green gram	78.2	1
2019	Green gram	76.8	1
2020	Green gram	76.4	1
2016	Red gram	59.9	2
2017	Red gram	37.0	2
2018	Red gram	53.7	2
2019	Red gram	51.6	2
2020	Red gram	53.2	2
2016	Horse gram	44.6	3
2017	Horse gram	12.3	3
2018	Horse gram	48.5	3
2019	Horse gram	53.5	3
2020	Horse gram	61.3	3

In the table above, we convert the Pulses varieties of green gram, red gram and horse gram into number 1, 2 and 3 for our convenience.

Table 7 : *Model Summary*

S	*sq*	*R-sq(adj)*	*R-sq(pred)*
7.06029	90.82%	89.93%	67.72%

The above table represent the model summary of R^2 value is 90.82% and R^2 value is 89.93%.

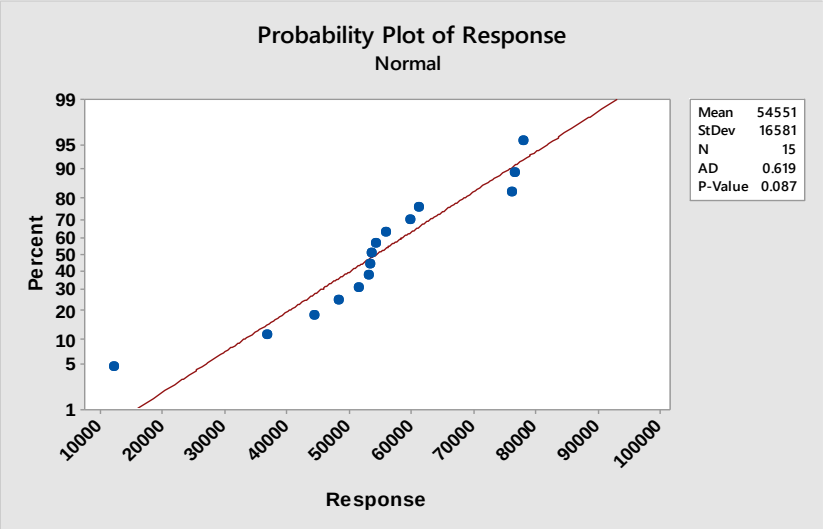


Figure 3: *Normal Plot*

This figure 3 is a probability plot of a response variable, showing that the data fit a normal distribution. It has a red line representing the expected normal distribution and blue dots representing the actual data points. The data points follow the reference line, indicating a reasonable fit to a normal distribution. The p-value is 0.139. Therefore, a p-value greater than 0.05 indicates that we have failed to reject the null hypothesis, and then the data conform to a normal distribution.

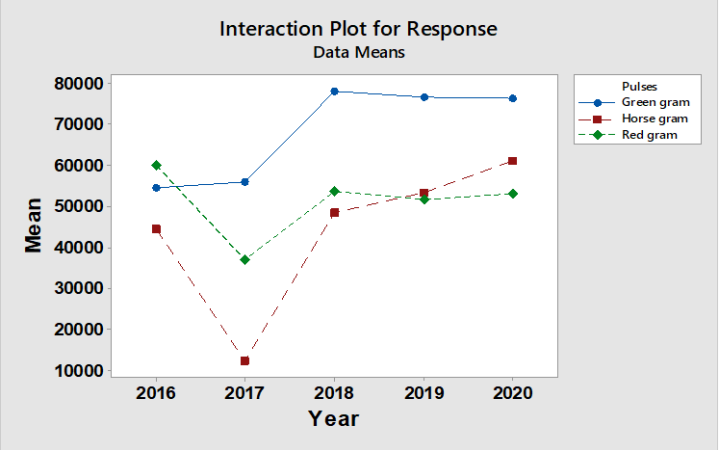


Figure 4: *Interaction Plot*

This figure 4 is a correlation graph for the response showing data for different types of pulses and different years. The graph showing the average values for each type of pulse has changed over the years, with significant interaction and trends across the different types.

Table 8 : *Coefficients*

Term	*Effect*	*Coef*	*SE Coef*	*T-Value*	*p-Value*	*VIF*
Constant		46.42	6.74	6.89	0.000	
Year	12.59	6.29	4.53	1.39	0.198	1.00
Pulsesvariety	-28.94	-14.47	3.92	-3.69	0.005	1.00
Year*Year	18.63	9.31	7.65	1.22	0.254	1.00
Pulsesvariety*Pulsesvariety	14.86	7.43	6.79	1.09	0.302	1.00
Year*Pulsesvariety	11.20	5.60	5.54	1.01	0.339	1.00

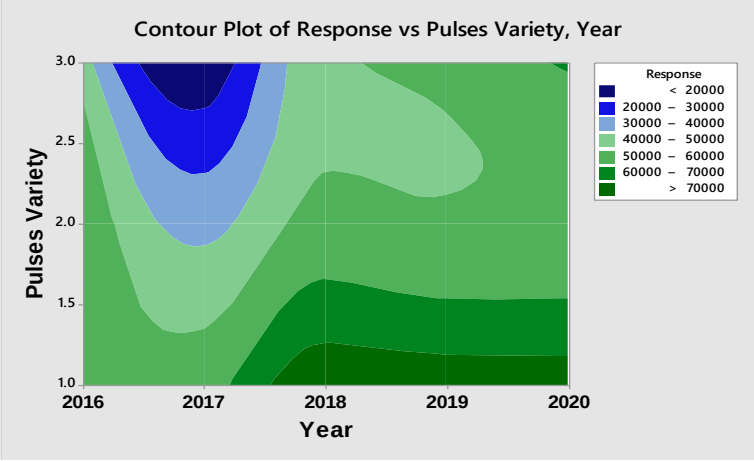


Figure 5: *Contour Plot*

The above figure 5 shows the relationship between the variables pulses variety and year, and the response variable is represented by different color gradients. These darker colors represent higher values in this graph.

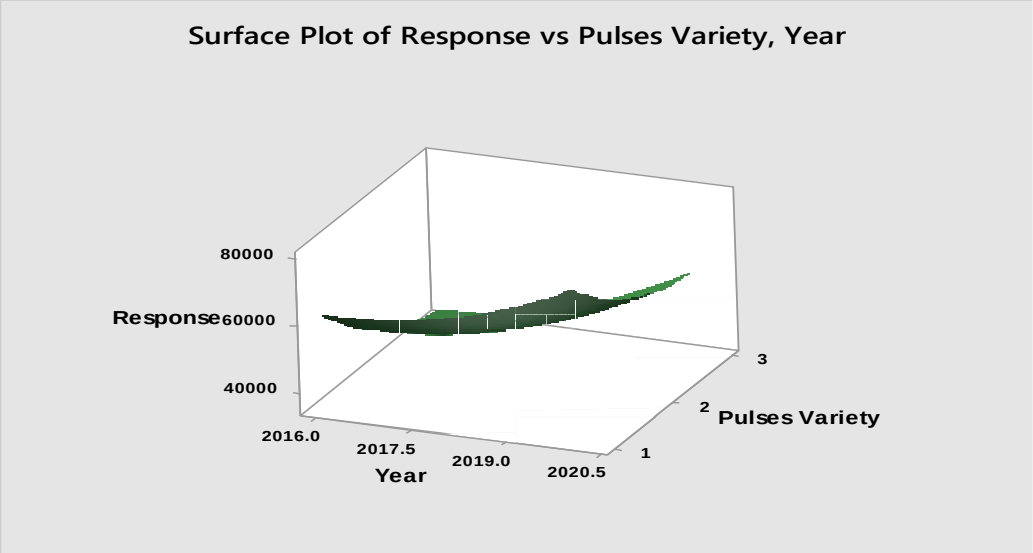


Figure 6: *3D Surface Plot*

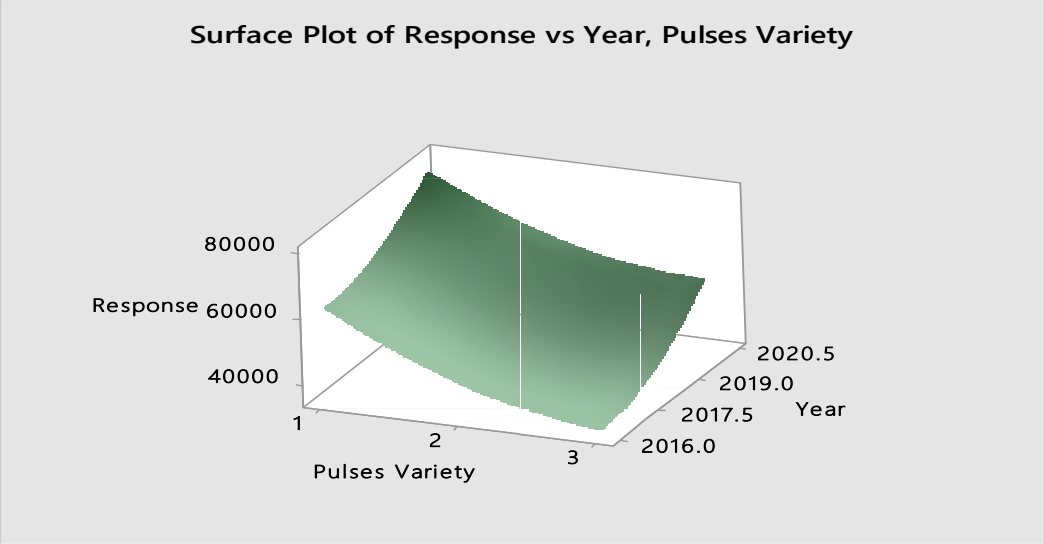


Figure 7: *3D Surface Plot*

The above figure 6 and 7 are represent the *3D* surface plot of response variable year and Pulses variety.

Table 9 : *ANOVA table for RBD layout*

SV	D*f*	*SS*	*MSS*	*F- Ratio*	*p-Value*
Variety	2	2277.8	1138.91	22.85	0.000
Year	4	1666.2	416.55	8.36	0.006
Error	8	687319021	85914878	-	-
Total	14	3849123322	-	-	-

Inference: Since the *p*-values of variety and years are less than 0.05, we reject the null hypothesis. Therefore, there is a significant difference in the yield of pulses across different years. However, the data is under control. Next, we forecasting for green gram, red gram and horse gram for next seven year (2027) are given below.

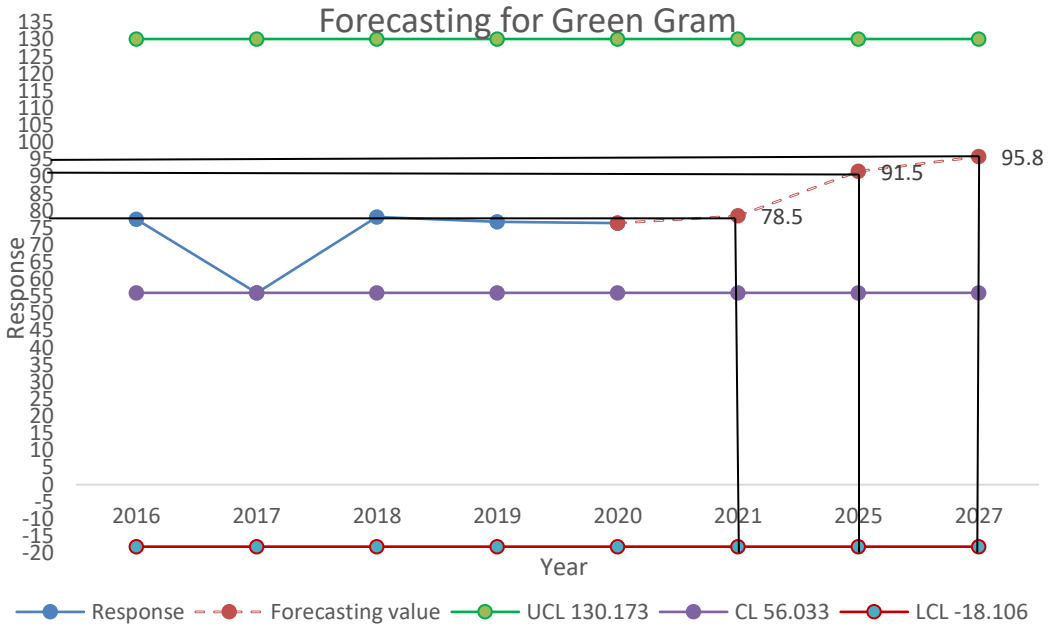


Figure 8: *Forecasting for Green Gram from 2021 to 2027*

The above figure 8 shows the green crop yield from 2016 to 2020. A predicted the future yields for the following years: 2021: 78.5, 2022: 86.1, 2023: 84.5, 2024: 88.0, 2025: 91.5, 2026: 94.1, and 2027: 95.8.

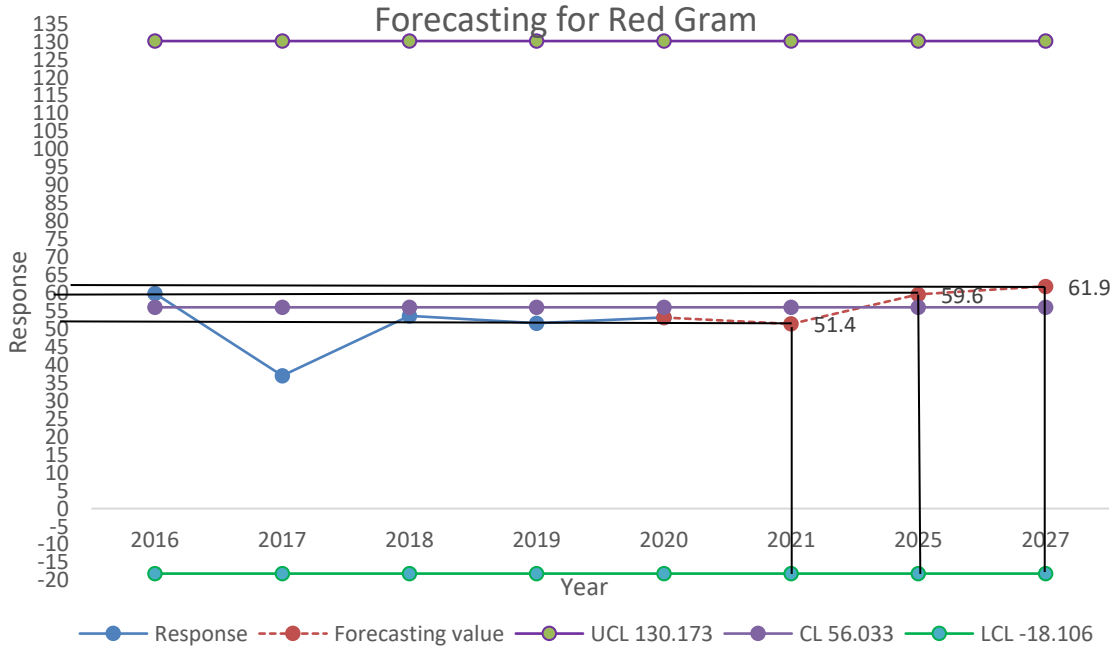


Figure**9:** *Forecasting for Red Gram from 2021 to 2027*

The above figure 9 shows the yield of red gram from 2016 to 2020. A predicted the subsequent yield for the following years: 2021: 51.4, 2022: 57.9, 2023: 56.0, 2024: 58.1, 2025: 59.6, 2026: 61.6 and 2027: 61.9.

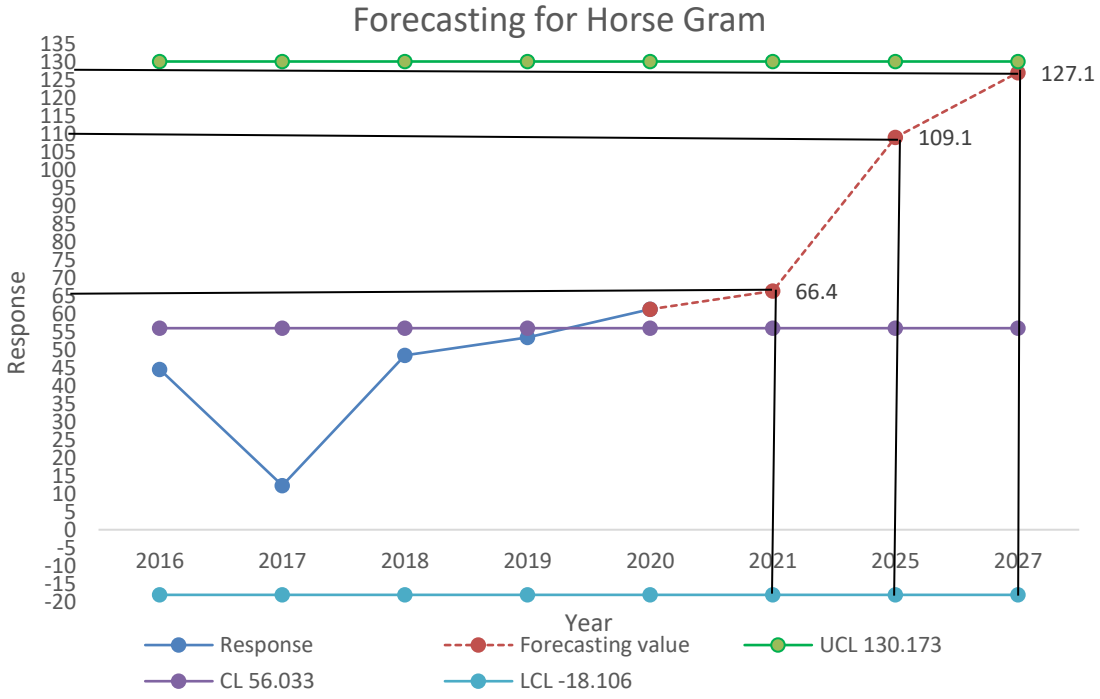


Figure 10: *Forecasting for Horse Gram from 2021 to 2027*

The above figure 10 shows the yield of the horse gram from 2016 to 2020. A predicted the yield for the following years: 2021: 66.4, 2022: 84.7, 2023: 88.4, 2024: 98.9, 2025: 109.1, 2026: 119.4, and 2027: 127.1.

V. Conclusion

In this paper, new methods were introduced to estimate a missing value in an *RBD* and to determine whether the data is in control or out of control using control charts, specifically 3σ and 6σ control charts, along with *RSM* in a numerical example. First, one missing value was estimated as $\hat{x} = 77.5$ using the *RBD* one-missing-value formula, and the original dataset was then fitted. Statistical analysis was carried out using statistical software, confirming a significant difference in the yield of pulses across different years. Also, the study provides future yield predictions for pulses up to the year 2027, offering valuable insights for agricultural planning and decision-making. The predicted yields are as follows: green gram: 2021: 78.5, 2022: 86.1, 2023: 84.5, 2024: 88.0, 2025: 91.5, 2026: 94.1, 2027: 95.8. red gram: 2021: 51.4, 2022: 57.9, 2023: 56.0, 2024: 58.1, 2025: 59.6, 2026: 61.6, 2027: 61.9. horse gram: 2021: 66.4, 2022: 84.7, 2023: 88.4, 2024: 98.9, 2025: 109.1, 2026: 119.4, 2027: 127.1. The integration of control charts and *RSM* proves to be an effective approach for handling missing values and optimizing responses in agricultural datasets. For future research, this methodology can be extended to other crop varieties and environmental conditions to enhance predictive accuracy. Furthermore, incorporating advanced machine learning techniques and hybrid statistical models may further refine missing value estimations and improve yield forecasting. Exploring the impact of climate change, soil conditions, and other agronomic factors alongside *RSM*-based estimations could provide a more comprehensive framework for agricultural yield analysis.

References

[1] Azadeh, A. et al (2013). Optimum estimation of missing values in randomized complete block design by genetic algorithm. *ScienceDirect Elsevier.* 37, 37-47, https://doi.org/10.1016/j.knosys.2012.06.014.

[2] Best, D.J and Rayner, J.C.W (2016). Analysis of ranked data in randomized blocks when there are missing values. *Journal of Applied Statistics*. 44(1), 16-23, https://doi.org/10.1080/02664763.2016.1158245.

[3] Pachamuthu, M. et al (2018). A statistical analysis of latin square designs through fuzzy ranking approach using median value. *International journal of agricultural & statistical sciences*. Vol. 17(1), pp. 163-170, https://connectjournals.com/03899.2021.17.163.

[4] Ghosh, D.K. et.al (2019). Use of control charts for conducting anova study. *International Journal of Agricultural and Statistical Science*. 15(1), 375-384, ISSN: 0973-1903.

[5] Pachamuthu, M. et al (2018). An Optimisation Method of Construction for Warping Copper Plates and Engines Using Complete Block Designs with Some Special Types of Graphs. *Advances in Materials Science and Engineering*. Volume 2023, Article ID 9002743, 12 pages https://doi.org/10.1155/2023/9002743.

[6] Naugraiya, M. and Drury, C.G. (2014). A fractional factorial screening experiment to determine factors affecting discrete part process control**.** *Theoretical Issues in Ergonomics Science*. 10:1, 1-17, http://dx.doi.org/10.1080/14639220802075832.

[7] Pachamuthu, M. and Shanmugasundram, V. (2017). Construction of control chart using fuzzy probabilistic approach for cotton sweater product**.** *Research Journal of Mathematical and Statistical Sciences*. Vol. 5(4), 1-11.

[8] Pachamuthu, M. et al (2018). Construction of 3-Sigma Control Chart Using Fuzzy Probabilistic Approach for Quality Evaluation. *Science and Technology*. Volume 4, Issue 5, ISSN: 2395-6011, Online ISSN: 2395-602X.

[9] Pachamuthu, M. and Mariappan, A. (2018). On the Construction of Statistical Quality Control Chart Using Fuzzy Probabilistic Approach. *Science and Technology*. Volume 4, Issue 5, ISSN: 2395-6011, Online ISSN: 2395-602X.

[10] Pachamuthu, M. and Ramya, A. (2019). On the construction of balanced incomplete block designs through factorization and coloring graphs using mutually orthogonal Latin square designs*. International journal of agricultural and statistical sciences*. Vol. 15(2).

[11] Pachamuthu, M. and Mariappan, A. (2021). A Statistical Analysis of Variance for One Way Classification with Addition Operations using Triangular Fuzzy Numbers. *International Journal of Statistics and Reliability Engineering*. Volume 8, Issue 1, 45-56.

[12] Richa Seth, et.al (2017). Missing plot techniques using regression analysis and its comparison with randomized block design missing plot techniques. *International Journal of Applied Mathematics and Statistical Sciences.* 6(4), 55-66. ISSN(P): 2319-3972; ISSN(E): 2319-3980.

[13] Subramani, J and Balamurali, B. (2012). Loss of efficiency in randomized block designs with two missing values*. Elixir International*. 49, 10042-10044.

[14] Tae-Yeon ChoGhosh, et.al (2011). Mixture-process variable experiments including control and noise variables within a split-plot structure. *International Journal of Quality Engineering and Technology.* 2(1), 1-28, https://doi.org/10.1504/IJQET.2011.038719.

[15] Pachamuthu, M. et al (2022). A statistical hypotheses technique of one-way analysis of variance using triangular fuzzy numbers. *Advances and Applications in Mathematical Sciences.* Volume 21, Issue 6, Pages 2989-3009.

RT&A, No 2 (84)
Volume 20, June 2025

RELIABILITY AND PROFIT ANALYSIS OF REPAIRABLE EDIBLE OIL PLANT

Ashish Kumar[1], and Geetanjali Sharma[2*]

•

[1, 2]Department of Mathematics and Statistics, Banasthali Vidyapith, Rajasthan
ak3151998@gmail.com, geetanjali.bu@gmail.com

*Corresponding Author

Abstract

In daily life, edible oil is necessary for making food to human. This paper described the reliability, availability and profit values of the edible oil plant by using regenerative point graphical technique. Generally, it contains four units such that bleacher (B), deodorizer (D), thermo fluid boiler (C) and pressured filter (A). The whole refinery plant work properly when all units work properly. Only unit (A) is in partially failed state as well as complete failed state and other remaining units are in complete failed state. The system works when unit is in operative state or partially failed state. The system fails when one unit completely fails. A technician is always available to repair the failed unit. In this paper, the failure time and repair time follows general distributions. The regenerative point graphical technique with stochastic process is used to explore the reliability measures such that mean time to system failure, availability of the system, and profit values.

Keywords: Base state, oil refinery, reliability, availability and profit values.

I. Introduction

Manufacturers must constantly innovate their products to keep up with the rising demand for their products, which is made feasible by optimizing their manufacturing processes. The MTSF, availability and profitability of the edible oil plant with priority in repair are discussed in this study by utilizing the regenerating point graphical technique under specific circumstances. Many researchers threw light on the various industrial systems using different methods. Balagurusamy [2] described the terms related to the systems meantime, failure, repair, redundancy, maintainability, and availability measures. Gupta et al. [5] examined the single unit system using base state and regenerative point graphical technique. The reliability analysis of a one-unit system with finite vacations was examined by Liu and Liu [13]. Aggarwal et al. [1] threw light on the behaviour of refining system of a sugar plant under serial processing with regenerative point graphical technique. Sadeghi and Roghanian [15] analyzed the reliability, availability and profit values of two unit warm standby system under imperfect switch. Kumar et al. [11] examined the availability and profit values of bread manufacturing system with regenerative point graphical technique. Kumar et al. [12] described the performance of industrial system under multistate failure. Thori and Goel [16] evaluated the stochastic behaviour of edible oil refinery.

Barak et al. [3] examined the reliability characteristics of milk plant using regenerative point graphical technique. Barak et al. [4] threw light on the performance of two unit cold standby system with availability of technician and refreshment approach under semi markov process and regenerative point technique. Kumar and Sharma [6] evaluated the two unit cold standby system under refreshment facility. Kumar et al. [8] descried the performance and reliability measures of

cold standby computer system under software upgrade and load recovery facility. Kumar et al. [9] analyzed the computer system hardware and software performance using Weibull distribution and regenerative point technique. Kumar et al. [10] examined the reliability measures of a two unit cold standby system subject to inspection before repair. Kumar and Sharma [7] described the reliability characteristics of two unit cold standby system with reboot and refreshment facilities. Rahul et al. [14] discussed on the availability and profit values of juice plant with regenerative point graphical technique.

II. System Assumptions

A refinery system is the combination of subunits and every subunit has its significant values in the refinery plant. To describe the edible oil refinery, there are following assumptions

- The edible oil refinery consists of four distinct units *A*, *B*, *C* and *D*.
- It is considered that units *A* is in a complete failed state through partial failure mode but unit *B*, *C* and *D* are in complete failed state.
- Unit *A* has pressured filters, and unit *B* has bleaching.
- Unit *C* has thermo fluid boiler, and unit D has deodorizer.
- A technician is always available with the system.
- Failure time and repair time are generally distributed and are independent.
- The unit works just like a brand-new one after repair.

III. System Notations

To explain the edible oil refinery, there are following notations

$i \xrightarrow{Sr} j$	r^{th} directed simple path from state '*i*' to state '*j*' where '*r*' takes the positive integral values for different directions from state '*i*' to state '*j*'.
$\xi \xrightarrow{s\,f\,f} i$	A directed simple failure free path from state ξ to state '*i*'.
$m-cycle$	A circuit (may be formed through regenerative or non regenerative / failed state) whose terminals are at the regenerative state '*m*'.
$m-\overline{cycle}$	A circuit (may be formed through the unfailed regenerative or non regenerative state) whose terminals are at the regenerative '*m*' state.
$q_{ij}(t)$	PDF of the first passage time from a regenerative state *i* to a regenerative state state *j* or to a failed state *j* without visiting any other regenerative state in (0,1].
$U_{k,k}$	Probability factor of the state '*k*' reachable from the terminal state '*k*' of '*k*' cycle.
$U_{\overline{k,k}}$	The probability factor of state '*k*' reachable from the terminal state '*k*' of $k\ \overline{cycle}$.
μ_i	Mean sojourn time spent in the state '*i*' before visiting any other states.
μ_i'	Total unconditional time spent before transiting to any other regenerative state while the system entered regenerative state '*i*' at t=0.
ξ	Base state of the system.
f_j	Fuzziness measure of the *j*-state.
η_i	Expected waiting time spent while doing a job given that the system entered to the regenerative state '*i*' at t=0.
$A / \overline{A} / a$	The unit is in the operative state/reduced state/failed state.
B / b	The unit is in the operative state/failed state.

C / c	The unit is in the operative state/failed state.
D / d	The unit is in the operative state/failed state.
λ_1 / λ_5	Fixed partial failure rate/complete failure rate of the unit A respectively.
λ_2 , λ_6	Fixed complete failure rate of the unit B.
λ_3 , λ_7	Fixed complete failure rate of the unit C.
λ_4 , λ_8	Fixed complete failure rate of the unit D.
w_1, w_5	Fixed repair rate of the partially failed/complete failed unit A respectively.
w_2 , w_6	Fixed repair rate of unit B after the complete failure.
w_3 , w_7	Fixed repair rate of unit C after the complete failure.
w_4 , w_8	Fixed repair rate of unit D after the complete failure.
○/ ⬭ ▭	Upstate/ reduced state/ failed state.

IV. Circuits Descriptions

Using the transition diagram, primary, secondary and tertiary circuits are calculated and used to find the base state of the system such that

Table 1: *Circuit Descriptions*

i	(C1)	(C2)	(C3)
0	(0,1,0), (0,2,0), (0,3,0), (0,4,0), (0,1,5,0)	(1,6,1), (1,7,1), (1,8,1)	Nil
1	(1,0,1), (1,6,1), (1,7,1), (1,8,1), (1,5,0,1)	(0,2,0), (0,3,0), (0,4,0)	Nil
2	(2,0,2)	(0,1,0), (0,3,0), (0,4,0)	Nil
3	(3,0,3)	(0,1,0), (0,2,0), (0,4,0)	Nil
4	(4,0,4)	(0,1,0), (0,2,0), (0,3,0)	Nil
5	(5,0,1,5)	(0,1,0), (0,3,0), (0,4,0), (1,6,1), (1,7,1), (1,8,1)	Nil
6	(6,1,6)	(1,0,1), (1,7,1), (1,8,1), (1,5,0,1)	Nil
7	(7,1,7)	(1,0,1), (1,6,1), (1,8,1), (1,5,0,1)	Nil
8	(8,1,8)	(1,0,1), (1,6,1), (1,7,1), (1,5,0,1)	Nil

where, $S_0 = ABCD$, $S_1 = \bar{A}BCD$, $S_2 = AbCD$, $S_3 = ABcD$
$S_4 = ABCd$, $S_5 = aBCD$, $S_6 = \bar{A}bCD$, $S_7 = ABcD$, $S_8 = \bar{A}BCd$

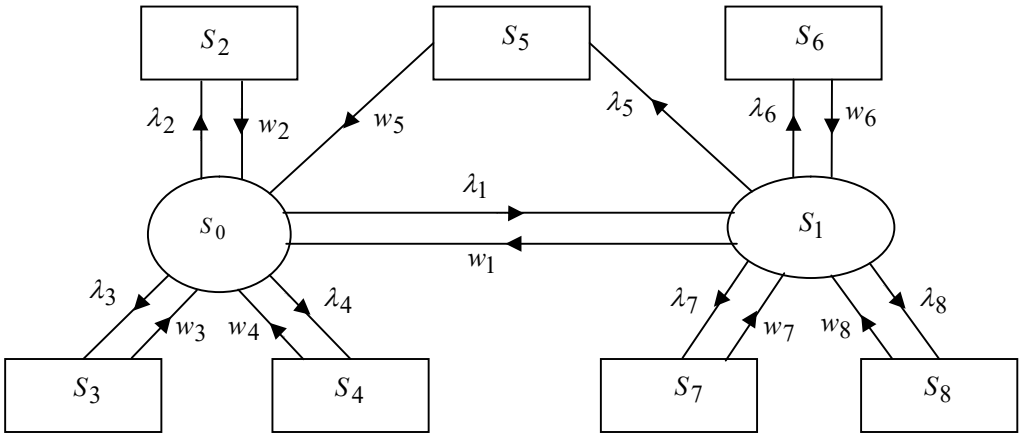


Figure 1 *State Transition Diagram*

V. Transition Probabilities

There are following transition probabilities

$$p_{0,1} = \lambda_1/(\lambda_1+\lambda_2+\lambda_3+\lambda_4),\ p_{0,2} = \lambda_2/(\lambda_1+\lambda_2+\lambda_3+\lambda_4)$$

$$p_{0,3} = \lambda_3/(\lambda_1+\lambda_2+\lambda_3+\lambda_4),\ p_{0,4} = \lambda_4/(\lambda_1+\lambda_2+\lambda_3+\lambda_4)$$

$$p_{1,0} = w_1/(w_1+\lambda_5+\lambda_6+\lambda_7+\lambda_8),\ p_{1,5} = \lambda_5/(w_1+\lambda_5+\lambda_6+\lambda_7+\lambda_8)$$

$$p_{1,6} = \lambda_6/(w_1+\lambda_5+\lambda_6+\lambda_7+\lambda_8),\ p_{1,7} = \lambda_5/(w_1+\lambda_5+\lambda_6+\lambda_7+\lambda_8)$$

$$p_{1,8} = \lambda_8/(w_1+\lambda_2+\lambda_3+\lambda_4+\lambda_5),\ p_{2,0} = p_{3,0} = p_{4,0} = p_{5,0} = p_{6,1} = p_{7,1} = p_{8,1} = 1 \quad (1)$$

It is verified that

$$p_{0,1} + p_{0,2} + p_{0,3} + p_{0,4} = 1,\ p_{1,0} + p_{1,5} + p_{1,6} + p_{1,7} + p_{1,8} = 1$$

VI. Mean Sojourn Time

For the particular state, it becomes

$$\mu_0 = 1/(\lambda_1+\lambda_2+\lambda_3+\lambda_4),\ \mu_1 = 1/(w_1+\lambda_5+\lambda_6+\lambda_7+\lambda_8),\ \mu_2 = 1/(w_2),\ \mu_3 = 1/(w_3)$$

$$\mu_4 = 1/(w_4),\ \mu_5 = 1/(w_5),\ \mu_6 = 1/(w_6),\ \mu_7 = 1/(w_7),\ \mu_8 = 1/(w_8) \quad (2)$$

VII. Evaluation of Parameters

Using the circuit table, '0' is used as the base state to calculate the reliability measures with the regenerative point graphical technique. The probability factors of all the reachable states from the base state '0' are given below

$$U_{0,0} = 1,\ U_{0,1} = \frac{\mu_1(w_1+\lambda_5+\lambda_6+\lambda_7+\lambda_8)^3}{\left[\begin{array}{l}(w_1+\lambda_2+\lambda_3+\lambda_4)(w_1+\lambda_3+\lambda_4+\lambda_5)\times \\ (w_1+\lambda_2+\lambda_4+\lambda_5)(\lambda_2+\lambda_3+\lambda_4+\lambda_5)\end{array}\right]}$$

$$U_{0,2} = \lambda_2/(\lambda_1+\lambda_2+\lambda_3+\lambda_4),\ U_{0,3} = \lambda_3/(\lambda_1+\lambda_2+\lambda_3+\lambda_4),\ U_{0,4} = \lambda_4/(\lambda_1+\lambda_2+\lambda_3+\lambda_4)$$

$$U_{0,5} = \frac{\mu_5 \lambda_1 (w_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5)^2}{\left[\begin{array}{c}(w_1 + \lambda_2 + \lambda_3 + \lambda_4)(w_1 + \lambda_3 + \lambda_4 + \lambda_5)\times \\ (w_1 + \lambda_2 + \lambda_4 + \lambda_5)(\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5)\end{array}\right]}$$

$$U_{0,6} = \frac{\mu_6 \lambda_1 (w_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5)^2}{\left[\begin{array}{c}(w_1 + \lambda_2 + \lambda_3 + \lambda_4)(w_1 + \lambda_3 + \lambda_4 + \lambda_5)\times \\ (w_1 + \lambda_2 + \lambda_4 + \lambda_5)(\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5)\end{array}\right]}$$

$$U_{0,7} = \frac{\mu_7 \lambda_1 (w_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5)^2}{\left[\begin{array}{c}(w_1 + \lambda_2 + \lambda_3 + \lambda_4)(w_1 + \lambda_3 + \lambda_4 + \lambda_5)\times \\ (w_1 + \lambda_2 + \lambda_4 + \lambda_5)(\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5)\end{array}\right]}$$

$$U_{0,8} = \frac{\mu_8 \lambda_1 (w_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5)^2}{\left[\begin{array}{c}(w_1 + \lambda_2 + \lambda_3 + \lambda_4)(w_1 + \lambda_3 + \lambda_4 + \lambda_5)\times \\ (w_1 + \lambda_2 + \lambda_4 + \lambda_5)(\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5)\end{array}\right]}$$

I. Mean Time to System Failure

The regenerative un-failed states (i=0, 1) to which the system can transit (with initial state 0) before entering to any failed state (using base state ξ=0) then MTSF becomes

$$T_0 = \left[\sum_{i=0}^{1} Sr\left\{\frac{\left\{pr\,(0 \xrightarrow{Sr\,(sff)} i)\right\}.\mu_i}{\prod_{k_1 \neq 0}\left\{1 - V_{k_1 k_1}\right\}}\right\}\right] \div \left[1 - \sum Sr\left\{\frac{\left\{pr\,(0 \xrightarrow{Sr\,(sff)} 0)\right\}}{\prod_{k_2 \neq 0}\left\{1 - V_{k_2 k_2}\right\}}\right\}\right]$$

$$T_0 = \frac{U_{0,0}\mu_0 + U_{0,1}\mu_1}{[1 - (1,0,1)]} \tag{3}$$

II. Availability of the system

The system is available for use at regenerative states j=0, 1 with ξ=0 then the availability of system is defined as

$$A_0 = \left[\sum_{j=0}^{1} Sr\left\{\frac{\left\{pr\,(0 \xrightarrow{Sr} j)\right\}.f_j.\mu_j}{\prod_{k_1 \neq 0}\left\{1 - V_{k_1 k_1}\right\}}\right\}\right] \div \left[\sum_{i=0}^{8} Sr\left\{\frac{\left\{pr\,(0 \xrightarrow{Sr} i)\right\}.\mu_i'}{\prod_{k_2 \neq 0}\left\{1 - V_{k_2 k_2}\right\}}\right\}\right]$$

$$A_0 = \frac{U_{0,0}\mu_0 + U_{0,1}\mu_1}{\left[\begin{array}{c}U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3 + U_{0,4}\mu_4 \\ + U_{0,5}\mu_5 + U_{0,6}\mu_6 + U_{0,7}\mu_7 + U_{0,8}\mu_8\end{array}\right]} \tag{4}$$

III. Busy Period of the Technician

A technician is always available with the system to repair the failed unit. The technician is busy due to repair of the failed unit at regenerative states j= 1, 2, 3, 4, 5, 6, 7, 8 with ξ = 0 then the fraction of time for which the technician remains busy is defined as

$$B_0 = \left[\sum_{j=1}^{8} Sr\left\{\frac{\left\{pr\,(0 \xrightarrow{Sr} j)\right\}.\eta_j}{\prod_{k_1 \neq 0}\left\{1 - V_{k_1 k_1}\right\}}\right\}\right] \div \left[\sum_{i=0}^{8} Sr\left\{\frac{\left\{pr\,(0 \xrightarrow{Sr} i)\right\}.\mu_i'}{\prod_{k_2 \neq 0}\left\{1 - V_{k_2 k_2}\right\}}\right\}\right]$$

$$B_0 = \frac{\begin{bmatrix} U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3 + U_{0,4}\mu_4 \\ + U_{0,5}\mu_5 + U_{0,6}\mu_6 + U_{0,7}\mu_7 + U_{0,8}\mu_8 \end{bmatrix}}{\begin{bmatrix} U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3 + U_{0,4}\mu_4 \\ + U_{0,5}\mu_5 + U_{0,6}\mu_6 + U_{0,7}\mu_7 + U_{0,8}\mu_8 \end{bmatrix}} \tag{5}$$

IV. Estimated number of visits made by the Technician

The technician visits at regenerative states j= 1 with ξ=0 then the number of visits by the technician is defined as

$$V_0 = \left[\sum_{j=1} Sr \left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} j) \right\}}{\prod k_1 \neq 0 \left\{ 1 - V_{k_1 k_1} \right\}} \right\} \right] \div \left[\sum_{i=0}^{8} Sr \left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} i) \right\} . \mu_i'}{\prod k_2 \neq 0 \left\{ 1 - V_{k_2 k_2} \right\}} \right\} \right]$$

$$V_0 = \frac{\left| U_{0,1}\mu_1 \right|}{\begin{bmatrix} U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3 + U_{0,4}\mu_4 \\ + U_{0,5}\mu_5 + U_{0,6}\mu_6 + U_{0,7}\mu_7 + U_{0,8}\mu_8 \end{bmatrix}} \tag{6}$$

V. Profit Analysis

The profit function may be used to do a profit analysis of the system and it is given by

$$P = E_0 A_0 - E_1 B_0 - E_2 V_0 \tag{7}$$

where, $E_0 = 2500$ (Revenue per unit uptime of the system)

$E_1 = 500$ (Cost per unit time for which technician is busy due to repair)

$E_2 = 200$ (Cost per visit of the technician)

VIII. Discussion

It is assumed that all $\lambda_i = \lambda$ for i=1, 2, 3, 4, 5, 6, 7, 8 and $w_j = w$ for j=1, 2, 3, 4, 5, 6, 7, 8. Tables 2, 3 and 4 describe the nature of mean time to system failure, availability and profit values of the edible

Table 2: *MTSF vs. Repair Rate* (w)

w ↓	λ=0.3	λ=0.4	λ=0.5
0.5	3.545455	3.244275	3.040826
0.55	3.588589	3.276736	3.079858
0.6	3.630952	3.308458	3.118812
0.65	3.672566	3.339483	3.157248
0.7	3.713458	3.36983	3.195122
0.75	3.753623	3.399519	3.232446
0.8	3.793103	3.428571	3.269231
0.85	3.831909	3.457008	3.305489
0.9	3.870056	3.484848	3.341232
0.95	3.907563	3.512111	3.376471

Table 3: *Availability vs. Repair Rate* (*w*)

w ↓	λ=0.3	λ=0.4	λ=0.5
0.5	0.629837	0.609587	0.547732
0.55	0.634798	0.614607	0.552793
0.6	0.639628	0.619499	0.557741
0.65	0.644332	0.624269	0.562582
0.7	0.648915	0.628925	0.567318
0.75	0.653381	0.633458	0.571952
0.8	0.657734	0.637887	0.576488
0.85	0.661985	0.642209	0.580929
0.9	0.666122	0.646438	0.585278
0.95	0.670164	0.650552	0.589538

Table 4: *Profit vs. Repair Rate* (*w*)

w ↓	λ=0.3	λ=0.4	λ=0.5
0.5	2456.69	2399.61	2035.15
0.55	2485.35	2429.18	2064.97
0.6	2513.15	2458.02	2094.16
0.65	2540.26	2486.16	2122.65
0.7	2566.68	2513.53	2150.56
0.75	2592.43	2540.28	2177.86
0.8	2617.54	2566.36	2204.59
0.85	2642.02	2591.83	2230.76
0.9	2665.89	2616.76	2256.38
0.95	2689.22	2640.98	2281.48

oil refinery having an increasing trend corresponding to increment in repair rate (*w*). In these tables, the values of parameters change such that λ=0.3, 0.4, 0.5 and *w*=0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95 respectively. When λ=0.3 changes into 0.4, 0.5 then MTSF, availability and profit values have decreasing trends.

IX. Conclusion

The performance of the edible oil refinery is discussed using the regenerative point graphical technique. The above tables explore that when the repair rate increases then the MTSF, availability and profit values also increase but when the failure rate increases then the MTSF, availability and profit values decrease. It is clear that RPGT is helpful for industries to analyze the behaviour of the products and components of a system.

X. Future Scope

It is observed that the role of the regenerative point graphical technique for the edible oil will be beneficial and also used by the management, manufacturers and the persons engaged in reliability

engineering and working on analyzing the nature and performance analysis of the system like soft drink, paper industry, etc.

References

[1] Aggarwal, A. K., Kumar, S., and Singh, V. (2017). Performance modeling of the serial processes in refining system of a sugar plant using RAMD analysis. *International Journal of System Assurance Engineering and Management*, *8*(Suppl 2), 1910-1922.

[2] Balagurusamy, E. (1984). *Reliability Engineering*. Tata McGraw-Hill Education.

[3] Barak, M. S., Garg, R., and Kumar, A. (2021). Reliability measures analysis of a milk plant using RPGT. *Life Cycle Reliability and Safety Engineering*, *10*(3), 295-302.

[4] Barak, M. S., Kumar, A., and Garg, R. (2023). Contribution of Refreshment Provided to the Server During His Job in the Repairable Cold Standby System. *Computational Intelligence in Sustainable Reliability Engineering*, 251-267.

[5] Gupta, V. K. (2011). Analysis of a single unit system using a base state. *Aryabhatta Journal of Mathematics and Informatics*, *3*(1), 59-66.

[6] Kumar, A., and Sharma, A. (2023). Profit analysis of repairable cold standby system under refreshment. *Reliability Theory and Applications*, *18*(4(76)), 604-612.

[7] Kumar, A., and Sharma, A. (2024). Profit analysis of repairable cold standby system subject to reboot facility under. *Reliability: Theory and Applications*, *19*(2 (78)), 594-603.

[8] Kumar, A., Garg, R., & Barak, M. S. (2023). Performance analysis of computer systems with Weibull distribution subject to software upgrade and load recovery. *Life Cycle Reliability and Safety Engineering*, *12*(1), 51-63.

[9] Kumar, A., Garg, R., & Barak, M. S. (2023). Performance of computer system with three types of failure using weibull distribution subject to hardware repair and software up-gradation. *International Journal of System Assurance Engineering and Management*, *14*(Suppl 1), 483-491.

[10] Kumar, A., Garg, R., & Barak, M. S. (2023). Reliability measures of a cold standby system subject to refreshment. *International Journal of System Assurance Engineering and Management*, *14*(1), 147-155.

[11] Kumar, A., Goel, P., and Garg, D. (2018). Behaviour analysis of a bread making system. *International Journal of Statistics and Applied Mathematics*, *3*(6), 56-61.

[12] Kumar, A., Ram, M., Pant, S., and Kumar, A. (2018). Industrial system performance under multistate failures with standby mode. *Modeling and Simulation in Industrial Engineering*, 85-100.

[13] Liu, R. and Liu, Z. (2011). Reliability analysis of a one-unit system with finite vacations. In *MSIE*, 248-252.

[14] Rahul, Yadav, M., and Kumar, H. (2024). Profit analysis of repairable juice plant. *Reliability: Theory and Applications*, *19*(1 (77)), 688-695.

[15] Sadeghi, M., & Roghanian, E. (2017). Reliability analysis of a warm standby repairable system with two cases of imperfect switching mechanism. *Scientia Iranica*, *24*(2), 808-822.

[16] Thori, S., & Goel, P. (2018). Stochastic analysis of edible oil refinery industry. *Int J Stat Appl Math*, *3*(1), 4-12.

TOUCH AND STEP VOLTAGE HAZARD ANALYSIS

E.S.Safiyev[1], S.Y.Shikhaliyeva[2]

•

Azerbaijan State Oil and Industry University, Baku, Azerbaijan
[1]elshad@safiyev.com; [2]saadaasmar@mail.ru

Abstract

This article analyzes the risks associated with electric shock, specifically focusing on the effects of touch and step voltage, which can lead to severe injury or death. Citing international studies, the research confirms that older individuals typically exhibit a body resistance between 1000-2000 Ω, with exposure to 230 V potentially causing dangerous currents. The study highlights key factors influencing shock severity, including current magnitude, duration, body resistance, and current path. Additionally, the article explores the understudied phenomenon of step voltage, which arises from lightning strikes or electrical faults, and poses significant risks to both humans and animals. Despite theoretical analysis, the lack of real-world data on step voltage's impact on the human body is emphasized. The paper presents a novel approach to evaluating and mitigating these risks by considering soil resistance, current flow, and physiological effects. The reliability of the studies cited is supported by extensive theoretical research, but the article calls for further investigation and the implementation of improved grounding systems and safety protocols to better protect high-risk environments, such as industrial sites and military installations. The findings contribute valuable insights into the understanding of lightning-induced hazards and propose new safety measures for minimizing step voltage dangers.

Key words: touch voltage, step voltage, lightning strikes, electrical resistance, grounding systems, lightning protection, electrical hazards, voltage distribution, nervous system damage, safety protocols, power surges, soil resistivity.

I. Introduction

Electric shock occurs when a person touches an electrically charged object and at the same time comes into contact with another surface that can conduct electric current to the ground. Common sources of electric shock include bare and damaged wires, electrical machines, appliances, conductive cables, and other electrical installations. Proper grounding and the use of electrical safety devices can help prevent electric shock.

According to international studies, 95% of older people have a total body resistance (impedance) in the range of 1000-2000 Ω. When a person with a body resistance of 1000 Ω is exposed to a voltage of 230 Volts, a current of up to 230 milliamperes can pass through the body [1-3]. This current poses a serious and potentially fatal risk (figure 1).

The curves in figure 1 are explained as follows, zone: AC-1 - generally has no effect on humans; AC-2 - generally is not dangerous for the human body; AC-3 - generally has no effect on the body. However, if the current lasts more than 2 seconds, there is a possibility of muscle contraction and respiratory difficulty; AC-4 - with increasing intensity and duration of the current, pathophysiological effects such as cardiac arrest, respiratory arrest and severe burns may occur. AC-4.1 - the probability of ventricular fibrillation of the heart is about 5%; AC-4.2 - the probability of ventricular fibrillation of the heart is about 50%; AC-4.3 - the probability of ventricular fibrillation of the heart is more than 50%.

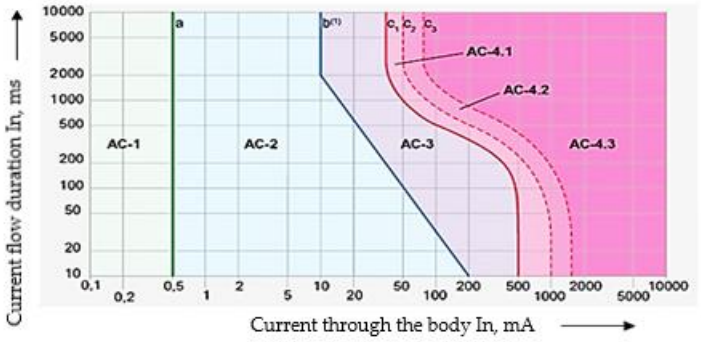


Figure 1. *Ampere-time curve showing the effect of alternating current with a frequency of 10-100 Hz on a person*

To understand the risk posed by electric current, the following five factors must be taken into account:

- the magnitude of the current;
- the duration of the current;
- the path it takes through the human body;
- the resistance of the body;
- frequency of the power supply.

According to the IEC 61140 standards, two types of dangerous contact are defined: direct contact and indirect contact [22]. Direct contact refers to a person who makes contact with a conductor that is normally under voltage. Indirect contact refers to a person who makes contact with any part of an electrical circuit that is not under voltage but is conductive as a result of accidental damage to the insulation or any other fault.

II. Formulation of the problem

The value of the maximum allowable current, depending on the duration, can be calculated analytically using the following expressions (1), (2).

$$I_b = \frac{0.116}{\sqrt{t_s}} \qquad \text{(body mass 50 kg)} \qquad (1)$$

$$I_b = \frac{0.157}{\sqrt{t_s}} \qquad \text{(body mass 70 kg)} \qquad (2)$$

Equations (1) and (2) were determined based on experiments conducted by Dalziel and it was estimated that 99.5% of the participants could withstand the shock without suffering from ventricular fibrillation. Dalziel's experiments were conducted for a contact time of 0.03÷3 sec. [6-9].

Main part. Touch voltage (U_{touch}) is the potential difference between an object under voltage (for example, the housing of electrical equipment) and the foot of a person touching this object. When a person touches the grounded housing of the equipment in contact with one of the phases, part of the grounding current passes through the person, and if the housing is not grounded, the grounding current passes completely through the person (unipolar touch) [4]. The value of the touch voltage for a person who is on the ground and touches a grounded housing under voltage can be considered as the difference between the potentials of his hand (housing) and his foot (ground) (figure 2). When the current is closed to the case of any equipment, the potential on the ground surface is distributed along a hyperbolic curve (figure 3 and 4).

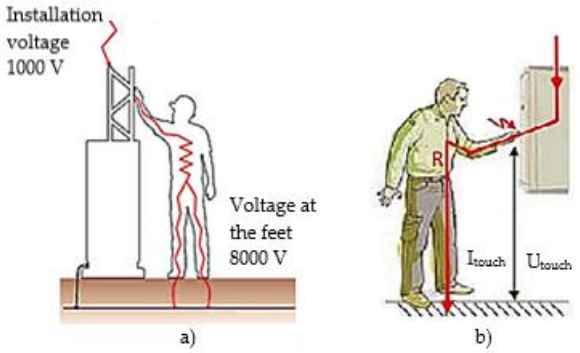


Figure 2. *Touch voltage, when the equipment is a)grounded and b) ungrounded*

In this case, touch voltage is the difference in potentials between the case of electrical equipment (for example, an electric motor) and the points of the ground surface touched by a human foot [5].

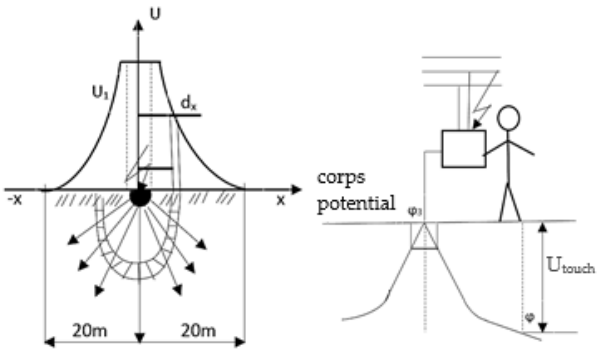


Figure 3. *Hemispherical grounding conductor and potential distribution curves on the ground surface*

The further the electrical equipment is from the earthing conductor, the greater the touch voltage U_{touch} to which the person is exposed, and vice versa, the closer to the earthing conductor, the lower the touch voltage (figure 4). Outside the current propagation zone (≥20 m), the touch voltage is equal to the voltage on the housing of the electrical equipment [11].

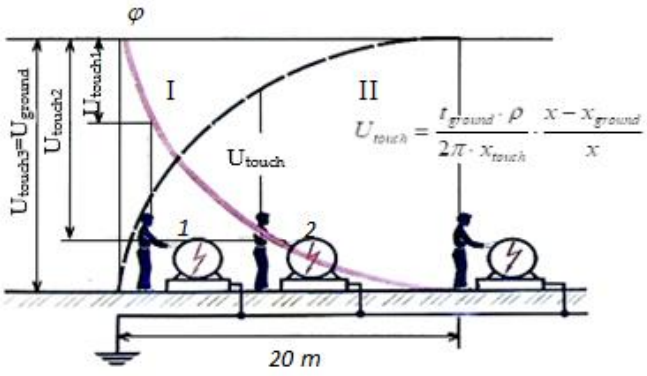


Figure 4. *Grounded electrical equipment and touch voltage.*
I – Potential distribution curve on the ground surface; II – Touch voltage curve

III. Problem solution

When a person standing on the ground touches a grounded body that is under voltage, the contact voltage can be calculated using the following expression (3) [12-15]:

$$U_{touch} = \varphi_{hand} - \varphi_{ground}, \tag{3}$$

where φ_{hand} is the potential of the hand and body (4):

$$\varphi_{hand} = U_{ground} = \frac{I_{ground} \cdot \rho}{2\pi \cdot x_{ground}}; \tag{4}$$

φ_{ground} is the potential of the ground on which the person is standing (5):

$$\varphi_{ground} = \frac{I_{ground} \cdot \rho}{2\pi \cdot x}; \tag{5}$$

p is the specific resistance of the ground. Taking into account the value of potentials (6):

$$U_{touch} = \frac{I_{ground} \cdot \rho}{2\pi} \cdot \left(\frac{1}{x_{ground}} - \frac{1}{x} \right) = \frac{I_{ground} \cdot \rho}{2\pi \cdot x_{ground}} \cdot \frac{x - x_{ground}}{x} = U_{ground} \cdot a_1 . \tag{6}$$

Here $\alpha_1=(x-x_{ground})/x$ is the coefficient of contact voltage for a hemispherical ground electrode (figure 2). A similar coefficient (β_1) is calculated for step voltage. These coefficients are determined experimentally for particularly complex group ground electrodes, which are distinguished by their hemispherical shape. This expression allows calculating the contact voltage without taking into account the resistance of the supporting surface of the foot (R_{foot}) and the resistance of the shoe (R_{shoe}). The influence of these resistances can be determined by the coefficient α_2. This coefficient takes into account the voltage drop across additional resistances in the human circuit (7):

$$a_2 = \frac{R_{body}}{R_{body} + R_{foot} + R_{shoe}} = \frac{R_{body}}{R_{total}}, \tag{7}$$

where R_{body} is the resistance of the human body, R_{total} is the total resistance.

The analog coefficient (β_2) calculated for step voltage is 4 times smaller than the coefficient α_2 calculated for touch voltage (8), i.e.:

$$a_2 = 4 \cdot \beta_2 . \tag{8}$$

The full expression for the touch voltage is (9):

$$U_{full.touch} = U_{touch} \cdot a_1 \cdot a_2 . \tag{9}$$

If there is a grounding conductor for electrical equipment (figure 4), when a person touches electrical equipment outside the zone where the current flows, the touch coefficient $a_1=1$, then $U_{touch}=U_{max}=U_{ground}$ (not taking into account the coefficient α_2).

Touching a housing that is under voltage and outside the field distribution is the most dangerous situation for a person.

The current flowing through a person exposed to touch voltage can be calculated as (10):

$$I_{body} = I_g \cdot \frac{R_{ground}}{R_{body}} \cdot a_1 \cdot a_2 , \quad (10)$$

where R_{ground} is the resistance of the grounding conductor.

In the indicated cases, a step voltage (U_{step}) also arises. The greatest step voltage arises near the grounding conductor, especially when a person stands on the grounding conductor with one foot. On the same equipotential line or outside the potential distribution area (x>20 m), the step voltage is equal to zero. The value of the step voltage is much less than the touch voltage [19].

Experience shows that:

- if the distance between the person servicing the electrical installation and the grounding conductor is equal to or greater than 20 m (figure 4, case 3 x≥20 m), then the potential in the base (ground) will be equal to zero. In this case, the touch voltage is maximum and equal to the grounding potential;

- if the distance x is zero (figure 3, case 1, x=0;), then the potential at the base is equal to the potential at the grounding conductor, and the touch voltage will be zero.

Thus, considering the step voltage and touch voltage, the following conclusions can be drawn:

- the exposure of a person to danger depends on his distance from the grounding conductor. As this distance decreases, the danger from step voltage increases, while the danger from touch voltage decreases, and vice versa, as the distance increases, the danger from step voltage decreases, while the danger from touch voltage increases;
- in general, touch voltage is more dangerous than step voltage, since the maximum value of the touch voltage coefficient is equal to unity amax=1, and the maximum value of the step voltage coefficient is less than unity βmax < 1. In addition, the flow of current through the touch voltage along the "hand-to-foot" path is more dangerous than the flow through the "foot-to-foot" path.

In order to reduce the danger of people from touch voltage, grounding conductors should be installed as close as possible to electrical installations in production areas [16-18].

In engineering standards, the touch distance for calculating the touch potential is taken as 1 m, in dangerous places - 2 m, and in especially dangerous places - even 3 m.

In industrial areas, a technical protection measure is taken to reduce the danger from touch voltage and step voltage at the same time - the potential of the core is equalized. For this, contour grounding with a group of grounding conductors in the form of a grid is used. Vertical grounding conductors are placed at a distance equal to or less than their lengths from each other, i.e. S≤L (figure 5).

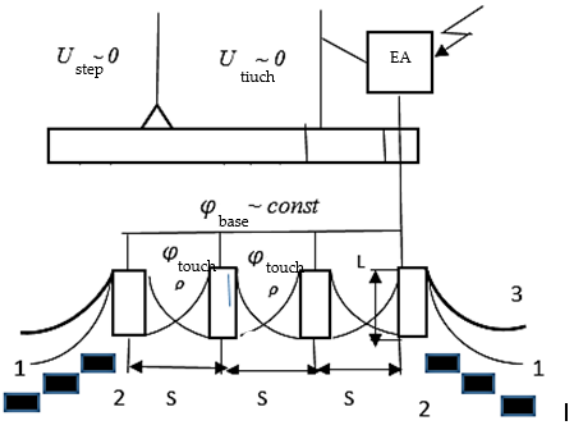


Figure 5. *Contour grounding with a group of earthing conductors*
1 – natural potential curves; 2 – metal tires; 3 – low-slope potential curves

When the phase is closed to the housing of an electrical installation (EA) connected to a group of earthing conductors, current begins to flow from them along mutually intersecting potential curves 1 (figure 5), their potentials are summed up to form the group potential sum φqr [7, 8]. As a result, at each point of the foundation of the production area, an almost identical potential is created, equal to the potential of the EA housing: $U_{base} \approx const$. Thus, the values of the touch and step voltages will be close to zero: $U_{touch} \approx 0$ and $U_{step} \approx 0$. To reduce the step voltage outside the contour, additional metal tires 2 are laid in the ground along the transitions, as a result of which the potential decrease goes along a more sloping curve [20].

In order to protect against various types of hazardous conditions, the following measures should be taken: do not assume that a fallen wire is safe just because it is in the ground or does not create a spark; don't assume that any insulated wires are just telephone, television or fiber optic cables; low-hanging wires are still live even if they are not touching the ground, so don't touch them; never approach a fallen electrical appliance or a downed power line. Always assume that it is live. Touching it can be fatal; electric current can travel in a circular pattern from the point of contact to the ground. The greater the distance from the center, the greater the voltage difference; never drive over downed power lines. Assume that they are live; if you are in a car and come into contact with a live power line, stay calm and do not get out unless the car is on fire. Call for help if possible; if you have to leave any equipment for fire or other safety reasons, try to jump completely clear, making sure not to touch the equipment or the ground at the same time.

To minimize the path of the electric current and avoid electric shock, land on both feet and walk away in small steps. Be careful to maintain your balance [12].

Step voltage is the potential difference between two parts of the earth. This physical effect has not been sufficiently studied, despite the fact that the surge voltage that occurs between parts of the earth due to lightning strikes, breaks and falls of power lines to the ground and other causes is very dangerous. Since the specific electrical resistance of the soil is not zero, the current in each case "flows" in the soil, creating a zone of high potential. If there is a conductor on the ground, an electrical short circuit may occur. Such a conductor can also be a person on the ground, in which case the current enters one leg and exits the other, turning the human body into an "electric charge". This is very dangerous, since high voltage can cause muscle paralysis. As a result, a person can fall to the ground on his hands, and with subsequent lightning discharges, the current can pass through the heart area, increasing the risk of cardiac arrest. When a person's head touches the ground, the risk of irreversible damage to the central nervous system increases. Step voltage is more dangerous for large horned animals, because the distance between the front and hind legs of these animals is very large, and accordingly, the voltage falling on this distance is also high. There are not many cases of death from step voltage in such animals. Note that if the earth's surface has a high resistance (rocky or dry sandy area, etc.), the voltage can instantly spread over a wider area when lightning strikes, and animals can die at a greater distance from the lightning strike. The specific electrical resistance of some parts of the earth's surface is shown in Table 1 [4-6].

Table 1. *Specific electrical resistance of some parts of the Earth's surface (Ω ·m)*

Sand	400 - 1000 and more
Sandy clay	150 - 400
Clay	8 - 70
Garden soil	40
Black soil	10 - 50
Peat	20
Stony clay	100
Limestone, coarse sand	1000 - 2000
Rock, stones	2000 - 4000
River water (plain)	10 - 80
Sea water	0,2
Tap water	5 - 60

Although there are algorithms for accurately determining the step voltage near lightning rods and tall objects, additional knowledge is needed to determine the safety zone. In most cases, lightning strike victims believed that they were not in danger and that their property was protected by lightning rods. Theory is very important, but the decisive parameter for practice is the distance from the arrester at which the step voltage is guaranteed to be safe. It is important to report on the installation of lightning rods and voltage distribution grids. Unfortunately, there is little research on this topic, and in this case, military research in the field of lightning is of great importance. This is due to the fact that in military units of different countries, in different climatic zones and weather conditions, lightning often strikes lightning rods in armories, near tents where equipment and military personnel are located. Accurate recording of these events allows us to collect useful statistics. The 2002 US Army TRADOC Lightning Protection Guidelines state that there have been deaths from power surges within 10-20 m of a lightning strike. In most cases, the lightning strike was not caused by the direct impact of a lightning discharge, but by a power surge and arc discharges from devices [7].

The calculation of the expected step voltage value is provided by a special warranty program and the corresponding service of specialists. Step voltage, as we have already mentioned, depends on the current strength and the specific resistance of the earth. Dangerous power surges can occur near the grounding of electrical equipment during an emergency shutdown. Therefore, the permissible resistance value of grounding devices and the specific resistance of the soil should be standardized [21-23].

As we have already mentioned, when loading a step, the leg muscles contract strongly, as a result of which a person falls to the ground, and as a result, the current begins to flow through new points of support - for example, from the hands to the feet, which can lead to death [18].

The likelihood of being struck by a power surge can be reduced by observing the rules of behavior in the danger zone. To do this, you must be free of potentials on the ground. Therefore, the distance between the legs should be reduced to a minimum and in no case should you touch the ground or various objects (steel supports, reinforced concrete walls, etc.) with your hands. In emergency situations, i.e., if lightning suddenly strikes a lightning rod or tall objects, if you are near them, you should move away in very short and tight steps ("walk like a goose"). Sometimes it is recommended to stand on one leg, but in this case, an unstable body position can lead to falling on your hand or instinctively grabbing an object. Due to the above physiological reasons, it is absolutely forbidden to allow current to flow along the head-to-leg or arm-to-leg circuit, since this can damage the brain and heart. Therefore, in such conditions, you must stand with your legs tightly pressed together. The voltage step depends on the resistance of different soil layers, although it can be roughly calculated. The short-circuit current in the event of a break in one phase of the high-voltage wire and a fall to the ground can be calculated using expression (11):

$$I_q = \frac{U_f}{\sqrt{3(R_0 + R_{cont})}}, \tag{11}$$

where U_f is the phase voltage, R_0 is the resistance of the neutral to the operating voltage (several Ω), R_{cont} is the resistance to the flow of current at the contact point (usually 12 Ω) [12].

The potential of the contact point at any point x on the earth's surface is given by expression:

$$\varphi_x = U_x = \frac{I_q \cdot \rho}{2 \cdot \pi \cdot x}, \tag{12}$$

where p is the resistance of the ground layer. Considering that ——=*k*=const. Obtain expression (13):

$$\varphi_x = U_x = \frac{k}{x}. \tag{13}$$

The expression (13) is a hyperbolic equation, meaning that the potential of points on the earth's surface varies according to a hyperbolic law.

When lightning strikes a lightning rod or the housing of electrical equipment falls under the AG, the potential of point t on the surface of the grounding device (voltage on the grounding device) will have a maximum value, which is shown in expression (14):

$$\varphi_t = \varphi_{\max} = U_t = \frac{I_q \cdot \rho}{2 \cdot \pi \cdot x_t} = I_t \cdot R_t, \quad (14)$$

where $R_t=p/2\cdot\pi\cdot x_t$ - is the resistance of the grounding conductor.

The step voltage between points x and x+a can be calculated according to expression (15):

$$U_{step} = \varphi_x - \varphi_{x+a} = \frac{k}{x} - \frac{k}{x+a} = k \cdot \frac{a}{x \cdot (x+a)} = \varphi_x \cdot \frac{a}{(x+a)} = \frac{I_q \cdot \rho \cdot a}{2 \cdot \pi \cdot x \cdot (x+a)}, \quad (15)$$

where p is the soil resistivity (several hundred Ω·m), x is the distance from the lightning strik point or the point where the YG wire falls, a is the length of the step [19-22].

Highly resistive stony, frozen, etc. types of soil are especially dangerous due to their low permeability. When lightning strikes them, very high voltages arise. A person can be struck by tens of thousands of volts of voltage in a normal stride [16].

Being near a building is also dangerous, because the step voltage can spread through the foundation. As can be seen from the graph (figure 2), at a lightning current of 30 kA, the step voltage at a distance of 5 m on a reinforced concrete foundation slab is 4.5 kV. Let us consider that this voltage is within the normative grounding resistance (4 Ω) [12].

In industrial areas, to reduce the danger of step voltage, technical protection measures are taken - they equalize the potential of the core. For this purpose, loop grounding with a grounding group in the form of a grid is used. Vertical grounding conductors are placed at a distance equal to or less than their length, i.e. S≤L (figure 6).

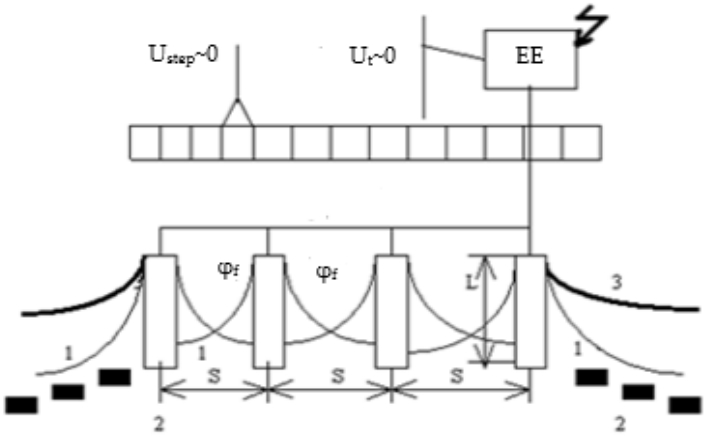


Figure 6. *Contour grounding with a group of grounding electrodes*

When a phase is shorted to the body of electrical equipment (EE) combined with a group of grounding electrodes, current 1 begins to flow from them along mutually intersecting potential curves (figure 6), their potential is collected and the group potential φ_f is collected. As a result, at each point of the structure of the production zone, a practically identical potential is created, equal to the foundation potential: $U_f\approx$const. Therefore, the value of the touch and step voltage will be close to zero: $U_{step}\approx0$ and $U_t\approx0$. To reduce the step voltage outside the circuit, additional metal

buses 2 are placed along the transitions to the ground, as a result of which the potential drop occurs along a flatter curve.

Thus, the step voltage still remains an additional dangerous effect of a lightning strike. The effect of such microsecond, multiple impulse voltages on the body has been poorly studied. Therefore, a special attitude should be paid to lightning protection and the study of the rules of behavior in risky conditions of a step voltage strike.

Modern technologies allow to significantly reduce the risk of step voltage. Some equipment has an automatic blocking function in case of damage to the electrical network. Such capabilities allow to significantly increase the level of safety and reduce the number of accidents at the enterprise.

In combination, methods of reducing step voltage give excellent results. Automated enterprises working with innovative equipment almost never encounter emergency situations.

The article studies touch and step voltage, a phenomenon caused by the distribution of electrical potential in the ground following lightning strikes or electrical faults [13-15].

The dangerous physiological effects of step voltage are highlighted, in particular its unpredictable and potentially fatal consequences for the human body. It is noted that even low voltages, such as 36 V, can cause serious injuries, including muscle paralysis, cardiac arrest, and damage to the nervous system. The lack of a clearly defined "safe" step voltage threshold is emphasized, indicating the complexity of human resistance and the impact of lightning-induced currents on the body.

The behavior of step voltage as it propagates through the ground is investigated as a function of soil resistivity.

It highlights how different soil types (e.g. rocky, sandy, or frozen soil) affect the step voltage hazard zone, resulting in different distances at which the voltage can affect both humans and animals.

Environmental factors such as soil type can influence the degree of hazard caused by step voltage. In areas with highly resistive soil, the voltage can spread over a larger area, increasing the risk to people and animals further from the lightning strike point.

The availability of theoretical algorithms for calculating step voltage highlights the lack of real-world data on lightning protection systems, especially in areas such as military installations and high-risk industrial environments.

Practical risk reduction measures are suggested, such as the use of grounding systems, metal meshes, and dielectric materials to create zones of equal potential. However, it highlights the need for better designed protective infrastructure to ensure safety.

Placing additional metal bars in the ground along the transitions to reduce the off-loop voltage step results in a steeper potential drop curve.

Overall, the article advances the discussion by highlighting the lack of research on the real-world effects of step voltage, both physiologically and in terms of its wider environmental effects. It suggests more in-depth study and more effective protective measures to prevent fatalities caused by lightning-related surges.

IV. Conclusions

In conclusion, the article provides a comprehensive analysis of the dangers posed by various hazardous electrical conditions, emphasizing the importance of precautionary measures to prevent fatal accidents. It underscores that both direct contact with electrical appliances or downed power lines, as well as indirect hazards such as step voltage, present significant risks to human safety. The article highlights the unpredictability and potentially deadly effects of step voltage, even at low voltages, and its capacity to affect both humans and animals, with the severity of the risk influenced by environmental factors such as soil resistivity.

The reliability of the conclusions drawn from the article is supported by the detailed exploration of the phenomenon of step voltage, the physiological effects it has on the body, and the

analysis of the factors influencing the hazard zone. However, the article also points out the lack of real-world data on the effectiveness of current lightning protection systems and suggests that further research is required to fully understand the environmental and physiological implications of step voltage, particularly in high-risk environments like military and industrial settings.

From a broader perspective, the article calls for a more refined approach to the design of protective measures, emphasizing the need for improved infrastructure, such as better grounding systems and dielectric materials, to minimize the risks associated with electrical hazards. The absence of clear safety thresholds for step voltage indicates a significant gap in the existing safety protocols, suggesting that more robust and scientifically informed safety guidelines are essential to mitigate the risks of electrical hazards.

Overall, the article provides valuable insights into the dangers of step voltage and other electrical hazards, urging for further research, better-designed protective systems, and enhanced public awareness to prevent fatalities and injuries caused by electrical accidents.

References

[1] Safiyev E.S., Ahmadov E.N., Rzayeva S.V. Step voltage hazard and reliable protection against it. 6th International Artificial Intelligence and Data Processing Symposium, 08-09 Sep. 2022, Inonu University, Malatya, Turkey, pp.176-181.

[2] Safiyev E.S., Ahmadov E.N., Aliyeva G.A. Touch voltage and reliable protection against it. "Modern problems and development prospects of electric power". International scientific and technical conference. 17-18 November 2022, Baku, pp.28-33.

[3] Semekhin Yu. G., Bondin V. И. Safety of life. M.-Berlin, izd. Direct-Media, 2020. – 412 c.

[4] Kotsch, J.; Prussak, B.; Morse, M.; Kohl, J. Exploration of the Theory of Electric Shock Drowning. In Proceedings of the 2018 IEEE IAS Electrical Safety Workshop (ESW), Fort Worth, TX, USA, 19–23 March 2018; pp. 1–3.

[5]Harid, N.; Bogias, A.C.; Griffiths, H.; Robson, S.; Haddad, A. A Wireless System for Monitoring Leakage Current in Electrical Substation Equipment. IEEE Access 2016, 4, 2965–2975.

[6] Wu, S.; Lin, S.-Y.; Guo, M.-F. Application of WE-AE-BP Method to Electric Shock Faults Identification in the Low-voltage Distribution Network. In Proceedings of the 2020 IEEE 1st China International Youth Conference on Electrical Engineering, Wuhan, China, 1–4 November 2020; pp. 1–6.

[7] Zhang, M.; Qin, P.; Chen, Y.; Jia, H.; Wang, Z.; Deng, W. Study on the Detection Method of Leakage in TN-C Area of Low-voltage Distribution Network. In Proceedings of the 2021 International Conference on Power System Technology, Fukuoka, Japan, 10–12 September 2021; pp. 378–381.

[8] Zhang, X.; Liu, Z.; Wang, Y.; Dou, Z.; Zhai, G.; Wei, Q. Residual Current Detection Method Based on Variational Modal Decomposition and Dynamic Fuzzy Neural Network. IEEE Access 2021, 9, 142925–142937.

[9] Dou, Z.; Han, Q.; Niu, X.; Peng, X.; Guo, H. An Adaptive Filter for Aeromagnetic Compensation Based on Wavelet Multiresolution Analysis. IEEE Geosci. Remote Sens. Lett. 2016, 13, 1069–1073.

[10] N.M. Pirieva, S.V. Rzaeva, S.N. Talibov, "Analysis of surge protection devices for electrical networks" Interscience: scientific journal, No. 43 (266). Part 3., Publishing house. "Internauka", pp. 14-17, Moscow, 2022.

[11] Li, K.; Lin, J.; Niu, F.; Wang, Y.; Li, Q.; Guo, Z.; Wu, Y. A Novel Fault Leakage Current Detection Method with Protection Deadzone Elimination. IEEE Trans. Instrum. Meas. 2021, 70, 3505009.

[12] Arrabe, R.G.; Granados, J.M.G.; Gomez, F.A.; Gaona, C.A.P. New Selective Directional Ground Fault Protection Method for Ungrounded Systems. IEEE Trans. Power Deliv. 2022, 37, 5234–5243.

[13] S. Shikhaliyeva, E. Safiyev, "Solving optimization problems in steady operation of regulated asynchronous motor", PRZEGLAD Elektrotechniczny, ISSN 0033-2097, R. 100 NR 10/2024, pp. 39-42, 2024.

[14] S.Y. Shikhaliyeva, "Increasing the capacıty of electrical transmission lines", MAS 19th International European Conference on mathematics, engineering, MSU, , Mingechaur, Azerbaijan, April 17-18, 2024. pp. 423-429.

[15] S.Y. Shikhaliyeva "Analysis of unexplained breaks in power transmission lines", International Journal on "Technical and Physical Problems of Engineering", Vo. 15, N2 2023, p. 198-206.

[16] S.Y. Shikhaliyeva, E.S.Safiyev "Extra high voltage on overhead power lines", The 20th International Conference on "Technical and Physical Problems of Engineering" International Organization of IOTPE, 2024, p. 76-82.

[17] S.A. Khanahmedova, S.Y. Shikhaliyeva, S.J. Alimamedova, S.M. Kerimova "Some issues of designing a hybrid electric machine", Dimensions 3 (3), 2023, p. 3-4.

[18] I.M. Marufov, S.Y. Shikhaliyeva "Analysis of the effect of temperature on solar panels and their cooling methods", Reliability: Theory & Applications 19 (4 (80)), 2024, p. 581-585.

[19] Ismayilova G.G., Mufidzade N.A., Khanahmedova S.A. "Determinatıon of insulation degradation level and location between windıngs of transformer", International Journal on "Technical and Physical Problems of Engineering" (IJTPE) Published by International Organization of IOTPE, vol. 16, pp. 127-134, 2024.

[20] Aliyev N.A., Ahmadov E.N., Khanahmedova S.A., Rzayeva S.R. "Synthesis of the Exact Parameters of the Electromagnetic Brake of a Wind Electric Installation", PRZEGLAD Elektrotechniczny, , R. 99 NR 10/2023, pp.120-124.

[21] S.V. Rzayeva, N.A. Ganiyeva, N.M.Piriyeva. "Modern approaches to electrical equipment diagnostics", İnternational Journal on "Technical and Physical Problems of Engineering" (IJTPE) - Issue 58, Volume 16, Number 1, March 2024 (Serial No: 0058-1601-0324), pp.182-189.

[22] G.S. Kerimzade "Structure of the monitoring and tracking electromechanical control system", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT. ISSN 0033-2097,R. 100. NR 07 /2024. Warszawa.. pp.295-297.

[23] M. Mitolo, P. Montazemi: "Electrical Safety in the Industrial Workplace: an IEC Point of View". IEEE Transactions On Industry Applications; November/December 2014.

RELIABILITY ESTIMATION AND PARAMETERS FOR THE LIFESPAN DISTRIBUTION BASED ON FAILURE TIME DATA FOR ELECTRICAL COMPONENTS

K. Geetha[1], S.N. Manoharan[2], A. Radhika[1], S. Mythily[1], and P. Manigandan[1*]

•

[1]Department of Statistics, Periyar University, Salem-11, Tamil Nadu, India.
[2]Associate Professor, Department of Computer Science and Engineering, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Avadi, Chennai, India.
geethakmanju2002@gmail.com, kritimanoharan@gmail.com, mythilyshanmugam03@gmail.com, radhisaran2004@gmail.com, srimanigandan95@gmail.com

Abstract

The objective of the present study is to analyse the failure times of electrical components in order to identify the optimal distribution model for their longevity. Fifteen components failure data were analysed using the Weibull, Lognormal, and Exponential lifespan distribution algorithms. The Akaike Information Criterion (AIC), Kolmogorov-Smirnov (KS) test, likelihood ratio test (LRT), and log-likelihood function were among the statistical tests used to estimate and assess the goodness-of-fit of each distribution's essential parameters. The results demonstrate that, with better LRT and AIC values, the exponential distribution best fits this electrical component failure times. Real-world issues are used to illustrate the importance of the suggested model.

Keywords: Weibull, Lognormal, Exponential distribution, Reliability Measures, electrical components

I. Introduction

Reliability theory and survival analysis both heavily rely on lifetime distributions. Weibull, log-normal, exponential, and gamma distributions are a few of the distributions that are frequently used to represent lifetime data. The reason why exponential and Weibull distributions are more often used than Gamma and log-Normal distributions is that the survival function of the latter two does not have a closed form. Studies on survival and reliability have recently made extensive use of mixture distributions. It has been receiving a lot of interest because it allows for the simultaneous modelling of numerous causes of failure and mixture models are more appropriate. Because survival mixture models are so flexible, they are a preferable option for analysing survival or reliability data when the data are thought to be diverse and a single parametric distribution might not be enough. Using power series distributions to compound the exponentiated half-logistic distribution, Esmaeili and Niaparast [1] established the exponentiated half logistics-power series. In order to meet the mean lifetime of the products under Birnbaum-Saunders and Weibull distributions, Aslam et al. [2] devised and created a variation of the sampling plan for multiple dependent states. Using the lognormal distribution, Shaheen et al. [3] created a control chart to

track variation under a repeated sampling approach. A new monitoring system for the intervals between events under exponential and gamma distributions was proposed by Shah et al. [4]. The Weibull-Lindly Distribution was introduced by Chacko et al. [5] to model failure rate data with a bathtub shape. By introducing a new parameter into the exponential distribution, Mahdavi and Kundu [6] created a novel alpha power transformation technique that yielded new probability distributions. The transmuted exponential distribution's performance rating was applied to a real-world dataset by Enahoro et al. [7]. Weibull distributions with two or three parameters (shape, scale, and location) can be used to estimate the reliability of a system or sub-system [8].

The two most common approaches used to estimate the values of these characteristics are graphic and analytical procedures. See Barreto-Souza et al. [10] for a new combination of an exponential and binomial distribution, which was part of the exponential power series (EPS) family of distributions created by Chahkandi and Ganjali [9]. The maintenance engineers can use the resulting best-fit data to assist them make strategic decisions about which essential components are most likely to cause machine failure [11]. The distribution that most closely matches the failure times is chosen for analysis [12]. The most popular distributions for modelling failure times are Weibull, Gamma, Exponential, and Lognormal, according to Haviaras [13]. Lawless [14] provides a thorough analysis of Statistical Models and Methods for Lifetime Data. According to [15], the most popular failure distributions used in reliability analysis—exponential, Weibull, normal, and lognormal are critically reviewed in this study. Studying the suggested parameter estimation techniques (Probability Plotting, Least Square Estimation, and Maximum Likelihood Estimation) requires a thorough understanding of the failure rate distributions. In order to determine which distribution best fits a real set of field data, the second section of the research aims to evaluate the benefits and drawbacks of each approach.

In this paper, a new lifetime distribution by considering a series system such that the components are Exponential, log-normal and Weibull distribution is introduced. This lifetime model was introduced for the following main reasons: (i) to propose a new flexible lifetime distribution that can be used to model lifetime data in a broader class of reliability problems; (ii) to extend the log-normal, exponential, and Weibull distributions; (iii) to assess the goodness of fit and estimate distribution parameters using statistical tests like the Kolmogorov-Smirnov and log-likelihood tests.

The following is how the remainder of the article goes: Section 2 introduces the revised distribution. Goodness-of-fit statistics and parameter estimation are covered in Section 3. Section 4 offers applications to actual data. The study concludes with some conclusions in Section 5.

II. Methods

I. Exponential Distribution

The Exponential distribution is commonly known as completely random failure time. It is widely used in the domain of reliability. Exponential distribution shows the "lack of memoryless property". It has only one parameter, which is the scale parameter. Exponential distribution is a constant failure rate and it does not suitable consistent failure rate, typically the components that fatigue or enhances its performance over time.

Recollected equations for this distribution are as:

$$\mathrm{f(f)} = \frac{1}{\theta} e^{-\frac{f}{\theta}} = \lambda e^{-\lambda f}, f \geq 0, \text{ where f is the time to failure.}$$

Reliability function denoted as:

$$\mathrm{R(f)} = e^{-\frac{f}{\theta}} = e^{-\lambda f}, f \geq 0, \text{Where } \theta = \frac{1}{\lambda} > 0 \text{ is an MTTF'S parameter and } \lambda \geq 0$$

is a constant failure rate. The failure rate for the exponential density function is:

$$\mathrm{h(f)} = \frac{f(f)}{R(f)} = \frac{1}{\theta} = \lambda$$

This distribution is favored due to its basic mathematical structure and quick manipulation. The long flat "intrinsic failure" part of the bathtub curve is well modelled by the exponential because of its constant failure rate characteristic. This encourages widespread use since a majority of systems and parts lose the majority of their lifespans in this area of the bathtub curve of the exponential. Inter-arrival times are a good fit for the exponential model. The exponential lifespan model can be used when these occurrences trigger failures.

II. Weibull Distribution

This distribution, which Wallodi Weibull first suggested in 1939, is mostly used in dependability to determine when items fail. An expanded form of the exponential distribution is the Weibull distribution. Forecasting the parts or system of the life span that are being studied is also helpful. Finding the best fit distributions is more commonly accomplished with the Weibull distribution. The cumulative distribution function (CDF) of the two parameter Weibull distribution is

$$F(f, a, b) = 1 - exp\left[-(\frac{f}{a})^b\right]$$

where f is the time to failure, a is the scale parameter, b is the shape parameter. Reliability refers to the likelihood that an item or thing will fulfil its intended function consistently throughout time under given conditions. The reliability function for the two-parameter Weibull distribution is as follows.
Reliability function denoted as:

$$R(f) = exp\left[-(\frac{f}{a})^b\right]$$

The Weibull failure rate function is defined as the number of failures per unit time that can expected to occur for the product. It is given as:

$$\lambda(f) = \frac{b}{a} \times (\frac{f}{a})^{b-1}$$

The two-parameter Weibull probability density function f(t) is given as

$$PDF = f(f) = \frac{b}{a} \times (\frac{f}{a})^{b-1} \times e^{-\left(\frac{f}{a}\right)^b}$$

$$CDF = F(f) = 1 - exp\left[-(\frac{f}{a})^b\right]$$

III. Lognormal Distribution

The lognormal distribution is a distribution of a variable whose logarithm is similar to the normal distribution. The lognormal distribution hypothesis was introduced by McAlister in 1879 and is used in many areas of medicine. The density function for the two-parameter lognormal distribution is:

$$f(f) = \frac{1}{af\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{logf - \theta}{a}\right)^2}$$

Where f is the time to failure,

Reliability function for Lognormal Distribution:

$$R(f) = P(T>f) = 1\text{-}F(f)$$

Where F(f) is the CDF of the lognormal Distribution.
Cumulative Distribution function is denoted as:

$$F(f) = P(T > f) = \phi(\frac{logf-\theta}{a})$$

Φ is the standard normal cumulative distribution function.
Reliability function:

$$R(f) = 1\text{-}\phi(\frac{logf-\theta}{a})$$

IV. Akaike Information Criteria

The Akaike information criteria (AIC), which are an estimator of the out-of-sample prediction error, help determine the relative quality of statistical models for a given collection of data. AIC estimates the proportionate amount of information lost by a particular model. It should be noted that the more information a model keeps, the better it is. When determining how much information the model loses, AIC takes into account the risk of both over-fitting and under-fitting. The formula for AIC that was employed to compare the parametric model is

$$AIC = -2LL + 2p$$

Where, p is the number of parameters in the model, LL is the log-likelihood and AIC values that are smaller indicate superior models for data adaptation.

IV. Kolmogorov-Simron Test

In reliability analysis, the Kolmogorov-Smirnov (K-S) test is used to determine whether a sample conforms to a particular distribution (Weibull, exponential, lognormal, etc.), which is essential for forecasting product reliability. The real data is shown to be closer to the fitted distribution by the smaller KS Statistic.

$$D = max \mid Fn(f) - F(f) \mid$$

$Fn(f)$ is the empirical distribution function and $F(f)$ is the cumulative distribution function of the proposed model.

III. Results and Discussions

The user needs to enter the data in a column containing the equipment failure times in order to begin the reliability study using the code. Figure 1 illustrates how a histogram relating to the frequency of these failures over time is created from these timings. Failure time data for a component must be supplied in order to run the RELPF function. The information must be organised in columns. In order to alter and analyse the data in order to suit the different distribution models, reliability analysis can be performed using R software. Packages like MASS, fidistrplus, and stats4 have been used for this study's setup and initialisation. Installing the MASS, fidistrplus, and stats4 packages from the RELPF function will allow you to estimate all parameters,

including shape and scale, based on the distribution that has to be computed. The parameters that the dependability function calculated for the distributions are shown in Table 1.

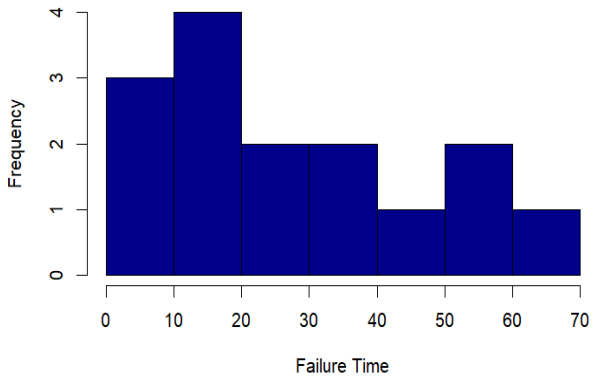


Figure 1: *Histogram generated by the software R of the failure time of electronic components.*

Table 1: *Values of the shape and scale parameters generated by the RELPF function of the failure time of electronic components*

Distributions	Shape parameter	Scale parameter
Lognormal	2.93059	1.02527
Exponential	0.03630	-
Weibull	1.30560	29.7645

From Table 1, in comparison to the other two distributions, the Lognormal distribution has a significant skew and spread, as shown by its scale parameter of 1.02527 as well as shape parameter of 2.93059. The exponential distribution has a modest scale (rate) parameter of 0.03630, indicating a continuous failure rate, due to its memoryless quality. The Weibull distribution's shape parameter of 1.30560 and scale parameter of 29.7645, which show a declining failure rate over time (since the shape is >1 but close to 1), make it more flexible than the other distributions for modelling life data.

Table 2: *Table obtained by the RELPF function showing the LRV, AIC, and Loglik results for the electrical component failure time*

Distributions	Loglik	AIC	LRT	p-value
Lognormal	-65.61742	135.23	2.8629	0.0906
Exponential	-64.73822	131.47	1.1045	0.2937
Weibull	-64.0202	132.04	1.4360	0.2307

The result is obtained in the Table 2 display that the Lognormal model has highest AIC = 135.23, the Weibull model was shown to be the second highest AIC = 132.04 and the Exponential model has the least AIC = 131.47.

Table 3: *Result of the KS test for the failure time of the electronic components.*

Distribution	Statistics KS	D-critic	Hypothesis
Lognormal	0.161	0.3511	It cannot be rejected
Exponential	0.155	0.3511	It cannot be rejected
Weibull	0.098	0.3511	It cannot be rejected

Table 3 indicate the all three models have D-statistics (0.161, 0.155, and 0.098, respectively) below the critical value (0.3511) when the three distributions (Lognormal, Exponential, and Weibull) are compared using the KS test. This means that none of the models can be dismissed. The Weibull distribution, which fits the data the best among them, has the lowest D-statistic (0.098), followed by the exponential and lognormal distributions. As a result, it seems that the Weibull distribution is the best fit for this dataset. The CDFM estimation approach is the best estimation method to produce the estimates of the Lognormal, Exponential, and Weibull parameters. The p value of the K–S test in Table 3 shows that the Lognormal, Exponential, and Weibull can well match this data set. Furthermore, the sample histogram and the fitted PDFs of the Weibull distributions using the parameter estimate findings in Table 3 are displayed in Figure 2.

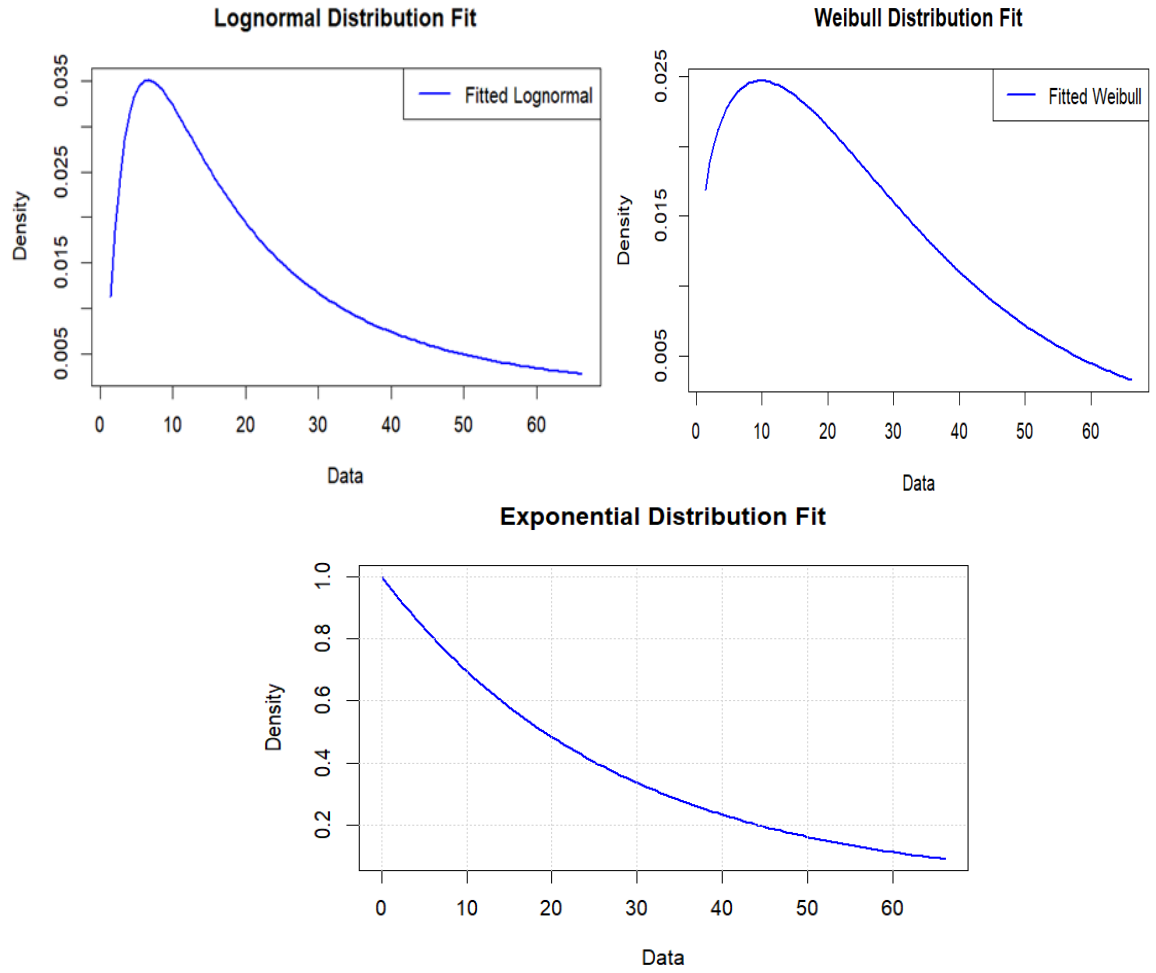


Figure 2: *Graphical presentation of failure rate functions of Exponential, Lognormal and Weibull*

IV. Conclusion

In this paper, a new lifetime distribution by considering a series system such that the components are Exponential, log-normal and Weibull distribution is introduced. The proposed distribution has successfully attained the density function features in the graphical analysis of the exponential, weibull, and lognormal distributions under different functions with varying parameter values. For the lifetime distribution, the statistical and mathematical properties are used. The lifespan distribution satisfies the conditions mentioned above. Using the maximum likelihood estimation method, the parameters of the lifespan distributions are computed. The Akaike Information Criterion (AIC), Kolmogorov-Smirnov (KS) test, likelihood ratio test (LRT), and log-likelihood

function were used to verify the goodness of fit test. We also looked at the application of real-time failure time electronic components to lifetime distribution. The random samples were generated from the Exponential, Weibull, and lognormal distributions. The Weibull distribution can better suit the failure time electronic components in actual life, according to the results mentioned above. In various fields of study, including reliability analysis, medical, engineering, and economics, among others, we anticipate that the suggested distribution would attract more extensive applications.

References

[1] Esmaeili, L. and Niaparast, M. (2022). New compounding lifetime distributions with applications to real data. *International Journal of Modeling, Simulation, and Scientific Computing*, 13(05): 2250038.

[2] Aslam, M. Jeyadurga, P. Balamurali, S. Azam, M. and Al-Marshadi, A. (2021). Economic determination of modified multiple dependent state sampling plan under some lifetime distributions. *Journal of Mathematics*, 2021(1): 7470196.

[3] Shaheen, U. Azam, M. and Aslam, M. (2020). A control chart for monitoring the lognormal process variation using repetitive sampling. *Quality and Reliability Engineering International*, 36(3): 1028-1047.

[4] Shah, M. T. Azam, M. Aslam, M. and Sherazi, U. (2021). Time between events control charts for gamma distribution. *Quality and Reliability Engineering International*, 37(2): 785-803.

[5] Chacko, V. M. Deepthi, K. S. Thomas, B. and Rajitha, C. (2018). Weibull-Lindley Distribution: A Bathtub Shaped Failure Rate Model. *Reliability: Theory and Applications*, 13(4): 9-20.

[6] Mahdavi, A. and Kundu, D. (2017). A new method for generating distributions with an application to exponential distribution. *Communications in Statistics-Theory and Methods*, 46(13): 6543-6557.

[7] Owoloko, E. A. Oguntunde, P. E. and Adejumo, A. O. (2015). Performance rating of the transmuted exponential distribution: an analytical approach. *SpringerPlus*, 4: 1-15.

[8] Dolas, D. R. Jaybhaye, M. D. and Deshmukh, S. D. (2014). Estimation the system reliability using Weibull distribution. *International Proceedings of Economics Development and Research*, 75(29): 144-148.

[9] Chahkandi, M. and Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Computational Statistics and Data Analysis*, 53(12): 4433-4440.

[10] Barreto-Souza, W. Santos, A. H. and Cordeiro, G. M. (2010). The beta generalized exponential distribution. *Journal of statistical Computation and Simulation*, 80(2): 159-172.

[11] Herbert, G. J. Iniyan, S. and Goic, R. (2010). Performance, reliability and failure analysis of wind farm in a developing country. *Renewable energy*, 35(12): 2739-2751.

[12] Agete, A. Ayalew, M. M. Admassu, S. and Dessie, Z. G. (2024). Prevalence and associated factors of teenage childbearing among Ethiopian women using semi-parametric and parametric proportional hazard and accelerated failure time models. *BMC Women's Health*, 24(1): 342.

[13] Yang, L. Cheng, J. Luo, Y. Zhou, T. Zhang, X. Shi, L. and Xu, Y. (2025). Domain adaptation via gamma, Weibull, and lognormal distributions for fault detection in chemical and energy processes. *The Canadian Journal of Chemical Engineering*, 103(1): 359-372.

[14] Lawless, J. F. Statistical Models and Methods for Lifetime Data. John Wiley and Sons, New York, 2003.

[15] Rausand, M. A. Hoyland, System Reliability Theory, John Wiley & Sons Inc. Publication, 2004.

COMPARATIVE STUDY OF REPAIRABLE JUICE PLANTS USING RPGT

Mohit Yadav[1], Vandana Swami[2], Naveen Kumar[3], Puneet Garg[4]

[1]Department of Mathematics, University Institute of Sciences, Chandigarh University, Mohali, Punjab
[2]SOET, Raffles University, Neemrana, Rajasthan
[3]Department of Applied Science and Humanities, Ganga Institute of Technology and Management, Kablana, Jhajjar
[4]Department of CSE-AI, KIET Group of Institutions, Delhi NCR, Ghaziabad

mohit.e15793@cumail.in, vandana.swami@rafflesuniversity.edu.in, ndeswal90@gmail.com, puneetgarg.er@gmail.com

Abstract

In this paper, comparative study of repairable juice plants is analyzed by using regenerative point graphical technique. Generally, juice contains calcium, vitamin, iron, etc. to give the refresh tests. There are multiple steps to store the juice at large levels such as storing, grinding, pasteurization, etc. In this paper, the performance and reliability measures of existing juice plant-I are discussed and compared with juice plane-II. Both juice plants have three distinct units. Unit A has washing and storage tank, unit B has grinding, blending, evaporation and pasteurization, and unit C has bottling, labeling and packing units. If any unit partially fails then the system works to a limited extent. A technician is always available to repair the failed unit. The system fails when one unit completely fails. In this paper, the failure time and repair time follow general distributions. The regenerative point graphical technique is used to explore the comparison of reliability measures.

Keywords: Base state, juice plant, availability, and profit values.

I. Introduction

Every industrialist wants to maximize their manufacturing processes to meet customer requests. This study uses the regenerating point graphical technique with semi Markov process to explore the mean time to system failure, availability and profit values of a juice plant with repair facility. Here, existing juice plant-I is compared with the juice plant-II discussed by Rahul and his coauthors. Balagurusamy [2] described the terms related to the system's meantime, failure, repair, redundancy, maintainability, availability and profit function. Wang et al. [19] analyzed the reliability and sensitivity of two unit system under repair conditions. Srinivasan and Subramanian [17] discussed on the reliability measures of warm standby system under repair facility. Teng et al. [18] discussed on the software and hardware system reliability and availability using stochastic process. Bhatti et al. [5] inspected the availability and profit values of the two unit system with two type of failure using discrete distribution.

Jyh-Bin et al. [6] threw light on the reliability parameters of repairable system using standby repair facility and reboot delay approach. Yuang and Meng [20] analyzed the reliability of two unit repairable redundant system under priority in use. The reliability analysis of a one-unit system with finite vacations was examined by Liu and Liu [13]. Kumar and Goel [12] evaluated the behaviour of two unit cold standby system under general distribution. Aggarwal et al. [1]

described the performance of serial process in refinery system of a sugar plant using regenerative point technique and semi-Markov process. Sadeghi and Roghanian [16] described the reliability parameters of a repairable system under imperfect switching. Barak et al. [3] evaluated the reliability, availability and profit values of repairable milk plant. Barak et al. [4] discussed on the role of refreshment in repairable two unit system. Kumar et al. [9] evaluated the performance of two unit cold standby system subject to software upgrade and load recovery facility. Kumar et al. [10] described the two unit computer system under hardware repair and software upgrade facility. Kumar et al. [11] analyzed the cold standby system subject to refreshment facility. Kumar and Sharma [7] analyzed the availability and profit function of two unit cold standby system using regenerative point technique. Kumar and Sharma [8] examined the reliability measures of two unit cold standby system subject to reboot delay facility. Reddy et al. [15] analyzed the reliability measures of utensil industry under repair facility. Rahul et al. [14] described the reliability measures of juice plants using regenerative point technique.

II. System Assumptions

To describe the juice plant-I, there are following assumptions

- The juice plant consists of three distinct units *A*, *B* and *C*.
- It is considered that units *A*, *B* and *C* may be in a complete failed state through partial failure mode.
- Unit *A* has washing and storage tank.
- Unit *B* has grinding, blending, evaporation and pasteurization.
- Unit *C* has bottling, labeling and packing units.
- Failure rate and repair rate are generally distributed and are independent.
- The repaired unit functions just like a brand-new one.

To describe the juice plant-II, all assumptions are same as juice plant-I with difference in

- It is considered that units *A* and *B* may be in a complete failed state through partial failure mode but unit *C* is in only partially failed state.

III. System Notations

To explain the juice plants, there are following notations

$i \xrightarrow{Sr} j$	r^{th} directed simple path from state '*i*' to state '*j*' where '*r*' takes the positive integral values for different directions from state '*i*' to state '*j*'.
$\xi \xrightarrow{s\,f\,f} i$	A directed simple failure free path from state ξ to state '*i*'.
$m-cycle$	A circuit (may be formed through regenerative or non regenerative / failed state) whose terminals are at the regenerative state '*m*'.
$m-\overline{cycle}$	A circuit (may be formed through the unfailed regenerative or non regenerative state) whose terminals are at the regenerative '*m*' state.
$U_{k,k}$	Probability factor of the state '*k*' reachable from the terminal state '*k*' of '*k*' cycle.
$U_{\overline{k,k}}$	The probability factor of state '*k*' reachable from terminal state '*k*' of $k\ \overline{cycle}$.
μ_i	Mean sojourn time spent in the state '*i*' before visiting any other states.
μ_i'	Total unconditional time spent before transiting to any other regenerative state while the system entered regenerative state '*i*' at t=0.
η_i	Expected waiting time spent while doing a job given that the system entered to the regenerative state '*i*' at t=0.
$A/\overline{A}/a$	The first unit is in the operative state/reduced state/failed state.

$B/\overline{B}/b$	The second unit is in the operative state/reduced state/failed state.
$C/\overline{C}/c$	The third unit is in the operative state/reduced state/failed state.
$\lambda_1, \lambda_2, \lambda_3$	Fixed partial failure rate of the unit A/B/C respectively.
$\lambda_4, \lambda_5, \lambda_6$	Fixed complete failure rate of the unit A/B/C respectively.
w_1, w_2, w_3	Fixed repair rate of the unit A/B/C after partial failure respectively.
w_4, w_5, w_6	Fixed repair rate of unit A/B/C after the complete failure respectively.
○ ⬭ ▭	Upstate/ reduced state/ failed state.

IV. Circuits Descriptions of Juice Plant-I

Primary, secondary and tertiary circuits are used to find the base state such that

Table 1: *Circuit Descriptions*

i	(C1)	(C2)	(C3)
0	(0,1,0), (0,2,0), (0,3,0) (0,1,4,0), (0,2,5,0), (0,3,6,0)	Nil	Nil
1	(1,0,1)	(0,2,0), (0,3,0)	Nil
2	(2,0,2)	(0,1,0), (0,3,0)	Nil
3	(3,0,3)	(0,1,0), (0,2,0)	Nil
4	(4,0,1,4)	(0,1,0), (0,2,0) (0,3,0), (1,0,1)	(2,0,2), (3,0,3)
5	(5,0,2,5)	(0,1,0), (0,2,0) (0,3,0), (2,0,2)	(1,0,1), (3,0,3)
6	(6,0,3,6)	(0,1,0), (0,2,0) (0,3,0), (3,0,3)	(1,0,1), (2,0,2)

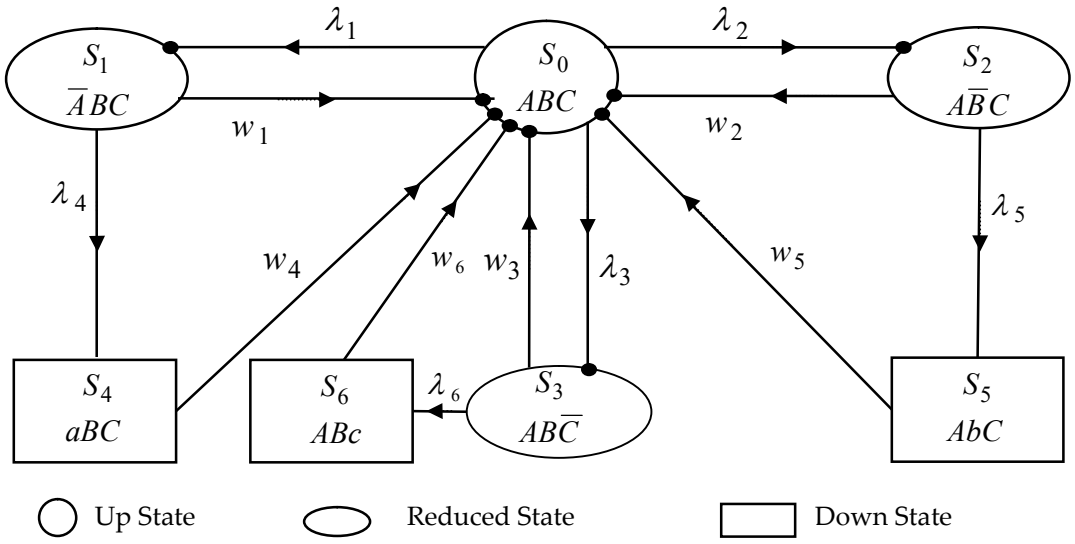


Figure 1 *State Transition Diagram of Juice Plant-I*

where, $S_0 = ABC$, $S_1 = \overline{A}BC$, $S_2 = A\overline{B}C$, $S_3 = AB\overline{C}$, $S_4 = aBC$, $S_5 = AbC$, $S_6 = ABc$

V. Transition Probabilities

There are following transition probabilities

$p_{0,1} = \lambda_1/(\lambda_1+\lambda_2+\lambda_3)$, $p_{0,2} = \lambda_2/(\lambda_1+\lambda_2+\lambda_3)$,

$p_{0,3} = \lambda_3/(\lambda_1+\lambda_2+\lambda_3)$, $p_{1,0} = w_1/(w_1+\lambda_4)$,

$p_{2,0} = w_2/(w_2+\lambda_5)$, $p_{2,5} = \lambda_5/(w_2+\lambda_5)$,

$p_{3,0} = w_3/(w_3+\lambda_6)$, $p_{3,6} = \lambda_6/(w_3+\lambda_6)$

$p_{4,0} = p_{5,0} = p_{6,0} = 1$ (1)

It has been conclusively established that

$p_{0,1} + p_{0,2} + p_{0,3} = 1$, $p_{1,0} + p_{1,4} = 1$, $p_{2,0} + p_{2,5} = 1$, $p_{3,0} + p_{3,6} = 1$

VI. Mean Sojourn Time

For the particular state, it becomes

$\mu_0 = 1/(\lambda_1+\lambda_2+\lambda_3)$, $\mu_1 = 1/(w_1+\lambda_4)$,

$\mu_2 = 1/(w_2+\lambda_5)$, $\mu_3 = 1/(w_3+\lambda_6)$,

$\mu_4(t) = 1/(w_4)$, $\mu_5 = 1/(w_5)$, $\mu_6 = 1/(w_6)$ (2)

VII. Evaluation of Parameters

Using the circuit table, '0' is used as the base state to calculate the reliability using the regenerative point graphical technique. The probability factors of all the reachable states from the base state '0' are given below

$U_{0,0} = (0,1,0) + (0,2,0) + (0,3,0) = 1$, $U_{0,1} = \dfrac{\lambda_1}{\lambda_1+\lambda_2+\lambda_3}$, $U_{0,2} = \dfrac{\lambda_2}{\lambda_1+\lambda_2++\lambda_3}$

$U_{0,3} = \dfrac{\lambda_3}{(\lambda_1+\lambda_2+\lambda_3)}$, $U_{0,4} = \dfrac{\lambda_1\lambda_4}{(\lambda_1+\lambda_2+\lambda_3)(w_1+\lambda_4)}$

$U_{0,5} = \dfrac{\lambda_2\lambda_5}{(\lambda_1+\lambda_2+\lambda_3)(w_2+\lambda_5)}$, $U_{0,6} = \dfrac{\lambda_3\lambda_6}{(\lambda_1+\lambda_2+\lambda_3)(w_3+\lambda_6)}$

I. Mean Time to System Failure

The regenerative un-failed states (i=0, 1, 2, 3) to which the system can transit (with initial state 0) before entering to any failed state (using base state ξ=0) then MTSF becomes

$$T_0 = \left[\sum_{i=0}^{3} Sr \left\{ \frac{\left\{ pr\,(0 \xrightarrow{Sr\,(sff)} i) \right\} . \mu_i}{\prod_{k_1 \neq 0} \left\{ 1 - V_{k_1 k_1} \right\}} \right\} \right] \div \left[1 - \sum Sr \left\{ \frac{\left\{ pr\,(0 \xrightarrow{Sr\,(sff)} 0) \right\}}{\prod_{k_2 \neq 0} \left\{ 1 - V_{k_2 k_2} \right\}} \right\} \right]$$

$$T_0 = \frac{U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3}{[1-(1,0,1)]} \quad (3)$$

II. Availability of the system

The system is available for use at regenerative states j=0, 1, 2, 3 with ξ=0 then the availability of system is defined as

$$A_0 = \left[\sum_{j=0}^{3} Sr \left\{ \frac{\left\{ pr\,(0 \xrightarrow{Sr} j) \right\}.f_j.\mu_j}{\prod_{k_1 \neq 0} \left\{ 1 - V_{k_1 k_1} \right\}} \right\} \right] \div \left[\sum_{i=0}^{6} Sr \left\{ \frac{\left\{ pr\,(0 \xrightarrow{Sr} i) \right\}.\mu_i'}{\prod_{k_2 \neq 0} \left\{ 1 - V_{k_2 k_2} \right\}} \right\} \right]$$

$$A_0 = \frac{U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3}{U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3 + U_{0,4}\mu_4 + U_{0,5}\mu_5 + U_{0,6}\mu_6} \tag{4}$$

III. Busy Period of the Technician

The Technician is busy due to repair of the failed unit at regenerative states *j*= 1, 2, 3, 4, 5, 6 with ξ = 0 then the fraction of time for which the server remains busy is defined as

$$B_0 = \left[\sum_{j=1}^{6} Sr \left\{ \frac{\left\{ pr\,(0 \xrightarrow{Sr} j) \right\}.\eta_j}{\prod_{k_1 \neq 0} \left\{ 1 - V_{k_1 k_1} \right\}} \right\} \right] \div \left[\sum_{i=0}^{6} Sr \left\{ \frac{\left\{ pr\,(0 \xrightarrow{Sr} i) \right\}.\mu_i'}{\prod_{k_2 \neq 0} \left\{ 1 - V_{k_2 k_2} \right\}} \right\} \right]$$

$$B_0 = \frac{U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3 + U_{0,4}\mu_4 + U_{0,5}\mu_5 + U_{0,6}\mu_6}{U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3 + U_{0,4}\mu_4 + U_{0,5}\mu_5 + U_{0,6}\mu_6} \tag{5}$$

IV. Estimated number of visits made by the Technician

The technician visits at regenerative states *j*= 1, 2, 3 with ξ=0 then the number of visits by the repairman is defined as

$$V_0 = \left[\sum_{j=1}^{3} Sr \left\{ \frac{\left\{ pr\,(0 \xrightarrow{Sr} j) \right\}}{\prod_{k_1 \neq 0} \left\{ 1 - V_{k_1 k_1} \right\}} \right\} \right] \div \left[\sum_{i=0}^{6} Sr \left\{ \frac{\left\{ pr\,(0 \xrightarrow{Sr} i) \right\}.\mu_i'}{\prod_{k_2 \neq 0} \left\{ 1 - V_{k_2 k_2} \right\}} \right\} \right]$$

$$V_0 = \frac{U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3}{U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3 + U_{0,4}\mu_4 + U_{0,5}\mu_5 + U_{0,6}\mu_6} \tag{6}$$

V. Profit Analysis

The profit values of the juice plant system depend on the system reliability measures such that availability of the system, busy period of the technician, and expected number of visits made by the technician. An expert technician is always available to perform the repair work.

The following profit function may be used to do a profit analysis of the system and it is calculated by

$$P = E_0 A_0 - E_1 B_0 - E_2 V_0 \tag{7}$$

where, $E_0 = 25000$ (Revenue per unit uptime of the system)

$E_1 = 500$ (Cost per unit time for which technician is busy due to repair)

$E_2 = 200$ (Cost per visit of the technician)

VIII. Discussion

Generally, juice plant fulfills the daily requirement of juice as beverage. In the given paper, there are discussed on reliability measures such that tables 2, 3 and 4 describe the nature of mean time to

system failure, availability and profit values of the system under repair facility. It is analyzed that when repair rate enhance from 0.4 to 0.85 then mean time to system failure also enhanced. When failure rate λ_1 change from 0.3 to 0.4, λ_2 changes from 0.4 to 0.5, and λ_3 changes from 0.25 to 0.3 under other parameters as constant then MTSF value decline. It is observed that when repair rate increase (0.4-0.85) then availability values also increased but when λ_1 increase (0.3 to 0.4), λ_2 increase (0.4 to 0.5), and λ_3 increase (0.25 to 0.3) under other parameters as constant then availability values decline.

It is also analyzed that profit values of system are increased with respect to increment in repair rate (0.4-0.85) but decline when failure rate λ_1(0.3 to 0.4), λ_2(0.4 to 0.5), and λ_3(0.25 to 0.3) enhanced under other parameters as constant.

Table 2: *MTSF vs. Repair Rate* (w)

w_2 ↓	λ_1=0.3, λ_2=0.4, λ_3=0.25 λ_4=0.35,λ_5=0.5, λ_6=0.55 w_1=0.4, w_3=0.5, w_4=0.5 w_5=0.6, w_6=0.65	λ_1=0.4	λ_2=0.5	λ_3=0.3
0.4	4.76082	4.413408	4.192308	3.96967
0.45	4.929006	4.585448	4.290541	4.045207
0.5	5.063985	4.728682	4.36747	4.103806
0.55	5.174709	4.849785	4.429348	4.150591
0.6	5.267176	4.953519	4.480198	4.188807
0.65	5.345557	5.043371	4.522727	4.220611
0.7	5.412844	5.121951	4.558824	4.247492
0.75	5.471236	5.191257	4.589844	4.27051
0.8	5.522388	5.252838	4.616788	4.290443
0.85	5.567568	5.307918	4.640411	4.307871

Table 3: *Availability vs. Repair Rate* (w)

w_2 ↓	λ_1=0.3, λ_2=0.4 λ_3=0.25, λ_4=0.35 λ_5=0.5, λ_6=0.55 w_1=0.4 w_3=0.5, w_4=0.5 w_5=0.6, w_6=0.65	λ_1=0.4	λ_2=0.5	λ_2=0.3
0.4	0.686403	0.670466	0.657986	0.648775
0.45	0.69355	0.678566	0.662905	0.638243
0.5	0.699041	0.684997	0.666652	0.630578
0.55	0.703393	0.690227	0.669601	0.624746
0.6	0.706926	0.694563	0.671982	0.620162
0.65	0.709852	0.698216	0.673945	0.616462
0.7	0.712315	0.701336	0.675592	0.613414
0.75	0.714416	0.704032	0.676992	0.610865
0.8	0.716231	0.706384	0.678198	0.608687
0.85	0.717813	0.708454	0.679247	0.606818

Table 4: *Profit vs. Repair Rate* (w)

w_2 ↓	λ_1=0.3, λ_2=0.4 λ_3=0.25, λ_4=0.35 λ_5=0.5, λ_6=0.55 w_1=0.4 w_3=0.5, w_4=0.5 w_5=0.6, w_6=0.65	λ_1=0.4	λ_2=0.5	λ_2=0.3
0.5	2722.902	2572.193	2533.318	2459.635
0.55	2788.908	2650.579	2583.419	2428.449
0.6	2838.719	2711.387	2620.861	2405.743
0.65	2877.66	2759.963	2649.913	2388.471
0.7	2908.947	2799.676	2673.117	2374.89
0.75	2934.638	2832.757	2692.08	2363.932
0.8	2956.115	2860.746	2707.868	2354.904
0.85	2974.335	2884.737	2721.218	2347.337
0.9	2989.99	2905.531	2732.655	2340.903
0.95	3003.585	2923.73	2742.563	2335.365

IX. Comparative study

Comparative study of juice plants described that mean time to system failure, availability of the system, and profit values of juice plant-I are greater than corresponding values of juice plant-II.
In juice plant-I, mean time to system failure lowest value is 3.96967 and highest value is 5.567568 but in juice plant-II, its lowest value is3.037974 and highest value is 4.019608.

In juice plant-I, availability lie between (0.648775 - 0.717813) but in juice plant-II, it lie between (0.58624-0.663876). The profit values of juice plant-I exists from 2459.63 to 3003.585 but in juice plant –II, it exists from 2333.81 to 2673.74.

X. Conclusion

The performance of the juice plant is discussed using the regenerative point graphical technique. The above tables explore that when the repair rate increases then the MTSF, availability and profit values also increase but when the failure rate increases then the MTSF, availability and profit values decrease. It is clear that RPGT is helpful for industries to analyze the behaviour of the products and components of a system.

It is observed that the role of the regenerative point graphical technique for the juice plant will be beneficial and also used by the management, manufacturers and the persons engaged in reliability engineering and working on analyzing the nature and performance analysis of the system like soft drink, paper industry, etc.

References

[1] Aggarwal, A. K., Kumar, S., and Singh, V. (2017). Performance modeling of the serial processes in refining system of a sugar plant using RAMD analysis. *International Journal of System Assurance Engineering and Management*, *8*(Suppl 2), 1910-1922.
[2] Balagurusamy, E. (1984). *Reliability Engineering*. Tata McGraw-Hill Education.

[3] Barak, M. S., Garg, R., and Kumar, A. (2021). Reliability measures analysis of a milk plant using RPGT. *Life Cycle Reliability and Safety Engineering*, *10*(3), 295-302.

[4] Barak, M. S., Kumar, A., and Garg, R. (2023). Contribution of Refreshment Provided to the Server During His Job in the Repairable Cold Standby System. *Computational Intelligence in Sustainable Reliability Engineering*, 251-267.

[5] Bhatti, J., Chitkara, A., and Bhardwaj, N. (2011). Profit analysis of two unit cold standby system with two types of failure under inspection policy and discrete distribution. *International Journal of science and engineering research*, *2*(2), 1-7.

[6] Jyh-Bin, K., Jyh-Wei, C., and Kuo-Hsiung, W. (2011). Reliability measures of a repairable system with standby switching failures and reboot delay. *Quality Technology and Quantitative Management*, *8*(1), 15-26.

[7] Kumar, A., & Sharma, A. (2023). Profit analysis of repairable cold standby system under refreshments. *Reliability Theory and Applications*, *18*(4 (76)), 604-612.

[8] Kumar, A., & Sharma, A. (2024). Profit analysis of repairable cold standby system subject to reboot facility under. *Reliability Theory and Applications*, *19*(2 (78)), 594-603.

[9] Kumar, A., Garg, R., & Barak, M. S. (2023). Performance analysis of computer systems with Weibull distribution subject to software upgrade and load recovery. *Life Cycle Reliability and Safety Engineering*, *12*(1), 51-63.

[10] Kumar, A., Garg, R., & Barak, M. S. (2023). Performance of computer system with three types of failure using weibull distribution subject to hardware repair and software up-gradation. *International Journal of System Assurance Engineering and Management*, *14*(Suppl 1), 483-491.

[11] Kumar, A., Garg, R., & Barak, M. S. (2023). Reliability measures of a cold standby system subject to refreshment. *International Journal of System Assurance Engineering and Management*, *14*(1), 147-155.

[12] Kumar, J. and Goel, M. (2016). Availability and profit analysis of a two-unit cold standby system for general distribution. *Cogent Mathematics*, 3(1): 1262937.

[13] Liu, R. and Liu, Z. (2011). Reliability analysis of a one-unit system with finite vacations. In *MSIE,* 248-252.

[14] Rahul, Yadav, M., and Kumar, H. (2024). Profit analysis of repairable juice plant. *Reliability Theory and Applications*, *19*(1 (77)), 688-695.

[15] Reddy, H. I., Yadav, M., and Kumar, H. (2024). Stochastic analysis of utensil industry subject to repair facility. *Reliability Theory and Applications*, *19*(2 (78)), 170-177.

[16] Sadeghi, M., and Roghanian, E. (2017). Reliability analysis of a warm standby repairable system with two cases of imperfect switching mechanism. *Scientia Iranica*, *24*(2), 808-822.

[17] Srinivasan, S. K., and Subramanian, R. (2006). Reliability analysis of a three unit warm standby redundant system with repair. *Annals of Operations Research*, *143*(1), 227-235.

[18] Teng, X., Pham, H., and Jeske, D. R. (2006). Reliability modeling of hardware and software interactions, and its applications. *IEEE Transactions on Reliability*, *55*(4), 571-577.

[19] Wang, K. H., Lai, Y. J., and Ke, J. B. (2004). Reliability and sensitivity analysis of a system with warm standbys and a repairable service station. *International Journal of Operations Research*, *1*(1), 61-70.

[20] Yuan, L., and Meng, X. Y. (2011). Reliability analysis of a warm standby repairable system with priority in use. *Applied Mathematical Modelling*, *35*(9), 4295-4303.

CLUSTERING ELECTRIC POWER SYSTEMS INTO ADEQUACY ASSESSMENT AREAS

Dmitry Krupenev[1], Denis Boyarkin[1], Nikolay Belyaev[2]

•

[1] Melentiev Energy Systems Institute, Russia.
[2] JSC «STC UPS», Russia.
krupenev@isem.irk.ru
denisboyarkin@isem.irk.ru
belyaev.na@yandex.ru

Abstract

The article discusses the issue of clustering electric power systems (EPS) into adequacy assessment areas to form equivalent models (EM) of power systems, with such models designed to calculate adequacy metrics. The level of detail of equivalent models is backed by the consideration of two properties of power systems being reliability and economic feasibility. To solve the problem posed above, the paper proposes to use a number of engineering approaches that take into account reliability and economic feasibility of power systems. The engineering approaches are supplemented with analytical methods. We present an analysis of analytical clustering methods as applied to the task of forming power system EMs. We review well-established clustering methods, such as k-means and the shortest path clustering, as well as methods of community detection in graphs by specified features, which prove the most suitable for solving the problem. Based on our analysis, we have found that for clustering of power system structure graphs, the most plausible result can be obtained by applying the Leiden method. The Experimental Study section presents the results of applying the methods we analyze here to form an EM of the test electric power system.

Keywords: electric power system, adequacy, assessment area, equivalent model, clustering, Leiden algorithm

I. Introduction

Electric power systems (EPS) are complex structures consisting of hundreds of thousands of different elements [1]. To analyze the balance reliability of such systems [2-4], as a rule, a technique based on the Monte Carlo method [5] is used, which is the most computationally efficient compared to others [6]. In adequacy assessment, due to the high dimensionality of the analyzed EPS and the resulting computational and analytical difficulties, EPS is clustered into reliability zones (RZ) and energy calculation models (ECM) are formed. In the practice of solving system analysis problems, various clustering methods have been developed [7,8], but there are no formalized clustering methods for adequacy assessment. When clustering for certain parts of the EPS, network restrictions that have little effect on the adequacy indicators are excluded. ECM formation is usually carried out before adequacy assessment, and the formed ECM is accepted as unchanged for the entire adequacy assessment period, but if necessary and computationally possible, EPS clustering into reliability zones can be carried out multiple times in a adequacy assessment cycle. The main reason for multiple clustering may be significant changes in the EPS parameters during the calculation period and, accordingly, the impact of these parameters on the

grid's capacity to transmit power. This may also be due to the fact that modern EPS have a complex structure with a wide variety of power equipment, in addition, EPS tend to enlarge and merge, which further complicates the structure and uncertainty in possible power flows.

In the practice of adequacy assessment, there are various approaches to EPS clustering. They are mainly based either on expert knowledge or on a simplified analysis of the balance situation in the EPS. The purpose of this article is to present a methodology for EPS clustering on the RZ, which has a mathematical formalization. The substantive problem to be solved can be formulated as follows: for a known EPS structure, characteristics of generating capacities and power consumption schedules, limitations on the capacity of transmission lines and electrical grid sections, organizational and economic conditions for the functioning of the EPS, it is necessary to determine the RZ boundaries for the formation of the ECM and further adequacy assessment.

II. Statement of the problem of clustering of EPS and formation of ECM based on methods of detecting communities in graphs

For the most efficient (fast and accurate) solution to the problem of clustering the EPS on the ZN and forming the ECM, it is necessary to develop an approach that will most fully reflect the influencing criteria and restrictions on the level of the EPS adequacy and will allow for the sufficiently fast and adaptive clustering of the EPS in the process of conducting research.

When building an algorithm of power system clustering into assessment areas, we need to obtain the function a:X→Z that assigns each node of the power system ($x \in X$) into an assessment area $z \in Z$. In the considered case, the set Z, i.e., the number of assessment areas, is unknown. The optimal number of assessment areas needs to be determined as part of the calculation process. It is worth noting that, as is always the case with clustering, Kleinberg's impossibility theorem holds for the problem at hand, which states that there is no optimal clustering algorithm. Each algorithm applied will introduce its own bias into the final solution. This will depend on the selected criterion and clustering metric.

The solution to the problem of clustering the EPS will be its division, satisfying the given optimality criterion. In the case under consideration, the optimality criterion is the minimization of the number of reliability zones:

$$\mathrm{Z} \underset{\rho}{\rightarrow} min. \tag{1}$$

The desire to minimize the number of reliability zones is natural, since this will lead to a decrease in the dimension of the ERM in the further analysis of balance reliability.

A limitation in solving the problem will be a limitation specified for the metric, which is formed from various features. In the case under consideration, the clustering metric should include both technical and economic features. Technical features include: installed capacity of generating equipment in EPS nodes; average value of the number of hours of use of installed capacity in EPS nodes; seasonal maximums and minimums of power consumption in EPS nodes; accident rate of power equipment; standards for scheduled repairs of power equipment; lengths of power transmission lines (PTL) between EPS nodes; transmission capacities of PTLs [9]. Economic features include: cost of commissioning generating and network facilities. The result of clustering will ultimately depend on the choice of metric. The metric should maximally reflect the specifics of the EPS operation and the influence of the taken into account features on clustering. One of the following metrics can be used as a measure of distance (degree of similarity): Euclidean distance or its square, Manhattan distance, Chebyshev distance, power distance, etc.

III. Overview of clustering algorithms for systems

As it stands now, quite a large number of algorithms of dataset clustering have been developed [10-12], each of which has its own unique features, advantages and shortcomings. The methods and adaptations of methods that allow working in a network of interconnected components, in other words graphs, stand out from the rest. The presence of relationships between the objects being clustered significantly affects the obtained result and imposes certain constraints. There are many graph-based clustering methods that can be used as part of network analysis. Some of them are modifications of established clustering methods such as k-means and spectral clustering and have adaptations that make them work with graphs.

Classification of clustering methods can be done based on the way the graph partition boundary is detected. The following classes of these methods can be distinguished:

1. Similarity-based methods. These methods detect clusters based on the degree of similarity between graph vertices. Such methods may use different similarity metrics, such as cosine similarity or Euclidean distance.

2. Distance-based methods. These methods detect clusters based on the distance between vertices of the graph. We may use different distance metrics for different methods for example Manhattan distance or Euclidean distance

3. Model-based methods. These methods detect clusters based on a probabilistic graph model. Such methods can be used to cluster graphs with a large number of vertices and/or edges.

4. Graph structure-based methods. These methods detect clusters based on the graph structure such as graph density, vertex centrality, etc.

Eeach class of the methods presented above has its own advantages and downsides, and the choice of a particular algorithm depends on the problem to be solved and the properties of the graph to be clustered.

Also, clustering methods can be classified on the basis of the approaches underpinning the clustering algorithm. Some of the most common methods for clustering on graphs include:

1. Similarity-based methods. These methods detect clusters based on the degree of similarity between graph vertices. Such methods may use different similarity metrics, such as cosine similarity or Euclidean distance.

1. The class of partitioning methods.

1.1. The k-means algorithm [11] is one of the most popular clustering methods belonging to the class of partitioning methods. It is based on dividing the vertices of the graph into a given number of clusters, so as to minimize the sum of squares of distances between vertices and centroids of clusters. The k-means algorithm can be applied when vertex properties are numeric or when it is possible to convert them to numeric.

1.2. The Shortest Path Clustering [12] builds clusters based on the distances between vertices of the graph. This method also belongs to the class of partitioning methods.

2. Hierarchical clustering methods [13] are a class of methods that enable creation of a hierarchical structure of clusters. These methods can be used to cluster graph vertices based on their properties by hierarchically grouping nearby vertices. The general idea behind the methods of this group is a sequential hierarchical decomposition of a set of objects. Depending on the direction of hierarchy construction the methods fall into two categories: divisive and agglomerative clustering. In the case of the agglomerative approach, the decomposition process starts with each object being an independent cluster. Then, at each iteration, pairs of nearby clusters are successively grouped into a common cluster. Iterations continue until all objects are grouped into a single cluster or until some stopping condition is met. Also, we have the divisive approach. The approach is contained all objects into a single cluster in the initial stage. Further

from single cluster the method form to smaller clusters while each object is in a separate cluster or the stopping condition is met. In what follows, we will review some of them.

2.1. The Louvain method [14] has graph vertices gradually clustered together to maximize connectivity inside clusters and minimize connectivity between clusters.

2.2. Spectral Clustering [15] is based on analysing the eigenvalues of the graph's adjacency matrix and allows partitioning the graph into clusters based on its structural information.

2.3. The Edge Betweenness Clustering [16] uses a measure of connectivity between clusters (between center vertices) to partition the graph into clusters.

3. Density-based clustering methods are based on analysing the density distribution of graph nodes in the feature space. They can be particularly useful if vertices of the graph are not numeric.

4. Network analysis methods [17] can also be used for clustering of graphs. Network methods are an approach to data analysis that are based on representing data as networks or graphs, which has an identical structure. The main direction the methods consist that the object space is separate into a cells forming finite number a network structure. All clustering operations are performed in the structure. The main advantage of methods of this group is the short execution time, which usually does not depend on the number of data objects.

4.1. The CLIQUE algorithm [18], adapted to clustering of high dimensional data, is one of the classical network algorithms. The method is based on the assumption that if in a multidimensional data space, the distribution of objects is not uniform (there are dense and sparse regions), then the projection of the dense region into a subspace of lower dimensionality will be a part of the dense region in this subspace.

4.2. Infomap method [19]. The method relies on information theory to determine the most likely communities in a graph.

We will focus specifically on the modularity-based algorithm [20], which maximizes the difference between the actual number of edges in a set and the expected number of such edges; various modifications and improvements of modularity based on heuristic methods include the annealing method [21], spectral algorithms [15], etc. [22]. One of the most commonly used algorithms for community detection in graphs is the Louvain algorithm [14]. As a result of our analysis of clustering methods, it was found that for the problem to be solved, which is clustering of power system into adequacy assessment areas, the most suitable is the Leiden algorithm [23].

IV. Clustering of power system into assessment areas based on the Leiden algorithm

The Leiden algorithm is an algorithm for community detection in large networks. The Leiden method is an iterative algorithm that works as follows (Fig. 1 visualizes the operation of the Leiden algorithm):

1. Initially, each node is considered as a separate cluster.
2. The algorithm sequentially aggregates nearby clusters by iteratively optimizing the clustering quality score.
3. The clustering quality score in the Leiden method is based on the distribution of objects between clusters and shows how similar objects within a cluster are to each other and how different are objects from different clusters.
4. To aggregate clusters, the Leiden method uses the so-called "modularity gain", which shows how much the clustering quality will improve when two clusters are aggregated.
5. Clusters are aggregated as long as further aggregation does not lead to degradation of the clustering quality score.
6. After all clusters are aggregated, the result of the algorithm is the final partitioning of objects into clusters.

The Leiden method is greedy in that it selects a locally optimal solution at each step. It is also relatively fast and performs well when clustering large graphs and networks.

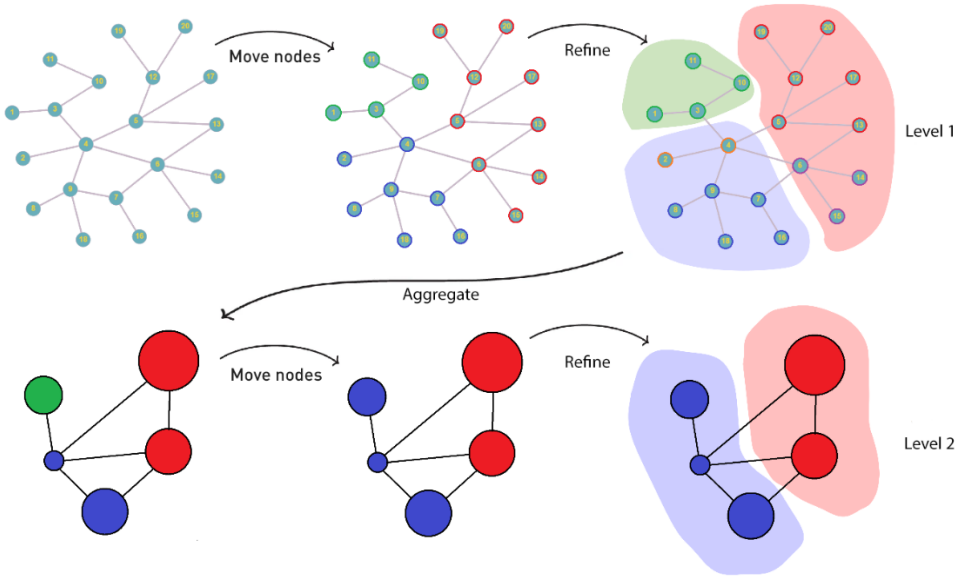


Figure 1: *A visualization of the operation of the Leiden algorithm.*

More formally speaking, the algorithm divides nodes into non-overlapping clusters to maximize the modularity score for each cluster. Modularity quantifies the quality of node assignment to clusters, i.e., how tightly connected the nodes in a cluster are compared to how connected they would be in a random network. The Leiden algorithm performs hierarchical clustering by greedily optimizing modularity, and the process is repeated in the compressed graph. The Leiden algorithm consists of three main steps:

1) Local moving of nodes into clusters based on the modularity score (metric), which is defined as:

$$Q = \frac{1}{2m}\sum_{ij}(A_{ij} - \frac{k_i k_j}{2m})\delta(c_i, c_j)\,; \tag{2}$$

where: A_{ij} - edge weight formed according to the features used; k_i and k_j - sum of weights of edges attached to nodes i and j; m - sum of all edge weights in the graph; c_i and c_j - clusters; δ - the Kronecker delta.

$$\delta(c_i, c_j) = \begin{cases} 1, c_i = c_j \\ 0, c_i \neq c_j \end{cases}. \tag{3}$$

The following expression is used to determine the modularity of the cluster c:

$$Q_c = \frac{\sum in}{2m} - (\frac{\sum tot}{2m})^2, \tag{4}$$

where: $\sum in$ - the sum of edge weights between nodes within the cluster c (each edge is counted twice); $\sum tot$ - the sum of all edge weights for nodes within the cluster.

In the iterative process, nodes are moved from one cluster to another. For each option of moving, the difference of modularities ΔQ is calculated:

$$\delta(c_i, c_j) = \begin{cases} 1, c_i = c_j \\ 0, c_i \neq c_j \end{cases}. \quad (5)$$

Based on the best result, the corresponding node is assigned to the corresponding cluster.

2) Cluster refinement, namely the identification of the clusters proposed at the first step. The clusters proposed at the first step can be divided into several clusters. The refinement phase does not follow a greedy approach and can aggregate a node with a randomly selected cluster, which increases the quality (modularity) score. This randomness allows the cluster space to be opened up more widely.

3) Aggregation and repetition of steps 1 and 2 until the quality of moving and aggregation of nodes cannot be improved.

V. Experimental studies

To evaluate the performance of the Leiden algorithm, experimental studies were performed on a test circuit of the Unified Power System (UPS) of Siberia, which includes the Altai, Buryat, Irkutsk, Krasnoyarsk, Kuzbass, Novosibirsk, Omsk, Tomsk, Khakassia and Transbaikal energy systems. For subsequent clustering, the UPS of Siberia was presented as an undirected graph, in which each of the vertices is characterized by maximum and minimum generation and load, the lines are characterized by length and voltage. The initial data for clustering the UPS of Siberia were adopted in accordance with [24]. Initial circuit of the UPS of Siberia during clustering consisted of 540 generating and loading nodes and 1025 power transmission lines.

Figure 2 shows the result of the Leiden clustering method (algorithm) in the form of a graphical representation of the division of the nodes of the Siberian UPS into reliability zones.

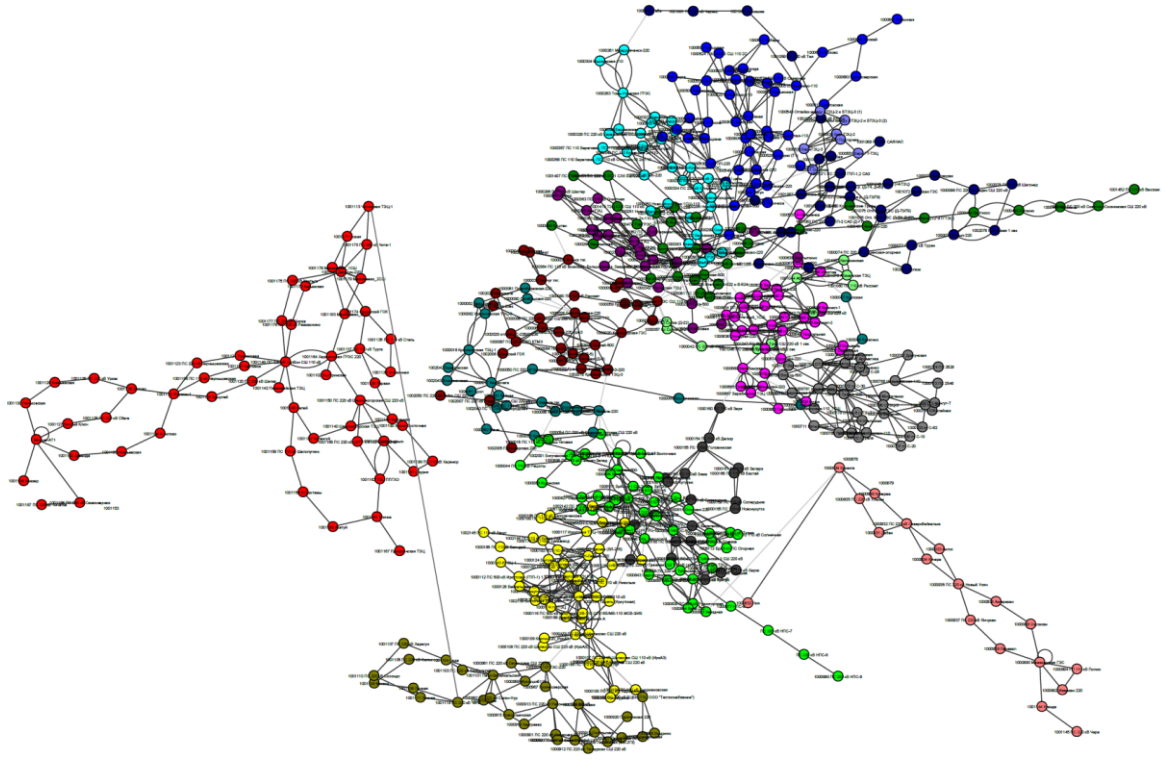


Figure 2: *Graphical representation of assignment of power system nodes to assessment areas based on the application of the Leiden algorithm.*

Figure 3 presents the same clustering but mapped to the power system circuit.

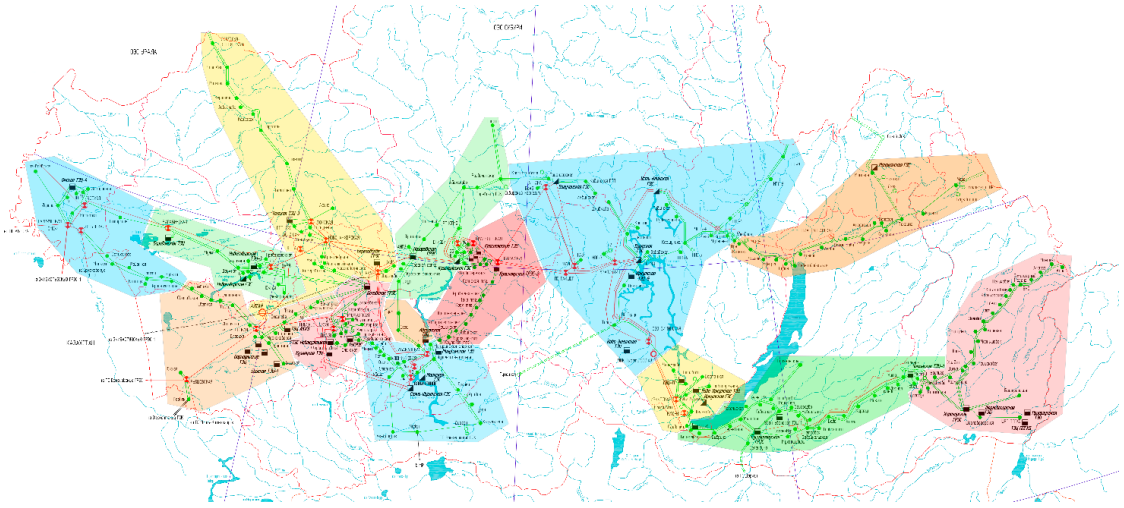


Figure 3: *The result of clustering the Siberian UPS into reliability zones based on the application of the Leiden algorithm (geographical referencing).*

As a result of clustering using the Leiden method, 14 reliability zones were obtained. The time spent on clustering was 1.16 s, which is sufficient for clustering in the process of assessing adequacy with the preparatory work with the initial data.

Analysis of the obtained clusters shows that the algorithm performed clustering by the weakest points of the EPS in terms of the density of interrelations and their quality. Therefore, the obtained reliability zones turned out to be maximally independent with the smallest number of connections between each other. Such a division in the subsequent assessment of adequacy will well determine the workload of connections, which were already identified at the clustering stage as bottlenecks.

VI. Conclusion

The presented study considers the problem of forming energy calculation models that are used in electric power systems adequacy assessment. An analysis of clustering methods on graphs was carried out and it was determined that one of the most suitable methods for clustering EPS into reliability zones is the Leiden algorithm.

Based on the results of the experimental calculations, it can be noted that in comparison with known engineering approaches to the formation of energy calculation models, the proposed approach gives a close result. Figure 3 shows that, based on the results of forming reliability zones, all the main restrictions on power transmission in the UPS of Siberia were taken into account. The boundaries of the formed reliability zones correlate with the corresponding controlled sections in the electric network, despite the fact that the controlled sections were not taken into account in any way in the proposed approach.

Thus, it can be concluded that with the help of the Leiden algorithm, it becomes possible to formalize the definition of reliability zones directly in the cycles of performing the corresponding calculations, which allows optimizing the solution of problems of analysis and synthesis of adequacy.

Funding

The research was supported by RSF (project No. 23-71-01090).

References

[1] Rudenko Y.N., Cheltsov M.B. Reliability and redundancy in electric power systems. "Nauka", Siberian branch, Novosibirsk. 1974.

[2] Kovalev G.F., Krupenev D.S., Lebedeva L.M. System reliability of the UES of Russia at the level of 2030. *Power stations*. 2011. No. 2. P. 44-47.

[3] Probabilistic Adequacy and Measures. Technical Reference Report Final, NERC, July, 2018.

[4] 2021 Long-Term Reliability Assessment. NERC. 2021.

[5] Kovalev G.F., Lebedeva L.M. Reliability of Power Systems. Springer-Verlag GmbH, 2019. DOI: 10.1007/978-3-030-18736-1.

[6] Krupenev D, Boyarkin D, Iakubovskii D. Improvement in the computational efficiency of a technique for assessing the reliability of electric power systems based on the Monte Carlo method. *Reliability Engineering & System Safety*. № 204. 2020. doi:10.1016/j.ress.2020.107171

[7] M. Rouhani, M. Mohammadi, M. Aiello. Soft clustering based probabilistic power flow with correlated inter temporal events. *Electric Power Systems Research*. № 204. 2022. 107677. https://doi.org/10.1016/j.epsr.2021.107677.

[8] Gheorghe G., Scarlatache F., Neagu B. Clustering in Power Systems. Applications. 2016.

[9] D.L. Faibisovich. Reference book on the design of electrical networks. Ed. 4th, revised. and additional M.: Publishing house of NTs ENAS. 2012.

[10] Mandel I. D. Cluster analysis. M.: Finance and statistics. 1988.

[11] MacQueen, J. Some methods for classification and analysis of multivariate observations. In Proc. 5th Berkeley Symp. On Math. Statistics and Probability, 1967.

[12] A. Sarkar, S. Ramalingam. Graph Clustering: A Review. 2020.

[13] Kaufman, L.; Rousseeuw, P.J. Finding Groups in Data: An Introduction to Cluster Analysis (1 ed.). New York: John Wiley, 1990.

[14] Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. J. Stat. Mech. Theory Exp. 10008. 6. 2008. https://doi.org/10.1088/1742-5468/2008/10/P10008.

[15] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. Phys. Rev. E 69. 026113. 2004.

[16] Ester, M., H. P. Kriegel, J. Sander, and X. Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, Portland, OR, AAAI Press. P. 226-231. 1996.

[17] F. Beligianni, G. Tsatsaronis, T. Palpanas. A Survey on Clustering Techniques for Graph Structured Data. 2019.

[18] Palla, G., Derényi, I., Farkas, I. et al. Uncovering the overlapping community structure of complex networks in nature and society. *Nature,* 435. P. 814–818. 2005. https://doi.org/10.1038/nature03607

[19] M. Rosvall, C. T. Bergstrom. Maps of information flow reveal community structure in complex networks. *PNAS*. 105. 1118. 2008.

[20] Orlov A.O., Chepovsky A.A. About modularity properties and actual adjustments of the blondel algorithm. *Vestnik NSU. Series: Information technologies*. №3. 2017.

[21] Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., and Teller E. Equation of State Calculations by Fast Computer Machines. *Chemical Physics,* 21. 6. June. 1953. P. 1087-1092.

[22] Lancichinetti, A., Fortunato, S. Community detection algorithms: A comparative analysis. *Phys. Rev.* E 80, 056117, 2009. https://doi.org/10.1103/PhysRevE.80.056117.

[23] Traag, V.A., Waltman, L., van Eck, N.J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* 9. 5233. 2019. https://doi.org/10.1038/s41598-019-41695-z.

[24] Order of the Ministry of Energy of Russia dated February 28, 2022 No. 146 "On approval of the scheme and program for the development of the Unified Energy System of Russia for 2022-2028".

ON CHARACTERIZATION OF 2KTH ORDER WEIGHTED MAXWELL BOLTZMANN DISTRIBUTION

[1*]Nuzhat Ahad , [1]S.P.Ahmad, [2]J.A.Reshi

•

[1]Department of Statistics, University of Kashmir, Srinagar, India.
[2]Department of Statistics, Govt. Degree College Pulwama , J&K, India.
[1*]nuzhatahad01@gmail.com , [1]sprvz@yahoo.com, [2] drjareshi@gmail.com

Abstract

Characterizing a probability distribution is essential in the fields of probability and statistics. This paper presents the characterization results of the 2Kth order weighted Maxwell Boltzmann distribution (KWMBD). The characterization results are based on the coefficient of variance, a straightforward relationship between two truncated moments, the conditional expectation of a specific function of the random variable, the hazard function, and the reverse hazard function. Notably, for these characterizations, the cumulative distribution function (CDF) does not need to have a closed form; instead, the CDF is derived from the solution of a first-order differential equation. This approach not only extends the applicability of the KWMBD but also establishes a crucial link between probability theory and differential equations. The resulting insights have significant implications for both theoretical research and practical applications in statistical modeling and analysis.

Keywords: Characterization, Truncated Moments, Hazard rate function, reverse hazard rate function.

I. Introduction

The 2Kth order weighted Maxwell Boltzmann distribution is a flexible continuous univariate probability distribution suitable for modeling datasets of decreasing increasing, increasing and constant behavior. The cumulative distribution function (CDF) and probability density function (PDF) of KWMBD is given by:

$$F(y,\alpha,k) = 1 - \frac{\Gamma\left(k+\frac{3}{2},\frac{y^2}{2\alpha^2}\right)}{\Gamma\left(k+\frac{3}{2}\right)} \qquad y>0,\alpha>0,k>-1.5. \tag{1}$$

$$f(y,\alpha,k) = \frac{y^{2(k+1)}\alpha^{-(3+2k)}e^{-\frac{y^2}{2\alpha^2}}}{2^{(k+\frac{1}{2})}\Gamma(k+\frac{3}{2})}, \qquad y>0,\alpha>0,k>-1.5. \tag{2}$$

The hazard function, reverse hazard function of KWMBD are respectively given by

$$\lambda(y,\alpha,k)=\frac{y^{2(k+1)}\alpha^{-(3+2k)}e^{-\frac{y^2}{2\alpha^2}}}{2^{\left(k+\frac{1}{2}\right)}\left\{\Gamma\left(k+\frac{3}{2},\frac{y^2}{2\alpha^2}\right)\right\}},\qquad y>0,\alpha>0,k>-1.5. \tag{3}$$

$$r(y,\propto,k)=\frac{y^{2(k+1)}\alpha^{-(3+2k)}e^{-\frac{y^2}{2\alpha^2}}}{2^{\left(k+\frac{1}{2}\right)}\left\{\Gamma\left(k+\frac{3}{2}\right)-\Gamma\left(k+\frac{3}{2},\frac{y^2}{2\alpha^2}\right)\right\}},\quad y>0,\alpha>0,k>-1.5. \tag{4}$$

The structural properties like mean, variance, and coefficient of variance are respectively given as

$$E(y)=\sqrt{2}\alpha\frac{\Gamma\left(\frac{2k+4}{2}\right)}{\Gamma\left(k+\frac{3}{2}\right)} \tag{5}$$

$$var(y)=\frac{2\propto^2}{\left(\Gamma\left(k+\frac{3}{2}\right)\right)^2}\left[\Gamma\left(\frac{2k+5}{2}\right)\Gamma\left(k+\frac{3}{2}\right)-\left(\Gamma\left(\frac{2k+4}{2}\right)\right)^2\right] \tag{6}$$

$$CV=\frac{\left[\Gamma\left(\frac{2k+5}{2}\right)\Gamma\left(k+\frac{3}{2}\right)-\left(\Gamma\left(\frac{2k+4}{2}\right)\right)^2\right]^{\frac{1}{2}}}{\Gamma\left(\frac{2k+4}{2}\right)} \tag{7}$$

II. Characterization

To develop a stochastic function for a specific problem, it is crucial to ensure that the function aligns with the theory of a particular underlying probability distribution. This necessitates the study of characterizations of the probability distribution. Characterization involves identifying certain distributional or statistical properties of a statistic that uniquely determine the associated probability distribution. Recently, the problem of characterizing a distribution has drawn considerable attention from researchers, leading to the development of various characterizations in multiple directions. For example, Glänzel [5] investigated characterizations of probability distributions by truncated moments, while Glänzel [6] delved into the consequences of such characterizations. Hamedani [7],[8] focused on characterizations of univariate continuous distributions. These characterizations can serve as a basis for parameter estimation of a probability distribution. Before applying a particular probability distribution model to fit real-world data, it is essential to confirm that the distribution meets the necessary requirements through its characterization.

Many scholars have used characterization theorems to describe both discrete and continuous probability distributions and to highlight the importance of characterizations in verifying the applicability and accuracy of probability distribution models in representing real-world data. Prominent examples include the characterization of the uniform distribution by Arslan et al. [1], the characterization of an exponential power life-testing distribution by Shakil et al. [10], the characterization of the Maxwell distribution based on truncated moments by Saghir et al. [11], the characterization and estimation of the Maxwell distribution by Dar et al. [4], the characterization of

the new Weibull-X family of distributions by Ahmad et al. [2], and the characterization of the exponential distribution by Bakouch et al. [3].

In this article, KWMBD is characterized through coefficient of variance theorem, ratio of truncated moments, conditional expectation of a specific function, hazard rate function, and reverse hazard rate function.

I. Characterization of KWMBD through Coefficient of variance

In this subsection we characterize the KWMBD by statistical measure (coefficient of variation) (CV)). It is a measure of relative variability, calculated as the ratio of the standard deviation to the mean.

$$CV = \frac{\sigma}{\mu}$$

Where, σ and μ are respectively the mean and standard deviation.

Theorem I: Let $Y_1, Y_2, .., Y_m$, be a independent random sample of size m drawn from KWMBD. Then, the square of sample coefficient of variation is asymptotically unbiased estimator of the square of population coefficient of variation.
Mathematically,

$$\lim_{m\to\infty} E\left(\frac{{S_m}^2}{\overline{Y_m}^2}\right) = \lim_{m\to\infty} E\left(\frac{S_m}{\overline{Y_m}}\right)^2 = \left(\frac{\sigma}{\mu}\right)^2,$$

Where $\overline{Y_m}$ and ${S_m}^2$ are respectively the mean and variance of the sample.
Proof: Since, the sample mean $(\overline{Y_m})$ is an unbiased estimator of population mean (μ) with variance $\frac{\sigma^2}{m}$. i.e, $E(Y_m) = \mu,\ var\ (Y_m) = \frac{\sigma^2}{m}$
Also, we have,

$$var\ (\overline{Y_m}) = E(\overline{Y_m}^2) - [E(\overline{Y_m})\ ]^2$$

$$E(\overline{Y_m}^2) = var\ (\overline{Y_m}\ ) + [E\overline{(Y_m})\ ]^2$$

Substituting (5) and (6), we get,

$$E\left(\overline{Y_m}^2\right) = \frac{2\propto^2}{m\left(\Gamma\left(k+\frac{3}{2}\right)\right)^2}\left[\Gamma\left(\frac{2k+5}{2}\right)\Gamma\left(k+\frac{3}{2}\right) - (1-m)\left(\Gamma\left(\frac{2k+4}{2}\right)\right)^2\right] \quad (8)$$

Also, $E\left({S_m}^2\right) = \sigma^2$, therefore, from (6) we have,

$$E\left({S_m}^2\right) = \frac{2\propto^2}{\left(\Gamma\left(k+\frac{3}{2}\right)\right)^2}\left[\Gamma\left(\frac{2k+5}{2}\right)\Gamma\left(k+\frac{3}{2}\right) - \left(\Gamma\left(\frac{2k+4}{2}\right)\right)^2\right] \quad (9)$$

Since ,$Y_1, Y_2, .., Y_m$ are independent, we can write,

$$E\left(\frac{{S_m}^2}{\overline{Y_m}^2}\right) = \frac{E\left({S_m}^2\right)}{E(\overline{Y_m}^2)}$$

Now, using (8) and (9) we obtain

$$E\left(\frac{{S_m}^2}{\overline{Y_m}^2}\right) = \frac{\Gamma\left(\frac{2k+5}{2}\right)\Gamma\left(k+\frac{3}{2}\right) - \left(\Gamma\left(\frac{2k+4}{2}\right)\right)^2}{\frac{1}{m}\Gamma\left(\frac{2k+5}{2}\right)\Gamma\left(k+\frac{3}{2}\right) - (1-\frac{1}{m})\left(\Gamma\left(\frac{2k+4}{2}\right)\right)^2}$$

Applying $\lim_{m\to\infty}$ on both sides we get,

$$\lim_{m\to\infty} E\left(\frac{{S_m}^2}{\overline{Y_m}^2}\right) = \frac{\Gamma\left(\frac{2k+5}{2}\right)\Gamma\left(k+\frac{3}{2}\right) - \left(\Gamma\left(\frac{2k+4}{2}\right)\right)^2}{\left(\Gamma\left(\frac{2k+4}{2}\right)\right)^2}$$

$$\lim_{m\to\infty} E\left(\frac{{S_m}^2}{\overline{Y_m}^2}\right) = \left\lfloor \frac{\sqrt{\Gamma\left(\frac{2k+5}{2}\right)\Gamma\left(k+\frac{3}{2}\right) - \left(\Gamma\left(\frac{2k+4}{2}\right)\right)^2}}{\Gamma\left(\frac{2k+4}{2}\right)} \right\rfloor^2 = \left(\frac{\sigma}{\mu}\right)^2$$

Hence the theorem.

II. Characterization based on ratio of truncated moments

This subsection explores the characterization of the KWMBD through the ratio of two truncated moments, utilizing Theorem II from Glänzel [5]. The result is valid even if the interval H is not closed, since the theorem's condition applies within the interior of H.

Theorem II: Let $(\Omega; F; P)$ be a given probability space, and let $H = [e, f]$ be an interval for some $e < f$, Where, $e = -\infty$ and $f = +\infty$ might as well be allowed. Also, let $Y: \Omega \to H$ be a continuous RV with the CDF $F(y)$ and $h_1(y)$ and $h_2(y)$ be two real functions defined on H and such that

$$E[h_1(y)|Y \geq y] = \vartheta(y)E[h_2(y)|Y \geq y] \qquad y\epsilon H$$

Is defined with some real function $\vartheta(y)$. Assume that $h_1(y)$, $h_2(y)\epsilon C^1(H)$, $\vartheta(y)\epsilon C^2(H)$, and $F(y)$ is a twice continuously differentiable and strictly monotone function on the set H. Finally, assume that the equation $\vartheta(y)h_2(y) = h_1(y)$ has no real solution in the interior of H. Then, $F(y)$ is uniquely determined by the functions $h_1(y)$, $h_2(y)$, and $\vartheta(y)$ as follows:

$$F(y) = \int_0^y C\left|\frac{\vartheta'(u)}{\vartheta(u)h_1(u) - h_2(u)}\right| e^{-S(u)}\, du$$

Where the function $s(y)$ is a solution of the differential equation

$$S(y) = \frac{\vartheta'(y)h_1(y)}{\vartheta(y)h_1(y) - h_2(y)}$$

And C is a constant such that $\int_H dF(y) = 1$.

Preposition I: Let $Y: \Omega\ (0, \infty)$ be a continuous random variable and let $h_1(y) = y^{-(2k+1)}$ and $h_2(y) = h_1(y)e^{-\frac{y^2}{2\alpha^2}}$ for > 0 . Then, the random variable has PDF $f(y)$ if and only if the function defined in Theorem II is of the form,

$$\vartheta(y) = \frac{1}{2}e^{-\frac{y^2}{2\alpha^2}}: \qquad y > 0 \qquad (10)$$

Or satisfying differential equation

$$\frac{\vartheta'(y)h_1(y)}{\vartheta(y)h_1(y) - h_2(y)} = \frac{y}{\alpha^2}; \qquad y > 0 \qquad (11)$$

Proof:

Necessity: Suppose the RV $Y \sim KWMBD(\alpha, k)$, with the CDF and PDF given by equations (1) and (2), respectively, then, we have,

$$[1 - F(y)]E[h_1(y)|Y \geq y] = qe^{-\frac{y^2}{2\alpha^2}}$$

And

$$[1 - F(y)]E[h_2(y)|Y \geq y] = \frac{q}{2}e^{-\frac{y^2}{\alpha^2}}$$

Where, q is constant.
Also, we have by theorem (II)

$$E[h_1(y)|Y \geq y] = E[h_2(y)|Y \geq y]\vartheta(y)$$

$$\Rightarrow \qquad \vartheta(y) = \frac{E[h_1(y)|Y \geq y]}{E[h_2(y)|Y \geq y]}$$

$$\vartheta(y) = \frac{1}{2}e^{-\frac{y^2}{2\alpha^2}}$$

We also have,

$$\vartheta(y)h_1(y) - h_2(y) = -\frac{1}{2}y^{-(2k+1)}e^{-\frac{y^2}{2\alpha^2}} = -\frac{1}{2}h_1(y)e^{-\frac{y^2}{2\alpha^2}}$$

$$\Rightarrow \qquad \frac{\vartheta'(y)h_1(y)}{\vartheta(y)h_1(y) - h_2(y)} = \frac{y}{\alpha^2}$$

Hence equation (11) holds

Sufficiency: If the function $\vartheta(y)$ satisfies the equation (11), then it follows that

$$S'(y) = \frac{\vartheta'(y)h_1(y)}{\vartheta(y)h_1(y) - h_2(y)} = \frac{y}{\alpha^2}$$

$$\Rightarrow \qquad S(y) = \frac{y^2}{2\alpha^2}$$

Thus from theorem (II), $F(x)$ is uniquely determined by the functions $h_1(y)$, $h_2(y)$ and $\vartheta(y)$ as follows:

$$F(y) = \int_0^y C\left|\frac{\vartheta'(y)}{\vartheta(y)h_1(y) - h_2(y)}\right| e^{-S(y)}\, dy$$

$$F(y) = \int_0^y C\frac{y^{2k+2}}{\alpha^2} e^{-\frac{y^2}{2\alpha^2}}\, dy$$

Where, C is normalizing constant such that $\int_\Omega f(y)dy = 1$

$$C^{-1} = \alpha^{(1+2k)}2^{(k+\frac{1}{2})}\Gamma(k+\frac{3}{2})$$

Thus,

$$F(y) = \int_0^y \frac{y^{2k+2}}{\alpha^{(2+2k)}2^{(k+\frac{1}{2})}\Gamma(k+\frac{3}{2})} e^{-\frac{y^2}{2\alpha^2}}\, dy$$

This is distribution function of KWMBD, thus RV $Y \sim KWMBD(\alpha, k)$.

Collorary I: Let $Y: \Omega(0, \infty)$ be a continuous random variable and let $h_1(y)$ be as in above proportion. The PDF of Y is (2) iff there exits $h_2(y)$ and $\vartheta(y)$ defined in theorem (II) satisfying differential equation

$$\frac{\vartheta'(y)h_1(y)}{\vartheta(y)h_1(y) - h_2(y)} = \frac{y}{\alpha^2}; \qquad y > 0 \qquad (12)$$

After simplification, we get ,

$$\vartheta'(y) = \frac{y}{\alpha^2}\left(\vartheta(y) - {h_1}^{-1}(y)h_2(y)\right)$$

Upon integrating and multiplying both sides by $e^{-\frac{y^2}{2\alpha^2}}$. It is easy to show that the general solution of differential equation is

$$\vartheta'(y) = e^{-\frac{y^2}{2\alpha^2}}\left(-\int \frac{y}{2\alpha^2} e^{-\frac{y^2}{2\alpha^2}} {h_1}^{-1}(y)h_2(y) + M\right)$$

Where, M is a constant. Note that a set of functions satisfying the differential equation (12) is given in Proposition I with $M = 0$. However, it should also be noted that there are other triplets $(h_1; h_1;\ \vartheta)$ satisfying the conditions of Theorem (II).

Proof: Same as preposition (I)

III. Characterization based on the conditional expectation of certain function of the random variable

In this subsection, we utilize a single function ϑ of Y to describe the distribution of Y based on the truncated moments of $\vartheta(y)$. This Theorem III, which appeared in Hamedani's earlier work [9], is presented here to aid in characterizing the KWMBD.

Theorem III: Let $Y: \Omega \to (e, f)$ be a continuous random variable with CDF $F(\mathrm{y})$. Let $\vartheta(\mathrm{y})$ be a differentiable function on (e, f) with $\lim_{y \to e^+} \vartheta(\mathrm{y})$. Then for $\delta \neq 1$,

$$E[\vartheta(\mathrm{y})|Y \geq y] = \delta\vartheta(\mathrm{y}) \qquad y\epsilon(\, e, f)$$

If and only if

$$\vartheta(\mathrm{y}) = [1 - F(y)]^{\frac{1}{\delta}-1} \qquad y\epsilon(\, e, f)$$

Preposition II: Let $Y: \Omega\ (\,0, \infty)$ be a continuous random variable with CDF (1). Let $\vartheta(\mathrm{y})$ be a differentiable function on $(\,0, \infty)$ with $\lim_{y \to 0^+} \vartheta(\mathrm{y})$,then for $\delta = \frac{1}{2}$

$$E[\vartheta(\mathrm{y})|Y \geq y] = \delta\vartheta(\mathrm{y}) \qquad y\epsilon(\, 0, \infty)$$

Iff

$$\vartheta(\mathrm{y}) = \frac{1}{2}\left[\frac{\Gamma\left(\mathrm{k} + \frac{3}{2}, \frac{y^2}{2\alpha^2}\right)}{\Gamma\left(k + \frac{3}{2}\right)}\right]^{\frac{1}{\delta}-1} \tag{13}$$

Proof:

Necessity: Suppose the RV $Y \sim KWMBD(\alpha\ , k)$, with the CDF and PDF given by equations (1) and (2), respectively, then, we have,

$$E[\vartheta(\mathrm{y})|Y \geq y] = \int_y^{\infty} \vartheta(\mathrm{y})\frac{f(y)}{P(Y \geq y)}dy$$

$$=> \qquad E[\vartheta(\mathrm{y})|Y \geq y] = \frac{1}{2}\int_y^{\infty} f(y)dy$$

$$=> \qquad E[\vartheta(\mathrm{y})|Y \geq y] = \frac{1}{2}\left[\frac{\Gamma\left(\mathrm{k} + \frac{3}{2}, \frac{y^2}{2\alpha^2}\right)}{\Gamma\left(k + \frac{3}{2}\right)}\right]^{\frac{1}{\frac{1}{2}}-1}$$

$$=> \qquad \vartheta(y) = \left[\frac{\Gamma\left(k+\frac{3}{2},\frac{y^2}{2\alpha^2}\right)}{\Gamma\left(k+\frac{3}{2}\right)}\right]^{\frac{1}{\delta}-1}$$

Hence, the necessary condition.

Sufficiency: If the function $\vartheta(y)$ satisfies the equation (13), then it follows that

$$\vartheta(y) = \frac{1}{2}\left[\frac{\Gamma\left(k+\frac{3}{2},\frac{y^2}{2\alpha^2}\right)}{\Gamma\left(k+\frac{3}{2}\right)}\right]^{\frac{1}{\delta}-1}$$

Also,

$$E[\vartheta(y)|Y \geq y] = \frac{1}{2}\left[\frac{\Gamma\left(k+\frac{3}{2},\frac{y^2}{2\alpha^2}\right)}{\Gamma\left(k+\frac{3}{2}\right)}\right]^{\frac{1}{\frac{1}{2}}-1}$$

$$=> \quad [1-F(y)] = \left[\frac{\Gamma\left(k+\frac{3}{2},\frac{y^2}{2\alpha^2}\right)}{\Gamma\left(k+\frac{3}{2}\right)}\right]$$

$$=> \quad F(y) = 1 - \frac{\Gamma\left(k+\frac{3}{2},\frac{y^2}{2\alpha^2}\right)}{\Gamma\left(k+\frac{3}{2}\right)}$$

This is distribution function of KWMBD, thus RV $Y \sim KWMBD(\alpha, k)$.

IV. Characterization in terms of hazard function

The hazard function $\lambda(y)$ of a twice-differentiable distribution function $F(y)$ satisfies a specific first-order differential equation given in theorem IV. For many univariate continuous distributions, this is the sole differential equation involving the hazard function. In this subsection, we provide a detailed characterization of the KWMBD using the hazard function.

Theorem IV: The RV Y has continuous density function $f(y)$ if and only if its hazard function $\lambda(y)$ satisfies the differential equation

$$\frac{f'(y)}{f(y)} = \frac{\lambda'(y)}{\lambda(y)} - \lambda(y)$$

Preposition III: Let $Y: \Omega\ (0, \infty)$ be a continuous random variable. The PDF of Y is (2) if and only if its hazard function $\lambda(y)$ satisfies the differential equation

$$\lambda'(y) - \frac{2(k+1)}{y}\,\lambda(y) = -\frac{1}{q\alpha^2} y^{2k+3} e^{-\frac{y^2}{2\alpha^2}} \tag{14}$$

Proof:

Necessity: If has PDF (2), then, we have

$$\lambda'(y) = -\frac{1}{q\alpha^2} y^{2k+3} \boldsymbol{e}^{-\frac{y^2}{2\alpha^2}} \left[\mathbf{1} - \frac{\mathbf{2(k+1)}\alpha^2}{y^2}\right]$$

Also,

$$\lambda'(y) - \frac{2(k+1)}{y}\,\lambda(y) = -\frac{1}{q\alpha^2} y^{2k+3} e^{-\frac{y^2}{2\alpha^2}}$$

Hence the necessary condition

Sufficiency: if its hazard function $\lambda(y)$ satisfies the differential equation (14)

$$y^{-(2k+2)}\left[\lambda'(y) - \frac{2(k+1)}{y}\,\lambda(y)\right] = -\frac{y}{q\alpha^2} e^{-\frac{y^2}{2\alpha^2}}$$

$$\Rightarrow \quad \frac{d}{dy}\left(y^{-(2k+2)}\lambda(y)\right) = -\frac{1}{q}\frac{d}{dy} e^{-\frac{y^2}{2\alpha^2}}$$

$$\Rightarrow \quad \lambda(y) = \frac{1}{q\alpha^2} y^{2k+2} e^{-\frac{y^2}{2\alpha^2}}$$

This is hazard rate function of KWMBD, thus RV $Y \sim KWMBD(\alpha, k)$.

V. Characterization in terms of the reverse hazard function

The reverse hazard function $r(y)$ of a twice-differentiable distribution function F is defined as $\frac{f(y)}{1-F(y)}$. In this subsection, we provide a characterization of the KWMBD using the reverse hazard function.

Preposition IV: Let $Y: \Omega\,(0, \infty)$ be a continuous random variable. The PDF of Y is (2) if and only if its reverse hazard function $r(y)$ satisfies the differential equation

$$r'(y) - \frac{2(k+1)}{y}\,r(y) = \frac{y^{2k+2}}{q\alpha^2}\frac{d}{dy}\left(\frac{e^{-\frac{y^2}{2\alpha^2}}}{\Gamma\left(k+\frac{3}{2}\right) - \Gamma\left(\mathrm{k}+\frac{3}{2}, \frac{y^2}{2\alpha^2}\right)}\right) \tag{15}$$

Proof:

Necessity: If has PDF (2), then, we have

$$r'(y) = \frac{y^{2k+2}}{q\alpha^2}\frac{d}{dy}\left(\frac{e^{-\frac{y^2}{2\alpha^2}}}{\Gamma\left(k+\frac{3}{2}\right) - \Gamma\left(\mathrm{k}+\frac{3}{2}, \frac{y^2}{2\alpha^2}\right)}\right) + \frac{1}{q\alpha^2}\left(\frac{e^{-\frac{y^2}{2\alpha^2}}}{\Gamma\left(k+\frac{3}{2}\right) - \Gamma\left(\mathrm{k}+\frac{3}{2}, \frac{y^2}{2\alpha^2}\right)}\right)(2k+1)y^{2k+2}$$

$$=> \quad r'(y) - \frac{2(k+1)}{y}\, r(y) = \frac{y^{2k+2}}{q\alpha^2}\frac{d}{dy}\left(\frac{e^{-\frac{y^2}{2\alpha^2}}}{\Gamma\left(k+\frac{3}{2}\right) - \Gamma\left(\mathrm{k}+\frac{3}{2},\frac{y^2}{2\alpha^2}\right)}\right)$$

Hence, the necessary part.

Sufficiency: if its reverse hazard function $r(y)$ satisfies the differential equation (15)

$$y^{-(2k+2)}\left[r'(y) - \frac{2(k+1)}{y}\, r(y)\right] = \frac{1}{q\alpha^2}\frac{d}{dy}\left(\frac{e^{-\frac{y^2}{2\alpha^2}}}{\Gamma\left(k+\frac{3}{2}\right) - \Gamma\left(\mathrm{k}+\frac{3}{2},\frac{y^2}{2\alpha^2}\right)}\right)$$

$$=> \quad \frac{d}{dy}\left(y^{-(2k+2)} r(y)\right) = \frac{1}{q\alpha^2}\frac{d}{dy}\left(\frac{e^{-\frac{y^2}{2\alpha^2}}}{\Gamma\left(k+\frac{3}{2}\right) - \Gamma\left(\mathrm{k}+\frac{3}{2},\frac{y^2}{2\alpha^2}\right)}\right)$$

$$=> \quad r(y) = \frac{y^{-(2k+2)} e^{-\frac{y^2}{2\alpha^2}}}{q\left(\Gamma\left(k+\frac{3}{2}\right) - \Gamma\left(\mathrm{k}+\frac{3}{2},\frac{y^2}{2\alpha^2}\right)\right)}$$

This is reverse hazard rate function of KWMBD, thus RV $Y \sim KWMBD(\alpha, k)$.

III. Conclusion

In this article, we studied the characterization results of 2Kth order weighted Maxwell Boltzmann distribution (KWMBD), an extension of the traditional Maxwell-Boltzmann distribution. We have characterized KWMBD via the coefficient of variance, ratio of truncated moments, the conditional expectation of a specific function of the random variable, hazard rate functions, and reverse hazard rate functions. Our principal findings are presented in Preposition I, II, III and IV. We found that, for these characterization results, the cumulative distribution function (CDF) does not need to have a closed form, instead, the CDF is derived from the solution of a first-order differential equation. These findings establish a crucial link between probability theory and differential equations, thereby extending the applicability of the KWMBD.

Refrences

[1] Arslan, G., Ahsanullah, M., & Bairamov, I. G. (2005). On some characteristic properties of the uniform distribution. *Advances on Models, Characterizations and Applications*, 153-160.

[2] Ahmad, Z., Elgarhy, M., & Hamedani, G. G. (2018). A new Weibull-X family of distributions: properties, characterizations and applications. *Journal of Statistical Distributions and Applications*, *5*(1), 5.

[3] Bakouch, H. S., Hussain, T., Tošić, M., Stojanović, V. S., & Qarmalah, N. (2023). Unit exponential probability distribution: Characterization and applications in environmental and engineering data modeling. *Mathematics*, *11*(19), 4207.

[4] Dar, A. A., Ahmed, A., & Reshi, J. A. (2018). Characterization and estimation of weighted maxwell-Boltzmann distribution. *Applied Mathematics & Information Sciences*, *12*(1), 193-202.

[5] Glänzel, W. (1987). A characterization theorem based on truncated moments and its application to some distribution families. In *Mathematical statistics and probability theory: volume B*

statistical inference and methods proceedings of the 6th Pannonian symposium on mathematical statistics, bad Tatzmannsdorf, Austria, September 14–20, 1986 (pp. 75-84). Springer Netherlands.

[6] Glanzel, W. (1990). Some consequences of a characterization theorem based on truncated moments. *Statistics*, *21*(4), 613-618.

[7] Hamedani, G. G. (2002). Characterizations of univariate continuous distributions. II. *Studia Scientiarum Mathematicarum Hungarica*, *39*(3-4), 407-424.

[8] Hamedani, G. G., & Hamedani, G. G. (2006). Characterizations of univariate continuous distributions, III. *Studia Scientiarum Mathematicarum Hungarica*, *43*(3), 361-385.

[9] Hamedani, G. G. (2013). On certain generalized gamma convolution distributions II (No. 484). technical report.

[10] Shakil, M., & Kibria, B. G. (2014). Two characterizations of an exponential power life-testing distribution. *Journal of Applied Statistical Science*, *22*(3/4).

[11] Saghir, A., & Hamedani, G. G. (2018). A GENERALIZED WEIGHTED MAXWELL DISTRIBUTION: PROPERTIES AND ESTIMATION. *Pakistan Journal of Statistics*, *34*(3).

RT&A, No 2 (84)
Volume 20, June 2025

CERTAIN STATIONARY POINT OUTCOMES FOR ASYMPTOTICALLY REGULAR MAPS IN INTERVAL-VALUED N-FUZZY METRIC SPACE

Heera Ahirwar[1], Kavita Shrivastava[2] , Harshit Khare[3]

•

[1,2,3]Department of Mathematics and Statistics
Dr. Harisingh Gour Vishwavidyalaya, Sagar, Madhya Pradesh, India – 470339
[1]heera.maths15@gmail.com, [2]kavita.rohit@rediff.com, [3]khareh412@gmail.com

Abstract

In this study, we introduce interval valued N-fuzzy metric space (IVNFM) and we verify certain popular fixed point results on the structure of interval valued N-fuzzy metric space through asymptotically regular mappings. Fixed point theorems are crucial in various mathematical and applied fields, providing solutions to equations where a function maps a point to itself. In this research, we focus on applying diverse contractive conditions specifically in the context of asymptotically regular mappings to achieve our objectives. By employing this type of mapping, we are able to verify and validate several well-established fixed point results within IVNFM. This approach not only supports the known results but also broadens their applicability, thereby enhancing the theoretical foundation of fixed point theorems in fuzzy metric spaces.

Keywords: Interval-valued N-fuzzy metric, common fixed point, asymptotically regular maps, altering distance function, commutative self mappings.

I. Introduction

The foundation of fuzzy mathematics is laid by [25] in 1965. In 1975, [10] proposed the idea of fuzzy metric space. [14] employed fuzzy metric space membership functions in 1983 to determine the fuzzy metric between two fuzzy sets; this approach differed from that suggested by [10]. Fuzzy metric space was reformulated in 1994 when [3] presented the modified concept of Continuous t-norm. Letter on, numerous authors and researchers generalized various type of metric spaces like fuzzy 2-metric space, D-metric space, G-metric space-metric space, Q(G)-metric space, D*-metric space M-metric spaces-metric space and have thoroughly and extensively examined a variety of issues pertaining to this space from a variety of perspectives, including compatible mapping, R-weakly computing mapping, Weak compatible mapping, CLR-Property, E.A. property, etc. and have produced new findings on fuzzy metric space. Based on an interval valued membership function, [25] first the idea of an interval-valued fuzzy set in 1975. By inspiring the concept of compatible maps, [6] presented the idea of compatible maps of type (A) in metric space and Banach space in 1993. Three different types of distances between two interval-valued fuzzy sets on real line R were employed by [11] in 2009. Using the concept of continuous interval-valued t-norm, it is possible to define interval-valued fuzzy metric space and describe the uncertainty of the distance between two points in a fuzzy metric space. In 2015, [12] introduced the new notion of

NM-fuzzy metric space, Pseudo N-fuzzy metric space and describes some of their properties and examples. In 2020 [5] proved some fixed point theorems for asymptotically regular maps in N-fuzzy b-metric space. [1] recently established some common fixed point solutions for mapping on interval-valued fuzzy metric space that is occasionally weakly compatible. This was demonstrated in 2023.

In this study, we provide solutions to equations where a function maps a point to itself. In this research, we focus on applying diverse contractive conditions specifically in the context of asymptotically regular mappings to achieve our objectives validate certain well-known fixed point results for asymptotically regular mapping in interval valued N-fuzzy metric space (IVFNM). Our results generalize and extend the results of the common fixed point theorem for intimate mapping that were published recently.

II. Basic Definitions

Definition 1[18] In a non empty set $\mho$, a mapping $R_n: \mho \to [I]$ is called an interval- valued fuzzy set on $\mho$. Collection of all interval-valued fuzzy sets on $\mho$ is denoted by IVF($\mho$). if $R_n \in IVF(\mho)$. , let $R_n(x) = [R_n^-, R_n^+]$, $R_n^-(x) \le R_n^+(x)$ for all $x \in \mho$, then the set $R_n^-: \mho \to [I]$ and $R_n^+: \mho \to [I]$ are called lower fuzzy set and upper fuzzy set of R_n and if $R^-(x) = R_n^+(x)$ then is called degenerate fuzzy set for all $x \in \mho$.

Definition 2[18]**.** A binary operation of the form is an interval valued t_{norm} is $*_I: [I] \times [I] \to [I]$ on [I] such that for all $\overline{u}, \overline{v}, \overline{w} \in [I]$ if satisfying following four conditions:

(1) Commutativity : $\overline{u}_{ivf} *_I \overline{v}_{ivf} = \overline{v}_{ivf} *_I \overline{u}_{ivf}$,
(2) Associativity: $[\overline{u}_{ivf} *_I \overline{v}_{ivf}] *_I \overline{w}_{ivf} = \overline{u}_{ivf} *_I [\overline{v}_{ivf} *_I \overline{w}_{ivf}]$,
(3) Monotonicity: $\overline{u}_{ivf} *_I \overline{v}_{ivf} \le \overline{u}_{ivf} *_I \overline{w}_{ivf}$, whenever $\overline{v}_{ivf} *_I \overline{w}_{ivf}$,
(4) Boundary condition: $\overline{u}_{ivf} *_I \overline{1} = \overline{u}_{ivf}$, $\overline{u}_{ivf} *_I \overline{0} = [u^-, u^+] *_I [0,1] = [0, u^+]$.

Note: Each interval valued t_{norm} satisfies some additional boundary conditions for all $\overline{u} \in [I]$:
$\overline{u} *_I \overline{0} = \overline{0} *_I \overline{u} = \overline{0}$,
$\overline{I} *_I \overline{u} = [0,1] *_I [u^-, u^+] = \overline{0}$,
$\overline{1} *_I \overline{u} = \overline{1}$.

Example: (a) $\overline{u} *_I \overline{v} = [u^-.v^-, u^+.v^+]$; (b) $\overline{u} *_I \overline{v} = [u^- \wedge v^-, u^+ \wedge v^+]$

Definition 3[18]. Let $\{\overline{u}_i\} = \{[u_i^-, u_i^+]\}, \mathbf{i} \in \mathbb{N}^+$ be a sequence of interval numbers in [I], $\overline{u} = [u^-, u^+] \in [I]$, if $\lim_{i\to\infty} u_i^- = u^-$ and $\lim_{i\to\infty} u_i^+ = u^+$ then the sequence $\{\overline{u}_i\}$ is convergent to $\overline{u}$ and denoted by $\lim_{i\to\infty} \overline{u}_i = \overline{u}$.

Definition 4 [1]. An interval valued t_{norm} $*_I$ is continuous iff it is continuous in its first component, i.e. for each $\overline{v} \in [I]$, if $\lim_{i\to\infty} \overline{u}_i = \overline{u}$, then

$\lim_{i\to\infty}(\overline{u}_i *_I \overline{v}) = (\lim_{i\to\infty} \overline{u}_i *_I \overline{v}) = \overline{u} *_I \overline{v}$, Where $\{\overline{u}_i\} \subseteq [\mathbf{I}], \overline{u} \in [I]$.

Definition 5 [5]. A triplet $(\mathfrak{X}, M, *)$ is called fuzzy metric space (FMS) if $\mathfrak{X}$ is an arbitrary set, $*$ is a continuous t_{norm} on [0,∞] and M is a fuzzy set on $\mathfrak{X}^2 \times (0,\infty)$ satisfying the following conditions :

(1) $M(x, y, t) > 0$;
(2) $M(x, y, t) = 1$ for all $t > 0$ iff $x = y$;
(3) $M(x, y, t) = M(y, x, t)$;
(4) $M(x, z, t_1 + t_2) \ge T(M(x, y, t_1), M(x, y, t_2))$; $\forall x, y, z \in \mathfrak{X}$ and $t_1, t_2 > 0$
(5) $M(x, y, t): [0,\infty) \to [0,1]$ is continuous;

Definition 6[5]. A triplet $(\mathfrak{X}, M, *)$ is called N-fuzzy metric space (N-FMS) if $\mathfrak{X}$ is an arbitrary set, $*$ is a continuous t_{norm} on [0,∞] and M is a fuzzy set on $\mathfrak{X}^3 \times (0,\infty)$ satisfying the following

conditions:

(1) $M(x, y, z, t) > 0$;

(2) $M(x, y, z, t) = 1$ for all $t > 0$ iff $x = y = z$;

(3) $M(x, y, z, t_1 + t_2 + t_3) \geq M(x, x, a, t_1) * M(y, y, a, t_2) * M(z, z, a, t_3)$;
$\forall x, y, z \in \mathfrak{X}$ and $t_1, t_2, t_3 > 0$

(4) $M(x, y, z, *): (0, \infty) \to [0,1]$ is continuous;

Definition 7[5]. A triplet $(\mathfrak{X}, N_{IVFNM}, *_I)$ is called interval valued N-fuzzy metric space (IVFMS) if $\mathfrak{X}$ is an arbitrary set, $*_I$ is a continuous interval valued t_{norm} on [I] and N_{IVFNM} is a fuzzy set on $\mathfrak{X}^3 \times (0, \infty)$ satisfying the following conditions :

(1) $N_{IVFNS}(x, y, z, t_{norm}) > \bar{0}$;

(2) $N_{IVFNM}(x, y, z, t_{norm}) = \bar{1}$ for all $t_{norm} > 0$ iff $x = y = z$;

(3) $N_{IVFNM}(x, y, z, a, t_1 + t_2 + t_3) \geq N_{IVFNM}(x, x, a, t_1) *_I N_{IVFNM}(y, y, a, t_2) *_I N_{IVFNM}(z, z, a, t_3)$;
$\forall x, y, z \in \mathfrak{X}$ and $t_1, t_2, t_3 > 0$

(4) $N_{IVFNM}(x, y, z, *_I): [0, \infty] \to (I]$ is continuous;

(5) $\lim_{t \to \infty} N_{IVFNM}(x, y, z, t_{norm}) = \bar{1}$; $\forall x, y, z \in \mathfrak{X}, t_{norm} > 0$,

Note: Every metric can induce an interval valued N-fuzzy metric space.

Definition 8[1]. Let $(\mathfrak{X}, N_{IVFNM}, *_I)$ is an IVFNM

(a) If $\beta > t_{norm} > 0$ then $N_{IVFNM}(x, y, z, t_{norm}) \leq N_{IVFNM}(x, y, z, \beta)$ for $x, y, z \in \mathfrak{X}$.

(b) A sequence $\{x_n\}$ in $\mathfrak{X}$ is referred to as a Cauchy sequence if for all $\bar{\epsilon} > \bar{0}$ and $t_{norm} > 0$ then there exists a $n_0 \in \mathbb{N}$ $\bar{\epsilon} N_{IVFNM}(x, y, z, t_{norm}) > 1 - \bar{\epsilon}$, for all $x, y, z, \geq n_0$.

(c) If every Cauchy sequence is convergent in $(\mathfrak{X}, N_{IVFNM}, *_I)$ then it is complete IVFNM.

Definition 9[1]. A mapping $\gamma: [I] \to [I]$ is called an altering distance function if

(a) γ is strictly decreasing and left continuous.

(b) $\gamma(\alpha) = 0$ if and only if $\alpha = \bar{1}$ i.e. $\lim_{\gamma \to I^-} \gamma(I) = 0$.

Definition 10[1]**.** Let two self mapping $\mathfrak{K}$ and $\mathfrak{L}$ on IVFNM $(\mathfrak{X}., V_{IVFNM}, *_I)$ then compatible if $\lim_{n \to \infty} N_{IVFNM}(\mathfrak{K}\mathfrak{L}x_n, \mathfrak{K}\mathfrak{L}x_n, \mathfrak{L}\mathfrak{K}x_n, t_{norm}) = \bar{1}$ for all $t_{norm} > 0$ whenever a sequence $\{x_n\}$ in $\mathfrak{X}$ provided $\lim_{n \to \infty} \mathfrak{K}x_n = \lim_{n \to \infty} \mathfrak{L}x_n = u$, for all $u \in \mathfrak{X}$.

Definition 11[1]. Let two self mapping $\mathfrak{K}$ and $\mathfrak{L}$ on IVFNM $(\mathfrak{X}, \mathfrak{M}_{IVFMS}, *_I)$ and a sequence $\{x_n\}$ in $\mathfrak{X}$ then $\mathfrak{K}$ is called asymptotically regular at a point $z_0 \in \mathfrak{X}$
if $\lim_{n \to \infty} N_{IVFNM}(\mathfrak{K}^n(z_0), \mathfrak{K}^n(z_0), \mathfrak{K}^{n+1}(z_0), t_{norm}) = \bar{1}$, for all $t_{norm} > 0$
and also the sequence $\{x_n\}$ is called asymptotically regular with respect to the pair $(\mathfrak{K}, \mathfrak{L})$
if $\lim_{n \to \infty} N_{IVFNM}(\mathfrak{K}(z_0), \mathfrak{K}(z_0), \mathfrak{L}(z_0), t_{norm}) = \bar{1}$ for all $t_{norm} > 0$.

III. Main Results

***Theorem 1*:** Let $\mathfrak{F}: \mathcal{X} \to \mathcal{X}$ be a self mapping of a complete IVFNM $(\mathcal{X}, N_{IVFNM}, *_I)$ and γ be the altering distance function then satisfying the following conditions:

$$\begin{aligned}\gamma\big(N(\mathfrak{F}(\varrho), \mathfrak{F}(\varrho), \mathfrak{F}(\sigma), t_{norm})\big) &\leq b_1\theta\big[\min\{\gamma\big(N(\varrho, \varrho, \mathfrak{F}(\varrho), t_{norm})\big), \gamma\big(N(\sigma, \sigma, \mathfrak{F}(\sigma), t_{norm})\big)\}\big] \\ &+ b_2\omega\big[\{\gamma\big(N(\varrho, \varrho, \mathfrak{F}(\varrho), t_{norm})\big).\gamma\big(N(\sigma, \sigma, \mathfrak{F}(\sigma), t_{norm})\big)\}\big] \\ &+ b_3\big[\{\gamma\big(N(\varrho, \varrho, \mathfrak{F}(\varrho), t_{norm})\big) + \gamma\big(N(\sigma, \sigma, \mathfrak{F}(\sigma), t_{norm})\big)\}\big] \\ &+ b_4\big[\{\gamma\big(N(\varrho, \varrho, \mathfrak{F}(\sigma), t_{norm})\big) + \gamma\big(N(\mathfrak{F}(\varrho), \mathfrak{F}(\varrho), \sigma, t_{norm})\big)\}\big] \\ &+ b_5\gamma\big(N(\varrho, \varrho, \sigma, t_{norm})\big)\end{aligned}$$

Where $\forall \varrho, \sigma \in \mathcal{X}$ and $t_{norm} > 0, b_i = b_i(\varrho, \sigma) \geq 0, i = 1,2 \ldots 5$ are such that for some arbitrarily fixed $\lambda_1 > 0, 0 < \lambda_2 < 1$,

$b_1 + b_2 \leq \lambda_1$ and $(b_3 + 2b_4 + b_5) \leq \lambda_2$ and $\theta, \omega: R^+ \to R^+$ are continuous functions at 0 and $\theta(0) = \omega(0) = 0$. If $\mathfrak{F}$ is asymptotically regular at some point $\varrho_0 \in \mathcal{X}$ then has a unique fixed point in $\mathcal{X}$.

Proof: Let a sequences $\{\varrho_n\}$ in $\mathcal{X}$ where $\varrho_0 \in \mathcal{X}$ and $\varrho_{n+1} = \mathfrak{F}(\varrho_n)$, for all $n \geq 0$,
Now if $n \geq 0$, $\varrho_{n+1} = \varrho_n$ then ϱ_n is a fixed point of $\mathfrak{F}$. suppose $\varrho_{n+1} \neq \varrho_n$ then to prove $\{\varrho_n\}$ is a Cauchy's sequence in $\mathcal{X}$.
Suppose to the contrary $\exists, 0 < \epsilon < 1, t_{norm} > 0$ and two sequences $\{u_n\}$ and $\{v_n\}$ such that $u_n > v_n > n$,

$$N_{IVFNM}(\varrho_{u_n}, \varrho_{u_n}, \varrho_{v_n}, t_{norm}) \leq 1-\epsilon,$$
$$N_{IVFNM}(\varrho_{u_n-1}, \varrho_{u_n-1}, \varrho_{v_n-1}, t_{norm}) > 1-\epsilon$$
$$N_{IVFNM}(\varrho_{u_n-1}, \varrho_{u_n-1}, \varrho_{v_n}, t_{norm}) > 1-\epsilon, \forall n \in \mathbb{N} \cup \{0\} \ldots\ldots \text{(a)}$$

Now we have

$$1-\epsilon \geq N_{IVFNM}(\varrho_{u_n}, \varrho_{u_n}, \varrho_{v_n}, t_{norm}) \geq N_{IVFNM}(\varrho_{u_n}, \varrho_{u_n}, \varrho_{u_n-1}, t_{norm}) *_I N_{IVFNM}(\varrho_{u_n}, \varrho_{u_n}, \varrho_{u_n-1}, t_{norm}) *_I N_{IVFNM}(\varrho_{v_n}, \varrho_{v_n}, \varrho_{u_n-1}, t_{norm})$$

$1-\epsilon \geq \lim_{n\to\infty} N_{IVFNM}(\varrho_{u_n}, \varrho_{u_n}, \varrho_{v_n}, t_{norm}) \geq (\bar{1} *_I \bar{1} *_I 1-\epsilon)$ (Since $\mathfrak{F}$ asymptotically regular at ϱ_0)

$\Rightarrow \lim_{n\to\infty} N_{IVFNM}(\varrho_{u_n}, \varrho_{u_n}, \varrho_{v_n}, t_{norm}) = 1-\epsilon$ …… (b)

Again,

$N_{IVFNM}(\varrho_{u_n}, \varrho_{u_n}, \varrho_{v_n-1}, t_{norm}) \geq$
$N_{IVFNM}(\varrho_{u_n}, \varrho_{u_n}, \varrho_{v_n}, t_{norm}) *_I N_{IVFNM}(\varrho_{u_n}, \varrho_{u_n}, \varrho_{v_n}, t_{norm}) *_I N_{IVFNM}(\varrho_{v_{n-1}}, \varrho_{v_{n-1}}, \varrho_{v_n}, t_{norm})$

$\lim_{n\to\infty} N_{IVFNM}(\varrho_{u_n}, \varrho_{u_n}, \varrho_{v_n-1}, t_{norm}) > 1-\epsilon$ ……(c)

By putting $\varrho = \varrho_{u_n-1}$, $\sigma = \varrho_{v_n-1}$ in theorem (3.1) then we have ,

$$\begin{aligned}
&\gamma\left(N(\mathfrak{F}(\varrho_{u_n-1}), \mathfrak{F}(\varrho_{u_n-1}), \mathfrak{F}(\varrho_{v_n-1}), t_{norm})\right) \\
&\leq b_1\theta\left[\min\left\{\gamma\left(N(\varrho_{u_n-1}, \varrho_{u_n-1}, \mathfrak{F}(\varrho_{u_n-1}), t_{norm})\right), \gamma\left(N(\varrho_{v_n-1}, \varrho_{v_n-1}, \mathfrak{F}(\varrho_{v_n-1}), t_{norm})\right)\right\}\right] \\
&+ b_2\omega\left[\left\{\gamma\left(N(\varrho_{u_n-1}, \varrho_{u_n-1}, \mathfrak{F}(\varrho_{u_n-1}), t_{norm})\right).\gamma\left(N(\varrho_{v_n-1}, \varrho_{v_n-1}, \mathfrak{F}(\varrho_{v_n-1}), t_{norm})\right)\right\}\right] \\
&+ b_3\left[\left\{\gamma\left(N(\varrho_{u_n-1}, \varrho_{u_n-1}, \mathfrak{F}(\varrho_{u_n-1}), t_{norm})\right) + \gamma\left(N(\varrho_{v_n-1}, \varrho_{v_n-1}, \mathfrak{F}(\varrho_{v_n-1}), t_{norm})\right)\right\}\right] \\
&+ b_4\left[\left\{\gamma\left(N(\varrho_{u_n-1}, \varrho_{u_n-1}, \mathfrak{F}(\varrho_{v_n-1}), t_{norm})\right) + \gamma\left(N(\mathfrak{F}(\varrho_{u_n-1}), \mathfrak{F}(\varrho_{u_n-1}), \varrho_{v_n-1}, t_{norm})\right)\right\}\right] \\
&+ b_5\gamma\left(N(\varrho_{u_n-1}, \varrho_{u_n-1}, \varrho_{v_n-1}, t_{norm})\right)
\end{aligned}$$

Since $\mathfrak{F}$ asymptotically regular at ϱ_0.

$$\begin{aligned}
&\gamma\left(N(\varrho_{u_n}, \varrho_{u_n}, \varrho_{v_n}, t_{norm})\right) \\
&\leq b_1\theta\left[\min\left\{\gamma\left(N(\varrho_{u_n-1}, \varrho_{u_n-1}, \varrho_{v_n}, t_{norm})\right), \gamma\left(N(\varrho_{v_{n-1}}, \varrho_{v_{n-1}}, \varrho_{v_n}, t_{norm})\right)\right\}\right] \\
&+ b_2\omega\left[\left\{\gamma\left(N(\varrho_{u_n-1}, \varrho_{u_n-1}, \varrho_{u_n}, t_{norm})\right).\gamma\left(N(\varrho_{v_n-1}, \varrho_{v_n-1}, \varrho_{v_n}, t_{norm})\right)\right\}\right] \\
&+ b_3\left[\left\{\gamma\left(N(\varrho_{u_n-1}, \varrho_{u_n-1}, \varrho_{u_n}, t_{norm})\right) + \gamma\left(N(\varrho_{v_n-1}, \varrho_{v_n-1}, \varrho_{v_n}, t_{norm})\right)\right\}\right] \\
&+ b_4\left[\left\{\gamma\left(N(\varrho_{u_n-1}, \varrho_{u_n-1}, \varrho_{v_n}, t_{norm})\right) + \gamma\left(N(\varrho_{u_n}, \varrho_{u_n}, \varrho_{v_n-1}, t_{norm})\right)\right\}\right] \\
&+ b_5\gamma\left(N(\varrho_{u_n-1}, \varrho_{u_n-1}, \varrho_{v_n-1}, t_{norm})\right)
\end{aligned}$$

Letting $n \to \infty$ and applying (a), (b) and (c) and altering distance

$$\gamma(1-\epsilon) \leq b_1\theta[\min\{\gamma(\bar{1}), \gamma(\bar{1})\}] + b_2\omega[\{\gamma(\bar{1}).\gamma(\bar{1})\}] + b_3[\{\gamma(\bar{1}) + \gamma(\bar{1})\}] + b_4[\{\gamma(1-\epsilon) + \gamma(1-\epsilon)\}] + b_5\gamma(1-\epsilon)$$
$$\gamma(1-\epsilon) \leq 2b_4\gamma(1-\epsilon) + b_5\gamma(1-\epsilon) < \gamma(1-\epsilon)$$

This is a contradiction.

Thus $\{\varrho_n\}$ is a Cauchy's sequence in $\mathcal{X}$. since complete IVFNM $(\mathcal{X}, N_{IVFNM}, *_I)$ then $\exists$, $p \in \mathcal{X}$ such that $\varrho_n \to p$.

$$\begin{aligned}\gamma\big(N(\mathfrak{F}(\varrho_n),\mathfrak{F}(\varrho_n),\mathfrak{F}(p),t_{norm})\big)&\le b_1\theta\big[\min\{\gamma\big(N(\varrho_n,\varrho_n,\mathfrak{F}(\varrho_n),t_{norm})\big),\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\}\big]\\&+b_2\omega\big[\{\gamma\big(N(\varrho_n,\varrho_n,\mathfrak{F}(\varrho_n),t_{norm})\big).\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\}\big]\\&+b_3\big[\{\gamma\big(N(\varrho_n,\varrho_n,\mathfrak{F}(\varrho_n),t_{norm})\big)+\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\}\big]\\&+b_4\big[\{\gamma\big(N(\varrho_n,\varrho_n,\mathfrak{F}(p),t_{norm})\big)+\gamma\big(N(\mathfrak{F}(\varrho_n),\mathfrak{F}(\varrho_n),\mathfrak{F}(p),t_{norm})\big)\}\big]\\&+b_5\gamma\big(N(p,p,\mathfrak{F}(\varrho_n),t_{norm})\big)\end{aligned}$$

$$\begin{aligned}\gamma\big(N(\varrho_{n+1},\varrho_{n+1},\mathfrak{F}(p),t_{norm})\big)&\le b_1\theta\big[\min\{\gamma\big(N(\varrho_n,\varrho_n,\varrho_{n+1},t_{norm})\big),\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\}\big]\\&+b_2\omega\big[\{\gamma\big(N(\varrho_n,\varrho_n,\varrho_{n+1},t_{norm})\big).\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\}\big]\\&+b_3\big[\{\gamma\big(N(\varrho_n,\varrho_n,\varrho_{n+1},t_{norm})\big)+\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\}\big]\\&+b_4\big[\{\gamma\big(N(\varrho_n,\varrho_n,\mathfrak{F}(p),t_{norm})\big)+\gamma\big(N(\varrho_{n+1},\varrho_{n+1},\mathfrak{F}(p),t_{norm})\big)\}\big]\\&+b_5\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\end{aligned}$$

Letting $n\to\infty$ then we have,

$$\begin{aligned}\lim_{n\to\infty}\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)&\le b_3\lim_{n\to\infty}\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)+2b_4\lim_{n\to\infty}\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\\&+b_5\lim_{n\to\infty}\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\end{aligned}$$

$$\lim_{n\to\infty}\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\le(b_3+2b_4+b_5)\lim_{n\to\infty}\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)$$

$$\{1-(b_3+2b_4+b_5)\}\lim_{n\to\infty}\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\le 0.$$

$$\lim_{n\to\infty}\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\le 0\,,\quad[\text{since }\{1-(b_3+2b_4+b_5)\}<1\,]$$

$$\Rightarrow\lim_{n\to\infty}\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)=0\Rightarrow\mathfrak{F}(p)=p.$$

If u be the another fixed point of $\mathfrak{F}$ in $\mathcal{X}$. then

$$\begin{aligned}\gamma\big(N(\mathfrak{F}(u),\mathfrak{F}(u),\mathfrak{F}(p),t_{norm})\big)&\le b_1\theta\big[\min\{\gamma\big(N(u,u,\mathfrak{F}(u),t_{norm})\big),\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\}\big]\\&+b_2\omega\big[\{\gamma\big(N(u,u,\mathfrak{F}(u),t_{norm})\big).\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\}\big]\\&+b_3\big[\{\gamma\big(N(u,u,\mathfrak{F}(u),t_{norm})\big)+\gamma\big(N(p,p,\mathfrak{F}(p),t_{norm})\big)\}\big]\\&+b_4\big[\{\gamma\big(N(u,u,\mathfrak{F}(p),t_{norm})\big)+\gamma\big(N(\mathfrak{F}(u),\mathfrak{F}(u),p,t_{norm})\big)\}\big]+b_5\gamma\big(N(u,u,p,t_{norm})\big)\end{aligned}$$

$$\begin{aligned}\Rightarrow\gamma\big(N(u,u,p,t_{norm})\big)&\le b_4\big[\{\gamma\big(N(u,u,p,t_{norm})\big)+\gamma\big(N(u,u,p,t_{norm})\big)\}\big]+b_5\gamma\big(N(u,u,p,t_{norm})\big)\\\Rightarrow\gamma\big(N(u,u,p,t_{norm})\big)&\le 2b_4\gamma\big(N(u,u,p,t_{norm})\big)+b_5\gamma\big(N(u,u,p,t_{norm})\big)\\\Rightarrow\gamma\big(N(u,u,p,t_{norm})\big)&\le(2b_4+b_5)\gamma\big(N(u,u,p,t_{norm})\big)\\\Rightarrow\{1-(2b_4+b_5)\}\gamma\big(N(u,u,p,t_{norm})\big)&\le 0.\\\Rightarrow\gamma\big(N(u,u,p,t_{norm})\big)&=0.\\\Rightarrow u&=p.\end{aligned}$$

Example: Let IVFNM $(\mathcal{X},N_{IVFNM},*_I)$ and $\mathcal{X}=[0,1]$ and a interval valued N-fuzzy set on $\mathcal{X}^3\times(0,\infty)$ to $[0,1]$ defined by $N_{IVFNM}(x,y,z,t_{norm})=\frac{t}{t+[|x-z|+|y-z|]}$, for all $x,y,z\in\mathcal{X}$ and $*_I$ is minimum t_{norm} define $\mathfrak{F}:\mathcal{X}\to\mathcal{X}$ by

$$\mathfrak{F}(x)=\begin{cases}\frac{x}{3}, & x\in[0,\frac{1}{2}]\\ \frac{1}{6}, & x\in[\frac{1}{2},1]\end{cases}\quad\text{and}\quad\gamma(\alpha)=1-\alpha,\ \alpha\in[0,1].$$

Let $\theta(u)=\sqrt{u}$ and $\omega(v)=v^2$, $u,v\in R^+$

Also let $b_1(\varrho,\sigma)=|\varrho-\sigma|$, $b_2(\varrho,\sigma)=|\varrho^2-\sigma^2|$, $b_3(\varrho,\sigma)=\begin{cases}\frac{1}{|x-y|}, & x\neq y\\ 0, & x=y\end{cases}$,

$b_4(\varrho,\sigma)=b_5(\varrho,\sigma)=\left|\frac{1-b_3(\varrho,\sigma)}{3}\right|$. Then all the conditions of theorem 3.1 verified. Hence $\mathfrak{F}$ has a unique fixed point.

Theorem 2: Let $\mathfrak{F}$, $\mathscr{L}:\mathcal{X}\to\mathcal{X}$ be a complete IVFNM $(\mathcal{X},N_{IVFNM},*_I)$ and γ be the altering distance function. If $\mathfrak{F}$, $\mathscr{L}$ is asymptotically regular at a point $\varrho_0\in\mathcal{X}$ both satisfy the inequality (1).

Moreover, if

$$\gamma\big(N(\mathfrak{F}(\varrho),\mathfrak{F}(\varrho),\mathcal{L}(\sigma),t_{norm})\big) \leq \lambda\big[\gamma\big(N(\varrho,\varrho,\sigma,t_{norm})\big)+\gamma\big(N(\varrho,\varrho,\mathfrak{F}(\varrho),t_{norm})\big)+\gamma\big(N(\sigma,\sigma,\mathcal{L}(\sigma),t_{norm})\big)\big]$$

Where $\forall\ \varrho,\sigma \in \mathcal{X}$, $\lambda<1$.Then $\mathfrak{F}$ and $\mathscr{L}$ has a unique common fixed point in $\mathcal{X}$.

Proof: By above theorem both $\mathfrak{F}$ and $\mathscr{L}$ have unique fixed point say u and v respectively.then it's satisfied above inequality:

$$\gamma\big(N(\mathfrak{F}(u),\mathfrak{F}(u),\mathcal{L}(v),t_{norm})\big) \leq \lambda\big[\gamma\big(N(u,u,v,t_{norm})\big)+\gamma\big(N(u,u,\mathfrak{F}(u),t_{norm})\big)+\gamma\big(N(v,v,\mathcal{L}(v),t_{norm})\big)\big]$$

$$\gamma\big(N(u,u,v,t_{norm})\big) \leq \lambda\big[\gamma\big(N(u,u,v,t_{norm})\big)+\gamma\big(N(u,u,u,t_{norm})\big)+\gamma\big(N(v,v,v,t_{norm})\big)\big]$$

By definition of altering distance

$$\big(N(u,u,v,t_{norm})\big) \leq \lambda\big(N(u,u,v,t_{norm})\big)$$

$$(1-\lambda)\big(N(u,u,v,t_{norm})\big) \leq 0$$

$\big(N(u,u,v,t_{norm})\big) \leq 0$, (since $\lambda<1$)

$\Rightarrow u=v$. Hence $\mathfrak{F}$ and $\mathscr{L}$ have unique fixed point.

Theorem 3: Let $\mathfrak{F}:\mathcal{X}\to\mathcal{X}$ be a complete IVFNM $(\mathcal{X},N_{IVFNM},*_I)$ and γ be the altering distance function. If $\mathfrak{F}$ is asymptotically regular at a point $\varrho_0 \in \mathcal{X}$. then satisfying the conditions:

$$\gamma\big(N(\mathfrak{F}(\varrho),\mathfrak{F}(\varrho),\mathfrak{F}(\sigma),t_{norm})\big)$$
$$\leq \alpha\ \mathrm{Min}\big[\gamma\big(N(\varrho,\varrho,\sigma,t_{norm})\big),\gamma\big(N(\mathfrak{F}(\varrho),\mathfrak{F}(\varrho),\varrho,t_{norm})\big),\gamma\big(N(\mathfrak{F}(\varrho),\mathfrak{F}(\varrho),\sigma,t_{norm})\big)\big]$$
$$+\beta\ \mathrm{Min}\big[\gamma\big(N(\varrho,\varrho,\sigma,t_{norm})\big),\gamma\big(N(\mathfrak{F}(\sigma),\mathfrak{F}(\sigma),\sigma,t_{norm})\big),\gamma\big(N(\varrho,\varrho,\mathfrak{F}(\sigma),t_{norm})\big)\big]$$

Where $\forall\ \varrho,\sigma \in \mathcal{X}, t_{norm}>0$, $\alpha,\beta>0$ such that $\alpha+\beta<1$ .Then $\mathfrak{F}$ have a unique common fixed point in $\mathcal{X}$.

Proof : Let a sequences $\{\varrho_n\}$ in $\mathcal{X}$ where $\varrho_0 \in \mathcal{X}$ and $\varrho_{n+1}=\mathfrak{F}(\varrho_n)$, for all $n \geq 0$,

Now if $n \geq 0$, $\varrho_{n+1}=\varrho_n$ then ϱ_n is a fixed point of $\mathfrak{F}$. suppose $\varrho_{n+1}\neq\varrho_n$ then to prove $\{\varrho_n\}$ is a Cauchy's sequence in $\mathcal{X}$.

Let $m,n \in \mathbb{N}\cup\{0\}$

$$\gamma\big(N(\mathfrak{F}(\varrho_n),\mathfrak{F}(\varrho_n),\mathfrak{F}(\varrho_m),t_{norm})\big)$$
$$\leq \alpha\ \mathrm{Min}\big[\gamma\big(N(\varrho_n,\varrho_n,\varrho_m,t_{norm})\big),\gamma\big(N(\mathfrak{F}(\varrho_n),\mathfrak{F}(\varrho_n),\varrho_n,t_{norm})\big),\gamma\big(N(\mathfrak{F}(\varrho_n),\mathfrak{F}(\varrho_n),\varrho_m,t_{norm})\big)\big]$$
$$+\beta\ \mathrm{Min}\big[\gamma\big(N(\varrho_n,\varrho_n,\varrho_m,t_{norm})\big),\gamma\big(N(\mathfrak{F}(\varrho_m),\mathfrak{F}(\varrho_m),\varrho_m,t_{norm})\big),\gamma\big(N(\varrho_n,\varrho_n,\mathfrak{F}(\varrho_m),t_{norm})\big)\big]$$

Since $\mathfrak{F}$ is asymptotically regular at a point $\varrho_0 \in \mathcal{X}$ taking $n,m\to\infty$ then

$$\lim_{n,m\to\infty}\gamma\big(N(\mathfrak{F}(\varrho_n),\mathfrak{F}(\varrho_n),\mathfrak{F}(\varrho_m),t_{norm})\big)=0$$

and $\lim_{n,m\to\infty}\big(N(\mathfrak{F}(\varrho_n),\mathfrak{F}(\varrho_n),\mathfrak{F}(\varrho_m),t_{norm})\big)=1$

Thus $\{\varrho_n\}$ is a Cauchy's sequence in $\mathcal{X}$. since IVFNM $(\mathcal{X},N_{IVFNM},*_I)$ complete

then $\exists$, $p \in \mathcal{X}$ such that $\varrho_n\to p$(say).

$$\gamma\big(N(\mathfrak{F}(\varrho_n),\mathfrak{F}(\varrho_n),\mathfrak{F}(p),t_{norm})\big)$$
$$\leq \alpha\ \mathrm{Min}\big[\gamma\big(N(\varrho_n,\varrho_n,p,t_{norm})\big),\gamma\big(N(\mathfrak{F}(\varrho_n),\mathfrak{F}(\varrho_n),\varrho_n,t_{norm})\big),\gamma\big(N(\mathfrak{F}(\varrho_n),\mathfrak{F}(\varrho_n),p,t_{norm})\big)\big]$$
$$+\beta\ \mathrm{Min}\big[\gamma\big(N(\varrho_n,\varrho_n,p,t_{norm})\big),\gamma\big(N(\mathfrak{F}(p),\mathfrak{F}(p),p,t_{norm})\big),\gamma\big(N(\varrho_n,\varrho_n,\mathfrak{F}(p),t_{norm})\big)\big]$$

$$\gamma\big(N(\varrho_{n+1},\varrho_{n+1},\mathfrak{F}(p),t_{norm})\big)$$
$$\leq \alpha\ \mathrm{Min}\big[\gamma\big(N(\varrho_n,\varrho_n,p,t_{norm})\big),\gamma\big(N(\varrho_{n+1},\varrho_{n+1},\varrho_n,t_{norm})\big),\gamma\big(N(\varrho_{n+1},\varrho_{n+1},p,t_{norm})\big)\big]$$
$$+\beta\ \mathrm{Min}\big[\gamma\big(N(\varrho_n,\varrho_n,p,t_{norm})\big),\gamma\big(N(\mathfrak{F}(p),\mathfrak{F}(p),p,t_{norm})\big),\gamma\big(N(\varrho_n,\varrho_n,\mathfrak{F}(p),t_{norm})\big)\big]$$

$\Rightarrow \lim_{n\to\infty}\gamma\big(N(\varrho_{n+1},\varrho_{n+1},\mathfrak{F}(p),t_{norm})\big)=0.$

$\Rightarrow \gamma\big(N(\mathfrak{F}(p),\mathfrak{F}(p),p,t_{norm})\big)=0.$

$\Rightarrow \mathfrak{F}(p)=p.$

Hence p is the unique fixed point.

Theorem 4: Let a complete IVFNM $(\mathcal{X},N_{IVFNM},*_I)$ and γ be the altering distance function and $\mathfrak{F}$ and $\mathcal{L}$ be two commutative self mappings on $\mathcal{X}$ such that

$$\gamma\big(N(\mathfrak{F}(\varrho),\mathfrak{F}(\varrho),\mathfrak{F}(\sigma),t_{norm})\big) \le \lambda_1\big[\gamma\big(N(\mathcal{L}(\varrho),\mathcal{L}(\varrho),\mathcal{L}(\sigma),t_{norm})\big)+\lambda_2\gamma\big(N(\mathcal{L}(\varrho),\mathcal{L}(\varrho),\mathfrak{F}(\sigma),t_{norm})\big)+\gamma\big(N(\mathcal{L}(\sigma),\mathcal{L}(\sigma),\mathfrak{F}(\sigma),t_{norm})\big)\big]$$

Where $\forall\ \varrho,\sigma \in \mathcal{X}$, $t_{norm} > 0$ and $\lambda_1 : \mathbb{R} \to [I).\lambda_1 > 0,\ \lambda_2 < 1$.

Moreover, if

(i) $\mathfrak{F}$, $\mathcal{L}$ is asymptotically regular at a point $\varrho_0 \in \mathcal{X}$.

(ii) $\mathfrak{F}(\mathcal{X}) \subseteq \mathcal{L}(\mathcal{X})$.

(iii) $\mathfrak{F}(\mathcal{X})$ or $\mathcal{L}(\mathcal{X})$ is a complete subspace of $\mathcal{X}$.

Then $\mathfrak{F}$ and $\mathcal{L}$ has a unique common fixed point in $\mathcal{X}$.

Proof: Since $\mathfrak{F}(\mathcal{X}) \subseteq \mathcal{L}(\mathcal{X})$. Let a sequence $\{g_n\}$ in $\mathcal{X}$ where $g_0 \in \mathcal{X}$ and $g_{n+1} = \mathfrak{F}(\varrho_n) = \mathcal{L}(\varrho_{n+1})$, $n \in \mathbb{N} \cup \{0\}$, for all $n \ge 0$, and also $\mathfrak{F}$, $\mathcal{L}$ are asymptotically regular at a point $\varrho_0 \in \mathcal{X}$.

$\lim_{n\to\infty} \gamma\big(N(g_n, g_n, g_{n+1}, t_{norm})\big) = 0$

Then to prove $\{g_n\}$ is a Cauchy's sequence in $\mathcal{X}$.

Suppose to the contrary $\exists, 0 < \epsilon < 1, t_{norm} > 0$ and two sequences $\{u_n\}$ and $\{v_n\}$ such that $u_n > v_n > n$,

$$N_{IVFNM}\big(\varrho_{u_n}, \varrho_{u_n}, \varrho_{v_n} t_{norm}\big) \le 1-\epsilon,$$
$$N_{IVFNM}\big(\varrho_{u_n-1}, \varrho_{u_n-1}, \varrho_{v_{n-1}} t_{norm}\big) > 1-\epsilon$$
$$N_{IVFNM}\big(\varrho_{u_n-1}, \varrho_{u_n-1}, \varrho_{v_n} t_{norm}\big) > 1-\epsilon,\ \forall n \in \mathbb{N} \cup \{0\} \ldots\ldots (a)$$

Following concept applied in theorem 3.4 then we can show that

$\lim_{n\to\infty} N_{IVFNM}\big(\varrho_{u_n}, \varrho_{u_n}, \varrho_{v_n} t_{norm}\big) = 1-\epsilon,$

Now

$$\gamma\Big(N\big(\mathfrak{F}(\varrho_{u_n}),\mathfrak{F}(\varrho_{u_n}),\mathfrak{F}(\varrho_{v_n}),t_{norm}\big)\Big) \le \lambda_1\Big[\gamma\Big(N\big(\mathcal{L}(\varrho_{u_n}),\mathcal{L}(\varrho_{u_n}),\mathcal{L}(\varrho_{v_n}),t_{norm}\big)\Big) + \lambda_2\gamma\Big(N\big(\mathcal{L}(\varrho_{u_n}),\mathcal{L}(\varrho_{u_n}),\mathfrak{F}(\varrho_{v_n}),t_{norm}\big)\Big) + \gamma\Big(N\big(\mathfrak{F}(\varrho_{v_n}),\mathfrak{F}(\varrho_{v_n}),\mathcal{L}(\varrho_{v_n}),t_{norm}\big)\Big)\Big]$$

$$\gamma\Big(N\big(g_{u_n+1}, g_{u_n+1}, g_{v_n+1}, t_{norm}\big)\Big) \le \lambda_1\Big[\gamma\Big(N\big(\mathcal{L}(\varrho_{u_n}),\mathcal{L}(\varrho_{u_n}),\mathcal{L}(\varrho_{v_n}),t_{norm}\big)\Big) + \lambda_2\gamma\Big(N\big(\mathcal{L}(\varrho_{u_n}),\mathcal{L}(\varrho_{u_n}),\mathfrak{F}(\varrho_{v_n}),t_{norm}\big)\Big) + \gamma\Big(N\big(\mathfrak{F}(\varrho_{v_n}),\mathfrak{F}(\varrho_{v_n}),\mathcal{L}(\varrho_{v_n}),t_{norm}\big)\Big)\Big]$$

Taking $n \to \infty$ and using above facts then we have

$\gamma(1-\epsilon) \le \lambda_1[\gamma(1-\epsilon)] < \gamma(1-\epsilon)$ a contradiction. Hence $\{g_n\}$ is a Cauchy's sequence in $\mathcal{X}$.

Since $\mathcal{L}(\mathcal{X})$ is a complete subspace of $\mathcal{X}$ then $\exists, h \in \mathcal{L}(\mathcal{X})$ such that $\lim_{n\to\infty} g_n = h$ and also for some $z \in \mathcal{X}$ then $\mathcal{L}(z) = h$.

Now

$$\gamma\big(N(\mathfrak{F}(z),\mathfrak{F}(z),g_{n+1},t_{norm})\big) = \gamma\big(N(\mathfrak{F}(z),\mathfrak{F}(z),\mathfrak{F}(g_n),t_{norm})\big) \le \lambda_1\big[\gamma\big(N(\mathcal{L}(z),\mathcal{L}(z),\mathcal{L}(g_n),t_{norm})\big)+\lambda_2\gamma\big(N(\mathfrak{F}(z),\mathfrak{F}(z),\mathcal{L}(z),t_{norm})\big)+\gamma\big(N(\mathfrak{F}(g_n),\mathfrak{F}(g_n),\mathcal{L}(g_n),t_{norm})\big)\big]$$

As $n \to \infty$

$$\gamma\big(N(\mathfrak{F}(z),\mathfrak{F}(z),h,t_{norm})\big) \le \lambda_1\big[\gamma\big(N(h,h,h,t_{norm})\big)+\lambda_2\gamma\big(N(\mathfrak{F}(z),\mathfrak{F}(z),h,t_{norm})\big)+\gamma\big(N(h,h,h,t_{norm})\big)\big]$$

$$\gamma\big(N(\mathfrak{F}(z),\mathfrak{F}(z),h,t_{norm})\big) \le \lambda_1\big[\lambda_2\gamma\big(N(\mathfrak{F}(z),\mathfrak{F}(z),h,t_{norm})\big)\big]$$

$$\Rightarrow (1-\lambda_1\lambda_2)\gamma\big(N(\mathfrak{F}(z),\mathfrak{F}(z),h,t_{norm})\big) \le 0$$

$$\Rightarrow \gamma\big(N(\mathfrak{F}(z),\mathfrak{F}(z),h,t_{norm})\big) = 0$$

$$\Rightarrow \mathfrak{F}(z) = h$$

$\Rightarrow \mathfrak{F}(z) = h = \mathcal{L}(z)$. i.e. h is a coincident point of $\mathfrak{F}$ and $\mathcal{L}$.

Now put $\varrho = \mathfrak{F}(z)$ and $\sigma = z$

$$\gamma\left(N(\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(z),t_{norm})\right)$$
$$\leq \lambda_1\left[\gamma\left(N(\mathcal{L}(\mathfrak{F}(z)),\mathcal{L}(\mathfrak{F}(z)),\mathcal{L}(z),t_{norm})\right)\right.$$
$$\left.+\lambda_2\gamma\left(N(\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(\mathfrak{F}(z)),\mathcal{L}(\mathfrak{F}(z)),t_{norm})\right)+\gamma\left(N(\mathfrak{F}(z),\mathfrak{F}(z),\mathcal{L}(z),t_{norm})\right)\right]$$

$$\gamma\left(N(\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(z),t_{norm})\right)$$
$$= \lambda_1\left[\gamma\left(N(\mathfrak{F}(\mathcal{L}(z)),\mathfrak{F}(\mathcal{L}(z)),\mathcal{L}(z),t_{norm})\right)\right.$$
$$\left.+\lambda_2\gamma\left(N(\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(\mathcal{L}(z)),t_{norm})\right)+\gamma\left(N(\mathfrak{F}(z),\mathfrak{F}(z),\mathcal{L}(z),t_{norm})\right)\right]$$
(since $\mathfrak{F}\mathcal{L}=\mathcal{L}\mathfrak{F}$)

$$\gamma\left(N(\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(z),t_{norm})\right)$$
$$= \lambda_1\left[\gamma\left(N(\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(z),t_{norm})\right)\right.$$
$$\left.+\lambda_2\gamma\left(N(\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(\mathfrak{F}(z)),t_{norm})\right)+\gamma\left(N(\mathfrak{F}(z),\mathfrak{F}(z),\mathfrak{F}(z),t_{norm})\right)\right]$$

$$\Rightarrow(1-\lambda_1)\gamma\left(N(\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(z),t_{norm})\right)=0$$
$$\Rightarrow\gamma\left(N(\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(z),t_{norm})\right)=0$$
$$\Rightarrow\gamma\left(N(\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(\mathfrak{F}(z)),\mathfrak{F}(z),t_{norm})\right)=0$$
$$\Rightarrow\mathfrak{F}(\mathfrak{F}(z))=\mathfrak{F}(z)$$

Hence $\mathfrak{F}(\mathfrak{F}(z))=\mathfrak{F}(z)$

In similar manner we can show $\mathcal{L}(\mathcal{L}(z))=\mathcal{L}(z)=h$.

Thus h is unique common fixed point of $\mathfrak{F}$ and $\mathscr{L}$ in $\mathcal{X}$.

Uniqueness easily can be shown.

Corollary: Let a complete IVFNM $(\mathcal{X},N_{IVFNM},*_I)$ and γ be the altering distance function and $\mathfrak{F}$ and $\mathcal{L}$ be two self mappings on $\mathcal{X}$ such that

$$\gamma(N(\mathfrak{F}(\varrho),\mathfrak{F}(\varrho),\mathfrak{F}(\sigma),t_{norm}))$$
$$\leq\lambda_1[\gamma(N(\mathfrak{F}(\varrho),\mathfrak{F}(\varrho),\mathcal{L}(\varrho),t_{norm}))+\gamma(N(\mathfrak{F}(\sigma),\mathfrak{F}(\sigma),\mathcal{L}(\sigma),t_{norm}))]$$
$$+\lambda_2[\gamma(N(\mathfrak{F}(\varrho),\mathfrak{F}(\varrho),\mathcal{L}(\sigma),t_{norm}))+\gamma(N(\mathcal{L}(\varrho),\mathcal{L}(\varrho),\mathcal{L}(\sigma),t_{norm}))]$$
$$+\lambda_3\gamma(N(\lambda_3(\sigma),\mathfrak{F}(\sigma),\mathcal{L}(\varrho),t_{norm}))$$

Where $\forall\,\varrho,\sigma\in\mathcal{X}$, $t_{norm}>0$ and $\lambda_1>0$, $\lambda_2,\lambda_3<1$, $\lambda_1+2\lambda_2+\lambda_3<1$.

Moreover, if

(iv) $(\mathfrak{F},\mathcal{L})$ is asymptotically regular at a point ϱ_0 and commutes with each others.

(v) $\mathfrak{F}(\mathcal{X})$ or $\mathcal{L}(\mathcal{X})$ is a complete subspace of $\mathcal{X}$ and Range$(\mathfrak{F})\subseteq$ Range$(\mathcal{L})$.

Then $\mathfrak{F}$ and $\mathscr{L}$ has a unique common fixed point in $\mathcal{X}$.

Example: Let IVFNM $(\mathcal{X},N_{IVFNM},*_I)$ and $\mathcal{X}=\{1,2,3,4,5\}$ and a interval valued N-fuzzy set on $\mathcal{X}^3\times(0,\infty)$ defined by $N_{IVFNM}(x,y,z,t_{norm})=\frac{t}{t+[|x-z|+|y-z|]}$, for all $x,y,z\in\mathcal{X}$ and $t_{norm}>0$. then $(\mathcal{X},N_{IVFNM},*_I)$ is complete interval valued N-fuzzy space w.r.t t_{norm} T(a , b)= min(a ,b), a ,b ∈ [0,1]. Let $\gamma(\alpha)=1-\alpha$ and define $\mathfrak{F},\mathcal{L}:\mathcal{X}\to\mathcal{X}$ by $\mathfrak{F}(1)=\mathfrak{F}(2)=2,\mathfrak{F}(3)=\mathfrak{F}(4)=\mathfrak{F}(5)=1$ and $\mathcal{L}(1)=\mathcal{L}(2)=2,\mathcal{L}(3)=\mathcal{L}(4)=\mathcal{L}(5)=1$. then $\mathfrak{F}$ and $\mathscr{L}$ satisfy theorem 3.3 and $\mathfrak{F}$ and $\mathscr{L}$ has a unique common fixed point in $\mathcal{X}$.

IV. Conclusion

In this research, we verify certain well-known fixed point conclusions using asymptotically regular mappings in interval valued fuzzy N-metric space (IVFMN). Our results generalize and extend the results of the common fixed point theorem for interval valued fuzzy N-metric space (IVFMN), which were published recently. For future research, these findings can be applied to solve fuzzy

differential equations and integral equations, which will be of great assistance to the researchers.

References

[1] Deshmukh, V., Jadon, B. P. S., Singh, R., and Malhotra, S. K. (2023). Fixed point results on interval valued fuzzy metric space using notation of pairwise compatible maps with application, *Asian Research Journal of mathematics,* 19(11), 104-114.

[2] Dolas, U. (2016). A Common Fixed point theorem in Fuzzy Metric space using E.A like property, *Ultra Scientist,* 28(1), 1-6.

[3] George, A. and Veeramani, P. (1994). On some results in fuzzy metric space, *Fuzzy sets and systems,* 64, 395-399.

[4] Goswami, N. and Patir, B. (2019). Fixed points theorems for asymptotically regular mappings in fuzzy metric spaces, *Korean J. Math.,* 27(4), 861-877.

[5] Hooda, H., Malik, A., and Vats, M. (2020). Some fixed points results for asymptotically regular maps in N-Fuzzy b-metric space, *Turkish Journal of computer and Mathematics education,* 11(01), 647-656.

[6] Jungck, G., Murthy P.P., and Cho, Y.J. (1993). Compatible mappings of type (A) and common fixed points, *Math. Japan,* 38(2), 381-390.

[7] Kaleva, O., Seikkala, S. (1984). On fuzzy metric space, *Fuzzy sets and systems,* 12(3), 215-229.

[8] Khan, M.S., Swaleh, M., and Sessa, S. (1984). Fixed point theorems by altering distances between the points, *Bull. Aust. Math. Soc.,* 30, 1-9.

[9] Kotuku, S., Sharma , S. and Tokgoz, H. (2007). A Fixed Point Theorem in Fuzzy metric Spaces, *Int. Journal of Math. Ana.,* 1 (18), 861– 872.

[10] Kramosil, I., and Michalek, J. (1975). Fuzzy Metric and statistical metric spaces, *Kybernetika,* 11(5), 336-344.

[11] Li, C. (2009). Distance between interval-valued fuzzy sets, In NAFIPS 2009-2009 *Annual Meeting of the North American fuzzy Information Processing Society*, IEEE, 1-3.

[12] Malviya, N. (2015). The N-fuzzy metric space and mapping with application, *Fasciculi Mathematici.*

[13] Nigam, P., and Pagey, S.S. (2012). Some fixed point theorems for a pair of asymptotically regular and compatible mappings in fuzzy 2-metric, *Int. j. Open Problems Comput. Maths*., 5(1), 77-84.

[14] Osman, MT Abu. (1983). Fuzzy metric spaces and fixed fuzzy sets theorem, *Bull. Malayasian Math.,* 6(1), 1-4.

[15] Popa, V. (2001). Some fixed point theorems for weakly compatible mappings, *Radovi Mathematics,* 10, 245-252.

[16] Purdhvi, K. (2014). A common fixed point theorem for asymptotically regular in cone metric spaces. *Asian Journal of fuzzy and applied mathematics,* 2(1), 12-16.

[17] Sastry, K.P.R., Naidu, V.S.R., Rao, I.H.N. and Rao, K.P.R. (1984). Common fixed point results for asymptotically regular mappings, *Indian J. Pure Appl. Math.,* 15(8), 849-854.

[18] Shen, Y., Li, H., and Wang, F. (2012). on interval valued fuzzy metric spaces, *International Journal of fuzzy systems,* 14(1), 35-44.

[19] Shende, S., et. al. (2021). Some fixed point theorems for asymptotically regular maps in N-Fuzzy metric space, *J. of Science and Technological researchers*, 3(3).

[20] Singh, B., and Jain, S. (2005). Semi Compatibility and fixed point theorems in fuzzy metric space using implicit relation, *Int. J. of Math. Sciences,* 2617-2629.

[21] Singh, U.S. and Singh, N. (2019). Common fixed point theorem for rational expression in fuzzy metric space, *Journal of the Gujrat research society,* 21(14).

[22] Wadhaw, K., and Dubey, A. (2017). Common Fixed point theorems using E.A like property in Fuzzy Metric spaces, *International J.of Math.,* 8(6), 204-210.

[23] Wadhwa, K., Dubey, H., and Jain, R. (2013). Impact of E.A like Property on Common fixed point theorems in Fuzzy Metric spaces, *J. of Adv. Stud. In Topol.*, 3(1), 609-614.

[24] Yadav, D. S., and Thakur, S.S. (2013). Common Fixed Point for R-Weakly Commuting Mapping in Fuzzy 2-Metric spaces, *Int. J. Contemp. Math. Sci.*, 8(13), 609-614.

[25] Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning –I, *Information Sciences*, 8(3), 199-249.

OPTIMIZATION OF PERISHABLE INVENTORY IN TWO-WAREHOUSE SYSTEM USING GENETIC ALGORITHMS WITH LOG-GAMMA DECAY AND NONLINEAR DEMAND

Garima Sethi[1], Ajay Singh Yadav[1,*], Chaman Singh[2]

•

[1]Department of Mathematics, SRM Institute of Science and Technology,
Delhi-NCR Campus, Ghaziabad, India, 201204
[2]Department of Mathematics, Acharya Narendra Dev College,
Delhi, India, 110019
ginusethi999@gmail.com, ajaysiny@srmist.edu.in, xyz@gmail.com
*Corresponding author

Abstract

This study presents a detailed inventory framework for managing perishable products distributed across two storage facilities: a capacity-constrained, company-owned primary warehouse and an auxiliary rented warehouse. The model accounts for inventory shortages with partial backlogging, where both the demand rate and backlogging rate are modeled as generalized exponentially decreasing functions of time and selling price. When the order quantity surpasses the storage limit of the primary warehouse, the excess inventory is stored in the rented facility. To achieve cost efficiency, a genetic algorithm (GA) is employed to optimize the primary decision variables, following the policy that items in the rented warehouse are depleted first. As a result, inventory in the rented facility diminishes due to both product deterioration and customer demand, whereas the primary warehouse stock is affected solely by deterioration. Once the inventory in the owned warehouse is exhausted after a finite period, partial backlogging of shortages begins. The model incorporates a log-gamma function to represent the deterioration rate and a quadratic function to describe demand, aiming to reflect realistic product dynamics. Numerical examples validate the effectiveness of the GA-based optimization strategy under diverse operating conditions. Additionally, a sensitivity analysis explores the influence of critical parameters on the optimal inventory policy and overall cost, offering practical insights for decision-makers.

Keywords: Inventory management, deteriorating items, two warehouses, Partial backlogging, nonlinear demand, Genetic algorithm.

1. Introduction

Efficient inventory control is vital for businesses handling perishable goods across multiple storage facilities. With shifting market trends and changing consumer behavior, there is an increasing demand for flexible models that can incorporate various real-world complexities. This study presents an advanced inventory optimization method leveraging a Genetic Algorithm (GA) to manage the distribution of deteriorating items between two distinct warehouses. The proposed model effectively tackles the issue of shortages by incorporating partial backlogging, wherein demand is dynamically influenced by both the selling price and the passage of time. Managing inventories of perishable goods requires a thorough understanding of the interplay between storage

capacity, product deterioration, and demand dynamics. When the order quantity surpasses the storage limit of the primary (owned) warehouse, the excess inventory is systematically redirected to a secondary (rented) facility. To optimize overall storage costs, a Genetic Algorithm (GA) is utilized to ensure that items from the rented warehouse are used first. This strategy facilitates a gradual depletion of the rented warehouse inventory over time, driven by both product decay and customer demand, whereas stock in the primary warehouse diminishes only due to deterioration. Once the inventory in the owned warehouse is fully exhausted after a specific time period, the system transitions into a shortage phase, during which partial backlogging may occur.

The proposed model is built on the premise that both the backlogging rate and demand follow generalized exponentially decreasing functions with respect to selling price (p) and time (t). This formulation captures the complexities of real-world inventory systems, where pricing strategies and time-dependent factors significantly influence consumer behavior and inventory decisions. The paper outlines the development of this model, emphasizing the role of the Genetic Algorithm (GA) in optimizing key aspects such as inventory control, surplus stock allocation, prioritization of withdrawals from the rented warehouse, and anticipation of shortage periods. To demonstrate the model™s practical utility, several numerical examples are presented, highlighting its flexibility and performance across varying operational conditions. Furthermore, a sensitivity analysis is performed to examine how changes in critical parameters affect the model™s outcomes, offering valuable insights into its reliability and real-world applicability.

In conclusion, this paper introduces a novel and flexible inventory management framework that utilizes Genetic Algorithm (GA) techniques to tackle the complexities of handling perishable goods stored across two distinct warehouses. By incorporating generalized exponential decreasing functions dependent on selling price and time, the model effectively addresses the dual challenges of cost optimization and shortage reduction in dynamic market settings. The following sections delve into simulation methodologies, sensitivity analyses, and the practical tools necessary for implementation. The objective is not only to develop a theoretical model but also to offer actionable insights into its applicability and resilience in real-world operations. Notably, the foundational concept of managing inventory across owned and rented warehouses was first explored by Hartley [38] in 1976.

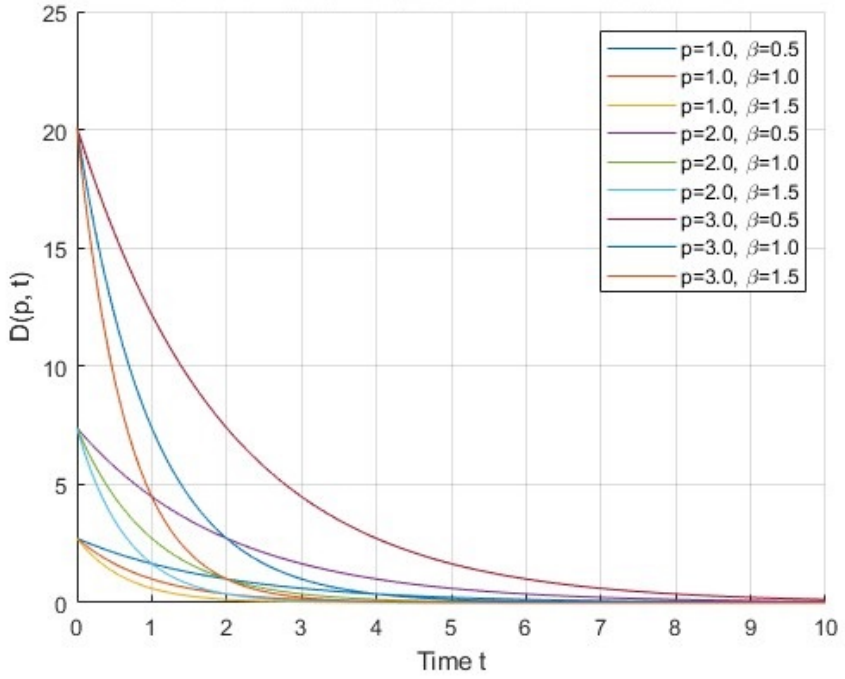


Figure 1: *Effect of time on demand for various p and β.*

2. Related Work

Supply Chain Management (SCM) involves the strategic coordination of production, storage, location, and transportation activities among various entities in the supply chain to optimize responsiveness and efficiency tailored to specific market needs. While much of the existing research has traditionally focused on non-perishable goods, a significant number of products are subject to deterioration over time. Additionally, inventory-related expenses"such as carrying and deterioration costs"tend to be higher in the Regional Warehouse (RW) compared to the Outlet Warehouse (OW). This is largely due to extra expenditures associated with maintenance, material handling, and advanced storage facilities. Since customer service operations take place exclusively at the OW, it's crucial that stock is consistently available there to ensure smooth business operations. As a result, inventory is initially stored in the OW, and any excess is held in the RW. To maintain full stock levels at the OW, inventory is continuously transferred from the RW, which is depleted first in the supply process.

The concept of the two-warehouse inventory system was initially introduced by Hartley [38] in 1976 through his work Operations Research"A Managerial Emphasis. In his formulation, the cost of transporting goods from the regional warehouse (RW) to the outlet warehouse (OW) was omitted. Sarma [1] later enhanced this model by including a fixed transportation cost from RW to OW, independent of the quantity moved. Building on this, Goswami and Chaudhuri [2] incorporated time-varying demand with a linear increase and made the transportation cost quantity-dependent, while also addressing scenarios with and without shortages. Further contributions by Pakkala and Achary [3"4] included finite replenishment rates and shortages within discrete and continuous time settings, though transportation costs were not considered in their models. Benkherouf [5], and Bhunia and Maiti [6"7], extended the framework to account for deteriorating items and time-dependent demand across dual storage locations. These earlier models generally assumed a continuous and direct release of inventory from both warehouses. Murdeshwar and Sathi [8], along with Pakkala and Achary [9], introduced a bulk release mechanism, where inventory is first transferred from RW to OW before being dispatched to customers. Despite these advancements, most models assumed an infinite planning horizon with demand parameters reset at the start of each cycle. Addressing this limitation, Kar et al. [10] proposed a two-warehouse model under a finite planning horizon with linearly time-dependent demand. Lee and Ma [11] later developed a heuristic-based approach for equal production cycle times under general time-dependent demand within a finite horizon. More recently, Lee and Hsu [12] extended this model by incorporating a finite replenishment rate and employing a variable production cycle time (VPCT) strategy to determine both the number and timing of production cycles over a finite planning period.

In further studies, Yadav and Kumar [13] investigated inventory control for electronic storage devices, integrating environmental concerns and network-based logistics. Yadav, A.S. [13-15] explored seven performance metrics in supply chain management aimed at enhancing inventory efficiency, using optimization tools like Genetic Algorithms (GA) and Particle Swarm Optimization (PSO). This research also addressed economic aspects of transportation and stock replenishment through computational intelligence techniques. Swami et al. [16-18] developed inventory strategies that incorporate factors such as payment delays and credit terms, presenting a depreciation model that relates to a variety of goods and services, particularly under inflationary conditions. Similarly, Gupta et al. [19-20] proposed multi-objective optimization frameworks using GA and PSO to balance supply sufficiency and inflation control, introducing models to assess scenarios of scarcity and low inflation. Singh et al. [21,22] examined inventory systems involving two different stock types, analyzing asset depreciation and storage costs using soft computing approaches. Their study also considered the effects of inflation and property risk in dual-inventory setups. Kumar et al. [23-25] focused on managing the supply of alcoholic beverages, developing environmentally friendly supply chain models and leveraging genetic algorithms for inventory and distribution improvements. Chauhan and Yadav [26-28] provided models for managing stock depreciation across two storage sites, utilizing vehicle routing and genetic inventory strategies to

address fluctuating demand and inflation across distribution networks. Pandey et al. [29] applied genetic and multi-particle techniques to enhance the management of industrial marble reserves. Meanwhile, Ahlawat et al. [30] studied SCM in the white wine industry, utilizing neural networks to optimize operations. Singh et al. [31] proposed policies for efficiently importing damaged goods and managing payment deferrals within two-warehouse frameworks.

Yadav et al. [32-33] focus on enhancing inventory management for perishable goods by incorporating green technology investments. Their study evaluates the influence of selling price, carbon emissions, and time-sensitive demand, aiming to develop sustainable and efficient inventory strategies. In a related study, Yadav, Yadav, and Bansal [34, 39], Sinha, et at. [35], Hazer [36], Negi and Singh [37] are applied an interval number approach to model a two-warehouse inventory system for perishable items, effectively addressing uncertainties in both demand and cost parameters. Their findings suggest that investments in preservation technologies not only mitigate product waste but also improve inventory utilization. Adak and Mahapatra [40] developed a multi-item inventory model incorporating reliability, time-dependent demand, and deterioration, while allowing shortages with partial backlogging to enhance decision-making under complex inventory scenarios. Mebarek-Oudina, F [41-44] are also used the advanced techniques for enhancement and some other factors. These models highlight the significance of managing degradation-related expenses to optimize inventory levels and enhance overall system performance.

3. Notations and assumptions

3.1. Notations

The following notations are used in this model.

Parameters	Descriptions
A	Ordering cost coefficient
h_1	Coefficient of holding cost of Rented Warehouse (RW).
h_2	Coefficient of holding cost of Owned Warehouse (OW).
C_P	Purchasing cost.
C_S	Shortage cost.
C_L	Coefficient of cost of lost sale.
θ	Constant rate of deterioration
C_D	Deterioration cost per unit.
R	Inflation factor
q_1	Positive height of inventory of (RW) with $I(t=0)$
q_2	Positive height of inventory of (OW) with $I(t=0)$
q_3	The Negative height of inventory with $I(t=T)$
Q	Total order quantity of order.
T	Total cycle time (Total cycle length).
t_1	The time where inventory height of rented Warehouse becomes zero.
t_2	The time where inventory height of Owned Warehouse becomes zero.
$I_1(t)$	The height of inventory in rented warehouse between time intervals $[0, t_1]$
$I_2(t)$	The height of inventory in owned warehouse between time intervals $[0, t_1]$
$I_3(t)$	Height of inventory in owned warehouse between the time intervals $[t_1, t_2]$
$I_4(t)$	Level of inventory in owned Warehouse between the time intervals $[t_2, T]$
PC	Cost of purchasing
HC	Cost of holding of inventory.
SC	Cost of shortage of the inventory.
LC	Cost of lost sale cost of inventory.
TAC	Present total Average cost.

3.2. Assumptions

The following assumptions are used in this paper.

1. The demand rate is generalized exponential function of selling price and time and taken as the following form : $D(p,t) = e^{p-\beta t};\ \ p > 0,\ \beta > 0.$
2. The decaying rate is $\theta(t) = \alpha\beta e^{\beta t};\ \ \alpha, \beta > 0.$
3. Infinite Time horizon is considered.
4. Lead time is zero with Infinite Replenishment rate is taken.
5. Warehouse (OW) has the limited space is allowed. On other hand the unlimited space area for rented warehouse has been permitted.
6. The holding cost (h_1) of the of Rented Warehouse is greater than the holding cost (h_2) of Owned Warehouse.
7. The charges for transportation and time between Rented Warehouse and Owned Warehouse are completely ignored.

4. FORMULATION AND SOLUTION OF THE MODEL

The issue which we have discussed here is the means by which retailers know whether or not to take a rented warehouse to hold the things. If the order quantity. The following is the formulation of the proposed model:

$$\frac{dI_r}{dt} + \alpha\beta e^{\beta t} I_r = -t^2, \qquad 0 \le t \le t_1 \tag{1}$$

$$\frac{dI_o}{dt} + \alpha\beta e^{\beta t} I_o = -t^2, \qquad 0 \le t \le t_1 \tag{2}$$

The solutions of differential equations (1), (2) are as follows:

$$I_r(t) = e^{-\theta(t_1 - t)} \left(\frac{K e^{a+bt_1}}{b+\theta} \right) - \frac{K e^{a+bt}}{b+\theta} \tag{3}$$

$$I_o(t) = q_2 e^{-\theta t} \tag{4}$$

Positive inventory level of rented warehouse with $I_r(0) = q_1$and Eqs. (5) given by

$$q_1 = e^{\theta t_1} \left(\frac{K e^{a+bt_1}}{b+\theta} \right) - \frac{K e^{a}}{b+\theta} \tag{5}$$

Negative inventory level with $I_o(T) = -q_3$ and Eqs. (8) given by

$$q_3 = \left(\frac{K e^{a+(b-\delta)T}}{b-\delta} \right) - \frac{K e^{a+(b-\delta)t_2}}{b-\delta} \tag{6}$$

Next, we calculate all the associated inventory costs as follow:

1. Ordering Cost (OC):
$$OC = A \tag{7}$$
2. Purchasing Cost (PC):
$$PC = C_P \left[e^{\theta t_1} \left(\frac{K e^{a+bt_1}}{b+\theta} \right) - \frac{K e^{a}}{b+\theta} + q_2 + \left(\frac{K e^{a+(b-\delta)T}}{b-\delta} \right) - \frac{K e^{a+(b-\delta)t_2}}{b-\delta} \right] \tag{8}$$

3. Shortage Cost (SC):

$$SC = -C_S \int_{t_2}^{T} I_o(t) \cdot e^{-Rt} dt$$

$$SC = -C_S \int_{t_2}^{T} \left[\frac{Ke^{a+(b-\delta)t_2}}{b+\delta} - \frac{Ke^{a+(b-\delta)t}}{b+\delta} \right] . e^{-Rt} dt$$

$$SC = -\frac{KC_S}{(b-\delta)} \left[\frac{e^{a+(b-\delta)t_2 - RT}}{-R} - \frac{e^{a+(b-\delta-R)T}}{b-\delta-R} - \frac{(b-\delta)e^{a+(b-\delta-R)t_2}}{R(b-\delta-R)} \right] \quad (9)$$

4. Lost sales Cost (LC):

$$LC = -C_L \int_{t_2}^{T} (1 - B(t)) \cdot D(t) \cdot e^{-Rt} dt$$

$$LC = -KC_L \left[\left(\frac{e^{a+(b-R)T}}{b-R} - \frac{e^{a+(b-\delta-R)T}}{b-\delta-R} \right) - \left(\frac{e^{a+(b-R)t_2}}{b-R} - \frac{e^{a+(b-\delta-R)t_2}}{b-\delta-R} \right) \right] \quad (10)$$

5. Holding Cost (HC):

$$HC = \left[h_1 \int_0^{t_1} I_r(t) . e^{-Rt} dt + h_2 \int_0^{t_1} I_o(t) \cdot e^{-Rt} dt + h_3 \int_0^{t_2} I_o(t) . e^{-Rt} dt \right]$$

$$= \left\{ \frac{h_1 k}{(b+\theta)} \left[\frac{-(b+\theta)e^{a+(b-R)t_1}}{(\theta+R)(b-R)} + \frac{e^{a+(b-R)t_1}}{(\theta+R)} + \frac{e^a}{(b-R)} \right] + \frac{h_2 q_2}{(R+\theta)} \left(1 - e^{-(R+\theta)t_1} \right) \right.$$

$$\left. + \frac{h_2 k}{(b+\theta)} \left[\frac{-(b+\theta)e^{a+(b-R)t_2}}{(\theta+R)(b-R)} + \frac{e^{a+(b+\theta)t_2-(\theta+R)t_1}}{(\theta+R)} + \frac{e^{a+(b-R)t_1}}{(b-R)} \right] \right\} \quad (11)$$

6. Deterioration Cost (DC):

$$DC = C_D \left[\theta \int_0^{t_1} I_r(t) \cdot e^{-Rt} dt + \theta \int_0^{t_1} I_o(t) \cdot e^{-Rt} dt + \theta \int_0^{t_2} I_o(t) \cdot e^{-Rt} dt \right]$$

$$= C_D \left\{ \frac{\theta k}{(b+\theta)} \left[\frac{-(b+\theta)e^{a+(b-R)t_1}}{(\theta+R)(b-R)} + \frac{e^{a+(b-R)t_1}}{(\theta+R)} + \frac{e^a}{(b-R)} \right] + \frac{\theta q_2}{(R+\theta)} \left(1 - e^{-(R+\theta)t_1} \right) \right.$$

$$\left. + \frac{\theta k}{(b+\theta)} \left[\frac{-(b+\theta)e^{a+(b-R)t_2}}{(\theta+R)(b-R)} + \frac{e^{a+(b+\theta)t_2-(\theta+R)t_1}}{(\theta+R)} + \frac{e^{a+(b-R)t_1}}{(b-R)} \right] \right\} \quad (12)$$

$$TAC = \frac{1}{T} \left[OC + PC + HC + DC - SC + LSC \right] \quad (13)$$

5. Genetic Algorithm Optimization (GAO) Methodology

Chromosomes (Genotypes): In a genetic algorithm, a potential solution to the optimization problem is encoded as a chromosome or genotype. The chromosome is typically represented as a string of symbols, often binary digits (0s and 1s), but it can be adapted to handle other representations.

Population: A population is a collection of individuals, where each individual represents a potential solution to the problem. The population evolves over generations, with each generation consisting of a set of individuals.

Fitness Function: The fitness function evaluates how well an individual solves the problem. It assigns a numerical value (fitness score) to each individual based on its performance. The goal is to maximize or minimize this fitness score, depending on the nature of the optimization problem.

Selection: Individuals are selected from the current population based on their fitness scores. High-fitness individuals have a higher probability of being selected. This mimics the process of

natural selection, where individuals with better adaptability have a higher chance of reproducing.
Crossover (Recombination): Crossover involves taking two parent individuals and creating new offspring by combining their genetic material. This is inspired by genetic recombination in biological reproduction. Different crossover techniques, such as one-point crossover or uniform crossover, can be used.
Mutation: Mutation introduces small random changes in an individuals chromosome. This helps to explore the search space more extensively, preventing the algorithm from getting stuck in local optima.
Termination Criteria: The algorithm stops when a certain condition is met, such as reaching a maximum number of generations, achieving a satisfactory solution, or a combination of factors.

6. Graphical Representation

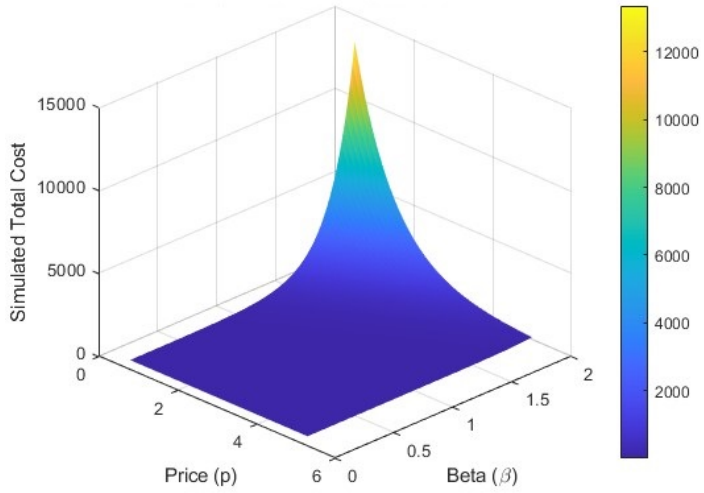


Figure 2: *Optimization with respect to decision variables.*

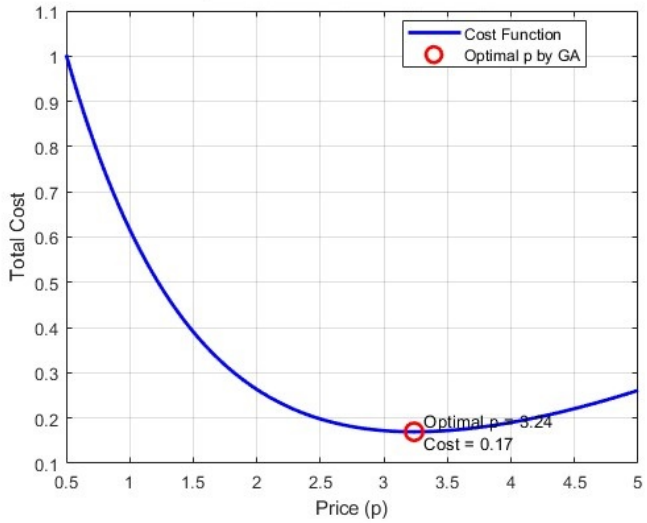


Figure 3: *Variation between total cost and price.*

The graphical depictions in this part provide a comprehensive grasp of the dynamics within the deteriorated inventory model of two-warehouses. Figure 2 and Figure 3 explains the link between demand rates and time, while Figure 2 illustrates how shortage impacts inventory management using different type of functions of time together with the adjustments needed to effectively balance demand, shortage, and time.

7. Conclusion

In conclusion, this paper introduces a novel Genetic Algorithm-based inventory management model tailored for businesses dealing with deteriorating items stored in two warehouses. The model addresses the complexities of shortages through a dynamic approach to partial backlogging, where demand is intricately linked to both selling price and time. The integration of the Genetic Algorithm provides a sophisticated optimization tool for strategic decision-making in the allocation and release of stock, emphasizing cost-effectiveness and adaptability to evolving market dynamics.

The models foundation on generalized exponential decreasing functions for backlogging rate and demand, considering selling price (p) and time (t), enhances its realism and applicability to real-world scenarios. This reflects the understanding that pricing dynamics and temporal considerations are integral factors influencing inventory management decisions. Throughout this paper, we have explored the models development, elucidating the genetic algorithms role in optimizing inventory management, handling excess stock, prioritizing rented warehouse releases, and predicting the onset of shortages. Numerical examples have been presented to showcase the practical implementation of the model, emphasizing its adaptability and efficacy in diverse business contexts. Furthermore, a comprehensive sensitivity analysis has been conducted to assess the models robustness and behavior under varying parameters. This analysis contributes valuable insights into the models performance and its suitability for different operational conditions.

In summary, the presented Genetic Algorithm-based inventory management model offers a sophisticated and adaptive solution for businesses seeking effective strategies in the face of deteriorating item inventories. By incorporating cutting-edge optimization techniques and considering influential factors such as selling price and time, the model provides a valuable framework for businesses to minimize storage costs and navigate the challenges of shortages in dynamic market environments. Future research may focus on further refining the model, expanding its applicability to diverse industries, and incorporating additional factors for a more comprehensive understanding of inventory dynamics.

References

[1] Sarma KVS (1987) A deterministic order-level inventory model for deteriorating items with two storage facilities. Eur J Oper Res 29:70“72
[2] Goswami A, Chaudhuri KS (1992) An economic order quantity model for items with two levels of storage for a linear trend in demand. J Oper Res Soc 43:157“167
[3] Pakkala TPM, Achary KK (1992a) Discrete time inventory model for deteriorating items with two warehouses. Opsearch 29:90“103
[4] Pakkala TPM, Achary KK (1992b) A deterministic inventory model for deteriorating items with two warehouses and finite replenishment rate. Eur J Oper Res 57:71“76
[5] Benkherouf L (1997) A deterministic order level inventory model for deteriorating items with two storage facilities. Int J Prod Econ 48:167“175
[6] Bhunia AK, Maiti M (1997) A deterministic inventory replenishment problem for deteriorating items with time-dependent demand and shortages for the finite time horizon. Opsearch 34(1):51“61
[7] Bhunia AK, Maiti M (1998) A two warehouses inventory model for deteriorating items with a linear trend in demand and shortages. J Oper Res Soc 49:287“292

[8] Murdeshwar TM, Sathi YS (1985) Some aspects of lot-size models with two levels of storage. Opsearch 22:255“262
[9] Pakkala TPM, Achary KK (1994) Two level storage inventory model for deteriorating items with bulk release rule. Opsearch 31:215“227
[10] Kar S, Bhunia AK, Maiti M (2001) Deterministic inventory model with two-levels of storage, a linear trend in demand and a fixed time horizon. Comput Oper Res 28:1315“1331
[11] Lee CC, Ma CY (2000) Optimal inventory policy for deteriorating items with two-warehouse and time dependent demand. Prod Plan Control 11(7):689“696
[12] Lee CC, Hsu SL (2009) A two-warehouse production model for deteriorating inventory items with timedependent demands. Eur J Oper Res 194:700“710
[13] Yadav, A.S. and Kumar, S. (2017) Electronic Components Supply Chain Management for Warehouse with Environmental Collaboration & Neural Networks. International Journal of Pure and Applied Mathematics, Volume 117 No. 17, 169-177.
[14] Yadav, A.S. (2017) Analysis of Seven Stages Supply Chain Management in Electronic Component Inventory Optimization for Warehouse with Economic Load Dispatch Using GA and PSO. Asian Journal Of Mathematics And Computer Research, volume 16 No.4, 208-219.
[15] Yadav, A.S. (2017) Analysis Of Supply Chain Management In Inventory Optimization For Warehouse With Logistics Using Genetic Algorithm International Journal of Control Theory And Applications, Volume 10 No.10, 1-12 .
[16] Yadav, A.S. (2017) Modeling and Analysis of Supply Chain Inventory Model with two-warehouses and Economic Load Dispatch Problem Using Genetic Algorithm. International Journal of Engineering and Technology, Volume 9 No 1, 33-44.
[17] Swami, A., Singh, S.R., Pareek, S. and Yadav, A.S. (2015) Inventory policies for deteriorating item with stock dependent demand and variable holding costs under permissible delay in payment. International Journal of Application or Innovation in Engineering & Management, Volume 4, Issue 2, 89-99.
[18] Swami, A., Pareek, S., Singh S.R. and Yadav, A.S. (2015) Inventory Model for Decaying Items with Multivariate Demand and Variable Holding cost under the facility of Trade-Credit. International Journal of Computer Application, 18-28.
[19] Swami, A., Pareek, S., Singh, S.R. and Yadav, A.S. (2015) An Inventory Model With Price Sensitive Demand, Variable Holding Cost And Trade-Credit Under Inflation. International Journal of Current Research, Volume 7, Issue, 06, 17312-17321.
[20] Gupta, K., Yadav, A.S., Garg, A. and Swami, A. (2015) A Binary Multi-Objective Genetic Algorithm & PSO involving Supply Chain Inventory Optimization with Shortages, inflation. International Journal of Application or Innovation in Engineering & Management, Volume 4, Issue 8, 37-44.
[21] Gupta, K., Yadav, A.S., Garg, A., (2015) Fuzzy-Genetic Algorithm based inventory model for shortages and inflation under hybrid & PSO. IOSR Journal of Computer Engineering, Volume 17, Issue 5, Ver. I , 61-67.
[22] Singh, R.K., Yadav, A.S. and Swami, A. (2016) A Two-Warehouse Model for Deteriorating Items with Holding Cost under Particle Swarm Optimization. International Journal of Advanced Engineering, Management and Science, Volume -2, Issue-6, 858-864.
[23] Singh, R.K., Yadav, A.S. and Swami, A. (2016) A Two-Warehouse Model for Deteriorating Items with Holding Cost under Inflation and Soft Computing Techniques. International Journal of Advanced Engineering, Management and Science, Volume -2, Issue-6, 869-876.
[24] Kumar, S., Yadav, A.S., Ahlawat, N. and Swami, A. (2019) Supply Chain Management of Alcoholic Beverage Industry Warehouse with Permissible Delay in Payments using Particle Swarm Optimization. International Journal for Research in Applied Science and Engineering Technology, Volume 7 Issue VIII, 504-509.
[25] Kumar, S., Yadav, A.S., Ahlawat, N. and Swami, A. (2019) Green Supply Chain Inventory System of Cement Industry for Warehouse with Inflation using Particle Swarm Optimization. International Journal for Research in Applied Science and Engineering Technology, Volume 7 Issue VIII, 498-503.

[26] Kumar, S., Yadav, A.S., Ahlawat, N. and Swami, A. (2019) Electronic Components Inventory Model for Deterioration Items with Distribution Centre using Genetic Algorithm. International Journal for Research in Applied Science and Engineering Technology, Volume 7 Issue VIII, 433-443.

[27] Chauhan, N. and Yadav, A.S. (2020) An Inventory Model for Deteriorating Items with Two-Warehouse & Stock Dependent Demand using Genetic algorithm. International Journal of Advanced Science and Technology, Vol. 29, No. 5s, 1152-1162.

[28] Chauhan, N. and Yadav, A.S. (2020) Inventory System of Automobile for Stock Dependent Demand & Inflation with Two-Distribution Center Using Genetic Algorithm. Test Engraining & Management, Volume 83, Issue: March - April, 6583 - 6591.

[29] Pandey, T., Yadav, A.S. and Medhavi Malik (2019) An Analysis Marble Industry Inventory Optimization Based on Genetic Algorithms and Particle swarm optimization. International Journal of Recent Technology and Engineering Volume-7, Issue-6S4, 369-373.

[30] Ahlawat, N., Agarwal, S., Pandey, T., Yadav, A.S., Swami, A. (2020) White Wine Industry of Supply Chain Management for Warehouse using Neural Networks Test Engraining & Management, Volume 83, Issue: March - April, 11259 - 11266.

[31] Singh, S. Yadav, A.S. and Swami, A. (2016) An Optimal Ordering Policy For Non-Instantaneous Deteriorating Items With Conditionally Permissible Delay In Payment Under Two Storage Management International Journal of Computer Applications, Volume 147 -No.1, 16-25.

[32] Yadav, A.S., Kumar, A., Yadav, K.K. et al. Optimization of an inventory model for deteriorating items with both selling price and time-sensitive demand and carbon emission under green technology investment. Int J Interact Des Manuf (2023). https://doi.org/10.1007/s12008-023-01689-8

[33] Yadav, K.K., Yadav, A.S. & Bansal, S. Interval number approach for two-warehouse inventory management of deteriorating items with preservation technology investment using analytical optimization methods. Int J Interact Des Manuf (2024). https://doi.org/10.1007/s12008-023-01672-3

[34] Krishan Kumar Yadav , Ajay Singh Yadav, and Shikha Bansal. "Optimization of an inventory model for deteriorating items assuming deterioration during carrying with two-warehouse facility" Reliability: Theory & Applications, vol. 19, no. 3 (79), 2024, pp. 442-459. doi:10.24412/1932-2321-2024-379-442-459

[35] Sinha, Amit Kumar, and Ankush Anand. "Optimizing supply chain network for perishable products using improved bacteria foraging algorithm." Applied Soft Computing 86 (2020): 105921.

[36] Hezer, Seda, and Yakup Kara. "Solving vehicle routing problem with simultaneous delivery and pick-up using an algorithm based on bacterial foraging optimization." (2013).

[37] Negi, Ashish, and Ompal Singh. "An inventory system with time-dependent linear demand for deteriorating items under uncertainty with a completely backlogged." Advances in Mathematical and Computational Sciences: Proceedings of The ICRTMPCS International Conference 2023. Walter de Gruyter GmbH & Co KG, 2024.

[38] Hartely, V.R., 1976. Operations Research"A Managerial Emphasis. Good Year, Santa Monica, CA, pp. 315"317.

[39] Yadav, Krishan Kumar, Ajay Singh Yadav, and Shikha Bansal. "OPTIMIZATION OF A TWO-WAREHOUSE INVENTORY MANAGEMENT FOR DETERIORATING ITEMS WITH TIME AND RELIABILITY-DEPENDENT DEMAND UNDER CARBON EMISSION CONSTRAINTS." Reliability: Theory & Applications 19.4 (80) (2024): 404-418.

[40] Adak, Sudip, and G. S. Mahapatra. "Effect of reliability on multi-item inventory system with shortages and partial backlog incorporating time dependent demand and deterioration." Annals of Operations Research 315.2 (2022): 1551-1571.

[41] Mebarek-Oudina, F. "Exploring Passive Heat Transfer Enhancement Techniques: Applications, Benefits, and Challenges." Babylonian Journal of Mechanical Engineering 2024 (2024): 122-127.

[42] Mebarek-Oudina, F., et al. "Thermal performance of MgO-SWCNT/water hybrid nanofluids in a zigzag walled cavity with differently shaped obstacles." Modern Physics Letters B (2025): 2550163.
[43] Kumar, M. Anil, et al. "The impact of Soret Dufour and radiation on the laminar flow of a rotating liquid past a porous plate via chemical reaction." Modern Physics Letters B 39.10 (2025): 2450458.
[44] Mebarek-Oudina, Fateh. "Numerical modeling of the hydrodynamic stability in vertical annulus with heat source of different lengths." Engineering science and technology, an international journal 20.4 (2017): 1324-1333.

DESIGNING AN ATTRIBUTE CONTROL CHART BASED ON EXPONENTIAL-RAYLEIGH DISTRIBUTION UNDER HYBRID CENSORING

T. KAVITHA* & M. GUNASEKARAN**

•

*PG & Research Department of Statistics
Salem Sowdeswari College (Women), Salem, Tamil Nadu, India.
sreesuryaa@gmail.com

**Department of Applied Research The Gandhigram Rural Institute (Deemed to be University), Dindigul, Tamil Nadu, India. gunasekaranm1988@gmail.com

Abstract

Statistical Process Control (SPC) is an approach to quality assurance that leverages statistical tools to analyse, monitor, and enhance processes. A control chart is an essential tool for observing process performance; it provides a visual mechanism to identify significant deviations caused by assignable factors. This chart compares the values of a quality attribute against specified control limits. Statistical distributions play a crucial role in parametric inferences and are commonly utilized to characterize real-world phenomena. Researchers have directed their attention towards the advancement of these widely used distributions to generate models that are progressively more realistic and versatile for data interpretation. The new Exponential-Rayleigh distribution can be derived by combining the cumulative distribution functions of both the Exponential and Rayleigh distributions. In this paper, we present a new attribute np control chart aimed at monitoring the median lifetime of products within a hybrid censoring scheme, based on the premise that the lifetimes of these products follow an Exponential-Rayleigh distribution. We determine the optimal parameters for the construction of the proposed control chart to achieve an average run length for the in-control process that is as near as possible to the designated average run length. Various sets of values are utilized to estimate the parameters of the control chart, and the performance of the developed control chart is analyzed through the average run length. Numerical examples are presented to illustrate the proposed control chart, and its applicability is validated with simulated data.

Keywords: exponential-Rayleigh distribution, attribute control chart, hybrid censoring scheme, average run length

1. Introduction

The control chart stands out as one of the most prevalent and effective methods for quality control, providing a visual representation of data collected from the manufacturing industry. It typically includes a central line (CL) that signifies the average, an upper line that marks the upper control limit (UCL), and a lower line that indicates the lower control limit (LCL). To enhance the monitoring of production processes in a variety of situations, many advanced control charting techniques have been introduced. Several varieties of control charts are similar and have been

designed to correspond with certain aspects of the quality attribute under consideration. For some products, lifetime is considered as a quality characteristic. Life tests are utilized to evaluate the manufacturing process of these products. The classification of a product as conforming or non-conforming is based on the results obtained from the life test. This testing approach is notably time-intensive due to the prolonged nature of the tests. Consequently, the use of censoring techniques is imperative and should not be disregarded. Some of the censoring techniques that are commonly employed in life testing include Type-I censoring, Type-II censoring, and hybrid censoring.

Attribute control charts, including the np control chart, are constructed on the premise that the fraction of non-conforming items is based on a normal distribution of the quality characteristic. Nevertheless, the actual distribution of these characteristics may not be normal. Therefore, the use of the current control chart in such scenarios could lead to misconceptions for industrial engineers and an increase in the occurrence of non-conforming items. Numerous studies focusing on the development of attribute control charts utilizing various lifetime distributions are documented in the literature. Notable contributions include [1] on the Weibull distribution, [2] addressing the Pareto distribution, and [3] exploring the Birnbaum-Saunders distribution. Further research by [4] revisits the Weibull distribution, while [5] examines the Inverse Rayleigh distribution. Reference [6] investigates the Burr X and XII, Inverse Gaussian, and Exponential distributions. Additional studies include [7] on the Lognormal distribution, [8] on the Exponentiated Half Logistic distribution, and [9] on the Dagum distribution. Reference [10] contributes to the literature on the Pareto distribution, while [11] discusses the Weibull-Pareto combination. Reference [12] focuses on the Exponentiated Exponential distribution, and [13] analyzes the Gamma distribution. Reference [14] presents findings on the Log-Logistic distribution, and [15] investigates the Weibull distribution. Reference [16] explores the Generalized Log Logistic distribution, while [17] examines the Generalized Exponential distribution. Reference [18] focuses on the Inverse Weibull distribution, and [19] studies the Length-Biased Weighted Lomax distribution. Reference [20] contributes insights on the Rayleigh distribution, and [21] investigates the Half Normal and Half Exponential Power distributions. Recently, [22] proposed an attribute np control chart that is based on the Exponentiated Exponential distribution, implemented in the framework of an accelerated life test with a hybrid censoring approach.

The Exponential and Rayleigh distributions are fundamental in the fields of life testing and reliability theory. Various methodologies have been employed to generalize these established distributions, enhancing their flexibility to better accommodate real-world data. One notable approach found in statistical literature for model development is the T-X family. Introduced by [23], the T-X families of distributions facilitate the creation of a new class of distributions that provide increased adaptability for modelling diverse data sets. Numerous researchers have derived different classes within these distributions, and several have proposed generalizations of standard distributions based on the T-X families. Recently, [24] presented a novel class of distribution known as the Exponential Rayleigh distribution, which utilizes the Rayleigh distribution as its foundational distribution and the Exponential distribution as its generator through the generator technique. This Exponential Rayleigh distribution exhibits greater flexibility compared to some related models. A review of the literature reveals a lack of research on control charts that incorporate life tests for non-normal distributions, specifically the Exponential-Rayleigh distribution with Hybrid censoring.

This study establishes an attribute control chart based on the assumption that an Exponential-Rayleigh distribution characterizes the lifetime, taking into account the influence of hybrid censoring. The control chart coefficient has been calculated, and the performance of the proposed chart has been analysed concerning the average run length (ARL). The design of the control chart is responsive to changes in the process median. The applicability of the proposed control chart is illustrated using simulated data. Section 2 focuses on the design of the new control chart and presents a procedure for deriving the control chart for a specified strength. The real-world application of the proposed control chart, supported by a simulation study, is discussed in Section 3. The final section offers a summary of the overall findings.

2. Construction of a Control Chart based on Exponential-Rayleigh Distribution

Consider "T" to represent the lifetime of the products manufactured in a production process, which conforms to an Exponential–Rayleigh distribution with parameters λ and β, valid for $t > 0$, as indicated by its density function.

$$f(t) = \lambda\beta t e^{\frac{\beta}{2}t^2}\{e^{-\lambda(e^{\frac{\beta}{2}t^2}-1)}\}, t, \lambda, \beta > 0 \tag{1}$$

The cumulative distribution function for the Exponential – Rayleigh distribution is expressed as follows:

$$F(t) = 1 - \{e^{-\lambda(e^{\frac{\beta}{2}t^2}-1)}\}, t, \lambda, \beta > 0 \tag{2}$$

The median lifetime of the product, as determined by the Exponential – Rayleigh distribution, is expressed as follows:

$$m = \sqrt{\frac{2}{\beta} log\{1 - \frac{log(\frac{1}{2})}{\lambda}\}} \tag{3}$$

where,

$$\beta = \frac{2}{m^2} log\{1 - \frac{log(\frac{1}{2})}{\lambda}\}$$

The expected average (median) lifetime of the product is indicated by m_0 when the process is stable. The termination time for the test, t_0, can be formulated in terms of the median of the in-control process, expressed as $t_0 = a * m_0$, with "a" representing a constant related to test termination. Let p be the probability of an item's failure before reaching the test termination time t_0, which can be defined as follows.

$$p_0 = 1 - \{e^{-\lambda(e^{\frac{\beta t_0^2}{2}}-1)}\} \tag{4}$$

The expression in equation (4) may be modified by representing the parameter β as a function of m.

$$p_0 = 1 - \{e^{-\lambda(e^{a^2 log\{1-\frac{log(\frac{1}{2})}{\lambda}\}}-1)}\} \tag{5}$$

We propose the following attribute control chart specifically for the Exponential – Rayleigh distribution, which is based on a hybrid censoring scheme that takes into account the number of products present in each subgroup.

- Randomly select a group of n products from the manufacturing process.
- Perform a life test on the chosen items, using t_0 as the test termination time.
- Record the number of items that have failed (denote this as D)
- End the life test either when time t_0 is reached or if the number of failures surpasses the upper control limit (UCL) before reaching t_0, depending on which event occurs first.
- If the number of failures is below the lower control limit (LCL) or above the UCL, the process should be deemed out of control. Conversely, if neither condition is satisfied, the production process is considered to be in control.

This procedure is identified as the np control chart, which focuses on plotting the number of failures rather than the defective fraction (p). The number of defective products adheres to a binomial distribution characterized by parameters n and p_0 in a stable process, where n represents the sample size and p_0 denotes the probability of a product failing before t_0, as specified in equation (5).

Therefore, the subsequent equations are utilized to ascertain the control limits for the proposed control chart.

$$UCL = np_0 + k\sqrt{np_0(1-p_0)} \tag{6}$$

$$LCL = max[0, np_0 - k\sqrt{np_0(1-p_0)}] \tag{7}$$

Here, "k" denotes the coefficient associated with control limits, and the failure probability p_0 of the product before the time t_0 is computed through equation (5).

In certain instances, the likelihood of product failure, denoted as p_0, may be indeterminate; therefore, the control limits for the number of defective products are established for practical applications as follows.

$$UCL = \bar{D} + k\sqrt{\bar{D}(1-\frac{\bar{D}}{n})} \tag{8}$$

$$LCL = max[0, \bar{D} - k\sqrt{\bar{D}(1-\frac{\bar{D}}{n})}] \tag{9}$$

where $\bar{D}$ represents the average number of defective products that arise during the production process across the entire subgroup.
The probability of the process being recognized as in control is indicated by p_{in}^0 and given as

$$p_{in}^0 = P(LCL \leq D \leq UCL|p_0)$$

$$p_{in}^0 = \sum_{d=[LCL]+1}^{[UCL]} \binom{n}{d} p_0^d (1-p_0)^{n-d} \tag{10}$$

The ARL is typically employed to assess the effectiveness of the control chart. The in-control average run length of the proposed control chart, referred to as ARL_0, is established by.

$$ARL_0 = \frac{1}{1 - [\sum_{d=[LCL]+1}^{[UCL]} \binom{n}{d} p_0^d (1-p_0)^{n-d}]} \tag{11}$$

Let us assume that the average (median) of the process has shifted from m_0 to $m_1 = f * m_0$, where f is a constant representing the degree of shift. Therefore, the failure probability of an item in a state of process instability denoted as p_1, is formulated as.

$$p_1 = 1 - \{e^{-\lambda(e^{(\frac{a}{f})^2 log\{1 - \frac{log(\frac{1}{2})}{\lambda}\}} - 1)}\} \tag{12}$$

The probability of the process being to be in control following the transition to m_1 is now established by

$$p_{in}^1 = P(LCL \leq D \leq UCL|p_1)$$

$$p_{in}^1 = \sum_{d=[LCL]+1}^{[UCL]} \binom{n}{d} p_1^d (1-p_1)^{n-d} \tag{13}$$

The ARL for the shifted process, which is considered out of control, is referred to as ARL_1 and can be determined by the following method:

$$ARL_1 = \frac{1}{1 - [\sum_{d=[LCL]+1}^{[UCL]} \binom{n}{d} p_1^d (1-p_1)^{n-d}]} \tag{14}$$

To evaluate the effectiveness of the proposed control chart in identifying process shifts, we will establish the optimal parameters LCL, UCL, a, and k based on the specified values of in-control ARL. In this study, we examined four scenarios of in-control ARL set at 200, 300, 400, and 500, maintaining a fixed sample size of n = 25. Additionally, we analyzed four sample sizes of 20, 25, 30, and 35, while keeping the in-control ARL constant at 370. The values of λ considered were 0.1, 1, and 5. The selection of optimal parameters aimed to ensure that the in-control ARL closely matched the predetermined ARL values, and the out-of-control ARLs were recorded for varying shift constants. The results, including the optimal parameters and their associated out-of-control ARLs, are detailed in Table 1 to Table 6, which reveal a trend of decreasing ARLs as the shift constant f is reduced.

Table 1: *ARLs when the process average (median) shifted for the fixed sample size n = 25*

$\lambda = 0.1$				
Parameter	$r_0 = 200$	$r_0 = 300$	$r_0 = 400$	$r_0 = 500$
LCL	0	0	0	0
UCL	12	12	13	13
a	0.7433	0.7986	0.7711	0.7820
k	3.3901	3.0418	3.4616	3.2941
$Shift(f)$	ARL	ARL	ARL	ARL
1.00	199.9331	299.6661	399.6790	499.7774
0.90	150.4326	22.2081	161.0164	104.8118
0.80	7.4186	2.0473	6.1044	4.4984
0.70	1.1601	1.0036	1.0781	1.0393
0.60	1.0000	1.0000	1.0000	1.0000
0.50	1.0000	1.0000	1.0000	1.0000
0.40	1.0000	1.0000	1.0000	1.0000
0.30	1.0000	1.0000	1.0000	1.0000
0.20	1.0000	1.0000	1.0000	1.0000
0.10	1.0000	1.0000	1.0000	1.0000

Table 2: *ARLs when the process average (median) shifted for the fixed sample size n = 25*

$\lambda = 1.0$				
Parameter	$r_0 = 200$	$r_0 = 300$	$r_0 = 400$	$r_0 = 500$
LCL	0	0	0	0
UCL	12	12	13	13
a	0.6069	0.635	0.6423	0.6566
k	3.3888	3.1694	3.4606	3.2932
$Shift(f)$	ARL	ARL	ARL	ARL
1.00	200.2535	300.3884	400.2921	500.3088
0.90	114.5821	190.5079	266.8718	340.8119
0.80	59.8863	29.2275	62.3957	42.6413
0.70	8.4136	4.9009	7.7912	5.8724
0.60	1.8474	1.4247	1.6791	1.4753
0.50	1.0233	1.0053	1.0110	1.0049
0.40	1.0000	1.0000	1.0000	1.0000
0.30	1.0000	1.0000	1.0000	1.0000
0.20	1.0000	1.0000	1.0000	1.0000
0.10	1.0000	1.0000	1.0000	1.0000

Table 3: *ARLs when the process average (median) shifted for the fixed sample size n = 25*

$\lambda = 5.0$				
Parameter	$r_0 = 200$	$r_0 = 300$	$r_0 = 400$	$r_0 = 500$
LCL	0	0	0	0
UCL	12	12	13	13
a	0.6671	0.6417	0.6041	0.6188
k	2.7768	3.0422	3.4614	3.2935
$Shift(f)$	ARL	ARL	ARL	ARL
1.00	200.0398	299.8027	399.8401	500.1496
0.90	44.3556	77.7719	218.1062	394.6173
0.80	9.7549	15.4763	91.5669	62.7539
0.70	2.7615	3.7734	13.3523	9.9351
0.60	1.2569	1.4351	2.7475	2.3034
0.50	1.0076	1.0197	1.1502	1.1001
0.40	1.0000	1.0000	1.0005	1.0000
0.30	1.0000	1.0000	1.0000	1.0000
0.20	1.0000	1.0000	1.0000	1.0000
0.10	1.0000	1.0000	1.0000	1.0000

Table 4: *ARLs when the process average (median) shifted for the fixed $r_0 = 370$*

$\lambda = 0.1$				
Parameter	$n = 20$	$n = 25$	$n = 30$	$n = 35$
LCL	0	0	1	2
UCL	12	13	15	18
a	0.8191	0.8322	0.8124	0.8552
k	3.2103	2.9882	2.9812	2.9722
$Shift(f)$	ARL	ARL	ARL	ARL
1.00	370.2829	369.972	370.1011	370.4807
0.90	106.7354	17.1442	28.1197	9.7886
0.80	4.4606	1.6244	1.8481	1.1557
0.70	1.0343	1.0004	1.0006	1.0000
0.60	1.0000	1.0000	1.0000	1.0000
0.50	1.0000	1.0000	1.0000	1.0000
0.40	1.0000	1.0000	1.0005	1.0000
0.30	1.0000	1.0000	1.0000	1.0000
0.20	1.0000	1.0000	1.0000	1.0000
0.10	1.0000	1.0000	1.0000	1.0000

Table 5: *ARLs when the process average (median) shifted for the fixed $r_0 = 370$*

$\lambda = 1.0$				
Parameter	$n = 20$	$n = 25$	$n = 30$	$n = 35$
LCL	0	0	1	2
UCL	12	13	15	18
a	0.7071	0.7257	0.7102	0.7594
k	3.2106	2.8440	3.0584	2.9379
$Shift(f)$	ARL	ARL	ARL	ARL
1.00	370.1107	370.1636	370.0825	369.7833
0.90	350.4508	57.8725	75.3012	41.1492
0.80	48.7114	9.0932	9.8727	5.2281
0.70	6.6761	2.1124	2.0630	1.3939
0.60	1.5712	1.0642	1.0485	1.0044
0.50	1.0072	1.0000	1.0000	1.0000
0.40	1.0000	1.0000	1.0000	1.0000
0.30	1.0000	1.0000	1.0000	1.0000
0.20	1.0000	1.0000	1.0000	1.0000
0.10	1.0000	1.0000	1.0000	1.0000

Table 6: *ARLs when the process average (median) shifted for the fixed $r_0 = 370$*

$\lambda = 5.0$				
Parameter	$n = 20$	$n = 25$	$n = 30$	$n = 35$
LCL	0	0	0	1
UCL	12	13	13	16
a	0.6713	0.6908	0.6060	0.6561
k	3.2101	2.9888	2.7447	2.8279
$Shift(f)$	ARL	ARL	ARL	ARL
1.00	370.3710	370.3930	370.1113	369.8262
0.90	200.7384	70.7588	80.2528	61.7222
0.80	74.9887	13.2549	14.8150	10.0207
0.70	12.2362	3.2301	3.4992	2.4188
0.60	2.7234	1.3122	1.3605	1.1419
0.50	1.1623	1.0092	1.0124	1.0012
0.40	1.0008	1.0000	1.0002	1.0000
0.30	1.0000	1.0000	1.0000	1.0000
0.20	1.0000	1.0000	1.0000	1.0000
0.10	1.0000	1.0000	1.0000	1.0000

An analysis of Table 1 to Table 6 reveals that the out-of-control ARL, specifically ARL_1, diminishes as the shift constant f is reduced. This indicates the proposed control chart's effectiveness in promptly identifying shifts in the process. To establish the optimal parameters for the design of this control chart, an algorithm can be employed, accommodating various combinations of the specified in-control ARL and sample size.

1. Define the parameters of the ARL, specifically r_0, λ, and the sample size n.
2. Assign the values $a = 0.001$ and $k = 0.001$.
3. Calculate the failure probability of the product (denoted as p_0) under the condition that the process is in control, utilizing equation (5). Additionally, compute the control chart parameters,

namely the Lower Control Limit (LCL) and Upper Control Limit (UCL), by inserting the values of n, p_0, and k into equations (6) and (7).

4. Determine the in-control ARL using equation (11) and compare the result with r_0. Repeat this procedure for various combinations of a and k until the in-control ARL closely approximates the specified ARL, r_0.

When an ARL is identified, the relevant parameters (LCL, UCL, a, k) are deemed the optimal parameters required. Following this, apply these optimal parameters in equation (14) to determine the out-of-control ARL, known as ARL_1.

3. Application of Proposed Control Chart

3.1. Real-Life Application 1

This section presents a design example to illustrate a practical application of the proposed control chart. Consider a manufacturer aiming to improve the quality of its product. It is established that the failure time of the product adheres to the Exponential-Rayleigh distribution with a parameter of $\lambda = 1$, and the desired median lifetime of the product is set at 1000 hours. A sample of size $n = 25$ will be drawn from each subgroup and will undergo a censored life test. The objective for the in-control ARL, denoted as r_0, is 370. According to Table 5, the parameters are established as $a = 0.7257$, $k = 2.8440$, with a lower control limit (LCL) of 0 and an upper control limit (UCL) of 13. Consequently, the manufacturer will implement the control chart as follows:

Step 1: A sample of 25 products will be selected from each subgroup, followed by a life test lasting 726 hours. The number of products that fail during this test will be recorded and referred to as "D."

Step 2: The life test will conclude either upon reaching 726 hours or if the number of failed products exceeds thirteen before the completion of 726 hours, whichever occurs first.

Step 3: If D exceeds 13, the process will be deemed out of control. Conversely, if the value of D falls within the range of 0 to 13, the process will be considered in control.

3.2. Real-Life Application 2

It is assumed that the lifetimes of the products follow an Exponential-Rayleigh distribution characterized by the parameter $\lambda = 5$. The product specifications are as follows: m_0 equals 1000 hours, r_0 is 500, and n is 25. The control chart parameters, as detailed in Table 3, are $k = 3.2935$, $a = 0.6188$, with a lower control limit (LCL) of 0 and an upper control limit (UCL) of 13. Consequently, the control chart was constructed as follows: For each subgroup, a sample of 25 products is selected and subjected to a life test lasting a minimum of 619 hours. The number of failures recorded during this testing phase (D) is then counted. If the value of D falls within the range of 0 to 13, the process is deemed to be in control; if not, it is considered out of control.

3.3. Simulation Study

In this section, the implementation of the proposed control chart is demonstrated through the use of simulated data. The data originates from an Exponential-Rayleigh distribution, assuming the process is in control with parameters $\lambda = 1$ and a desired average (median) lifetime of 1000 hours. Each sample batch consists of a random sample of size $n = 20$. The first fifteen samples are generated from the in-control process, while the following fifteen samples are derived from a shifted process with $f = 0.60$. From Table 5, with $n = 20$ and $ARL_0 = 370$, the constants are identified as $k = 3.2106$, $a = 0.7071$, UCL = 0, and LCL = 12. The number of products exhibiting a

lifetime below 707 hours is recorded ("D") and are: 7, 3, 4, 6, 9, 4, 6, 9, 5, 6, 4, 6, 5, 5, 7, 14, 7, 13, 9, 11, 14, 17, 13, 12, 16, 13, 12, 12, 14, 13. The values of the nonconforming items are plotted against two control limits (LCL = 0 and UCL = 12) in Figure 1.

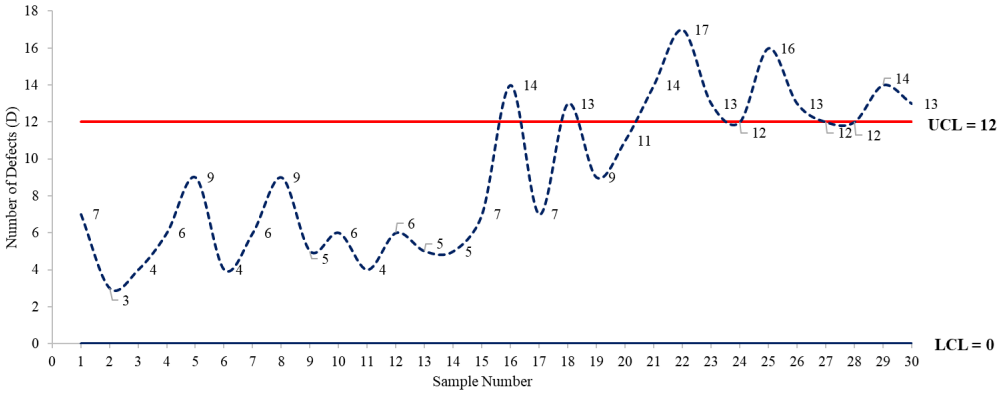


Figure 1: *Control Chart for Simulation Data*

4. Discussion

This paper presents a novel methodology for an attribute control chart designed to assess the lifetime of products. Life tests are conducted using a hybrid censoring approach, which assumes an Exponential-Rayleigh distribution for product lifetimes. The primary objective of the proposed control chart is to maintain the median lifetime of the product as a key quality criterion. This newly developed control chart is highly versatile and can effectively monitor the longevity of quality products. Accompanying tables are provided for practical industrial application and are illustrated using simulated data generated through *R* software based on the Exponential-Rayleigh distribution. The efficacy of the proposed control chart is evaluated in terms of *ARLs* for various shift constants (f). It is important to highlight that employing the hybrid censoring scheme for life testing can significantly reduce both the time and costs associated with sampling inspections. Furthermore, the developed attribute control chart holds promise for adaptation to various other statistical distributions in future research endeavours.

References

[1] Muhammad Aslam and Chi-Hyuck Jun (2015). Attribute Control Chart for the Weibull Distribution under Truncated Life Tests. *Quality Engineering*, 27(3): 283 – 288.

[2] Aslam, M, Nasrullah Khan and Chi–Hyuck Jan (2016a). A control chart for time truncated life tests using Pareto distribution of the second kind. *Journal of Statistical Computation and Simulation*, 86(11): 2113 – 2122.

[3] Aslam, M, Arif, O. H and Jun, C (2016b). An Attribute Control Chart based on the Birnbaum-Saunders Distribution using Repetitive Sampling. *IEEE Access*, 5: 9350 – 9360.

[4] Muhammad Aslam, Osama H. Arif and Chi-Hyuck Jun (2017). An Attribute Control Chart for a Weibull distribution under Accelerated Hybrid Censoring. *PLoS ONE*, 12(3):e0173406: 1 – 11.

[5] C. Nanthakumar and T. Kavitha (2017). Design of Attribute Control Chart based on Inverse Rayleigh Distribution under Type–I Censoring. *International Journal of Scientific Research in Mathematics and Statistical Sciences*, 4(6): 17 – 22.

[6] Ambreen Shafqat, Jaffer Hussain, Amjad D. Al-Nasser and Muhammad Aslam (2018). Attribute Control Chart for some popular distributions. *Communication in Statistics – Theory and Methods*, 47(8): 1978 – 1988.

[7] Arif, O. H, Muhammad Aslam and K.A. Butt (2018). Attribute Control Chart for a Lognormal distribution under Accelerated Time Censoring. *Journal of Computational and Theoretical Nanoscience*, 15(3): 919 – 923.

[8] Srinivasa Rao, G (2018). A Control Chart for Time Truncated Life Tests Using Exponentiated Half Logistic Distribution. *Applied Mathematics & Information Sciences: An International Journal*, 12(1): 125 – 131.

[9] Srinivasa Rao Gadde, A.K. Fulment and P.K Josephat (2019). Attribute Control Charts for Dagum distribution under truncated life tests. *Life cycle Reliability and Safety Engineering*, 8: 329 – 335.

[10] S. Balamurali and P. Jeyadurga (2019). An Attribute np Control Chart for Monitoring Mean Life using Multiple Deferred State Sampling based on Truncated Life Tests. *International Journal of Reliability, Quality and Safety Engineering*, 26(1): 1950004-1 – 1950004-18.

[11] Enayat M. Abd Elrazik (2020). Attribute control charts for the New Weibull Pareto Distribution under truncated life tests. *Journal of Statistics Applications & Probability*, 9(1): 43 – 49.

[12] T. Kavitha and M. Gunasekaran (2020). Construction of control chart based on Exponentiated Exponential distribution. *Journal of Engineering Sciences*, 11(4): 979 – 985.

[13] Ahmad Ibrahim Shawky, Muhammad Aslam and Khushnoor Khan (2020). Multiple dependent state sampling-based charts using Belief Statistic under Neutrosophic Statistics. *Journal of Mathematics*, 2020(7680286): 1 – 14.

[14] Srinivasa Rao Gadde, Edwin Paul (2020). Time Truncated Control Chart using Log- Logistic Distribution. *Biometrics & Biostatistics International Journal*, 9: 76 – 82.

[15] Ali Hussein Al-Marshadi, Ambreen Shafqat, Muhammad Aslam and Abdullah Alharbey (2021). Performance of a New Time-Truncated Control Chart for Weibull Distribution under Uncertainty. *International Journal of Computational Intelligence Systems*, 14(1): 1256 – 1262.

[16] K. Rosaiah, G. Srinivasa Rao and S. V. S. V. S. V. Prasad (2021). A Control Chart for Time Truncated Life Tests using type-II Generalized Log Logistic Distribution. *Biometrics & Biostatistics International Journal*, 10(4): 138 – 143.

[17] O. A. Adeoti, P. Ogundipe (2021). A Control Chart for the generalized exponential distribution under time truncated life test. *Life Cycle Reliability and Safety Engineering*, 10: 53 – 59.

[18] Ayman Baklizi and Sawsan Abu Ghannam (2022). An attribute control chart for the inverse Weibull distribution under truncated life tests. *Heliyon*, 8(e11976): 1 – 5.

[19] Srinivasa Rao Gadde and Amer Ibrahim Al-Omari (2022). Attribute control charts based on TLT for Length-Biased Weighted Lomax Distribution. *Journal of Mathematics*, 2022(3091850): 1 – 15.

[20] Olatunde Adebayo Adeoti and Gadde Srinivasa Rao (2022). Attribute control chart for Rayleigh distribution using Repetitive sampling under truncated life test. *Journal of Probability and Statistics*, 2022(8763091): 1 – 11.

[21] Muhammad Naveed, Muhammad Azam, Nasrullah Khan, Muhammad Aslam, Muhammad Saleem and Muhammad Saeed (2023). Control Chart for Half Normal and half Exponential power Distributed Process. *Scientific Reports*, 13(8632): 1 – 20.

[22] Gunasekaran M (2024). A New Attribute Control Chart based on Exponentiated Exponential Distribution under Accelerated Life Test with Hybrid Censoring. *Reliability: Theory & Applications*, 19, 4(80): 850 – 860.

[23] A. Alzaatreh, C Lee and F Famoye (2013). A New Method for Generating Families of Continuous Distribution. *Metron*, 71(1): 63 – 79.

[24] Kawsar Fatima and S. P. Ahmad (2017). Statistical Properties of Exponential Rayleigh Distribution and its Applications to Medical Science and Engineering. *International Conference on Recent Innovations in Science, Agriculture, Engineering and Management, University College of Computer Applications Guru Kashi University, Bathinda, Punjab*, 20th November 2017: 491 – 506. ISBN:978-93-86171-80-1

A NEW LENGTH BIASED EXPONENTIAL-EXPONENTIAL DISTRIBUTION AND ITS APPLICATIONS TO CANCER DATA

Latha J[1], M Vijayakumar[2], Shibu D S[3]

•

[1]Research Scholar, Department of Statistics,
Annamalai University, Annamalai Nagar, Tamil Nadu, India,
Assistant Professor, Department of Statistics, Christ Nagar College, Maranalloor,
Thiruvananthapuram, Kerala, India.
[2]Department of Statistics, Annamalai University,
Annamalai Nagar, Tamil Nadu, India.
[3]Department of Statistics, University of Kerala,
Thiruvananthapuram, Kerala, India
[1]lathalekshmi.j@gmail.com, [2]mvkumar1975@gmail.com, [3]statshibu@gmail.com

Abstract

The application of weighted distribution is used in many fields of real life such as medicine, reliability, ecology and so on, so we try to contribute in this field through this research. In this paper we propose a new length biased form of the Exponential-Exponential distribution called Length Biased Exponential-Exponential distribution (LBEED) and derive its statistical properties. The statistical properties including the moments, moment generating function, characteristic function, reliability functions and entropy measures are discussed. The method of maximum likelihood and method of moment was employed to estimate the parameter. The utility and adaptability of the proposed model are demonstrated by real-world data sets.

Keywords: exponential-exponential distribution, hazard function, length-biased distribution, maximum likelihood estimation, renyi entropy.

I. Introduction

The exponential distribution is the one that draws the greatest attention from researchers in the literature because of its broad variety of applications. Standard distribution can be understood with the help of weighted distributions, which also give distributions more flexibility. The form of the distribution of recorded observations can be influenced by ascertainment methods, as [5] examined. This is introduced and formulated by [11] in relation to statistical data modelling when standard distributions were found to be in appropriate. In addition to offering methods for expanding distributions for more flexibility in fitting a dataset, weighted distributions can be used to improve understanding. The length biased Rani distribution with survival data was introduced by [15]. The length biased weighted Lomax distribution with outliers was examined by [7]. Weighted distributions are used in many research in the domains of biology, branching processes, and reliability. According to [6], the idea of length biased distribution has several uses in the

biomedical field, including family history, illness survival, intermediate events, and the latency period of AIDS caused by blood transfusion.

The weighted exponential families of distributions are introduced with the assistance of [17]. Weighted distributions come in two varieties: size-biased and length. The length biased Burr XII distribution and length-biased weighted exponentiated inverted Weibull distribution were discussed in [3] and [12]. The length-biased exponential distribution for estimating Bayesian reliability is discussed and addressed in [8]. Many writers have used the idea of weighted distributions for various reasons: [10] looked at several general models that result in weighted distributions with weight functions that do not have to be limited by one and researched length biased (size biased) sampling with relevance to wildlife populations and human families. The Length biased quasi-Sujatha distribution with properties and applications to bladder cancer data is addressed by [16]. Recently, [2] examined and studied the length biased powered inverse Rayleigh distribution and its applications. Also [4] proposed a weighted intervened exponential distribution as a lifetime model. A new method for generating families of continuous distributions is presented in [1] and is used in [9] and motivated to find a new distribution which is a length biased form.

The probability density function (PDF) of the Exponential-Exponential distribution (EED) by [1] and [9] is given as

$$f(x) = \theta^2 e^{-\theta^2 x}, x > 0, \theta > 0 \quad (1)$$

Assume X is a non-negative random variable with probability density function (pdf) f (x). Let W(x) be the weight function which is a non-negative function, then the probability density function of the weighted random variable X_w is given by:

$$f_w(x) = \frac{W(x) f(x)}{E(w(x))}, x > 0$$

where w (x) be a non-negative weight function and $E(w(x)) = \int w(x) f(x)\, dx < \infty$.

The PDF of the length biased distribution is defined by

$$f_L(x; \theta) = \frac{x f(x;\theta)}{E(X)}, -\infty < x < \infty \quad (2)$$

where $\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$, $-\infty < x < \infty$.

II. The Length Biased Exponential-Exponential Distribution (LBEED)

Let X be a random variable with PDF given in (1), respectively, then by using the transformation given in (2), and replacing the expected value as given in the formula; the PDF of the suggested length biased Exponential-Exponential distribution (LBEED) is defined by,

$$f(x; \theta) = x\theta^4 e^{-\theta^2 x}, x > 0, \theta > 0 \quad (3)$$

where θ is the scale parameter.

The CDF of the is length biased Exponential-Exponential distribution (LBEED) is given by,

$$F(x; \theta) = 1 - (x\,\theta^2 + 1) e^{-\theta^2 x}, x > 0, \theta > 0 \quad (4)$$

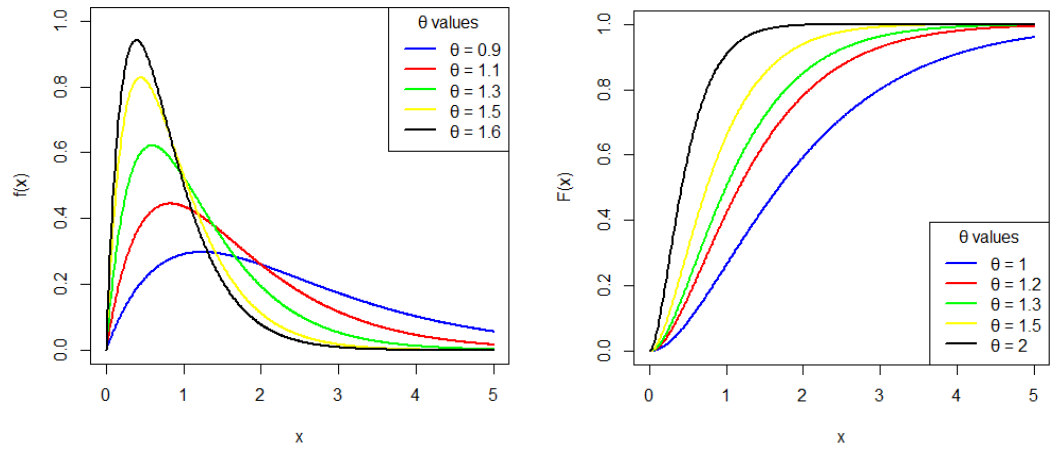


Figure 1: *PDF (left) and CDF (right) plots of LBEED for different values of the parameter θ.*

III. Reliability Analysis

I. Survival Function

This section deals with the survival function and hazard function of the proposed Length-biased Exponential-Exponential distribution.

$$S(x;\theta) = 1 - F(x;\theta)$$

$$S(x;\theta) = (x\,\theta^2 + 1)\,e^{-\theta^2 x}, \quad x > 0, \theta > 0 \tag{5}$$

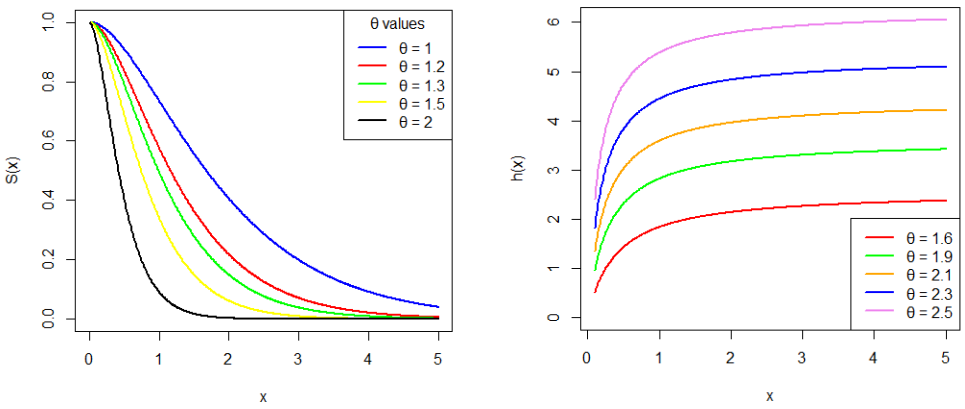


Figure 2: *Survival function (left) and Hazard function (right) of the LBEED for different values of the parameter θ.*

II. Hazard Function

The hazard function or failure rate of the proposed distribution is given by

$$H(x;\theta) = \frac{f(x;\theta)}{S(x;\theta)}$$

$$H(x;\theta) = \frac{x\theta^4}{x\theta^2+1} \quad (6)$$

III. Reverse Hazard Function

The reverse hazard function of the Length-biased Exponential-Exponential distribution is given by,

$$h_r(x) = \frac{f(x;\theta)}{F(x;\theta)}$$

$$h_r(x) = \frac{x\,\theta^4 e^{-\theta^2 x}}{1-(x\,\theta^2+1)e^{-\theta^2 x}} \quad (7)$$

IV. Statistical Properties

The structural properties of length biased exponential- exponential distribution are included in this section.

I. Moments and related measures

Assume that X is the random variable that follows a length-biased Exponential-Exponential distribution (LBEED), then the rth order moment, that is $E(X^r)$ can be obtained as,

$$\mu_r' = E(X^r) = \int_0^\infty x^r f(x;\theta)\,dx$$

$$\mu_r' = \int_0^\infty x^r\, x\,\theta^4 e^{-\theta^2 x}\,dx$$

$$= \int_0^\infty x^{r+1}\,\theta^4 e^{-\theta^2 x}\,dx$$

Put $\theta^2 x = u,\ \theta^2\,dx = du$

$$\mu_r' = \frac{\theta^2}{\theta^{2(r+1)}}\int_0^\infty u^{r+1}\,e^{-u}$$

$$\mu_r' = \frac{\Gamma(r+2)}{\theta^{2r}} \quad (8)$$

The mean and variance can be obtained by substituting r= 1, 2, 3 and 4

$$\mu_1' = \frac{2}{\theta^2}$$

$$\mu_2' = \frac{6}{\theta^4}$$

$$\mu_3' = \frac{24}{\theta^6}$$

$$\mu_4' = \frac{120}{\theta^8}$$

Variance, $\mu_2 = \frac{2}{\theta^4}$ (9)

II. Skewness and Kurtosis

The coefficient of skewness is defined as $\sqrt{\beta_1} = \frac{{\mu_3}^2}{{\mu_2}^{\frac{3}{2}}}$ and is given by $\sqrt{\beta_1} = \frac{\frac{4}{\theta^6}}{\left(\frac{2}{\theta^4}\right)^{\frac{3}{2}}}$.

The coefficient of kurtosis gives the flatness or peakedness of the curve and it is defined as $\beta_2 = \frac{\mu_4}{{\mu_2}^2}$ and is given by, $\beta_2 = \frac{\frac{24}{\theta^8}}{\left(\frac{2}{\theta^4}\right)^2}$

III. Mode

The mode of the length biased exponential-exponential distribution is given by $x = \frac{1}{\theta^2}$.

IV. Moment Generating function and Characteristic function

The moment generating function and Characteristic function of length-biased Exponential-Exponential distribution are included in the section.

$M_x(t) = E\,(e^{tx})$

$= \int_0^\infty e^{tx}\, x\, \theta^4 e^{-\theta^2 x} dx$

$M_x(t) = \frac{\theta^4}{(\theta^2 - t)^2}, |t| < \theta^2$ (10)

Similarly, the Characteristic function is obtained as

$\phi_t(x) = E\,(e^{itx}) = \frac{\theta^4}{(\theta^2 - it)^2}$ (11)

V. Entropy Measures

The entropy of a random variable X is a measure of variation of the uncertainty. It plays important role in various fields, including probability and statistics, physics communication theory and economics by quantifying diversity and unpredictability. A system with higher entropy indicates greater uncertainty and more uniform outcomes.

I. Renyi Entropy

The Renyi entropy is also important in quantum information, where it can be used as a measure of entanglement. For a given probability distribution, Renyi entropy is given by,

$H_\alpha(f(x,\theta)) = \frac{1}{1-\alpha} \log \int_0^\infty [f(x,\theta)]^\alpha\, dx$

$\int_0^\infty [f(x,\theta)]^\alpha\, dx = \int_0^\infty x^\alpha\, \theta^{4\alpha} e^{-\theta^2 x\alpha}\, dx$

Put $u = x\alpha\theta^2$, $du = \alpha\theta^2 dx$

$\int_0^\infty [f(x,\theta)]^\alpha \, dx = \frac{\theta^{4\alpha}}{(\alpha\theta^2)^{\alpha+1}} \ \Gamma(\alpha+1)$

$H_\alpha(f(x,\theta)) = \frac{1}{1-\alpha}\log\left[\frac{\theta^{4\alpha}}{(\alpha\theta^2)^{\alpha+1}} \ \Gamma(\alpha+1)\right]$

$$H_\alpha(f(x,\theta)) = \frac{1}{1-\alpha}\left[4\alpha\log\theta + \log\Gamma(\alpha+1) - (\alpha+1)\log\alpha\theta^2\right] \quad (12)$$

II. Tsallis Entropy

A generalization of Boltzmann-Gibbs (B-G) statistical mechanics initiated by Tsallis has focused a great deal to attention. This generalization of B-G statistics was proposed firstly by introducing the mathematical expression of Tsallis entropy [14] for a continuous random variable is defined as follows:

$$S_\lambda = \frac{1}{\lambda-1}\left[1 - \int_0^\infty f(x)^\lambda \, dx\right]$$

$$S_\lambda = \frac{1}{\lambda-1}\left[1 - \int_0^\infty \left(x\,\theta^4\, e^{-x\theta^2}\right)^\lambda dx\right]$$

on simplification we get,

$$S_\lambda = \frac{1}{\lambda-1}\left[1 - \frac{\Gamma(\lambda+1)}{\lambda^{\lambda+1}\theta^2}\right] \quad (13)$$

VI. Bonferroni and Lorenz Curves

The Bonferroni and Lorenz curves have applications not only in economics to study income and poverty, but also in other fields like reliability, demography, insurance and medical science. The Bonferroni curve of the LBEED is defined as:

$$B(p) = \frac{1}{p\mu}\int_0^q x\, f(x)dx = \frac{1}{p\mu}\left[\mu - \int_q^\infty x\, f(x)dx\right]$$

$$B(p) = \frac{1}{p}\left[1 - \frac{1}{2}\left[e^{-\theta^2 q}\left[(\theta^2 q)^2 + 2(\theta^2 q) + 2\right]\right]\right] \quad (14)$$

The Lorenz curve of the LBEED is defined as:

$$L(p) = \frac{1}{\mu}\int_0^q x\, f(x)dx = \frac{1}{\mu}\left[\mu - \int_q^\infty x\, f(x)dx\right]$$

$$L(p) = \frac{\theta^2}{2}\left[\frac{2}{\theta^2} - e^{-\theta^2 q}\left[(\theta^2 q)^2 + 2(\theta^2 q) + 2\right]\right] \quad (15)$$

VII. Fisher's Information

The Fisher information determine the amount of information contained in an observable random variable X about an unknown distribution parameter upon which the probability of X depends. The Fisher's information is defined as,

$$I\ (\theta) = \ -E\left(\frac{d^2}{d\theta^2}\log\big(f(x)\big)\right)$$

The Fisher's information of Length biased exponential-exponential distribution is given by $I(\theta) = \frac{4}{\theta^2}$, as θ increases the Fisher's value decreases.

VIII. Order Statistics

Order statistics make their appearance in many statistical theory and practice. We know that if $X_{(1)}$, $X_{(2)}$,…, $X_{(n)}$ denotes the order statistics of a random sample X_1, X_2,…, X_n from a continuous population with cdf $F_x(x)$ and pdf $f_x(x)$, then the pdf of r^{th} order statistics $X_{(r)}$ is given by,

$$f_{x(r)}(x) = \frac{n!}{(r-1)!\,(n-r)!}\cdot f_x(\mathrm{x}).\,[F_x(\mathrm{x})]^{r-1}\,[1-\ F_x(\mathrm{x})]^{n-r}$$

Using (3) and (4) in equation (16), we will obtain the probability density function of rth order statistics of length biased exponential-exponential distribution as

$$f_{x(r)}(x) = \frac{n!}{(r-1)!(n-r)!}\ x\ \theta^4 e^{-\theta^2 x}\left[1-(x\ \theta^2+1)e^{-\theta^2 x}\right]^{r-1}\left[(x\ \theta^2+1)e^{-\theta^2 x}\right]^{n-r} \tag{16}$$

The probability density function of first order statistic $X_{(1)}$ of length biased exponential-exponential distribution can be obtained as

$$f_{x(1)}(x) = \ n\ x\ \theta^4 e^{-\theta^2 x}\left[(x\ \theta^2+1)e^{-\theta^2 x}\right]^{n-1} \tag{17}$$

Similarly, the probability density function of higher order statistic $X_{(n)}$ of length biased exponential-exponential distribution can be obtained as

$$f_{x(n)}(x) = n\ x\ \theta^4 e^{-\theta^2 x}\left[1-(x\ \theta^2+1)e^{-\theta^2 x}\right]^{n-1} \tag{18}$$

IX. Parameter Estimation

In this section we estimate the parameter of length biased exponential-exponential distribution using two different methods such as maximum likelihood and method of moments.

I. Maximum Likelihood Estimation

In this section, the parameter estimation of length biased exponential-exponential distribution using maximum likelihood method is included. Let X_1, X_2, …, X_n be a random sample of size n from the length biased exponential-exponential distribution, then the likelihood function is given by,

$$L(x,\theta) = \prod_{i=1}^{n} f(x,\theta)$$

$$L(x,\theta) = \theta^4 \prod_{i=1}^{n} x\, e^{-\theta^2 x}$$

The log likelihood function is given by

$$\log L(x,\theta) = 4n \log\theta - \theta^2 \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \log x \tag{19}$$

Differentiating the equation (20) partially w.r.to θ and equating to zero, we get the following system of equation

$$\frac{4n}{\theta} - 2\theta \sum_{i=1}^{n} x_i = 0$$

The maximum likelihood estimator $\hat{\theta}$ of θ is given by

$$\hat{\theta} = \sqrt{\frac{2n}{\sum_{i=1}^{n} x_i}}.$$

II. Method of Moments

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the length biased exponential-exponential distribution. Equating population mean to the corresponding sample mean, the moment estimator $\widehat{\theta}$ of θ of LBEED is given by $\hat{\theta} = \sqrt{\frac{2}{\bar{x}}}$.

X. Applications

The real-life application of length biased exponential-exponential distribution is included in this section and show that the length biased exponential-exponential distribution (LBEED) can be a better model than the exponential-exponential distribution (EED) and length biased exponential distribution (LBED). The data set is given below.

In order to compare the performance, length biased exponential- exponential distribution with exponential-exponential distribution. We are using the Criterion values like AIC (Akaike information criterion), AICC (corrected Akaike information criterion), BIC (Bayesian information criterion) and HQIC (Hannan-Quinn Information Criterion). The better distribution corresponds to lesser AIC, AICC and BIC values. The formulae for calculation of AIC, BIC, AICC and HQIC are;

AIC = $2k - 2\log L$, AICC = $AIC + \frac{2k(k+1)}{n-k-1}$, BIC = $k\log n - 2\log L$ and

HQIC = $2k \log(\log(n)) - 2\log L$

where n is the sample size, K is the number of parameters, and L denotes the likelihood function. Any probability model having smaller value of AIC, BIC and $-LogL$ being the best model to fit the data set.

Data Set 1: The following data set represents the survival times (in months) of 32 patients of melanoma studied by [13].

3.25, 3.5, 4.75, 4.75, 5, 5.25, 5.75, 5.75, 6.25, 6.5, 6.5, 6.75, 6.75, 7.78, 8, 8.5, 8.5, 9.25, 9.5, 9.5, 10, 11.5, 12.5, 13.25, 13.5, 14.25, 14.5, 14.75, 15, 16.25, 16.25, 16.5.

Table 1: *The measures, AIC, CAIC, BIC, and HQIC for the melanoma (Skin Cancer) data*

Distribution	MLE	-Log L	AIC	BIC	AICC	HQIC
LBEED	θ = 0.4619	94.47	190.93	192.40	191.06	191.42
EED	θ = 0.3266	103.621	209.241	210.707	209.375	209.727
LBED	θ = 9.376	106.827	215.654	217.120	215.787	216.140

Data Set 2: The data set reported by Efron24 represent the survival times of a group of 44 patients suffering from Head and Neck cancer disease and treated using a combination of radiotherapy and chemotherapy (RT+CT).
12.20, 23.56, 23.7, 25.9, 31.98, 37, 41.35, 47.38, 55.46, 58.36, 63.47 ,68.46, 78.3, 74.5, 81.43, 84, 92, 94, 110, 112, 119, 127, 130, 133, 140, 146, 155, 159, 173, 179, 194, 195, 209, 249, 281, 319, 339, 432, 469, 519, 633, 725, 817, 1776.

Table 2: *The measures, AIC, CAIC, BIC, and HQIC for the Head and Neck cancer data*

Distribution	MLE	-Log L	AIC	BIC	AICC	HQIC
LBEED	θ = 0.0945	54.9225	111.8449	111.4884	111.9401	110.2765
EED	θ = 0.0669	147.3185	296.637	296.2805	296.7322	295.0686
LBED	θ= 53.7	296.19	594.38	596.14	594.48	594.93

AIC, BIC, AICC, -Log L and HIQC values are lower in the length biased exponential-exponential distribution than in the exponential-exponential distribution, according to Table 1 and Table 2. Our conclusion is that the length biased exponential-exponential distribution provides a better match than the exponential-exponential distribution.

XI. Conclusion

In the present paper, the new distribution called length biased exponential-exponential distribution has been derived. The subject distribution is generated by using the length biased technique and taking the one parameter exponential-exponential distribution as the base distribution. Its various statistical properties including its mean, variance, survival function, hazard rate function, reverse hazard rate function, moment generating function, characteristic function, order statistics, entropies, Bonferroni and Lorenz curves have been discussed. Its parameters have also been estimated through the method of maximum likelihood estimation and method of moments has been observed. The proposed model is very flexible because its density function has many shapes; it can be right skewness, decreasing, and unimodal. Finally, a newly proposed length biased exponential-exponential has been investigated with real life data set to examine its ability and superiority.

References

[1] Alzaatreh, A., Lee, C. and Famoye, F. (2013), "A new method for generating families of continuous distributions", *Metron* 71(1), 63-79.

[2] Mustafa, A. and Khan, M. I. (2022). The length-biased powered inverse Rayleigh distribution with applications. *J. Appl. Math. & Informatics*, 40(1-2), 1-13.

[3] Abd-Elfattah, A. M., Mahdy, M. and Ismail, G. (2021). Length biased Burr- XII distribution: properties and application. *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, 56(1), 222-244.

[4] Bhat, V. A. and Pundir, S. (2023): Weighted intervened exponential distribution as a lifetime distribution, *Reliability: Theory & Applications*, 2(73), 253-266.

[5] Fisher R. A., The effects of methods of ascertainment upon the estimation of frequencies (1934). *Annals of Eugenics*.6:13-25.

[6] Gupta, R.C. and Kirmani, S. (1990). The role of weighted distributions in stochastic modelling, *Communications in Statistics, Theory and Methods*. 19, 3147- 3162.

[7] Karimi, H. and Nasiri, P. (2018). Estimation parameter for length biased weighted Lomax distributions in the presence of outliers. *Mathematical and Computational Applications*, doi:10.3390/mca23010009.

[8] Mathew, J. and George, S. (2020). Length biased exponential distribution as a reliability model: a Bayesian approach, *Reliability: Theory & Applications*, 3(58):84-91.

[9] Ogunwale, O.D., Ayeni, T.M. (2022) The New Exponential-Exponential Distribution: Theory and Properties. *Quest Journals- Journal of Research in Applied Mathematics*, 8(8) :13-19.

[10] Patil, G. P. and Rao, C. R. (1978). Weighted distributions and Size biased sampling with applications to wildlife populations and human families. *Biometrics*, 34, 179-189.

[11] Rao C. R., (1965) On discrete distributions arising out of method of ascertainment, in classical and Contagious Discrete, G.P. Patiled; *Pergamum Press and Statistical publishing Society*, Calcutta;. p. 320-332.

[12] Saghir, A., Tazeem, S. and Ahmad, I. (2016). The length-biased weighted exponentiated inverted Weibull distribution. *Cogent Mathematics*, 3(1), 1-18.

[13] Susarla V, Vanryzin JV. Empirical Bayes estimation of a distribution (Survival) function from right Censored observations. *The Annals of Statistics*, 1978. doi:10.1214/aos/ 1176344249, 6(4), 740-75.

[14] Tsallis, C. (1988), Possible generalization of boltzmann-gibbs statistics. *Journal of Statistical Physics*, 52, 479-487

[15] Vijayakumar, M., Anwar Shameem, Z. A. and Punathumparambath, B. (2020). Length biased Rani distribution with survival data analysis. *High Technology Letters*, 26(6), 559-573.

[16] Subramanian, C. and Shenbagaraja, R. (2020). Length biased quasi-Sujatha distribution with properties and applications to bladder cancer data. *Journal of Information and Computational Science*, 10(1), 493- 506.

[17] Zubair Ahmad, Gholamhossien Hamedani, and Mohammed Elgarhy (2019). The Weighted Exponentiated Family of Distributions: Properties, Applications and Characterizations. *JIRSS* Vol. 19, No. 01, pp 209-228

A NEW GENERALIZATION OF AREA-BIASED XGAMMA DISTRIBUTION WITH PROPERTIES AND ITS APPLICATION TO CANCER DATA

Athira D V[1,2], P. Pandiyan[3], and D.S. Shibu[4]

•

Research Scholar, Department of Statistics, Annamalai University
[2]Assistant Professor, Department of Statistics, Christ Nagar College Maranalloor
[3]Professor, Department of Statistics, Annamalai University
[4]Professor , Department of Statistics, University College
[1,2]athiramanuphd@gmail.com, [3]pandianau@gmail.com , [4]statshibu@gmail.com

Abstract

In this paper, we proposed a new model, "Area Biased Xgamma Distribution", a generalization of the Xgamma distribution. The various distributional properties survival function, hazard function, mean residual life ,statistical properties, moments, moment generating function and characteristic function, Bonferroni and Lorenz curve, entropy of the new model were studied. The maximum likelihood estimation method is used to estimate the parameter of the area-biased Xgamma distribution. Finally, using three cancer data sets for the illustration and application of the new model.

Keywords: area biased Xgamma distribution, moments, maximum likelihood estimation, order statistics, renyi entropy, reliability analysis

I. Introduction

Xgamma distribution (XGD) [16] is a newly developed lifetime distribution introduced by Sen. S., Maiti S. S. & Chandra N. in 2016.The Xgamma distribution is a special combination of exponential (θ) and gamma (θ, α) distributions. The exponential and gamma distributions are well known and have some special structural properties [3]. Sen. S. and others in 2016 investigated the weighted Xgamma distribution (WXGD), a relatively novel statistical model for positive-valued data and survival analysis. The length-biased Xgamma distribution (LBXG) was produced as a special case of the density function of the Xgamma distribution [15]. Fisher developed the concept of weighted distribution [9] in 1934 to describe ascertainment biases in the data, and later on, Rao [12] reinterpreted it as a unified theory for problems when the observations fail in non-experimental, non-replicated, and non-random situations. In 2015 Dey. S. Ali, S., & Park, C. are introduced Weighted exponential distribution properties and different methods of estimation [5]. The weighted distribution concept defined as $f_w(x) = \frac{w(x)f(x)}{E(w(x))}$,where w(x) is the non-negative weight function and f(x) is probability function. We have many choices of weight functions [4]. If the weight function w(x)=x is a length-biased distribution, then w(x)=x^2 is an area-biased distribution. In [9], Length and area biased exponentiated Weibull distribution. In [17], established on a length and area-biased

Maxwell distributions. In [11], a proposal was made on area-biased weighted Weibull distribution. Recently [1], discussed the characterization and estimation of area-biased quasi-Akash distribution. In [13], proposed probability and survival analysis of cancer patients

In this paper, we introduced a new generalization of area-biased Xgamma distribution, which is the extension of the Xgamma distribution. The article is divided into several sections and is structured as follows: The alternative form of the Area-biased Xgamma distribution is introduced in Section 2. The survival properties are studied in Section 3. The statistical properties such as moments, moment generating-functions, and characteristic functions are described in Section 4. In Section 5, order statistics of the distribution are derived. The likelihood ratio test is discussed in Section 6. Bonferroni and Lorenz curves and entropy are discussed in Sections 7 and 8, respectively. Methods of estimating parameters are discussed in section 9. In section 10, three actual cancer data sets are analyzed to show the application of the weighted area-biased Xgamma distribution; finally, Section 11 concludes.

II. Area Biased Xgamma distribution

Let X be a random variable follow the Xgamma distribution with scale parameter θ; then the probability density function (pdf) and cumulative distribution function (cdf) of the Xgamma distribution, respectively,

$$f(x) = \frac{\theta^2}{(1+\theta)}\left(1+\frac{\theta}{2}x^2\right)e^{-\theta x} \qquad x>0, \theta>0 \tag{1}$$

And

$$F(X) = 1 - \frac{1+\theta+\theta x+\frac{\theta^2 x^2}{2}}{(1+\theta)}\, e^{-\theta x} \qquad x>0, \theta>0 \tag{2}$$

The probability density function of the weighted random variable X_w is given by

$$f_w(x) = \frac{w(x)f(x)}{E(w(x))} \qquad x>0$$

We have considered the weight function $w(x)=x^2$; then the probability function of area-biased distribution is given

$$f_\alpha(x) = \frac{x^2 f(x)}{E(x^2)} \tag{3}$$

where, $E(x^2) = \int_0^\infty x^2 f(x)dx$

$$E(x^2) = \frac{2(\theta+6)}{\theta^2(1+\theta)} \tag{4}$$

Substituting (1) and (4) in (3) we get the probability density function of area-biased Xgamma distribution

Definition: A continuous random variable X is said to follow the area-biased Xgamma (ABXG) distribution with parameter θ if its probability density function(pdf)

$$f_a(x) = \frac{x^2\theta^4}{2(\theta+6)}\left(1+\frac{\theta}{2}x^2\right)e^{-\theta x} \qquad x>0\,, \theta>0 \tag{5}$$

It is denoted by X~ABXGD(θ)

The cumulative density function (cdf) X is given by

$$F_a(x) = \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)} \qquad ; x > 0\,, \theta > 0 \tag{6}$$

Where, $\gamma(3,\theta x)$ and $\gamma(5,\theta x)$ are lower incomplete gamma functions.

The plots of the corresponding pdf and cdf of the area-biased Xgamma distribution are shown in **Figure 1** and **Figure 2**. It is possible to see from **Figure 1** that the area-biased Xgamma distribution is a positively skewed distribution.

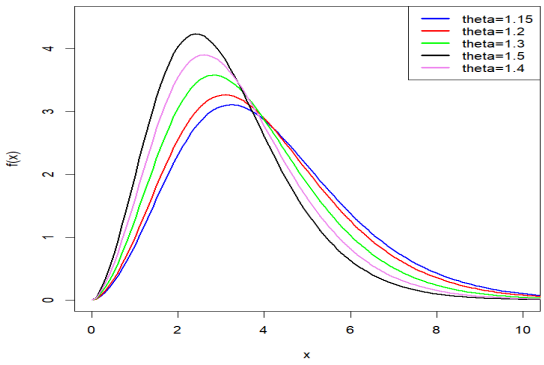


Figure 1: *probability density function of Area-biased Xgamma distribution*

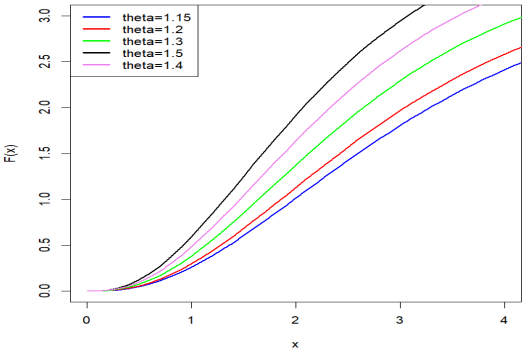


Figure 2: *cumulative distribution function area-biased Xgamma distribution*

III. Reliability analysis

In this section, we will cover the survival function, hazard function, reverse hazard function, cumulative hazard function, odds rate, Mills ratio, and, mean residual life for the area-biased Xgamma distribution**.**

I. Survival Function

Let X be a random variable with pdf f(x) and cdf F(x). The general formula for the survival function is $S(x) = 1 - F_a(x)$. Then the survival function, or the reliability function of the area-biased Xgamma distribution, is given by

$$S(x) = 1 - \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)} \qquad \text{x} > 0 \tag{7}$$

II. Hazard function

The corresponding hazard function / failure rate of the area-biased Xgamma distribution is provide as

$$h(x) = \frac{f_a(x)}{S(x)}$$

$$h(x) = \frac{\frac{x^2\theta^4}{2(\theta+6)}\left(1 + \frac{\theta}{2}x^2\right)e^{-\theta x}}{1 - \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)}}$$

after simplification, we get

$$h(x) = \frac{2x^2\theta^4\left(1+\frac{\theta}{2}x^2\right)e^{-\theta x}}{4(\theta+6) - 2\theta\,\gamma(3,\theta x) - \gamma(5,\theta x)} \tag{8}$$

III. Reverse Hazard function

Reverse hazard function of area-biased Xgamma distribution is indicated as

$$h_r(x) = \frac{f_a(x)}{F_a(x)}$$

$$h_r(x) = \frac{2x^2\theta^4\left(1+\frac{\theta}{2}\,x^2\right)e^{-\theta x}}{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)} \tag{9}$$

IV. Odds Rate Function

Odds Rate function of Area-biased Xgamma distribution is given by

$$O(x) = \frac{F_a(x)}{1 - F_a(x)}$$

$$O(x) = \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6) - \theta\gamma(3,\theta x) - \gamma(5,\theta x)} \tag{10}$$

V. Cumulative Hazard Function

Cumulative hazard function of Area-biased distribution is expressed as

$$H(x) = -\ln(1 - F_a(x))$$

$$H(x) = -\ln\left(1 - \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)}\right)$$

VI. Mills Ratio

The mills ratio of Area-biased area distribution is provided by

$$\text{Mills ratio} = \frac{1}{\mathrm{h_r(x)}}$$

$$\text{Mills ratio} = \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{2x^2\theta^4\left(1+\frac{\theta}{2}\,x^2\right)e^{-\theta x}} \tag{11}$$

VII. Mean residual life

The mean residual life (MRL) at time t is the expected remaining life given that the random variable x survived up to time t.

$$MRL(x) = \frac{1}{S(x)} \int_x^{\infty} t\, f_a(t)\, dt - x$$

after simplification, we get

$$MLR(x) = \frac{2\theta\Gamma(4,\theta t) + \Gamma(6,\theta t)}{S(x)\, 4\theta\, (\theta + 6)} \tag{12}$$

Where S(x) is the survival function (7).

Figures 3 and 4 are hazard function plot and survival function plot of the area-biased Xgamma distribution respectively.

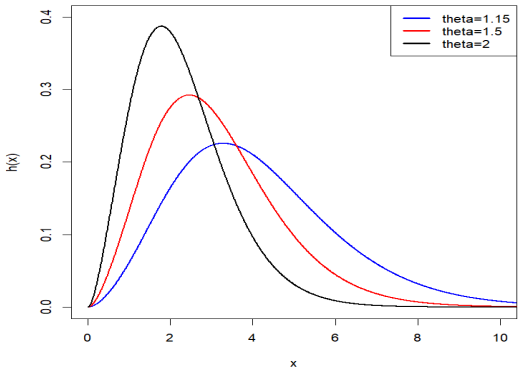


Figure 3: *Hazard function of area-biased Xgamma distribution*

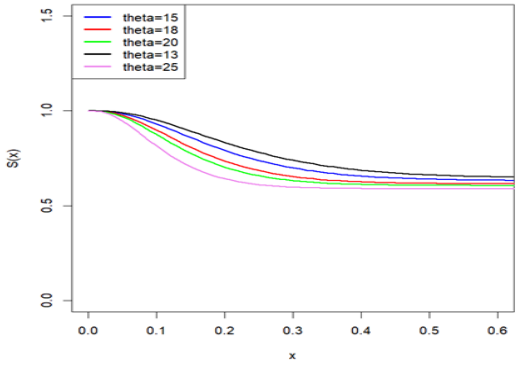


Figure 4: *survival function area-biased Xgamma distribution*

IV. Statistical Properties

In this section, discuss the different statistical properties of the area-biased Xgamma distribution that are moments, moment-generating functions, and characteristic functions.

I. Moments

Let the random variable X follow the area-biased Xgamma distribution with parameters θ; then the r^{th} order moment $E(X^r)$ of X about the origin can be obtained as

$$E(X^r) = \mu_r{}' = \int_0^{\infty} x^r f_a(x) dx$$

$$E(X^r) = \int_0^{\infty} x^r \frac{x^2\theta^4}{2(\theta+6)} \left(1 + \frac{\theta}{2} x^2\right) e^{-\theta x}\, dx$$

$$E(X^r) = \frac{\theta^4}{2(\theta+6)} \int_0^{\infty} x^2\, x^r \left(1 + \frac{\theta}{2} x^2\right) e^{-\theta x}\, dx \tag{13}$$

After simplification (13) get

$$E(X^r) = \mu_r{}' = \frac{2\theta\, (r+2)! + (r+4)!}{4\, (\theta+6)\, \theta^r} \tag{14}$$

Putting r=1,2,3 and 4 in equation (14), obtain the first four moments of the area-biased Xgamma distribution.

$$E(X^1) = \mu_1{}' = \frac{3\,\theta + 30}{\theta\,(\theta + 6)} \tag{15}$$

$$E(X^2) = \mu_2{}' = \frac{12\,\theta + 180}{\theta^2(\theta + 6)} \tag{16}$$

$$E(X^3) = \mu_3{}' = \frac{60\,\theta + 1260}{\theta^3(\theta + 6)} \tag{17}$$

$$E(X^4) = \mu_4{}' = \frac{360\theta + 10080}{\theta^4(\theta + 6)} \tag{18}$$

Property 1: Mean of the area-biased Xgamma distribution

$$mean = \frac{3\,\theta + 30}{\theta\,(\theta + 6)}$$

Property 2: Variance of the area-biased Xgamma distribution is given by

$$\text{variance} = \mu_2{}' - (\mu_1{}')^2$$

$$\text{variance} = \frac{3\theta^2 + 72\theta + 180}{\theta^2(\theta + 6)^2}$$

Property 3: Standard Deviation

$$\sigma = \frac{\sqrt{3\theta^2 + 72\theta + 180}}{\theta\,(\theta + 6)}$$

Property 4: Coefficient Of Variation

$$C.V = \left(\frac{\sigma}{\mu}\right) 100$$

$$C.V = \frac{\sqrt{3\theta^2 + 72\theta + 180}}{3\theta + 30}\ 100$$

II. Moment Generating Function and Characteristic Function

In this section, we discuss the moment-generating function and characteristic function of area-biased Xgamma distribution.

Theorem 1:
Let X be an area-biased Xgamma random variable with parameters θ; then the MGF of X is

$$M_X(t) = \frac{\theta^4(\theta - t)^2 - 6\theta^5}{(\theta + 6)(\theta - t)^5}$$

Proof: Let X~ABXG(θ),then the moment-generating function of X is obtained by

$$M_X(t) = E(e^{tx}) = \int_0^\infty e^{tx} f_a(x)dx \tag{19}$$

$$M_X(t) = E(e^{tx}) = \int_0^{\infty} e^{tx} \frac{x^2\theta^4}{2(\theta+6)} \left(1 + \frac{\theta}{2} x^2\right) e^{-\theta x} \, dx$$

$$M_X(t) = \frac{\theta^4}{2(\theta+6)} \int_0^{\infty} e^{-x(\theta-t)} x^2 \left(1 + \frac{\theta}{2} x^2\right) dx$$

After simplification we get the moment-generating function

$$M_X(t) = \frac{\theta^4(\theta-t)^2 - 6\theta^5}{(\theta+6)(\theta-t)^5} \tag{20}$$

Similarly, the characteristic function area-biased Xgamma distribution can be obtained as

$$\phi_X(t) = M_X(it)$$

$$\phi_X(t) = \frac{\theta^4(\theta-it)^2 - 6\theta^5}{(\theta+6)(\theta-it)^5} \tag{21}$$

V. Order statistics

In this section, we derived the distributions of order statistics from the area-biased Xgamma distribution. Let $X_{(1)}, X_{(2)}, X_{(3)}, \ldots . X_{(n)}$ be the order statistics of the random sample$X_1, X_2, X_3, \ldots . X_n$ selected from the area-biased Xgamma distribution. Then the probability density function of the r^{th}order statistics X(r) is defined as

$$f_{X(r)}(x) = \frac{n!}{(r-1)!\,(n-r)!} f_X(x)[F_X(x)]^{r-1}[1-F_X(x)]^{n-r} \tag{22}$$

Using equations (5) and (6) in equation (22), we get the probability density function of r^{th} order statistics of the area-biased Xgamma distribution.

$$f_{x(r)}(x) = \frac{n!}{(r-1)!\,(n-r)!} \left(\frac{x^2\theta^4}{2(\theta+6)}\left(1+\frac{\theta}{2}x^2\right)e^{-\theta x}\right)\left(\frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)}\right)^{r-1}\left(1 - \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)}\right)^{n-r} \tag{23}$$

Put r=1 in equation (23), we get first order statistics $X_{(1)}$=mini ($x_1, x_2, \ldots\ldots x_n$) and m=n, we get nth order statistics $X_{(n)}$=max ($x_1, x_2, \ldots\ldots x_n$) respectively.

$$f_{X(1)}(x) = \frac{nx^2\theta^4}{2(\theta+6)} \left(1+\frac{\theta}{2}x^2\right)e^{-\theta x}\left(1 - \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)}\right)^{n-1}$$

$$f_{X(n)}(x) = \frac{nx^2\theta^4}{2(\theta+6)} \left(1+\frac{\theta}{2}x^2\right)e^{-\theta x}\left(\frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)}\right)^{n-1}$$

VI. Likelihood Ratio Test

In this section, derive the likelihood ratio test from the area-biased Xgamma distribution. The likelihood-ratio test is a hypothesis test that compares two competing statistical models' goodness of fit. Let $X_1, X_2, \ldots, X_n$ be a random sample from the area biased samade distribution. To test the

hypothesis

$$H_0 : f(x) = f(x:\theta) \qquad \text{against} \qquad H_1 : f(x) = f_a(x:\theta)$$

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^{n} \frac{f_a(x:\theta)}{f(x:\theta)}$$

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^{n} \frac{\theta^2(1+\theta)}{2(\theta+6)}\, {x_i}^2$$

$$\Delta = \frac{L_1}{L_0} = \left(\frac{\theta^2(1+\theta)}{2(\theta+6)}\right)^n \prod_{i=1}^{n} {x_i}^2$$

We should reject the null hypothesis, if

$$\Delta = \left(\frac{\theta^2(1+\theta)}{2(\theta+6)}\right)^n \prod_{i=1}^{n} {x_i}^2 > k$$

Or equivalently, reject the null hypothesis

$$\Delta^* = \prod_{i=1}^{n} {x_i}^2 > k\left(\frac{2(\theta+6)}{\theta^2(1+\theta)}\right)^n$$

$$\Delta^* = \prod_{i=1}^{n} {x_i}^2 > k^* \ \ where\ k^* = k\left(\frac{2(\theta+6)}{\theta^2(1+\theta)}\right)^n$$

For a large sample size n, 2 log Δ is distribution as chi-square variates with one degree of freedom. Thus, we rejected the null hypothesis when the probability value is given by $p(\Delta^* > \alpha^*)$,where $\alpha^* = \prod_{i=1}^{n} {x_i}^2$ is less than the level of significance and $\prod_{i=1}^{n} {x_i}^2$is the observed value of the statistics Δ^*.

VII. Bonferroni and Lorenz Curves

In this section, we derived the Bonferroni and Lorenz curves from the area-biased Xgamma distribution. The Bonferroni and Lorenz curves are powerful tools in the analysis of distributions and have applications not only in economics but also in many fields, such as reliability analysis, medicine, insurance, and income. The Bonferroni is defined as

$$B(p) = \frac{1}{p{\mu_1}'} \int_0^q x\, f_a(x)dx$$

And

$$L(p) = \frac{1}{{\mu_1}'} \int_0^q x\, f_a(x)dx$$

Where

$${\mu_1}' = E(X) = \frac{3\,\theta + 30}{\theta\,(\theta+6)}$$

and $q = F^{-1}(p)$

$$B(p) = \frac{\theta\,(\theta+6)}{p(3\,\theta+30)} \int_0^q x \frac{x^2\theta^4}{2(\theta+6)} \left(1 + \frac{\theta}{2}\, x^2\right) e^{-\theta x}\, dx$$

$$B(p) = \frac{\theta^5\,(\theta+6)}{p(3\theta+30)2(\theta+6)} \int_0^q x^2 \left(1+\frac{\theta}{2}\,x^2\right) e^{-\theta x}\,dx$$

After simplification, we get

$$B(p) = \frac{(\theta+6)[\theta^2\gamma(3,\theta q) + \frac{\theta}{2}\gamma(5,\theta q)}{2p(3\theta+30)(\theta+6)} \tag{24}$$

Where $L(p) = pB(p)$

$$L(p) = \frac{(\theta+6)[\theta^2\gamma(3,\theta q) + \frac{\theta}{2}\gamma(5,\theta q)}{2(3\theta+30)(\theta+6)} \tag{25}$$

Figure 5 show the Lorence curve of area-biased Xgamma distribution

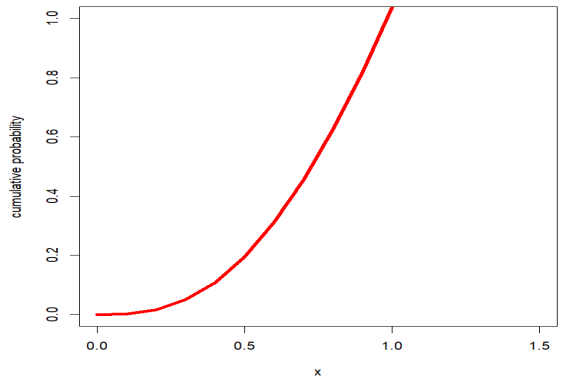


Figure 5: *Lorence curve of Area biased Xgamma distribution*

VIII. Entropies

In this section, we derived the entropies of the area-biased Xgamma distribution. The idea of entropy is essential in many domains, including probability, statistics, physics, communication theory, and economics. Entropies measure a systems diversity, uncertainty, or randomness. Entropy of a random variable X is a measure of variation in uncertainty.

I. Shannon Entropy

Shannon entropy of the area-biased Xgamma distribution is defined as

$$S_\tau = -\int_0^\infty f_\alpha(x)\ln(f_\alpha(x))dx \qquad \tau > 0, \tau \neq 1$$

$$S_\tau = -\int_0^\infty \frac{x^2\theta^4}{2(\theta+6)}\left(1+\frac{\theta}{2}\,x^2\right)e^{-\theta x}\ln\left(\frac{x^2\theta^4}{2(\theta+6)}\left(1+\frac{\theta}{2}\,x^2\right)e^{-\theta x}\right)dx$$

II. Renyi Entropy

The Renyi entropy is an essential diversity in ecology and statistics. It is also useful in quantum information, where it can be used to quantify entanglement. The Renyi entropy of the area-biased

Xgamma distribution is given by

$$R(\gamma) = \frac{1}{1-\gamma}\ln\int_0^\infty f_a(x)^\gamma dx$$

Where $\gamma > 0 \; and \; \gamma \neq 1$

$$R(\gamma) = \frac{1}{1-\gamma}\ln\int_0^\infty \left(\frac{x^2\theta^4}{2(\theta+6)}\left(1+\frac{\theta}{2}\,x^2\right)e^{-\theta x}\right)^\gamma dx$$

$$R(\gamma) = \frac{1}{1-\gamma}\ln\left(\frac{\theta^4}{2(\theta+6)}\right)^\gamma \int_0^\infty \left(x^2\left(1+\frac{\theta}{2}\,x^2\right)e^{-\theta x}\right)^\gamma dx$$

IX. Estimations of Parameter

The parameter estimation of the area-biased Xgamma distribution is obtained in this section. There are two estimation methods used in this section to estimate unknown parameters: the method of moment and maximum likelihood estimation.

I. Method of moment

Let $X_1, X_2 \ldots\ldots X_n$ be a random sample taken from area-biased Xgamma distribution with parameter θ. By the method of moment

$$\mu_r{}' = m_r{}'$$

Hence, we get

$$\mu_1{}' = m_1{}'$$

$$\frac{3\,\theta+30}{\theta\,(\theta+6)} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

After simplification ,we get

$$\bar{x}\,\theta^2 + \theta(6\bar{x}-3) - 30 = 0 \qquad (26)$$

Since the equation (26) is complicated, we can use the gmm package [10] from the R statistical software [14] to solve the equation (26).

II. Maximum likelihood estimation

Let $x_1, x_2 \ldots\ldots x_n$ be the random sample of size n taken from area-biased Xgamma distribution with parameter θ; then the likelihood function is

$$L(x) = \prod_{i=1}^{n} f_a(x)$$

$$L(x) = \prod_{i=1}^{n}\left(\frac{{x_i}^2\theta^4}{2(\theta+6)}\left(1+\frac{\theta}{2}\,{x_i}^2\right)e^{-\theta x_i}\right)$$

$$L(x) = \frac{\theta^4}{2(\theta+6)}\prod_{i=1}^{n}\left({x_i}^2\left(1+\frac{\theta}{2}x_i^2\right)e^{-\theta x_i}\right)$$

The log-likelihood function is given by

$$\ln L = n\ln\theta^4 - n\,ln2 - ln\,(6+\theta) + \sum_{i=1}^{n}\ln x_i^2 + \sum_{i=1}^{n}\ln(1+\frac{\theta}{2}x_i^2) - \theta\sum_{i=1}^{n}x_i \tag{27}$$

The normal equation is

$$\frac{\partial \ln L}{\partial\theta} = \frac{4n}{\theta} - \frac{n}{\theta+6} + \sum_{i=1}^{n}\frac{x_i^2}{2(1+\frac{\theta}{2}\,x_i^2)} - \sum_{i=1}^{n}x_i = 0 \tag{28}$$

The numerical method is used to solve the equations (28). We used the optimx package [10] from R statistical software [14] to obtain the maximum likelihood estimator of the area-biased Xgamma distribution.

X. Applications

In this section, we have considered three real data sets for the purpose of showing that the distribution of area-biased Xgamma distribution shows a better fit over Xgamma distribution and area-biased Aradhana distribution.

The Akaike information criterion (AIC), Bayesian information criterion (BIC), Akaike information criterion corrected (AICC), Hannan-Quinn information criterion (HQIC), and -2logl are used for model selection. It can be evaluated by using the formula as follows:

$$AIC = 2k - 2logL \qquad AICC = AIC + \frac{2k(k+1)}{(n-k-1)} \qquad BIC = klog(n) - 2logL$$

$$HQIC = 2klog(\log(n)) - 2logL$$

Where n is the sample size, k is the number of parameters, and -2logL is the maximal value of the log likelihood function. After the calculation of AIC, AICC, HQIC, and -2logL, the model with the minimum value is chosen as the best model to fit the data.

Data set 1: Breast cancer

The data set is the age of 155 patients with breast cancer from June to November 2014, (Breast Tumours Early Detection Unit, Benha Hospital University, Egypt).[8] Data are as follows

46, 32, 50, 46, 44, 42, 69, 31, 25, 29, 40, 42, 24, 17, 35, 48, 49, 50, 60, 26, 36, 56, 65, 48, 66, 44, 45, 30, 28, 40, 40, 50, 41, 39, 36, 63, 40, 42, 45, 31, 48, 36, 18, 24, 35, 30, 40, 48, 50, 60, 52, 47, 50, 49, 38, 30, 52, 52, 12, 48, 50, 45, 50, 50, 50, 53, 55, 38, 40, 42,42, 32, 40, 50, 58, 48, 32, 45, 42, 36, 30, 28, 38, 54, 90, 80, 60, 45, 40, 50,50, 40, 50, 50, 50, 60, 39, 34, 28, 18, 60, 50, 20, 40, 50, 38, 38, 42, 50, 40, 36, 38, 38, 50, 50, 31, 59, 40, 42, 38, 40, 38, 50, 50, 50, 40, 65, 38, 40, 38, 58, 35, 60, 90, 48, 58, 45, 35, 38, 32, 35, 38, 34, 43, 40, 35, 54, 60, 33, 35, 36, 43, 40, 45, 56

Table 1: *The summary of Brest cancer data set*

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
12.00	36.00	42.00	43.65	50.00	90.00

Data set 2: **blood cancer (leukemia)**

The following real lifetime (in years) data set consists of 40 patients suffering from blood cancer (leukemia) reported from one of the ministries of health hospitals in Saudi Arabia (see Abouammah et al. [2]). Data Set 2 is given by:

0.315, 0.496, 0.616,1.145, 1.208, 1.263, 1.414, 2.025, 2.036, 2.162, 2.211, 2.37, 2.532, 2.693, 3.858, 2.805, 2.91, 2.912, 3.192, 3.263, 3.348, 3.348, 3.427, 3.499, 3.534, 3.767, 3.751, 3.986, 4.049, 4.244, 4.323, 4.381, 4.392, 4.397, 4.753, 4.929, 4.973, 5.074, 5.381

Table 2: *The summary of blood cancer data set*

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.315	2.187	3.348	3.102	4.146	5.381

Data set 3: Head and Neck cancer data

The following real life data set of survival time of 44 patients diagnosed by head and Neck cancer disease are available in [6]. Data set 2 given by:

12.20, 23.56, 23.74, 25.87, 31.98, 37, 41.35, 47.38, 55.46, 58.36, 63.47, 68.46, 78.26, 74.47, 81.43, 84, 92, 94, 110, 112, 119, 127, 130, 133, 140, 146, 155, 159, 173, 179, 194, 195, 209, 249, 281, 319, 339, 432, 469, 519, 633, 725, 817, 1776

Table 3: *The summary of Head and Neck cancer data set*

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
12.20	67.21	128.50	223.48	219.00	1776.00

Table 4: *MLEs AIC, BIC, AICC, HQIC and -2log $\mathcal{L}$ of the fitted distribution for the given data set 1*

Distribution	MLE	-2logL	AIC	BIC	AICC	HQIC
Area biased Xgamma	$\hat{\theta}$=1.4815838 (0.1101315)	139.7668	141.7668	143.4303	141.8749	142.3637
Xgamma	$\hat{\theta}$=0.73392828 (0.07638865)	155.4983	157.4983	159.1619	157.6064	158.0952
Area biased Aradhana	$\hat{\theta}$=1.4311764 (0.1038796)	141.485	143.485	145.1485	143.5931	144.0819

Table 5: *MLEs AIC, BIC, AICC, HQIC and -2log $\mathcal{L}$ of the fitted distribution for the given data set2*

Distribution	MLE	-2logL	AIC	BIC	AICC	HQIC
Area biased Xgamma	$\hat{\theta}$=0.100000 (0.006751215)	1554.055	1556.055	1557.84	1556.151	1556.717
Xgamma	$\hat{\theta}$=0.100000 (0.008736029)	1790.128	1792.128	1793.128	1792.223	1792.79
Area biased Aradhana	$\hat{\theta}$=0.100000 (0.006741885)	1555.373	1557.373	1559.157	1557.468	1558.035

Table 6: *MLEs AIC, BIC, AICC, HQIC and -2log $\mathcal{L}$ of the fitted distribution for the given data set3*

Distribution	MLE	-2logL	AIC	BIC	AICC	HQIC
Area biased Xgamma	$\hat{\theta}$=0.113843128 (0.004098989)	1262.845	1264.845	1267.888	1264.871	1266.081
Xgamma	$\hat{\theta}$=0.100000 (0.004664461)	1418.24	1420.24	1423.284	1420.266	1421.476
Area biased Aradhana	$\hat{\theta}$=0.113273264 (0.004069012)	1262.981	1264.981	1268.024	1265.007	1266.217

From tables 4, 5, and 6 it can be clearly observed and seen from the results that the area-biased Xgamma distribution has the lesser AIC, BIC, AICC, HQIC, -2logL, and values as compared to the Xgamma and area-biased Aradhana, which indicates that the area-biased Xgamma distribution better fits than the Xgamma and area-biased Aradhana distributions. Hence, it can be concluded that the area-biased Xgamma distribution leads to a better fit over the other distributions.

XI. Conclusion

Area-Biased Xgamma distribution is a generalization of Xgamma distribution, was derived, and studied the different properties.Which is the lifetime model for a real-life data set. In the section 3, 4, 5, & 6 discussing the statistical property area-biased Xgamma distribution. The effectiveness and applicability of the suggested model is demonstrated by an examination of three real cancer data sets and compared with the Xgamma distribution and area-biased Aradhana distribution. The result indicates that the area-biased Xgamma distribution is more flexible and practical than the Xgamma distribution and area-biased Aradhana distribution.

Reference

[1] Ade, R. B., Teltumbade, D. P., & Tasare, P. W. (2021). Charecterizaton And Estimation of Area Biased Quasi Akash Distribution. *International Journal of Mathematics Trends and Technology-IJMTT*, *67*

[2]Atikankul, Y., Thongteeraparp, A., Bodhisuwan, W., & Volodin, A. (2020). The length-biased weighted Lindley distribution with applications. Lobachevskii Journal of Mathematics, 41, 308-319

[3] Bakoban, R. A. (2010). A study on mixture of exponential and exponentiated gamma distributions. Advances and Applications in Statistical Sciences, 2(1), 101-127.

[4] Bartoszewicz, J. (2009). On a representation of weighted distributions. Statistics & probability letters, 79(15), 1690-1694

[5] Dey, S., Ali, S., & Park, C. (2015). Weighted exponential distribution: properties and different methods of estimation. Journal of Statistical Computation and Simulation, 85(18), 3641-3661.

[6] Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. Journal of the American statistical Association, 83(402), 414-425

[7] Fisher, R. A. (1934). The effects of methods of ascertainment upon the estimation of frequencies. Annals of Eugenics, 6:13-25.

[8] Hassan, A. S., Elbatal, I., & Hemeda, S. E. (2016). Weibull quasi-Lindley distribution and its statistical properties with applications to lifetime data. International Journal of Applied Mathematics and Statistics, 55(3), 63-80..

[9] Patil, G. P. and Rao, C. R. (1978). Weighted distributions and Size biased sampling with applications to wildlife populations and human families. Biometrics, 34:179-189

[10] P. Chauss e, Computing generalized method of moments and generalized empirical likelihood with R, J. Stat. Software 34 (11), 1 35 (2010).

[11] Perveen, Z. Munir, M. Ahmed, Z. and Ahmad, M. (2016). On area biased weighted weibull distribution. Sci. Int. (Lahore), 28:3669- 3679

[12] Rao, C. R. and Patil, G. P. (1965). On discrete distributions arising out of methods of ascertainment, in Classical and Contagious Discrete Distribution. Pergamon Press and Statistical Publishing Society, Calcutta, 320-332.

[13] Ray, M., & Shanker, R. (2024). A PROBABILITY MODEL FOR SURVIVAL ANALYSIS OF CANCER PATIENTS. Reliability: Theory & Applications, 19(3 (79)), 78-94.

[14]RCoreTeam,R:ALanguageandEnvironmentforStatisticalComputing,RFoundationforStatistical Computing (Vienna, Austria, 2018). https://www.R-project.org/.Accessed 2019

[15] Sen, S., Chandra, N., & Maiti, S. S. (2017). The weighted xgamma distribution: properties and application.

[16] Sen, S., Maiti, S. S., & Chandra, N. (2016). The xgamma distribution: statistical properties and application. Journal of Modern Applied Statistical Methods, 15(1), 38.

[17] Sharma, VK. Dey, S. Singh, SK. and Manzoor, U. (2017). On length and area-biased Maxwell distributions. Communications in Statistics Simulation and Computation, 47:1506

APPLICATION OF TYPE-2 FUZZY TOPSIS METHOD FOR ESTIMATING RENEWABLE ENERGY SOURCES

Kamala Aliyeva
•
Azerbaijan State Oil and Industry University
kamala.aliyeva@asoiu.edu.az

Abstract

As global energy demands rise and the adverse effects of climate change become increasingly evident, the need for clean, sustainable energy sources has never been more critical. Renewable energy options, which harness natural processes like sunlight, wind, water, and geothermal heat, offer promising solutions to reduce dependence on fossil fuels and mitigate environmental damage. Estimating renewable energy options involves evaluating various energy sources to determine their suitability based on several criteria. Renewable energy potential varies significantly across different regions, influenced by factors such as natural resources, technological development, and policy support. Some regions are naturally more suited for certain types of renewable energy due to their climate, geography, and resources. The efficiency and cost-effectiveness of renewable technologies can vary based on how advanced technology is in each region. Government policies and incentives play a huge role in determining how much potential renewable energy can be developed. The key objective is to choose the most viable renewable energy sources for a given region or application by considering factors like efficiency, flexibility, and reliability. The TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) method helps decision-makers identify the preferred alternatives by evaluating the relative proximity of each alternative to an ideal solution. Unlike traditional type-1 fuzzy numbers, type-2 fuzzy numbers provide an additional layer of uncertainty by representing a fuzzy set with both primary and secondary membership functions. This allows capturing more detailed and complex uncertainties. Type-2 fuzzy numbers can be integrated into TOPSIS to handle the inherent uncertainty and imprecision in decision-making, particularly in the context of renewable energy options, where data is often uncertain or vague. This paper offers the integration of spectrum selection optimization algorithms based on the evaluation of the characteristics of various renewable energy options using the TOPSIS fuzzy method.

Keywords: Renewable energy options, type-2 fuzzy numbers, multi-criteria decision making, TOPSIS method

I. Introduction

As global energy demands rise and the adverse effects of climate change become increasingly evident, the need for clean, sustainable energy sources has never been more critical. Renewable energy options, which harness natural processes like sunlight, wind, water, and geothermal heat, offer promising solutions to reduce dependence on fossil fuels and mitigate environmental damage. The shift toward renewable energy is not only a necessity for environmental sustainability but also presents opportunities for economic growth and energy security. Among the most widely recognized renewable energy sources, solar energy uses photovoltaic cells to convert sunlight

directly into electricity. Solar power is abundant, particularly in regions with high levels of sunlight, making it an attractive option for decentralized power generation. Technology has advanced significantly, with the cost of solar panels decreasing over the past decade, making solar power more affordable and accessible to both individual and large-scale projects. Wind power harnesses the kinetic energy of wind through turbines to generate electricity. Onshore and offshore wind farms are becoming an increasingly common sight across the globe. Wind energy is particularly effective in areas with strong and consistent wind currents, such as coastal regions and open plains. Hydropower, one of the oldest and most established renewable energy sources, generates electricity by harnessing the energy of flowing water. Large-scale hydropower projects typically involve dams that store water in reservoirs, which are then released to drive turbines. Smaller, run-of-river projects have also been developed to reduce the environmental impacts associated with damming rivers. Geothermal energy taps into the Earth's natural heat, often through hot springs, geysers, or geothermal reservoirs found deep beneath the surface. This energy source can be used for heating or converting it into electricity. Geothermal power plants are particularly effective in regions with significant tectonic activity. Biomass energy comes from organic materials such as wood, agricultural residues, and waste. Biomass can be burned directly for heat or converted into biofuels like ethanol and biodiesel for transportation. Biomass energy is considered carbon-neutral because the carbon dioxide released during combustion is offset by the carbon dioxide absorbed by plants during their growth. Ocean energy encompasses both tidal and wave energy, harnessing the power of the oceans to generate electricity. Tidal energy relies on the rise and fall of tides, while wave energy captures the motion of the ocean's surface. Although still in the early stages of development, ocean energy holds significant potential. As the world confronts the challenges of climate change, energy security, and environmental degradation, the need to transition to renewable energy sources has become more urgent. The selection of the most suitable renewable energy option depends on a variety of factors such as cost, efficiency, flexibility, reliability, environmental impact, technological maturity, etc. However, these factors often involve uncertainties that make decision-making difficult. Governments, businesses, and individuals must work together to promote the growth of renewable energy, develop innovative solutions to overcome challenges, and ensure that future generations can enjoy the benefits of a healthier, more sustainable planet.

MCDM (Multi-Criteria Decision Making) techniques are widely used to evaluate and rank renewable energy alternatives based on multiple criteria [1]. Traditional decision-making methods like Weighted Sum Model (WSM) [2], Analytic Hierarchy Process (AHP) [3], and Elimination and Choice Expressing Reality (ELECTRE) [4] have been applied to renewable energy selection. These methods are effective in aggregating various factors, but they often fail to handle the vagueness and uncertainty associated with real-world data. For instance, cost estimations, environmental impact assessments, and efficiency measures of renewable energy systems are typically subject to uncertainty, making it difficult to apply crisp values in traditional methods. The use of fuzzy sets in MCDM allows decision-makers to express uncertainty in criteria, where each value is represented by a fuzzy number instead of a precise number [5]. Type-1 fuzzy numbers are commonly used to model such uncertainties, but they still do not fully capture the complexity of the decision-making process, especially when the degree of uncertainty itself is uncertain [6]. This limitation has prompted the exploration of type-2 fuzzy sets, which provide a more sophisticated representation of uncertainty. TOPSIS is a popular MCDM method for ranking alternatives based on their distance from an ideal and a negative ideal solution. Hwang and Yoon (1981) introduced TOPSIS as a technique to determine the best option by minimizing the distance from the ideal solution and maximizing the distance from the negative ideal solution [7]. The method involves constructing a decision matrix, normalizing the data, calculating the weighted normalized decision matrix, and then computing the separation measures (i.e., Euclidean distance) from the ideal and negative ideal solutions [8]. In the context of renewable energy, TOPSIS has been applied in several studies to compare and evaluate different energy sources (e.g., solar, wind, hydropower) based on criteria such as cost, efficiency, environmental impact, and reliability. Sengul et al. applied fuzzy TOPSIS for the

evaluation of renewable energy alternatives in Turkey, ranking solar, wind, and biomass energy based on economic, environmental, and technical factors [9]. Similarly, Kaya and Kahraman used the fuzzy TOPSIS method to assess the sustainability of various renewable energy technologies [10]. However, one limitation of TOPSIS in its classical form is its inability to handle uncertain or imprecise data. This is particularly important when evaluating renewable energy options, as many of the data points used in these evaluations are subject to variability. This has led to the incorporation of fuzzy logic into TOPSIS, with researchers proposing fuzzy-TOPSIS models to address these issues [11]. Goumas and Lygerou used the PROMETHEE method to assess the sustainability of various renewable energy technologies [12]. Similarly, Ramanathan employed fuzzy AHP to assess the suitability of various renewable energy resources (e.g., wind, solar, biomass) [13]. It used fuzzy numbers to express vagueness in evaluating factors like environmental impact, social acceptance, and availability of resources. Levitin et all. used genetic algorithm for open-loop distribution system design [14]. Mamlook and et all. used neuro-fuzzy program approach for evaluating electric power generation systems [15]. The integration of TOPSIS with type-2 fuzzy numbers enhances the robustness of the decision-making process by allowing for more precise handling of uncertainty [16].

Type-2 fuzzy TOPSIS provides more reliable results compared to traditional methods, especially when the decision data are uncertain or imprecise. This methodology is widely used for multi-criteria decision-making problems, where a set of alternatives needs to be ranked based on multiple conflicting criteria in deeply uncertain conditions. The application of the TOPSIS method combined with type-2 fuzzy numbers offers a powerful framework for evaluating and selecting renewable energy options. This approach offers more reliable decision-making than traditional methods by tackling the uncertainty in energy costs, efficiency, and environmental impact data. The literature shows increasing interest in using this hybrid approach for selecting renewable energy, with promising outcomes in several case studies. This paper explores how the TOPSIS method, when applied with type-2 fuzzy numbers, can improve decision-making in renewable energy selection. Paper is structured as follows. In section 2 literature review is given. Section 2 represents the major definitions of type-2 fuzzy numbers and the TOPSIS steps that are utilized for estimation problems. In section 3, the TOPSIS methodology using type-2 fuzzy numbers is applied to solve the problem of estimating the renewable energy options. Conclusion represents the main results developed in this paper.

II. Preliminaries

Definition 1. A type-1 fuzzy set A in the universe X is characterized by a membership function and is expressed in figure 1 [17]:

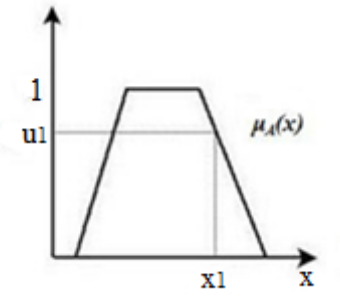


Figure 1: *Type-1 fuzzy set*

$$A = \left\{ \left(x, \mu_A(x) \right) \mid \forall x \in X, \mu_A(x) \in [0,1] \right\} \quad (1)$$

Definition 2. A type 2 fuzzy set $\tilde{A}$ is characterized by a fuzzy membership function $\mu_{\tilde{A}}(x,u)$ in the universe X where $x \in X$ and $u \in J_X, J_X \subseteq [0,1]$ is shown in figure 2 [17]:

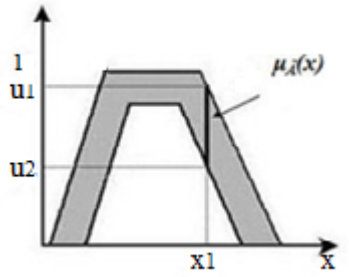


Figure 2: *Type-2 fuzzy set*

$$\tilde{A} = \left\{ \left(x, u, \mu_{\tilde{A}}(x,u) \right) \mid \forall x \in X, u \in [0,1] \right\} \tag{2}$$

Definition 3. In classical mathematics, functions operate on precise (crisp) inputs and produce crisp outputs. But in fuzzy logic, inputs and/or outputs can be fuzzy sets (i.e., sets with degrees of membership between 0 and 1). The extension principle provides a way to apply a function $f : X \to Y$ to fuzzy inputs and obtain a fuzzy output. Suppose that $\tilde{A} = \sum_i \mu_{\tilde{A}}(x_i) / x_i$ and $\tilde{B} = \sum_j \mu_{\tilde{B}}(x_j) / x_j$ are fuzzy sets included in the set X. According to Zadeh extension principal [18]

$$\tilde{A} * \tilde{B} = \left(\sum_i \mu_{\tilde{A}}(x_i) / x_i \right) * \left(\sum_j \mu_{\tilde{B}}(x_j) / x_j \right) = \sum_{i,j} \left(\mu_{\tilde{A}}(x_i) \wedge \mu_{\tilde{B}}(x_j) \right) / \left(x_i * x_j \right) \tag{3}$$

Extension principle allows to extend classical function f from a set X to a fuzzy set A on X. For each element x in A, the degree of membership in the fuzzy set $f(A)$ is determined by applying the function f to the degree of membership of x in A.

$$\mu_{f(A)}(y) = \sup_{x \in X, f(x)=y} \mu_A(x) \tag{4}$$

This means the membership function of the fuzzy set $f(A)$ at y is maximum membership degree of all elements x in A such that $f(x) = y$. The Extension Principle also applies to set operations. The union of two fuzzy sets A and B is given below.

$$\mu_{A \cup B}(x) = \max\left(\mu_A(x), \mu_B(x) \right) \tag{5}$$

The intersection is given below.

$$\mu_{A \cap B}(x) = \min\left(\mu_A(x), \mu_B(x) \right) \tag{6}$$

The complement of a fuzzy set is given below.

$$\mu_{\neg} A(x) = 1 - \mu_A(x) \tag{7}$$

Definition 4. If $\mu_{\tilde{A}}(x)$ and $\mu_{\tilde{B}}(x)$ are membership degrees of type-2 fuzzy sets $\tilde{A}$ and $\tilde{B}$ of the fuzzy set $J \in [0,1]$, then

$$\mu_{\tilde{A}}(x) = f(u_1)/u_1 + f(u_2)/u_2 + \cdots + f(u_n)/u_n = \sum_i f(u_i)/u_i, \quad u_i \in J \tag{8}$$

$$\mu_{\tilde{B}}(x) = g(w_1)/w_1 + g(w_2)/w_2 + \cdots + g(w_m)/w_m = \sum_j f(w_j)/w_j, \quad w_j \in J \tag{9}$$

where f and g are membership functions.

Definition 5. The union and intersection of type-2 fuzzy sets are defined as below [18]:

$$\begin{aligned} &\tilde{A} \cup \tilde{B} \Leftrightarrow \mu_{\tilde{A}\cup\tilde{B}}(x) = \mu_{\tilde{A}}(x) \cup \mu_{\tilde{B}}(x) = \left(\sum_i f(u_i)/u_i\right) \cup \left(\sum_i f(w_j)/w_j\right) = \\ &= \sum_{i,j} \left(f(u_i) \vee g(w_j)\right)/(u_i \vee w_j) \end{aligned} \tag{10}$$

$$\begin{aligned} &\tilde{A} \cap \tilde{B} \Leftrightarrow \mu_{\tilde{A}\cap\tilde{B}}(x) = \mu_{\tilde{A}}(x) \cap \mu_{\tilde{B}}(x) = \left(\sum_i f(u_i)/u_i\right) \cap \left(\sum_i f(w_j)/w_j\right) = \\ &= \sum_{i,j} \left(f(u_i) \wedge g(w_j)\right)/(u_i \wedge w_j) \end{aligned} \tag{11}$$

Definition 6. The defuzzification process for type-2 fuzzy numbers is more complicated than for type-1 fuzzy sets due to the two-dimensional nature of type-2 fuzzy sets. However, the most common defuzzification methods for type-2 fuzzy numbers are represented below.

Center of gravity (COG) method: Calculates the centroid of the type-2 fuzzy number considering both the primary and secondary membership functions.

$$\tilde{x} = \frac{\int_X \int_0^1 x \cdot \mu_A(x,u)\,du\,dx}{\int_X \int_0^1 \mu_A(x,u)\,du\,dx} \tag{12}$$

where $\int_X$ is the integration over the universe of discourse X, $\int_0^1 x$ is the integration over the secondary membership values, $\mu_A(x,u)$ is the membership function of the type-2 fuzzy number at x, u. Formula gives the crisp value x, which represents defuzzified center of the type-2 fuzzy set.

Alpha-cut method: Cuts the fuzzy number at different levels of α, treats each cut as a type-1 fuzzy set, defuzzifies each cut, and then averages the results.

$$\tilde{x}_\alpha = \frac{\int_X x \cdot \mu A_\alpha(x)\,dx}{\int_X \mu A_\alpha(x)\,dx} \tag{13}$$

Weighted average of the defuzzified values for each α-cut is taken, where the weight is based on the membership grade α:

$$\tilde{x} = \int_0^1 \tilde{x}_\alpha \cdot \alpha\, d\alpha \tag{14}$$

Mean of Maxima (MOM) method: Computes the average of the points where the membership function is maximized. For type-2 fuzzy numbers, this involves finding the values of x that correspond to the maximum membership degree for each possible secondary membership u, and then averaging these values.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{15}$$

where x_i are the values corresponding to the maximum degree of membership at each u.

III. Case Study Example: Renewable energy sources estimation problem

Suppose that there are three possible alternatives - A (wind), B (hydropower), C (solar) for estimation renewable energy sources with three criteria's: C_1 - efficiency (refers to how efficient each energy source is in converting the natural resource into usable energy), C_2 - flexibility (refers to the ability of an energy system or resource to adapt to changing demand, supply, or operational conditions), C_3 - reliability (the consistency and dependability of each energy source in providing power, especially considering factors like intermittency and weather dependence). Wind energy (A) is a key player in the transition to a cleaner, more sustainable energy future. While it faces challenges like intermittency, space requirements, and environmental impact concerns, its advantages - such as zero emissions, low operating costs, and abundant availability - make it one of the most promising renewable energy sources. Technological advancements in turbine efficiency, energy storage, and offshore wind development will likely overcome many of these challenges, making wind energy an even more integral part of the global energy mix in the coming years. Hydropower (B) remains a cornerstone of the renewable energy landscape due to its reliability, efficiency, and scalability. While it faces challenges, such as environmental impacts and high upfront costs, it offers significant benefits, including clean energy generation, energy storage capabilities, and long operational lifespans. As technology advances, hydropower is becoming more efficient and less intrusive, ensuring its role in the global transition to a sustainable and low-carbon energy future. Solar energy (C) is an essential part of the global transition to a sustainable, cleaner future. While it faces challenges such as intermittency and space requirements, advances in technology and policy support are making solar energy more viable and affordable. As the cost of solar energy continues to decrease and storage solutions improve, solar power will play an increasingly significant role in reducing dependence on fossil fuels and mitigating climate change.

The efficiency (C_1) of converting the resources (wind, sunlight, water flow, etc.) into usable energy is crucial. For example, modern photovoltaic solar panels can have efficiencies ranging from 15% to 22%, while wind turbines have efficiencies of 35-45%. Flexibility (C_2) in renewable energy selection refers to how adaptable and scalable an energy source is, both in terms of installation and integration into an existing grid or energy system. Flexibility also includes the ability to adjust power output based on changing demands, technological advancements, and regional or local conditions. Hybrid systems combining different energy sources, for example, wind and solar or integrating with traditional power sources can offer greater flexibility in maintaining stable energy output. The reliability (C_3) of an energy source depends on its ability to provide consistent power. Wind and solar can be intermittent (i.e., depending on weather), while biomass energy may be more reliable due to its consistent fuel source. Weights of criteria are w_1 = (0.2,0.3,0.4,0.4; 0.3,0.4,0.5;0.5), w_2 = (0.4,0.5,0.6,0.6; 0.5,0.6,0.7;0.7), w_3 = (0.4,0.5,0.6,0.6; 0.5,0.6,0.7;0.7). Linguistic values of type-2 fuzzy sets of linguistic terms are shown below:

Absolutely low - (0,0,0;0), Very low - (0.1,0.2,0.3;0.3), Low - (0.2,0.3,0.4;0.4), Slightly low - (0.25,0.35,0.45;0.45), Medium - (0.3,0.4,0.5;0.5), Slightly high - (0.4,0.5,0.6;0.6), High - (0.5,0.6,0.7;0.7), Very high - (0.6,0.7,0.8;0.8), Absolutely high - (0.7,0.8,0.9;0.9).

The detailed steps of the solution of estimation problem type-2 fuzzy TOPSIS modification are shown below [19]:

Step 1: Determining a decision matrix. A decision matrix is essentially a table where each alternative

is evaluated based on several criteria. Determining a decision matrix using linguistic terms and corresponding type-2 fuzzy numbers based on the decision maker's preferences involves a structured approach for evaluating alternatives under uncertainty. The idea is to represent the decision maker's qualitative judgments expressed in linguistic terms and convert them into fuzzy numbers that account for uncertainty and vagueness, allowing for a more flexible and accurate decision-making process. Decision matrix applying type-2 fuzzy numbers given in table 1.

Table 1: *Type-2 fuzzy decision matrix*

	C_1	C_2	C_3
A	(0.4,0.5,0.6;0.6), (0.5,0.6,0.7;0.7)	(0.3,0.4,0.5;0.5), (0.4,0.5,0.6;0.6)	(0.5,0.6,0.7;0.7) (0.6,0.7,0.8;0.8)
B	(0.6,0.7,0.8;0.8), (0.7,0.8,0.9;0.9)	(0.5,0.6,0.7;0.7), (0.6,0.7,0.8;0.8)	(0.5,0.6,0.7;0.7) (0.6,0.7,0.8;0.8)
C	(0.3,0.4,0.5;0.5), (0.5,0.6,0.7;0.7)	(0.3,0.4,0.5;0.5) 0.5,0.6,0.7;0.7)	(0.6,0.7,0.8;0.8), (0.7,0.8,0.9;0.9)

Step 2: Determining the weighted decision matrix. The weighted decision matrix is used to evaluate and compare different alternatives across multiple criteria, considering both the relative importance (weights) of each criterion values of the performance scores. Determining the weighted decision matrix $V=\left(v_{ij}\right)_{n\times m}$ by employing the next formula is represented in table 2.

$$v_{ij}=w_j r_{ij}; i=1,\cdots,n;\ j=1,\cdots,m \tag{16}$$

Table 2: *Weighted decision matrix*

	C_1	C_2	C_3
A	(0.08,0.15,0.24;0.24), (0.15,0.24,0.35;0.35)	(0.12,0.2,0.3;0.3), (0.2,0.3,0.42;0.42)	(0.2,0.3,0.42;0.42), (0.3,0.42,0.56;0.56)
B	(0.12,0.21,0.32;0.32), (0.21,0.32,0.45;0.45)	(0.2,0.3,0.42;0.42), (0.3,0.42,0.56;0.56)	(0.2,0.3,0.42;0.42), (0.3,0.42,0.56;0.56)
C	(0.06,0.12,0.2;0.2), (0.15,0.24,0.35;0.35)	(0.12,0.2,0.3;0.3) 0.25,0.36,0.49;0.49)	(0.24,0.35,0.48;0.48), (0.35,0.48,0.63;0.63)

Step 3: Determining positive and negative optimal solutions: In the TOPSIS methodology, the concept of the positive ideal solution (PIS) and the negative ideal solution (NIS) play a very important role in ranking the options. These two concepts represent the best and worst possible outcomes for each criterion. The PIS (A^*) is the best possible set of values for all criteria where possible. It represents the ideal or most desirable scenario for each alternative, based on the given decision criteria. For benefit criteria where higher values are better, the PIS will have the maximum values for each criterion. For cost criteria where lower values are better, the PIS will have the minimum values. The NIS (A^-) is the worst possible set of values for all criteria. It represents the least desirable scenario for each alternative. For benefit criteria, the NIS will have the minimum values for each criterion. For cost criteria, the NIS will have the maximum values. So, PIS is the ideal best - closest to what we would want to achieve, NIS is the ideal worst - farthest from what we want. Positive and negative optimal solutions are calculated as below.

$$A^*=\left\{v_1^*,\cdots,v_m^*\right\}=\left\{\max v_{ij}^* \mid j\in\Omega_b,\quad \min v_{ij}^* \mid j\in\Omega_b\right\} \tag{17}$$

$$A^-=\left\{v_1^-,\cdots,v_m^-\right\}=\left\{\max v_{ij}^- \mid j\in\Omega_b,\quad \min v_{ij}^- \mid j\in\Omega_b\right\} \tag{18}$$

$A^* = \{ (0.12,0.21,0.32;0.32),(0.21,0.32,0.45;0.45);(0.2,0.3,0.42;0.42),(0.3,0.42,0.56;0.56);$
$(0.24,0.35,0.48;0.48),(0.35,0.48,0.63;0.63) \}$

$A^- = \{ (0.06,0.12,0.2;0.2),(0.15,0.24,0.35;0.35);(0.12,0.2,0.3;0.3)0.25,0.36,0.49;0.49);$
$(0.2,0.3,0.42;0.42),(0.3,0.42,0.56;0.56) \}$

Step 4: Determining separation measure and proximity to the optimal solution. The separation measure refers to the distance between each alternative and the positive ideal solution (PIS) and the negative ideal solution (NIS). The separation measure quantifies how far each alternative is from the best and worst possible outcomes. The smaller the distance to the PIS (the ideal best), the better the alternative. Conversely, the larger the distance to the NIS (the ideal worst), the better. The Euclidean distance formula is commonly used to calculate the separation between an alternative and the ideal solutions. Separation measures and proximity to the optimal solutions are determined by employing Euclidean distance as shown below [19]:

$$D^* = \sqrt{\sum_{i=1}^{m} \left(v_{ij} - v_j^*\right)^2}; i = 1,\cdots,n \tag{19}$$

$$D^- = \sqrt{\sum_{i=1}^{m} \left(v_{ij} - v_j^-\right)^2}; i = 1,\cdots,n \tag{20}$$

Negative and positive separation measure for each option calculated as below:

$D_A^* = (0.03,0.02,0.1;0.01,0.16,0.034)$ $D_A^- = (0.016,0.025,0.036;0.0025,0.0036,0.0049)$

$D_B^* = (0.016,0.025,0.036;0.025,0.036,0.049)$ $D_B^- = (0.014,0.018,0.028;0.061,0.01,0.015)$

$D_C^* = (0.01,0.018,0.028;0.061,0.01,0.015)$ $D_C^- = (0.014,0.018,0.028;0.0025,0.0036,0.0049)$

Step 5: Ranking the options by computing the relative closeness of each alternative to the optimal solution. After calculating the separation measures, proximity to the optimal solution is represented by the closeness coefficient. The Closeness coefficient quantifies the relative closeness of each alternative to the ideal solution. It is a ratio that measures how near an alternative is to the positive ideal (PIS) and how far it is from the negative ideal (NIS). Closeness coefficient (CC) of the option is determined by applying calculations shown below:

$$CC_i = \frac{D_i^-}{D_i^* + D_i^-}; i = 1,\cdots,n \tag{21}$$

where, D_i^- is separation distance of alternative i to the negative ideal solution (NIS), D_i^* is separation distance of alternative i to the positive ideal solution (PIS).
The relative closeness for each option calculated as below:

$$C_A^* = \frac{(0.016,0.025,0.036;0.0025,0.0036,0.0049)}{(0.03,0.02,0.01;0.01,0.016,0.034)+(0.016,0.025,0.036;0.0025,0.0036,0.0049)} =$$
$$= \frac{(0.016,0.025,0.036;0.0025,0.0036,0.0049)}{(0.046,0.045,0.046;0.0185,0.0286,0.0389)} = (0.3478,0.5556,0.7826;0.0643,0.1254,0.2649)$$

$$C_B^* = \frac{(0.014, 0.018, 0.028; 0.061, 0.01, 0.015)}{(0.016, 0.025, 0.036; 0.025, 0.036, 0.049) + (0.014, 0.018, 0.028; 0.061, 0.01, 0.015)} =$$

$$= \frac{(0.014, 0.018, 0.028; 0.061, 0.01, 0.015)}{(0.030, 0.043, 0.064; 0.086, 0.046, 0.064)} = (0.2188, 0.4186, 0.9333; 0.9531, 0.2174, 0.2344)$$

$$C_C^* = \frac{(0.01, 0.018, 0.028; 0.061, 0.01, 0.015)}{(0.01, 0.018, 0.028; 0.061, 0.01, 0.015) + (0.014, 0.018, 0.028; 0.061, 0.01, 0.015)} =$$

$$= \frac{(0.01, 0.018, 0.028; 0.061, 0.01, 0.015)}{(0.024, 0.036, 0.056; 0.086, 0.046, 0.064)} = (0.25, 0.5, 1.1667; 0.3906, 0.7826, 0.5698)$$

Comparing all three alternatives (wind, hydropower, and solar energy) we determine that option wind energy (A) source is the best alternative.

IV. Conclusion

Renewable energy options refer to the various sources of energy that are naturally replenished on a human timescale and have minimal environmental impacts compared to traditional fossil fuels. These energy sources are crucial for reducing greenhouse gas emissions, improving energy security, and contributing to sustainability. Estimation renewable energy options require a thorough assessment of available resources, technological capabilities, economic viability, environmental impacts, social factors etc. By considering factors like energy potential, costs, storage needs, and regulatory frameworks, it can be identified the most suitable renewable energy options for a specific area or project. Estimating renewable energy options using type-2 fuzzy numbers is a powerful way to handle uncertainty and imprecision in renewable energy decision-making - especially when expert opinions or environmental data are vague, conflicting, or incomplete. Using the TOPSIS method with type-2 fuzzy numbers is a great way to rank renewable energy options under uncertainty. Type-2 fuzzy TOPSIS helps when expert judgments are imprecise or ambiguous, as is common in energy planning. This approach is especially helpful in multi-stakeholder, data-uncertain environments like energy planning, policy development, or rural electrification. For making decisions on renewable energy sources estimation in this article is applied fuzzy TOPSIS methodology with type-2 fuzzy numbers to consider high uncertainty in estimation problem. In this paper three options A (wind), B (hydropower), C (solar) with three criteria's C_1 - efficiency, C_2 - flexibility, and C_3 - reliability are used for renewable energy options estimation. By employing offered method it was identified that wind energy (A) source is more appropriate for selection.

V. Discussion

The decision-making process for selecting renewable energy sources is inherently complex and involves several challenges due to the numerous factors that must be considered. These challenges arise from the interplay of environmental, social, and economic factors that influence the feasibility and sustainability of renewable energy projects. Given the complexity and uncertainty involved in selecting renewable energy sources, Multi-Criteria Decision-Making (MCDM) methods are crucial for a systematic and comprehensive evaluation process. MCDM methods like type-2 fuzzy TOPSIS, can model and account for uncertainties in decision-making. For instance, fuzzy logic methods can deal with imprecise or incomplete data, which is common in renewable energy planning, especially when predicting variables such as energy production or costs over time. By using fuzzy sets, decision-makers can incorporate varying degrees of uncertainty into their evaluations. Type-2 fuzzy logic introduces a second layer of fuzziness to the membership functions. This means that rather than assigning a single value to represent an element's degree of membership, type-2 fuzzy sets assign a range of values for each input. In the case of renewable energy, this allows

the model to represent uncertainties in environmental variables, such as cloud cover or wind speed, that are difficult to quantify precisely. For example, instead of assigning a crisp membership value of 0.8 for solar panel efficiency, a type-2 fuzzy set could represent a range of possible values, for example, between 0.75 and 0.85, reflecting the uncertainty in weather conditions over time. In the context of renewable energy, TOPSIS is useful because it allows decision-makers to consider multiple factors, such as cost, efficiency, reliability, environmental impact, and availability of resources, while providing a clear ranking of the alternatives. By comparing the alternatives to the ideal and anti-ideal solutions, it helps identify the most suitable energy source for a specific region or project. While the type-2 fuzzy TOPSIS method offers significant advantages in handling uncertainty and providing more accurate results in complex decision-making scenarios, it also comes with several challenges. These challenges primarily revolve around the complexity of constructing decision matrices, the computational requirements for processing large datasets, and the difficulty in defining appropriate membership functions.

The type-2 fuzzy TOPSIS method for renewable energy estimation is a powerful tool, but there is significant room for future improvements. Integrating this method with machine learning for predictive accuracy, optimization algorithms for better decision-making, and real-time data for dynamic and adaptive responses could substantially enhance its applicability and usefulness. Additionally, improving its ability to handle multiple uncertainties and developing user-friendly decision support systems would make it more accessible and effective for a broader range of decision-makers in the renewable energy sector. These improvements will lead to more accurate, efficient, and responsive decision-making, ultimately helping to accelerate the transition to sustainable and reliable energy systems. The type-2 fuzzy TOPSIS method holds immense potential for improving decision-making in renewable energy planning by accounting for uncertainty and providing a more nuanced evaluation of alternatives. However, there are several exciting avenues for future research and potential improvements that can enhance its applicability, accuracy, and efficiency. Machine learning (ML) models, particularly supervised learning algorithms like regression or classification, can be integrated with the type-2 fuzzy TOPSIS method to predict or estimate fuzzy membership values based on historical or real-time data. For example, predictive models could be developed to estimate the efficiency of solar panels, wind turbines, or other renewable energy sources based on environmental factors such as weather patterns, geographical features, and technological advancements. Machine learning can improve the prediction accuracy of fuzzy membership functions by learning from historical data, reducing the subjective assumptions often involved in manually defining the fuzziness levels. For example, rather than relying on expert-defined fuzzy sets, ML algorithms could automatically learn the optimal fuzzy sets based on real-world observations. Optimization algorithms, such as genetic algorithms (GA), particle swarm optimization (PSO), or simulated annealing, can be integrated with the type-2 fuzzy TOPSIS method to automate and optimize the selection of the best renewable energy sources. These algorithms can help identify the optimal energy portfolio that minimizes costs, maximizes efficiency, or meets other specific goals e.g., environmental sustainability. In addition to environmental factors like weather and geography, renewable energy systems are influenced by factors such as policy shifts e.g., subsidies or tariffs, economic conditions e.g., fuel price fluctuations, and technological innovations e.g., breakthroughs in storage technology. Incorporating these uncertainties into the decision-making process can make the type-2 fuzzy TOPSIS method more robust. Future research could extend the type-2 fuzzy TOPSIS method to include not just environmental uncertainties but also economic and policy-related uncertainties. Fuzzy logic could be used to model uncertain variables such as the potential for future price changes of renewable energy technologies or the impact of upcoming government regulations on energy production. By incorporating various uncertainty scenarios e.g., future energy prices, technological advancements, policy shifts, type-2 fuzzy TOPSIS could be used for scenario-based decision-making, allowing decision-makers to explore different potential futures and make choices that are resilient to a wide range of possible outcomes.

References

[1] Beccali, M., Cellura, M., and Mistretta, M. (2003). Decision-making in energy planning. Application of the Electre method at regional level for the diffusion of renewable energy technology. *Renewable Energy*, 28(13): 2063-2087.

[2] Saravanan V., Ramachandran M., and Malarvizhi M. (2022). Selection of Photovoltaic Devices Using Weighted Sum Method, *Renewable and Nonrenewable Energy*, 1 (2): 67-73.

[3] Heo E., Kim J., & Boo K. (2010). Analysis of the assessment factors for renewable energy dissemination program evaluation using fuzzy AHP. *Renewable Sustain Energy Review*, 14: 2214-2220.

[4] Beccali M., Cellura M., and Ardente D. (1998). Decision making in energy planning: the ELECTRE multi-criterial analysis approach compared to a FUZZY-SETS methodology, *Energy Convers Management*, 39: 1869-1881.

[5] Zheng G., Jing Y., Huang H., Shi G., and Zhang X. (2010). Developing a fuzzy analytic hierarchical process model for building energy conservation assessment, *Renewable Energy*, 35: 78-87.

[6] Karahman C., and Kaya I. (2010). A fuzzy multicriteria methodology for selection among energy alternatives, *Expert Systems Applications*, 37: 6270-6281.

[7] Hwang, C., and Yoon, K. (1981). Multiple Attribute Decision Making: Methods and Applications, *Springer-Verlag*, New York.

[8] Aliyeva K.R. (2024). Research on Group Decision Making in Capital Investment Based on Fuzzy Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), *Journal of Fuzzy Extension and Applications*, 5(3): 313–329.

[9] Şengül U., Eren M., Shiraz S., and Gezder V. (2015). Fuzzy TOPSIS method for ranking renewable energy supply systems in Turkey, *Renewable Energy*, 75: 617-625.

[10] Kaya T., and Kahraman. (2011). Multicriteria decision making in energy planning using a modified Fuzzy TOPSIS methodology. Expert Systems Applications, 38(6): 6577-6585.

[11] Aliyeva, K.R. (2018). Facility location problem by using Fuzzy TOPSIS Method, *B-quadrat Verlags*, 55-59.

[12] Goumas M., and Lygerou V. (2000). An extension of the PROMETHEE method for decision-making in fuzzy environment: ranking of alternative energy exploitation projects. *European Journal Operational Research*, 123: 606–613.

[13] Ramanathan R., and Ganesh L. (1995). Energy alternatives for lighting in households: an evaluation using an integrated goal programming-AHP model, *Energy*, 20(1): 63–72.

[14] Levitin G., and Elmakis D. (1995). Genetic algorithm for open-loop distribution system design, *Electric Power Systems Research*, 32(2): 81–87.

[15] Mamlook R., Akash B., and Mohsen M. (2001) A neuro-fuzzy program approach for evaluating electric power generation systems. Energy, 26:619–632.

[16] Aliyeva, K.R. (2023). Applying Type-2 Fuzzy TOPSIS Method to Selection of Facility Location. *Lecture Notes in Networks and Systems*, 610, LNNS, 662–668.

[17] Zadeh, L. (1996). Fuzzy Sets, Fuzzy Logic, Fuzzy Systems, *World Scientific Press*, p.840.

[18] Zadeh L. (1975). The concept of a linguistic variable and its application to approximate reasoning. Information Science, 8 (3) pp. 199-249. DOI: https://doi.org/10.1007/978-981-16-5348-3_51

[19] Hwang, C., Lai, Y., and Liu, T.Y. (1993). A new approach for multiple objective decision making". Computers and Operational Research. 20 (8): 889–899.

A STATISTICAL ANALYSIS OF VARIANCE FOR ONE-WAY CLASSIFICATION USING TRAPEZOIDAL FUZZY TECHNIQUES AND RANKING METHODS

Vanitha. R[1], Pachamuthu. M[1*], Kirthik Vairamariappan. A[2] ,
Kokkilambigai. S[1] and Sona. S[1]

•

[1] Department of Statistics, Periyar University, India.
[2]Department of Statistics, Government Arts and Science College, India.

vanithapai50@gmail.com, pachamuthu@periyaruniversity.ac.in

Abstract

The Analysis of Variance (ANOVA) is a powerful tool for testing significance, with its primary purpose being to assess the homogeneity of several means. Variation is inherent in nature. The total variation in any set of numerical data arises from multiple causes, which are classified as assignable and chance causes. The variation due to assignable causes can be detected and measured, whereas variation due to chance is beyond human control. An ambiguous number is a normal number whose exact value is uncertain. Fuzziness occurs when the boundaries of information are unclear or imprecise. Sometimes, observational data may not be recorded, particularly during natural disasters. This paper presents a statistical analysis of variance for one-way classification using trapezoidal fuzzy techniques and the ranking method, demonstrated through a numerical example.

Keywords: *ANOVA*, Fuzzy Sets, Trapezoidal Fuzzy Numbers, Ranking Method, Decision Rule.

I. Introduction

The *ANOVA* is a powerful statistical tool for tests of significance. The term *ANOVA* was introduced by Prof. R. A. Fisher in 1920's to deal with problems in the analysis of agrochemical data. Variance is inherent in nature. The total variation in any set of numerical data is due to a number of causes which may be classified as assignable and chance causes. The variation due to assignable cause can be detected and measured whereas the variation due to assignable causes is beyond the control of human hand and cannot be traced separately. In general, *ANOVA* studies mainly the homogeneity of populations by separating the total variance into its various components. That is, this technique is to test the difference among the means of populations by studying the amount of variation within each of the samples relative to the amount of variance between the samples. Samples under employing in *ANOVA* model are assumed to be drawn from 'normal populations of equal variances. The variation of each value around its own grand mean should be independent for each value. The *ANOVA* is a powerful statistical tool for test of significance. The test of significance based on t - distribution is an adequate procedure only for testing the significance of the difference between two sample means. In a situation we have to study more than two samples to consider at a time an alternative producer is needed for testing the

hypothesis that all the samples are drawn from same population. For example, six fertilizers are applied to five plots are given. We may be interested in finding out whether the effect of these fertilizers on the yield is significant different. In other words, test whether the samples have same population. The answer to this problem is provided by technique of *ANOVA*. In sometimes, the observations data is not recorded during a natural disaster. Therefore, fuzzy analysis is almost inevitable. The fuzzy logic was introduced in 1965 by Lofti A. Zadeh [15] in his research paper *Fuzzy Sets*. He is considered as the father of fuzzy logic. The theory of fuzzy logic Zadeh [15] gives an inference method under cognitive uncertainty, computational neural networks and fuzzy logic offer exciting benefits such as learning adaptation, fault tolerance, parallelism and generalization. Generally, our impression of the real world is pervaded by concepts that do not have precisely defined boundaries. For example; many, tall, much, larger than, smaller than, superior to, inferior to, younger than etc. And these adjectives are true only to some degree and they are false to some extent as well. These concepts can be called fuzzy or vague concepts; a human brain works with them, while computers may not do it because of the reason with a string of 0's (false) and 1's (true). Imprecision can be associated with quantitative data and qualitative data. In traditional hypothesis testing, the sample observations are crisp and a statistical test leads to a binary decision is either accept or reject the null hypothesis. However, in real life the data sometimes cannot be recorded or collected precisely in these situations fuzzy logic is most inevitable to the statistical analysis of hypothesis testing. In 1965s Zadeh [15] was introduced fuzzy sets as an extension of the set with classical notations. Fuzzy sets theory allows the estimated membership function in intervals belongs to [0, 1]. In the first case, due to technical problems, the response variable cannot be measured properly. So, in this case, the data cannot be clearly recorded with the exact numbers and the measurement errors are computed linguistically to justify the required tolerance. The second phenomenon is that the response variable is presented in terms of linguistic forms such as a special linguistic or variance reports. As for his products, they are not counted. In both of the above cases, the observed variable of the fuzziness often occurs the data can be represented by the concept of fuzzy sets for analyzing the test (Zadeh [16]. Liou and Wang has discussed the ranked fuzzy numbers with total integral value [2], [3] and [6]. Wang et.al. has presented the method for centroid formulae for a generalized fuzzy number [10]. Iuliana Carmen Barbacioru have discussed the statistical hypotheses testing using membership fuzzy numbers [5]. Salim Rezvani have analyzed the ranking function with trapezoidal fuzzy numbers [11],[12] and [13]. Wang have derived some different approach for ranking trapezoidal fuzzy numbers [7],[8] and [10]. Thorani et al. has discussed the ranking function of a trapezoidal fuzzy number with some modifications [9]. S. Abbas bandy and B. Asady has derived the nearest trapezoidal fuzzy number to a fuzzy quantity [1]. S. Parthiban and B. Gajiwardhan has proposed a one factor ANOVA model using trapezoidal fuzzy numbers by using α-cut interval method [4] and [14]. S. Chandrasekaran and A. Tamilmani has proposed the arithmetic operations of fuzzy numbers using alpha cut interval method [17]. H.C. Wu [18] has proposed the technique, ambiguous data as well, given the vague assumptions of the tests were imprecise data, along with two hypothesis tests replacing *ANOVA* models crisp data, models and results, getting after using the results obtained in terms of the provisions of the proposed decision of the population receive the decision. M. Mashinchi et.al. [19] has derived the *ANOVA* based on fuzzy observations and extended into a one-way *ANOVA*, which would be fuzzy numbers rather than real numbers Wang arrived some different approach for ranking TZFNs [10]. In this paper, propose a statistical analysis of variance for one-way classification with trapezoidal fuzzy techniques using ranking method.

II. Methods

I. Fuzzy Set

The fuzzy set theory was developed to handle the vagueness inherent in nature. Zadeh (1965) [15] extended the notion of binary membership to accommodate various degree of membership on the

real continuous interval, where the endpoints of 0 and 1 refers to no membership and full membership respectively just as the indicator function does for crisp sets, but where the infinite number of values in between the endpoints 0 and 1 can represent various degrees of membership for an element in some set in the universe of discourse. The sets in the universe of discourse that can accommodate degrees of membership were termed by Zadeh (1965) [15] as fuzzy sets.

II. α-Cut

A fuzzy set A in X and any actual number $\alpha \in [0,1]$, then the α-cut α-level or cut worthy set of $\tilde{A}$ denoted by A_α is the crisp set $A_\alpha = \{x \in X: \mu_A(x) \geq \alpha\}$. The strong α-cut $\tilde{A}_{\alpha+}$ is the crisp set $A_{\alpha_+} = \{x \in X: \mu_A(x) > \alpha\}$. For example, let A be a fuzzy set whose membership function is

$$\mu_A(x) = \begin{cases} \dfrac{x-a}{b-a}; & a \leq x \leq b \\ \dfrac{c-x}{c-b}; & b \leq x \leq c \end{cases}$$

To find the α-cut of A, first set $\alpha \in [0,1]$ to both left and right reference functions of A. That is $\alpha = \frac{x-a}{b-a}$ and $\alpha = \frac{c-x}{c-b}$. Expressing x in terms of α, have $x = a + (b-a)\alpha$ and $x = c - (c-b)\alpha$, which gives the α-cut of A is $A_\alpha = [a + (b-a)\alpha;\ c - (c-b)\alpha]$. Similarly, the *TZFN*s to both left and right reference functions of A.

III. Crisp Set

The support of a fuzzy set A defined on X is a crisp set defined as $\text{Supp}(A) = \{x \in X: \mu_A(x) > 0\}$

IV. h-level of a Fuzzy Set

The h-level of a fuzzy set $\tilde{A}$, denoted by $h(\tilde{A})$ is the supremum (maximum) of the membership grades obtained by any element in the set $h(\tilde{A}) = \sup_{\text{xix}} \mu_{\tilde{A}}(x)$.

V. Degrees of Membership in Fuzzy Logic

Let us consider the fuzzy set $A, A = \{(x, \mu_A(x)) \mid x \in X\}$ where $\mu_A(x)$ is called the membership function for the fuzzy set A. X is referred to as the universe of discourse. The membership function associates each element $x \in X$ with a value in the interval [0,1] . In fuzzy sets, each element is mapped to [0,1] by membership function. That is, $\mu_A: X \in [0,1]$, where [0,1] means real numbers between 0 and 1 (including 0,1). Consequently, fuzzy set is with vague boundary set comparing with crisp set. The fuzzy set A can be alternatively denoted as follows: If X is discrete then $A = \sum \mu_A(x_i)/x_i$ and if X is continnous then $A = \int \mu_A(x)/x$. Here, $\mu_A(x)$ is the membership function. Value of this function is between 0 and 1. This value represents the degree of membership (membership value) of element x inset A. The members of a fuzzy set are members to some degree, known as a membership grade or degree of membership. The membership grade is the degree of belonging to the fuzzy set. The larger the number (in [0,1]) the more the degree of belonging. (This is not a probability). The translation from x to $\mu_A(x)$ is known as fuzzification. In the fuzzy theory, fuzzy set A of universe X is defined by a function $\mu_A(x)$ called the membership function of set A. An already discussed this point. $\mu_A(x): X \to [0,1]$, where $\mu_A(x) = 1$ it x is totally in A; $\mu_A(x) = 0$ it x is not in A; $0 < \mu_A(x) < 1$ it x is partly in A. This set allows a continutum of possible choices. For any element x of universe X, membership function $\mu_A(x)$ equals the degree to which x is an element of set A. This degree, a value between 0 and 1, represents the degree of membership, also

called membership value, of element x inset A.

VI. Generalized Triangular Fuzzy Numbers (*GTFN*s)

A Generalized Triangular Fuzzy Numbers (*GTFN*s) $\tilde{A} = (a, b, c; w)$ is described as any fuzzy subset of the real line $\mathbb{R}$, whose membership function $\mu_A(x)$ satisfies the following conditions:

- $\mu_A(x)$ is a continuous mapping from $\mathbb{R}$ to the closed interval $[0, w]; 0 \leq w \leq 1$
- $\mu_A(x) = 0$, for all $x \in (-\infty, a]$
- $\mu_L(x) = L_A(x)$ is strictly increasing on $[a, b]$
- $\mu_A(x) = w$, for all $[b], w$ is a constant and $0 \leq w \leq 1$
- $\mu_R(x) = R_{\hat{A}}(x)$ is strictly decreasing on $[b, c]$
- $\mu_A(x) = 0$, for all $x \in [c, \infty]; a, b, c$ are real numbers such that $a \leq b \leq c$.

A *GTFN*s is denoted by $\tilde{A} = (a, b, c; w)$ and is defined with the membership function,

$$\mu_{\hat{A}}(x) = \begin{cases} 0 & ; \quad x \leq a \\ \dfrac{w(x-a)}{(b-a)} & ; \quad a \leq x \leq b \\ \dfrac{w(c-x)}{(c-b)} & ; \quad b \leq x \leq c \\ 0 & ; \quad c < x \end{cases}$$

VII. Triangular Fuzzy Numbers (*TrFZ*)

If $w = 1$ in *GTFN*s have Normalized Triangular Fuzzy Numbers (*NTFN*s) $\tilde{A} = (a, b, c, 1)$, simply denoted by $\tilde{A} = (a, b, c)$ and is defined with the membership function,

$$\mu_{\tilde{A}}(x) = \begin{cases} 0 & ; x \leq a \\ \dfrac{(x-a)}{(b-a)} & ; a \leq x \leq b \\ \dfrac{(c-x)}{(c-b)} & ; b \leq x \leq c \\ 0 & ; c < x \end{cases}$$

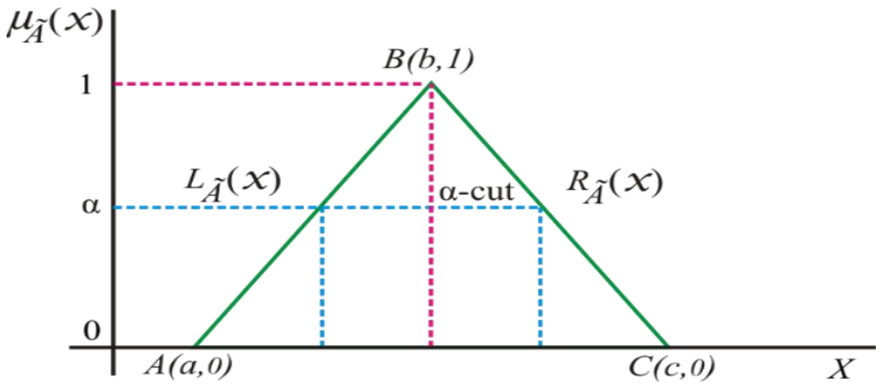


Figure 1: *Normalized Triangular Membership Function*

VIII. Normalized Trapezoidal Fuzzy Numbers (*NTZFN*s)

A Normalized Trapezoidal Fuzzy Numbers (*NTZFN*s) $\tilde{A} = (a, b, c, d; w)$ is described as any fuzzy

subset of the real line $\mathbb{R}$, whose membership function $\mu_A(x)$ satisfies the following conditions:

- $\mu_A(x)$ is a continuous mapping from $\mathbb{R}$ to the closed interval $[0,1]$; $0 \le w \le 1$,
- $\mu_A(x) = 0$, for all $x \in (-\infty, a]$,
- $\mu_L(x) = L_{\hat{A}}(x)$ is strictly increasing on $[a, b]$,
- $\mu_A(x) = 1$, for all $b \le x \le c$,
- $\mu_R(x) = R_{\hat{A}}(x)$ is strictly decreasing on $[c, d]$,
- $\mu_{\hat{A}}(x) = 0$, for all $x \in [d, \infty]$; a, b, c, d are real numbers such that $a < b \le c < d$.

It is known that for a *NTZFN*s. $\tilde{A} = (a, b, c, d; 1)$, there exist four numbers $a, b, c, d \in \mathbb{R}$ and two functions $L_{\hat{A}}(x), R_{\hat{A}}(x): \mathbb{R} \to [0,1]$, where $L_{\hat{A}}(x)$ and $R_{\hat{A}}(x)$ are non-decreasing and nonincreasing functions respectively. The functions $L_{\hat{A}}(x)$ and $R_{\hat{A}}(x)$ are also called the left and right side of the fuzzy number $\tilde{A}$ respectively. The membership function is defined as follows:

$$\mu_A(x) = \begin{cases} 0; & \text{otherwise} \\ \left(\frac{x-a}{b-a}\right); & a \le x \le b \\ 1; & b \le x \le c \\ \left(\frac{d-x}{d-c}\right); & c \le x \le d \\ 0; & \text{otherwise} \end{cases}$$

Similarly, it is known that for a *NTZFN*s. $\tilde{A} = (a, b, c, d; 1)$, there exist four numbers $a, b, c, d \in \mathbb{R}$ and two functions $L_{\hat{A}}(x), R_{\hat{A}}(x): \mathbb{R} \to [0,1]$, where $L_{\hat{A}}(x)$ and $R_{\hat{A}}(x)$ are nondecreasing and non-increasing functions respectively and its membership function is defined as follows:

$$\mu_{\tilde{A}}(x) = L_{\tilde{A}}(x) = (x-a)/(b-a) \text{ for } a \le x \le b \text{ ; } 1 \text{ for } b \le x \le c$$

$\mu_{\hat{A}}(x) = R_{\hat{A}}(x) = (x-c)/(d-c)$ for $c \le x \le d$ and 0 otherwise }. The functions $L_{\hat{A}}(x)$ and $R_{\tilde{A}}(x)$ are also called the left and right side of the fuzzy number $\tilde{A}$ respectively. In this proposed work, the assume that $\int_{-\infty}^{\infty} \tilde{A}(x)dx < +\infty$ and it is known that the α - cut of a fuzzy number is $\tilde{A}_\alpha = \{x \in \mathbb{R} / \mu_{\tilde{A}}(x) \ge \alpha\}$, for $\alpha \in (0,1]$ and $\tilde{A}_0 = cl\left(\bigcup_{\alpha \in (0,1]} \tilde{A}_\alpha\right)$, according to the definition of a fuzzy number, it is seen at once that every α - cut of a fuzzy number is a closed interval. Hence, for a fuzzy number $\tilde{A}$, we have $\tilde{A}(\alpha) = \left[\tilde{A}_L(\alpha), \tilde{A}_U(\alpha)\right]$ where $\tilde{A}_L(\alpha) = \inf\{x \in \mathbb{R}: \mu_{\tilde{A}}(x) \ge \alpha\}$ and $\tilde{A}_U(\alpha) = \sup\{x \in \mathbb{R}: \mu_{\tilde{A}}(x) \ge \alpha\}$. The left and right sides of the fuzzy number $\tilde{A}$ are strictly monotone, obviously, $\tilde{A}_L(\alpha)$ and $\tilde{A}_U(\alpha)$ are inverse functions of $L_{\tilde{A}}(x)$ and $R_{\tilde{A}}(x)$ respectively. The normal TZFNs $\tilde{A} = (a, b, c, d)$ and the corresponding α-cut is defined by

$$\tilde{A}_\alpha = [a + \alpha(b-a), d - \alpha(d-c)]; \alpha \in [0,1]$$

(i) If w = 1 in $GTZFN$ s, have a $NTZFN$ A$\tilde{A} = (a, b, c, d; 1)$ and can be simply denoted by $\tilde{A} = (a, b, c, d)$ and defined with the membership function,

$$\mu_{\hat{A}}(x) = \begin{cases} 0 & ; x \le a \\ L_A(x) = \dfrac{(x-a)}{(b-a)} & ; a \le x \le b \\ 1 & ; b \le x \le c \\ R_A(x) = \dfrac{(x-d)}{(d-c)} & ; c \le x \le d \\ 0 & ; d < x \end{cases}$$

where $a, b, c, d \in \mathbb{R}$ with $a < b < c < d$.

$$\mu_A(x) = \begin{cases} 0 & ; x \le a \\ L_A(x) = \dfrac{(x-a)}{(b-a)} & ; a \le x \le b \\ 1 & ; b \le x \le c \\ R_A(x) = \dfrac{(x-d)}{(d-c)} & ; c \le x \le d \\ 0 & ; d < x \end{cases}$$

where $a, b, c, d \in \mathbb{R}$ with $a < b < c < d$.

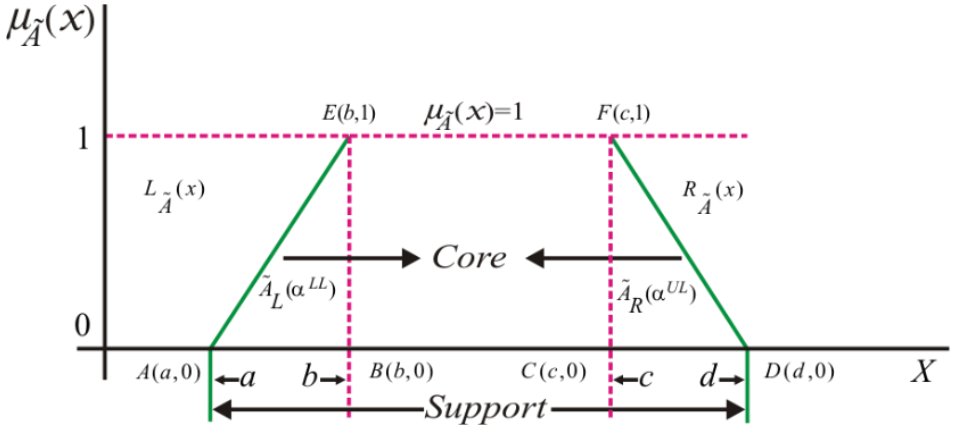


Figure 2: *Normalized Trapezoidal Membership Function*

(ii) If $w = 1$ in *GTZFNs*, have a *NTZFNs* $\tilde{A} = (a, b, c, d; 1)$ and can be simply denoted by $\tilde{A} = (a, b, c, d)$ dind uefined with the membership function is

$$\mu_A(x) = \begin{cases} 0 & ; x \le a \\ L_A(x) = \dfrac{(x-b)}{(b-a)} & ; a \le x \le b \\ 1 & ; b \le x \le c \\ R_A(x) = \dfrac{(x-c)}{(d-c)} & ; c \le x \le d \\ 0 & ; d < x \end{cases}$$

where $a, b, c, d \in \mathbb{R}$ with $a < b < c < d$

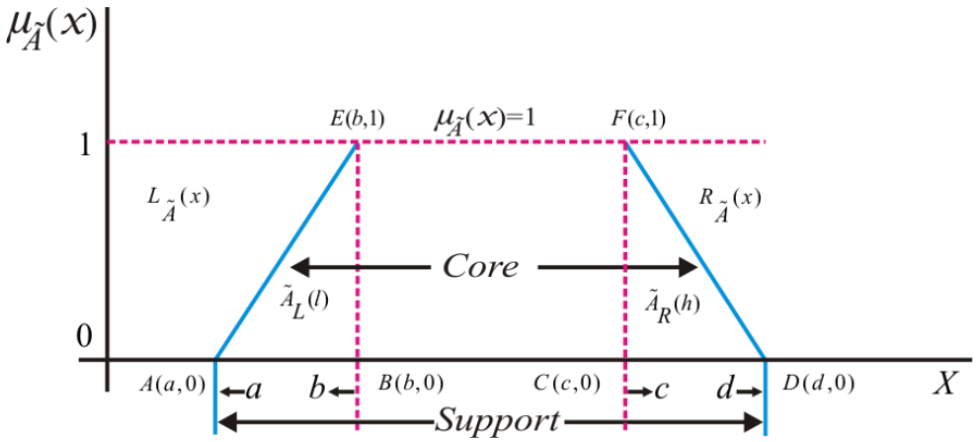


Figure 3: *Normalized Trapezoidal Membership Function*

The normal *TZFN*s $\tilde{A} = (a, b, c, d)$ and the corresponding without h level is defined by

$$\tilde{A} = [b - (b - a)l,\ c + (d - c)h];\ 0 \le h, l \le 1$$

IX. Trapezoidal Fuzzy Numbers (TrZFNs)

The membership function of a TZFNs $\widehat{A}$ is defined by

$$\mu_{\tilde{A}}(x) = \begin{cases} \dfrac{-(a-x)}{(b-a)}; a \le x \le b \\ 1; b \le x \le c \\ \dfrac{-(d-x)}{(c-d)}; c \le x \le d \end{cases}$$

Where, $a \le b \le c \le d$. A $TrFN$ becomes triangular fuzzy number if it satisfies b = c. In terms of $\alpha -$ cut interval, Tr FN is defined as follows:

$$\tilde{A} = [a + (b - a)\alpha, d - (d - c)\alpha]; 0 \le \alpha \le 1, \text{where, } a \le b \le c \le d$$

X. Wang's Centroid Point and Ranking Method

Wang et al. [14] has found that the centroid formulae proposed and have led to some misapplications such as by Liou, T. S [6]. They presented the correct method for centroid formulae for a generalized fuzzy number $\tilde{A} = (a, b, c, d; w)$ as;

$$(\bar{x}_0, \bar{y}_0) = \left[\left(\frac{1}{3}(a + b + c + d) - \left(\frac{dc - ab}{(d+c)-(a+b)}\right)\right), \left(\frac{w}{3}\right)\left(1 + \left(\frac{c-b}{(d+c)-(a+b)}\right)\right)\right] \quad (1)$$

and the ranking function associated with $\tilde{A}$ is

$$R(\tilde{A}) = \sqrt{\bar{x}_0^2 + \bar{y}_0^2}$$

For a normalized *TZFN*s, we put $w = 1$ in equations (1) so we have,

$$(\bar{x}_0, \bar{y}_0) = \left[\left(\frac{1}{3}(a + b + c + d) - \left(\frac{dc - ab}{(d+c)-(a+b)}\right)\right), \left(\frac{1}{3}\right)\left(1 + \left(\frac{c-b}{(d+c)-(a+b)}\right)\right)\right]$$

and the ranking function associated with $\tilde{A}$ is

$$R(\tilde{A}) = \sqrt{\bar{x}_0^2 + \bar{y}_0^2}$$

Let $\tilde{A}_i$and $\tilde{A}_j$be two fuzzy numbers, $(i)\ R(\tilde{A}_i) > R(\tilde{A}_j)$ then $\tilde{A}_i > \tilde{A}_j$ $(ii)\ R(\tilde{A}_i) < R(\tilde{A}_j)$ then $\tilde{A}_i < \tilde{A}_j$and $(iii) R(\tilde{A}_i) = R(\tilde{A}_j)$ then$\tilde{A}_i = \tilde{A}_j$.

XI. A Statistical Analysis of Variance for One-Way Classification with *TZFN*s using Ranking Method

The general linear model is

$$y_{ij} = \mu + \alpha_i + e_{ij}; \quad i = 1,2,\dots,k; \quad j = 1,2,\dots,n_i$$

The above general linear model of crisp *ANOVA* is classified into two interval fuzzy *ANOVA* models is Lower Limit Model (*LLM*) and Upper Limit Model (*ULM*). The *TZFN*s *LLM* is: $\left(\tilde{y}_{ij}\right)_l^{LL} = (\tilde{\mu})_l^{LL} + (\tilde{a}_i)_l^{LL} + \left(\tilde{e}_{ij}\right)_l^{LL}$, where $\left(\tilde{y}_{ij}\right)_l^{LL}$ is the j^{th} observations of the i^{th} treatment; $(\tilde{\mu})_l^{LL}$ is the general mean effect which is fixed; $(\tilde{a}_i)_l^{LL}$ is the fixed effect due to i^{th} treatment; and $\left(\tilde{e}_{ij}\right)_l^{LL}$ is the random error effect.

Null hypothesis: $\tilde{H}_0^{LL}: \tilde{\mu}_1^{LL} = \tilde{\mu}_2^{LL} = \cdots = \tilde{\mu}_r^{LL}$ against $\tilde{H}_1^{LL}: \tilde{\mu}_1^{LL} \neq \tilde{\mu}_2^{LL} \neq \cdots \neq \tilde{\mu}_r^{LL}$.

The *LLM* and *ULM* from the null hypothesis of acceptance levels of the direction in the pessimistic and optimistic of *TZFN*s. Using trapezoidal fuzzy *LLM* formula Is $b_{ij} - (b_{ij} - a_{ij})l$, where $0 \le i \le k, 0 \le j \le n_i$ and *ULM* formula is $c_{ij} + \left(d_{ij} - c_{ij}\right)h$, where $0 \le i \le k, 0 \le j \le n_i$.

The layout of *LLM* using *TZFN*s without -h level concept can be assigned as follows:

$$\begin{matrix} b_{11} - (b_{11} - a_{11})l, b_{12} - (b_{12} - a_{12})l, \dots, b_{1n_i} - \left(b_{1n_i} - a_{1n_i}\right)l \\ b_{21} - (b_{21} - a_{21})l, b_{22} - (b_{22} - a_{22})l, \dots, b_{2n_i} - \left(b_{2n_i} - a_{2n_i}\right)l \\ \vdots \quad\quad \vdots \quad\quad \dots \quad\quad \vdots \\ b_{1i} - (b_{1i} - a_{1i})l, b_{2i} - (b_{2i} - a_{2i})l, \dots, b_{in_i} - \left(b_{in_i} - a_{in_i}\right)l \\ \vdots \quad\quad \vdots \quad\quad \dots \quad\quad \vdots \\ b_{k1} - (b_{k1} - a_{k1})l, b_{k2} - (b_{t2} - a_{k2})l, \dots, b_{kn_i} - \left(b_{tn_i} - a_{kn_i}\right)l \end{matrix}$$

Then the required sum of squares for LLM is given below:

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} \left[\left(\tilde{y}_{ij}^2\right)_l^{LL}\right] = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left[b_{ij} - \left(b_{ij} - a_{ij}\right)l\right]^2 \text{ where } \left[\left(\tilde{y}_{ij}^2\right)_l^{IL}\right] = \left[b_{ij} - \left(b_{ij} - a_{ij}\right)l\right]^2$$

$$(cf)_l^{LL} = \frac{\left[(\tilde{y}^2)_l^{LL}\right]}{N}$$

$$(SS_T)_l^{LL} = \sum_{i=1}^{k} \sum_{j=1}^{n} \left[\left(\tilde{y}_{ij}^2\right)_l^{LL}\right] - \frac{\left[(\tilde{y}^2)_l^{LL}\right]}{N} \sim (N-1)\ df$$

$$(SS_{\tilde{T}})_l^{LL} = \sum_{i=1}^{k} \frac{\left[(\tilde{y}_i^2)_l^{LL}\right]}{n_i} - \frac{\left[(\tilde{y}^2)_l^{LL}\right]}{N} \sim (k-1)\ df$$

$$(SS_E)_l^{IL} = (SS_T)_l^{IL} - (SS_{T\tau})_l^{IL} \sim (N-k)df$$

$$\left(MS_{T_r}\right)_l^{IL} = \frac{(SS_{T\tau})_l^{LL}}{(k-1)}$$

$$(MS_E)_l^{IL} = \frac{(SS_E)_l^{IL}}{(N-k)}$$

$$\left(\tilde{F}_{TT}\right)_l^{IL} = \frac{(MS_{T\tau})_l^{IL}}{(MS_E)_l^{IL}}$$

The fuzzy *ANOVA* one-way table for *LLM* shown below;

Table 1: *ANOVA Table for ULM of Trapezoidal Fuzzy ANOVA One-Way*

SV	df	SS	MSS	$\tilde{\mathrm{F}}^{\mathrm{LL}} - \mathrm{Ratio}$
Treatments	$(k-1)$	$(SS_{T\tau})_l^{IL}$	$(MS_{T\tau})_l^{IL}$	$\frac{(MS_{T\tau})_l^{IL}}{(MS_E)_l^{LL}}$
Error	$(N-k)$	$(SS_E)_l^{LL}$	$(MS_E)_l^{IL}$	-
Total	$(N-1)$	$(SS_T)_l^{LL}$	-	-

Similarly, the trapezoidal fuzzy ULM is linear model is: $\left(\tilde{y}_{ij}\right)_h^{UI} = (\tilde{\mu})_h^{UL} + (\tilde{a}_i)_h^{UL} + \left(\tilde{e}_{ij}\right)_h^{UL}$ in which $\left(\tilde{y}_{ij}\right)_h^{UL}$ is the j^{th} observations of the i^{th} treatment; $(\tilde{\mu})_h^{UL}$ is the general mean effect which is fixed; $(\tilde{a}_i)_l^{LL}$ and $(\tilde{a}_i)_h^{UL}$ is the fixed effect due to i^{th} treatment; and $\left(\tilde{e}_{ij}\right)_h^{UL}$ is the random error effect; $i = 1,2,\dots,k$ and $j = 1,2,\dots,n_i$.

Null hypothesis: $\tilde{H}_0^{UI}: \tilde{\mu}_1^{UL} = \tilde{\mu}_2^{UL} = \cdots = \tilde{\mu}_r^{UI}$ against $\tilde{H}_1^{UL}: \tilde{\mu}_1^{UI} \neq \tilde{\mu}_2^{UI} \neq \cdots \neq \tilde{\mu}_r^{UL}$.

Similarly, the layout of *ULM* using *TZFN*s without *h*- level concept can be assigned as follows:

$$\begin{pmatrix} c_{11} + (d_{11} - c_{11})h, c_{12} + (d_{12} - c_{12})h, \dots, c_{1n_i} + \left(d_{1n_i} - c_{1n_i}\right)h \\ c_{21} + (d_{21} - c_{21})h, c_{22} + (d_{22} - c_{22})h, \dots, c_{2n_i} + \left(d_{2n_i} - c_{2n_i}\right)h \\ \vdots \quad \vdots \quad \dots \quad \vdots \\ c_{i1} + (d_{i1} - c_{i1})h, c_{i2} + (d_{i2} - c_{i2})h, \dots, c_{in_i} + \left(d_{in_i} - c_{in_i}\right)h \\ \vdots \quad \vdots \quad \dots \quad \vdots \\ c_{k1} + (d_{k1} - c_{k1})h, c_{k2} + (d_{k2} - c_{k2})h, \dots, c_{kn_i} + \left(d_{tn_i} - c_{kn_i}\right)h \end{pmatrix}$$

Then the required sum of squares for *ULM* is given bel

$$\sum_{i=1}^{k}\sum_{j=1}^{n}\left[\left(\tilde{y}_{ij}^2\right)_h^{UL}\right] = \sum_{i=1}^{k}\sum_{j=1}^{n}\left[c_{ij} + \left(d_{ij} - c_{ij}\right)h\right]^2 \left[\left(\tilde{y}_{ij}^2\right)_h^{UL}\right] = \left[c_{ij} + \left(d_{ij} - c_{ij}\right)h\right]^2$$

$$(cf)_h^{UL} = \frac{\left[\left(\tilde{y}^2\right)_h^{UL}\right]}{N} = \frac{\left[\left(G^2\right)_h^{UL}\right]}{N}$$

$$(SS_T)_h^{UL} = \sum_{i=1}^{k}\sum_{j=1}^{n}\left[\left(\tilde{y}_{ij}^2\right)_h^{UL}\right] - \frac{\left[\left(\tilde{y}^2\right)_h^{UL}\right]}{N} \sim (N-1)df$$

$$(SS_{T\tau})_h^{UL} = \sum_{i=1}^{k}\frac{\left[\left(\tilde{y}_i^2\right)_h^{UL}\right]}{n_i} - \frac{\left[\left(\tilde{y}_{\cdot}^2\right)_h^{UL}\right]}{N} \sim (k-1)df$$

$$(SS_E)_h^{UL} = (SS_T)_h^{UL} - (SS_{T\tau})_h^{UL} \sim (N-k)df$$

$$(MS_{T\tau})_h^{UL} = \frac{(SS_{T\tau})_h^{UL}}{(k\text{-}1)}$$

$$\left(\tilde{F}_{TF}\right)_h^{UL} = \frac{(MS_{T\tau})_h^{UI}}{(MS_E)_h^{UI}}$$

The fuzzy *ANOVA* one-way table for *ULM* shown below;

Table 2*: ANOVA Table for ULM of Trapezoidal Fuzzy ANOVA One-Way*

SV	df	SS	MSS	$\tilde{F}^{UL} - \text{Ratio}$
Treatments	$(k-1)$	$(SS_{TT})_h^{UL}$	$(MS_{T\tau})_h^{UL}$	$\frac{(MS_{T\tau})_h^{UL}}{(MS_E)_h^{UL}}$
Error	$(N-k)$	$(SS_E)_h^{UI}$	$(MS_E)_h^{UL}$	-
Total	$(N-1)$	$(SS_T)_h^{UL}$	-	-

Decision Rules for LLM and ULM

- If $(MS_{Tr})_l^{LL} > (MS_E)_l^{LL}$ or $\left(\tilde{F}_{T_r}\right)_l^{LL} = \frac{(MS_{T_r})_l^{LL}}{(MS_k)_i^{LL}} < F_T$, where $\left(\tilde{F}_{Tr}\right)_l^{IL}$ is the calculated value and F_T is the tabulated value of F for $(k-1),(N-k)$ df at α% loss, then the null hypothesis $\tilde{H}_0^{LL}$ is accepted, otherwise the alternative hypothesis $\tilde{H}_1^{LL}$ is accepted for all $0 \le l^{II} \le 1$.
- If $(MS_{TV})_h^{UL} > (MS_E)_h^{UL}$ or $\left(\tilde{F}_{T_r}\right)_h^{tL} = \frac{(MS_T)_k^{UL}}{(MS_E)_h^{ULL}} < F_T$, where $\left(\tilde{F}_{TT}\right)_h^{UL}$ is the calculated value and F_T is the tabulated value of F for $(k-1),(N-k)$ dif at α% lo.s, then the null hypothesis $\tilde{H}_0^{UI}$ is accepted, otherwise the alternative hypothesis $\tilde{H}_1^{UI}$ is accepted for all $0 \le h^{ul} \le 1$.

Note: If when l=0 or h=0, that is, centre level, the null hypothesis of the *ANOVA* model is accepted at 5% l.o.s. This study is about the acceptance levels of null hypothesis and its directions of pessimistic and optimistic. In this method, the notions of pessimistic degree and optimistic degree are not used in without h-level concept and the decision obtained.

III. Results

I. Application

The data was collected for yields of paddy on Salem District of Tamil Nadu in India. In the yields of four different types (Y_1, Y_2, Y_3, Y_4) in kilograms with three replications of paddy (*ADT, IR, ASD*) each by using *ANOVA* method. Stipulation we cannot record the precise number of yields in kilograms in a sample due to approximately an unexpected situation happen, there the proposed fuzzy data for number yields in kilograms. The trapezoidal fuzzy data is performed below;

Table 3: *Observations*

Yields in kilograms (i)	Varieties of paddy (j)		
	ADT	IR	ASD
Y_1	8, 10, 11, 14	11, 13, 15, 18	-
Y_2	7, 9, 12, 15	8, 9, 11, 14	8, 10, 12, 14
Y_3	12, 14, 16, 18	12, 13, 15, 19	14, 16, 19, 23
Y_4	12, 13, 16, 19	19, 20, 22, 24	-

To test whether there is any significant difference in the varieties of paddy performance of the yields in kilograms.

Let's take the initial set value (8,10,11,14). using LLM formula is $b-(b-a)la=8, b=10$, $10-(10-8)l, (10-2l)$. Similarly, using the same procedure get the remaining values as given below:

Table 4: *LLM using TZFNs without h Level*

Yields in kilograms (i)	Varieties of paddy (j)			
	ADT	IR	ASD	
Y_1	$10-2l$	$13-2l$	-	$23-4l$
Y_2	$9-2l$	$9-l$	$10-2l$	$28-5l$
Y_3	$14-2l$	$13-l$	$16-2l$	$43-5l$
Y_4	$13-l$	$20-l$	-	$33-2l$
				$127-16l$

$\tilde{H}_0^{LL}$: The varieties of paddy do not differ significantly with respect to the yields in kilograms The sum of squares is given below;

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\tilde{y}_{ij}^2\right)_l^{LL}=\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left[b_{ij}-\left(b_{ij}-a_{ij}\right)l\right]^2=(10-2l)^2+\cdots+(20-l)^2$$

$$(cf)_k^{UL}=\frac{[(\tilde{y}_{-}^2)_h^{LL}]}{N}=\frac{(127-16l)^2}{10}=25.6l^2-406.4l+1612.9$$

$$(SS_r)_j^{LL}=\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left[\left(\tilde{y}_{ij}^2\right)_l^{LL}\right]-\frac{[(\tilde{y}_{-}^2)_l^{LL}]}{N}=(28l^2-398l+1721)-(25.6l^2-406.4l+1612.9)$$

$$(SS_T)_l^{LL}=2.4l^2+8.4l+108.1$$

$$(SS_{TT})_l^{IL}=\sum_{i=1}^{k}\frac{[(\tilde{y}_i^2)_l^{LU}]}{n_i}-\frac{[(\tilde{y}^2)_l^{LL}]}{N}=\left\{\left(\frac{(23-4l)^2}{2}+\frac{(28-5l)^2}{3}+\frac{(43-5l)^2}{3}+\frac{(33-2l)^2}{2}\right)\right\}$$
$$-(25.6l^2-406.4l+1612.9)$$

$$(SS_E)_l^{LL}=1.33l^2-3.33l+34.33$$

$$(MS_{TT})_l^{IL}=\frac{\left(SS_{T_f}\right)_l^{IL}}{(4-1)}=\frac{1.37l^2+11.73l+73.77}{3}=0.36l^2+3.91l+24.59$$

$$(MS_E)_l^{IL}=\frac{(SS_E)_l^{IL}}{(10-4)}=\frac{1.33l^2-3.33l+34.33}{6}=0.22l^2-0.56l+5.72$$

$$\left(\tilde{F}_{Tr}\right)_l^{LL}=\frac{(MS_{Tr})_l^{IL}}{(MS_E)_l^{LL}}=\frac{0.36l^2+3.91l+24.59}{0.22l^2-0.56l+5.72}$$

$$\text{Put } l = 0, \left(\tilde{F}_{TT}\right)_l^{LL} = \frac{0.36l^2 + 3.91l + 24.59}{0.22l^2 - 0.56l + 5.72} = 4.3, \ldots, l = 1, \left(\tilde{F}_R\right)_l^{LL} = 5.36.$$

The sum of squares is summarized given below;

Table 5: *ANOVA Table for LLM of Fuzzy ANOVA One Way*

SV	df	SS	MSS	$\tilde{F}^{LL} -$ Ratio
Treatments	3	$1.07l^2 + 11.73l + 73.77$	$0.36l^2 + 3.91l + 24.59$	$\frac{0.36l^2 + 3.91l + 24.59}{0.22l^2 - 0.56l + 5.72}$
Error	6	$1.33l^2 - 3.33l + 34.33$	$0.22l^2 - 0.56l + 5.72$	-
Total	9	$2.4l^2 + 8.4l + 108.1$	-	-

Decision Rules for LLM : Since, the $(\tilde{F})_i^{LL} > F_T$, where $F_T = 4.76$ is the table value at 5% at 0.5 with (3,6) dif Then, the null hypothesis $\tilde{H}_0^{LL}$ of the LLM is accepted for all $0 \leq l^{LL} \leq 0.42$; otherwise, the null hypothesis $\tilde{H}_0^{UL}$ is rejected for all $043 \leq l^{IL} \leq 1$. Therefore, the varieties of paddy do not differ significantly with respect to the yields in kilograms. Let's take the initial set value (8,10,11, 14). using ULM formula is $c + (d - c)h\, c = 11, d = 14, 11 + (14 - 11)h, 11 + 3h$. Similarly, using the same procedure get the remaining values as given below

Table 6: *ULM using TZFNs without h Level*

Yields in kilograms (i)	Varieties of paddy (j)			
	ADT	IR	Total	
Y1	$11 + 3h$	$15 + 3h$	-	$26 + 6h$
Y2	$12 + 3h$	$11 + 3h$	$12 + 2h$	$35 + 8h$
Y3	$16 + 2h$	$15 + 4h$	$19 + 4h$	$50 + 10h$
Y4	$16 + 3h$	$22 + 2h$	-	$38 + 5h$
				$149 + 29h$

$\tilde{H}_0^{UL}$: The varieties of paddy do not differ significantly with respect to the yields in kilograms The sum of squares is given below;

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\tilde{y}_{ij}^2\right)_h^{uL} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left[c_{ij} + \left(d_{jy} - c_{ij}\right)h\right]^2 = (11 + 3h)^2 + \cdots + (22 + 2h)^2$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\tilde{y}_{ij}^2\right)_h^{iL} = 89h^2 + 862h + 2337$$

$$(S_T)_h^{UL} = \sum_{l=1}^{k}\sum_{j=1}^{n_i}\left[\left(\tilde{y}_{ij}^2\right)_h^{eL}\right] - \frac{[(\tilde{y}^2)_h^{ULL}]}{N}$$

$$(SS_T)_h^{UL} = 4.9h^2 - 2.2h + 116.9$$

$$(SS_{Tr})_h^{ULL} = 1.07h^2 + 1.8h + 81.57$$

$$(SS_E)_h^{UL} = (4.9h^2 - 2.2h + 116.9) - (1.07h^2 + 1.8h + 81.57)$$

$$(SS_E)_h^{UL} = 3.83h^2 - 4h + 35.33$$

$$(MS_{TT})_h^{UL} = \frac{(SS_{T\tau})_h^{UL}}{(4-1)} = \frac{1.07h^2 + 1.8h + 81.57}{3} = 0.36h^2 + 0.6h + 27.19$$

$$(MS_E)_h^{UL} = \frac{(SS_E)_h^{UI}}{(10-4)} = \frac{3.83h^2 - 4h + 35.33}{6} = 0.64h^2 - 0.67h + 5.89$$

$$\left(\tilde{F}_{V_r}\right)_h^{UL} = \frac{\left(MS_{r_r}\right.}{(MS_t)_h^{UL}} = \frac{0.36h^2 + 0.6h + 27.19}{0.64h^2 - 0.67h + 5.89}$$

Put $h = 0, \left(\tilde{F}_{T\tau}\right)_h^{UL} = \frac{0.36(0)^2+0.6(0)+27.19}{0.64(0)^2-0.67(0)+5.89} = 4.62 \ldots h = 1\left(\tilde{F}_R\right)_l^{LL} = 4.8.$

The sum of squares is summarized given below;

Table 7: *ANOVA Table for ULM of Fuzzy ANOVA One Way*

SV	df	SS	MSS	$\tilde{\mathrm{F}}^{\mathrm{UL}} -$ Ratio
Treatments	3	$1.07h^2 + 1.8h + 81.57$	$0.36h^2 + 0.6h + 27.19$	$\frac{0.36h^2 + 0.6h + 27.19}{0.64h^2 - 0.67h + 5.89}$
Error	6	$3.83h^2 - 4h + 35.33$	$0.64h^2 - 0.67h + 5.89$	-
Total	9	$4.9h^2 - 2.2h + 116.9$	-	-

Decision Rules for *ULM*: Since, the $(\tilde{F})_k^{UL} > F_T$, where $F_T = 4.76$ is the table value at 5% l. o. s with (3,6) *df.* Therefore, the null hypothesis $\tilde{H}_0^{UL}$ of the *ULM* is accepted for all $0 \leq h'^{IN} \leq 0.30$, otherwise the null hypothesis $\hat{H}_0^{UL}$ is rejected for all $0.31 \leq h^{JN} \leq 1$. Hence, the varieties of paddy do not differ significantly with respect to the yields in kilograms.

IV. Conclusion

This study applied fuzzy *ANOVA* techniques using both the *LLM* and *ULM* to evaluate the performance of different paddy varieties in terms of yield. The findings indicate that the null hypothesis is accepted across all levels of significance (α %) for both models, suggesting that the yields of the paddy varieties do not differ significantly. The null hypothesis $\tilde{H}_0^{LL}$ is accepted for all $0 \leq l^{lL} \leq 0.42$ and $\tilde{H}_0^{\mathrm{ULL}}$ is accepted for all $0 \leq h^{\mathrm{LL}} \leq 0.30$. Therefore, the fuzzy null hypothesis $\tilde{H}_0$ of the fuzzy *ANOVA* one-way model is accepted. It says that the varieties of paddy performance of the yields in kilograms are equal only if $0 \leq l^{tL} \leq 0.42$ and $0 \leq h^{iL} \leq 0.30$. But some specific levels of the maximum level of pessimistic value is 0.42 and the maximum level of optimistic value is 0.30. Finally, it is inference that there is no relation between the different yields in kilograms and paddy varieties as decision rules of the conclusions reached by both *LLM* and *ULM*. Furthermore, the analysis highlights specific fuzzy levels, with a maximum pessimistic value of 0.42 and a maximum optimistic value of 0.30, indicating no substantial variation in yield across the varieties. The consistency of the conclusions drawn from both *LLM* and *ULM* strengthens the inference that paddy varieties yield similar performance in kilograms. These results underscore the utility of fuzzy *ANOVA* in agricultural studies and provide a robust framework for analysing performance variability under uncertainty. Future research may explore the application of these fuzzy models to other crops and environmental conditions to validate their broader applicability. For future works, one can use the approach of this paper to extend other experimental designs such as two-way ANOVA, $LSD, BIBD, PBIBD$, etc.

References

[1] Abbasbandy, S., and Asady, B. (2004). The nearest trapezoidal fuzzy number to a fuzzy quantity. *Applied mathematics and computation*, 156(2): 381-386.

[2] Abbasbandy, S., and Amirfakhrian, M. (2006). The nearest approximation of a fuzzy quantity in parametric form. *Applied Mathematics and Computation*, 172(1): 624-632.

[3] Abbasbandy, S., and Amirfakhrian, M. (2006). The nearest trapezoidal form of a generalized left right fuzzy number. *International Journal of Approximate Reasoning*, 43(2): 166-178.

[4] Buckley, J. J., and Jowers, L. J. (2008). Monte Carlo methods in fuzzy optimization. *Berlin: Springer.*

[5] Barbacioru, I. C. (2012). Statistical hypothesis testing using fuzzy linguistic variables. *Fiability & Durability/Fiabilitate si Durabilitate*. (1): 336-342.

[6] Liou, T. S., and Wang, M. J. J. (1992). Ranking fuzzy numbers with integral value. *Fuzzy sets and systems*, 50(3): 247-255.

[7] Salahshour, S., et al. (2011). Ranking fuzzy numbers using fuzzy maximizing-minimizing points. In *Proceedings of the 7th conference of the European Society for Fuzzy Logic and Technology* (pp. 763-769). Atlantis Press.

[8] Rezvani, S., and Molani, M. (2014). Representation of trapezoidal fuzzy numbers with shape function. *Ann. Fuzzy Math. Inform*, 8(1): 89-112.

[9] Thorani, Y. L. P. et al. R. (2012). Ordering generalized trapezoidal fuzzy numbers. *Int. J. Contemp. Math. Sciences*, 7(12): 555-573.

[10] Wang, Y. M., et al. (2006). On the centroids of fuzzy numbers. Fuzzy sets and systems, 157(7): 919-926.

[11] Viertl, R. (2011). Statistical methods for fuzzy data. *John Wiley and Sons*.

[12] Viertl, R. (2006). Univariate statistical analysis with fuzzy data. *Computational Statistics & Data Analysis*, *51*(1): 133-147.

[13] Rezvani, S. (2013). Ranking generalized trapezoidal fuzzy numbers with Euclidean distance by the incentre of centroids. *Mathematica Aeterna*, 3(2): 103-114.

[14] Parthiban, S., and Gajivaradhan, P. (2016). One-factor ANOVA model using trapezoidal fuzzy numbers through alpha cut interval method. *Annals of Pure and Applied Mathematics*, 11(1): 45-61.

[15] Zadeh, L. A. (1996). On fuzzy algorithms. *In fuzzy sets, fuzzy logic, and fuzzy systems: selected papers By Lotfi A Zadeh*.

[16] Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning—I. *Information sciences*, 8(3): 199-249.

[17] Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning—I. *Information sciences*, 8(3): 199-249.

[18] Wu, H. C. (2007). Analysis of variance for fuzzy data. *International Journal of Systems Science*, 38(3): 235-246.

[19] Nourbakhsh, M., et al. (2013). Analysis of variance based on fuzzy observations. *International Journal of Systems Science*, 44(4): 714-726.

RT&A, No 2 (84)
Volume 20, June 2025

ON A CLASS OF INDEFINITE KENMOTSU MANIFOLDS ADMITTING QUARTER-SYMMETRIC METRIC CONNECTION

K. L. Sai Prasad[1], S. Sunitha Devi[2,*]

•

[1] Department of Mathematics, Gayatri Vidya Parishad College of Engineering for Women, Visakhapatnam, 530 048, INDIA

[2,*] Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur District, Andhra Pradesh, 522 302, INDIA

klsprasad@yahoo.com[1] sunithamallakula@yahoo.com[2,*]

Abstract

In this present paper, a class of almost contact metric manifolds equipped with an indefinite metric, termed indefinite Kenmotsu manifolds (known as, ϵ-Kenmotsu), is considered that accepts a connection of quarter-symmetric. We obtain some typical identities for the Riemannian curvature tensor and the Ricci tensor of indefinite Kenmotsu manifolds with respect to the quarter-symmetric metric connection. Furthermore, we derive the scalar curvature of locally symmetric indefinite Kenmotsu manifold which admits a quarter-symmetric metric connection. Finally in this article, in particular, we study and have shown the non-existence of generalised recurrent, ϕ-recurrent and the pseudo-symmetric indefinite Kenmotsu manifolds with respect to quarter-symmetric metric connection.

Keywords: Quarter-symmetric metric connection, indefinite Kenmotsu manifold, locally symmetric manifold, generalised recurrent manifold, ϕ-recurrent manifold, pseudo-symmetric manifold

1. Introduction

In 1969, Takahashi [16] introduced an almost contact manifold equipped with an associated indefinite metric and explored some geometrical properties of almost contact manifolds (particularly, Sasakian manifolds) with indefinite metrics [1]. Later on, it was shown by Xufeng and Xiaoli [20] that those manifolds are real hypersurface of indefinite Kaehlerian manifolds.

On the other hand, in 1972, Kenmotsu [8] established a new class of almost contact manifold known as Kenmotsu manifolds. A Kenmotsu manifold admitting an indefinite metric is termed as an ϵ-Kenmotsu manifold, which was proposed by De and Sarkar [4] in 2009. They showed that the existence of new structure on an indefinite metrics influences the curvatures.

Since the index of the metric generates variety of vector fields such as space-like, time-like and light-like, the study of indefinite structures on manifolds becomes very interesting and of great importance. In view of this, these manifolds have been widely studied by many geometers such as Haseeb, Khan and Siddiqi [7], Rajendra Prasad and Shashikant Pandey [11], Siddiqi and Chaubey [12], Venkatesha and Vishnuvardhana [19] and obtained several results on these manifolds. Recently the authors Sunitha and Sai Prasad established a new class of indefinite almost paracontact metric manifolds, termed (ϵ)-para Kenmotsu manifolds [14].

Let ∇ be a linear connection in an n-dimensional differentiable manifold M. The torsion tensor T and the curvature tensor R of ∇, respectively given by

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y], \tag{1}$$

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z. \tag{2}$$

The connection ∇ is said to be symmetric if its torsion tensor T vanishes, otherwise it is non-symmetric. The connection ∇ is said to be metric connection if there is a Riemannian metric g in M such that $\nabla g = 0$, otherwise it is non-metric. It is well known that a linear connection is symmetric and also metric if the connection is of Levi-Civita [7].

A linear connection ∇ in an n-dimensional differentiable manifold is said to be quarter-symmetric [6] if its torsion tensor T is of the form

$$T(X,Y) = \eta(Y)\phi(X) - \eta(X)\phi(Y), \tag{3}$$

where η is a 1-form and ϕ is a tensor field of type (1, 1). If we change ϕX by X then the connection is said to be semi-symmetric metric connection [21]. Both quarter-symmetric and semi-symmetric metric connections are widely studied by several researchers, for instance [10, 13, 15, 17, 18].

A non-flat n-dimensional Riemannian manifold $M, n > 3$, is called generalised recurrent [3] if its curvature tensor R satisfies the condition

$$(\nabla_X R)(Y,Z)W = \alpha(X)R(Y,Z)W + \beta(X)[g(Z,W)Y - g(Y,W)Z], \tag{4}$$

where α and β are two 1-forms, $(\beta \neq 0)$. If $\beta = 0$ and $\alpha \neq 0$ then M is called recurrent.

A non-flat n-dimensional Riemannian manifold $M, n > 3$, is called ϕ-recurrent [5] if its curvature tensor R satisfies the condition

$$\phi^2((\nabla_X R)(Y,Z)W) = \alpha(X)R(Y,Z)W. \tag{5}$$

A non-flat n-dimensional Riemannian manifold $M, n > 3$, is called pseudo-symmetric [2] if there exists a 1-form α on M such that

$$\begin{aligned}(\nabla_X R)(Y,Z,W) = {} & 2\alpha(X)R(Y,Z)W + \alpha(Y)R(X,Z)W + \alpha(Z)R(Y,X)W \\ & + \alpha(W)R(Y,Z)X + g(R(Y,Z)W,X)A,\end{aligned} \tag{6}$$

where $X, Y, Z \in \chi(M)$ are arbitrary vector fields and $A \in \chi(M)$ is the vector field, given by $g(X,A) = \alpha(X)$. If $\nabla R = 0$ then the manifold M is said to be locally symmetric [9].

In the present paper, we study indefinite Kenmotsu manifolds which accepts a connection of quarter-symmetric. The present paper is organised as follows:

Section 2 is equipped with some prerequisites about indefinite Kenmotsu manifolds. In Section 3, We obtain some typical identities for the Ricci and Riemannian curvature tensors of indefinite-Kenmotsu manifolds with respect to the quarter-symmetric metric connection. In Section 4, we derive the scalar curvature of locally symmetric indefinite Kenmotsu manifold which admits a quarter-symmetric metric connection. In Section 5, we study and have shown the non-existence of generalised recurrent, ϕ-recurrent and the pseudo-symmetric indefinite Kenmotsu manifolds with respect to quarter-symmetric metric connection.

2. Preliminaries

An n-dimensional differentiable manifold (M, g) is said to be an indefinite almost contact metric manifold if it admits a (1, 1) tensor field ϕ, a structure vector field ξ, a 1-form η and an indefinite metric g such that

$$\phi^2 X = -X + \eta(X)\xi, \tag{7}$$

$$\eta(\xi) = 1, \tag{8}$$

$$g(\xi, \xi) = \epsilon, \tag{9}$$

$$\eta(X) = \epsilon g(X, \xi), \tag{10}$$

$$g(\phi X, \phi Y) = g(X, Y) - \epsilon\eta(X)\eta(Y); \tag{11}$$

for every $X, Y \in \chi(M)$, and $\chi(M)$ is a collection of differentiable vector fields on M. Since the structure vector field ξ which has been vector field that is either space-like or time-like, and the rank of that tensor filed ϕ is $(n-1)$, in this case, (ϵ) is either 1 or -1.

If $g(X,\ Y)$ is positive definite, that is

$$d\eta(X,\ Y) = g(X,\ \phi Y), \tag{12}$$

the manifold $M(\phi,\ \xi,\ \eta,\ g,\ \epsilon)$ is referred as an indefinite almost contact metric manifold. Evidently, on M, we have

$$\phi\ \xi = 0, \qquad \eta(\phi X) = 0. \tag{13}$$

If an indefinite contact metric manifold satisfies

$$(\nabla_X \phi)Y = -g(X, \phi Y)\xi - \epsilon\eta(Y)\phi X, \tag{14}$$

the manifolds M is said to be an indefinite Kenmotsu manifold [4]. An indefinite almost contact metric manifold is an indefinite Kenmotsu if and only if

$$\nabla_X \xi = \epsilon(X - \eta(X)\xi). \tag{15}$$

Moreover, the curvature tensor R, the Ricci tensor S and the Ricci operator Q in an indefinite Kenmotsu manifold M with respect to the Levi-Civita connection satisfies [4] the following :

$$(\nabla_X \eta)Y = g(X, Y) - \epsilon\eta(X)\eta(Y), \tag{16}$$

$$R(X, Y)\xi = \eta(X)Y - \eta(Y)X, \tag{17}$$

$$R(\xi, X)\xi = X - \eta(X)\xi, \tag{18}$$

$$\eta(R(X, Y)Z) = \epsilon[g(X, Z)\eta(Y) - g(Y, Z)\eta(X)], \tag{19}$$

$$S(X, \xi) = -(n-1)\eta(X), \tag{20}$$

$$Q\xi = -\epsilon(n-1)\xi, \text{ where } g(QX, Y) = S(X, Y). \tag{21}$$

It yields

$$S(\phi X, \phi Y) = S(X, Y) + \epsilon(n-1)\eta(X)\eta(Y). \tag{22}$$

An indefinite Kenmotsu manifold M is said to be η-Einstein manifold if its Ricci tensor is of the form

$$S(X, Y) = ag(X, Y) + b\eta(X)\eta(Y), \tag{23}$$

where a and b are scalar functions of ϵ.

3. Expression of $\widetilde{R}$(X, Y)Z in terms of R(X, Y)Z

Let M be an n-dimensional indefinite Kenmotsu manifold and ∇ be the Levi-Civita connection on M. The quarter-symmetric metric connection $\widetilde{\nabla}$ on M is given by

$$\widetilde{\nabla}_X Y = \nabla_X Y + \eta(Y)X - g(X,Y)\xi. \tag{24}$$

The curvature tensor $\widetilde{R}$ of M with respect to the quarter-symmetric metric connection $\widetilde{\nabla}$ is defined by

$$\widetilde{R}(X,Y)Z = \widetilde{\nabla}_X\widetilde{\nabla}_Y Z - \widetilde{\nabla}_Y\widetilde{\nabla}_X Z - \widetilde{\nabla}_{[X,Y]}Z. \tag{25}$$

By Virtue of (24) and (25), we have

$$\begin{aligned}\widetilde{R}(X,Y)Z =& R(X,Y)Z + (2+\epsilon)\big[g(X,Z)Y - g(Y,Z)X\big] \\ & (1+\epsilon)\big[g(Y,Z)\eta(X) - g(X,Z)\eta(Y)\big]\xi \\ & (1+\epsilon)\big[\eta(Y)X - \eta(X)Y\big]\eta(Z),\end{aligned} \tag{26}$$

which is a relation between the curvature tensor of M with respect to the quarter-symmetric metric connection $\widetilde{\nabla}$ and Levi-Civita cpnnection ∇.

Then from (26) and (17), we get

$$\widetilde{R}(X,\xi)Y = (1+\epsilon)\big[g(X,Y)\xi - \eta(Y)X\big], \tag{27}$$

$$\widetilde{R}(X,Y)\xi = (1+\epsilon)\big[\eta(X)Y - \eta(Y)X\big], \tag{28}$$

and

$$\widetilde{R}(\xi,X)\xi = (1+\epsilon)\big[X - \epsilon\eta(X)\xi\big]. \tag{29}$$

By taking the inner product of (26) with W, we have

$$\begin{aligned}\widetilde{R}(X,Y,Z,W) = & R(X,Y,Z,W) + (2+\epsilon)\big[g(X,Z)g(Y,W) - g(Y,Z)g(X,W)\big] \\ & + (1+\epsilon)\big[g(Y,Z)\eta(X) - g(X,Z)\eta(Y)\big]g(\xi,W) \\ & + (1+\epsilon)\big[\eta(Y)g(X,W) - \eta(X)g(Y,W)\big]\eta(Z).\end{aligned} \tag{30}$$

On the contraction of the above expression with respect to X and W, we get

$$\widetilde{S}(Y,Z) = S(Y,Z) + g(Y,Z)\big[(\epsilon+2)(\epsilon-n)+2\big] + \eta(Y)\eta(Z)\big[(1+\epsilon)(n-2\epsilon)\big], \tag{31}$$

where $\widetilde{S}$ and S are the Ricci tensors of the connections $\widetilde{\nabla}$ and ∇, respectively. It shows that in an indefinite Kenmotsu manifold, the Ricci tensor of the quarter-symmetric metric connection is not symmetric.

Further on contracting (31) with regard to Y and Z, results in

$$\widetilde{r} = r + n\big[(\epsilon+2)(\epsilon-n)+2\big] + \big[(1+\epsilon)(n-2\epsilon)\big], \tag{32}$$

where $\widetilde{r}$ and r are the scalar curvatures of the connections $\widetilde{\nabla}$ and ∇, respectively.
From the above, we state the following theorem:

Theorem 1. For an indefinite Kenmotsu manifold M which admits the quarter-symmetric metric connection $\widetilde{\nabla}$, for every $X, Y, Z \in \chi(M)$, we obtain the following:

(a) The curvature tensor $\widetilde{R}$ is given by (26)
(b) The Ricci tensor $\widetilde{S}$ is given by (31)
(c) $\widetilde{R}(X,Y,Z,W) + \widetilde{R}(X,Y,W,Z) = 0$
(d) $\widetilde{R}(X,Y,Z,W) + \widetilde{R}(Y,X,W,Z) = 0$
(e) $\widetilde{S}(Y,\xi) = (1-n)(1+\epsilon)\eta(Y)$
(f) $\widetilde{r} = r + n\big[(\epsilon+2)(\epsilon-n)+2\big] + \big[(1+\epsilon)(n-2\epsilon)\big]$
(g) The Ricci tensor $\widetilde{S}$ is not symmetric.

4. Locally symmetric indefinite Kenmotsu manifolds with respect to quarter-symmetric metric connection

In this section, we consider locally symmetric indefinite Kenmotsu manifold with respect to the quarter-symmetric metric connection $\widetilde{\nabla}$.

Theorem 2. Let M be a locally symmetric indefinite Kenmotsu manifold with respect to the quarter symmetric metric connection $\widetilde{\nabla}$. Then the scalar curvature of the Levi-Civita connection of M is $r = (1-n)[1-(n-3)\epsilon]$.

Proof. Assume that M is a locally symmetric indefinite Kenmotsu manifold with respect to quarter symmetric metric connection $\widetilde{\nabla}$. Then $(\widetilde{\nabla}_X \widetilde{R})(Y,Z)W = 0$.

By a suitable contraction of this equation, we get

$$\widetilde{\nabla}_X(\widetilde{S}(Z,W)) = \widetilde{\nabla}_X \widetilde{S}(Z,W) - \widetilde{S}(\widetilde{\nabla}_X Z, W) - \widetilde{S}(Z, \widetilde{\nabla}_X W) = 0.$$

For $W = \xi$ in the above equation, we have

$$\widetilde{\nabla}_X(\widetilde{S}(Z,\xi)) = \widetilde{\nabla}_X \widetilde{S}(Z,\xi) - \widetilde{S}(\widetilde{\nabla}_X Z, \xi) - \widetilde{S}(Z, \widetilde{\nabla}_X \xi) = 0.$$

By making use of (15), (20), (24) and (31), we get

$$-S(Z,X) + g(Z,X)[n\epsilon - 1 - 2\epsilon] + \eta(X)\eta(Z)\big[(1-n)(\epsilon-1) - (1+\epsilon)(n-2\epsilon)\big] = 0. \tag{33}$$

Then on contracting the above equation over X and Z, we obtain

$$r = (1-n)[1-(n-3)\epsilon], \tag{34}$$

which completes the proof of the theorem. ■

5. Non-existence of certain kinds of indefinite Kenmotsu manifolds with respect to the quarter-symmetric metric connection

Theorem 3. There is no generalised recurrent indefinite Kenmotsu manifold with respect to quarter-symmetric metric connection $\widetilde{\nabla}$.

Proof. Suppose that there exists a generalised recurrent indefinite Kenmotsu manifold M with respect to quarter-symmetric metric connection $\widetilde{\nabla}$. Then from (4), we have

$$(\widetilde{\nabla}_X \widetilde{R})(Y,Z)W = \alpha(X)\widetilde{R}(Y,Z)W + \beta(X)\big[g(Z,W)Y - g(Y,W)Z\big], \tag{35}$$

for all vector fields $X, Y, Z, W \in \chi(M)$.

For $Y = W = \xi$ in (35), we have

$$(\widetilde{\nabla}_X \widetilde{R})(\xi,Z)\xi = \alpha(X)\widetilde{R}(\xi,Z)\xi + \beta(X)\big[\epsilon\eta(Z)\xi - \epsilon Z\big]. \tag{36}$$

Now by using (30) and (17), we have, from the above equation

$$(\widetilde{\nabla}_X \widetilde{R})(\xi,Z)\xi = \big[Z - \eta(Z)\xi\big]\alpha(X) + \epsilon\big[Z - \eta(Z)\xi\big]\big[\alpha(X) - \beta(X)\big]. \tag{37}$$

On the other hand, in view of (15), (17), (24), (27), (28) and (29), we have

$$(\widetilde{\nabla}_X \widetilde{R})(\xi,Z)\xi = 0, \tag{38}$$

and hence from the equations (37) and (38), we have

$$\big[Z - \eta(Z)\xi\big]\alpha(X) + \epsilon\big[Z - \eta(Z)\xi\big]\big[\alpha(X) - \beta(X)\big] = 0. \tag{39}$$

Replacing Z by ϕZ in (39) we get

$$\phi Z\ \alpha(X) + \epsilon\ \phi Z\big[\alpha(X) - \beta(X)\big] = 0, \tag{40}$$

which implies $\alpha = \beta = 0$. This contradicts the condition $\beta \neq 0$. Therefore the statement of the theorem follows. ■

Theorem 4. There is no ϕ-recurrent indefinite Kenmotsu manifold with respect to the quarter-symmetric metric connection $\widetilde{\nabla}$.

Proof. Suppose that there exists a ϕ-recurrent indefinite Kenmotsu manifold with respect to the quarter-symmetric metric connection $\widetilde{\nabla}$. Then from (5), we have

$$\phi^2\left[(\widetilde{\nabla}_X\widetilde{R}(Y,Z)W\right] = \alpha(X)\widetilde{R}(Y,Z)W,$$

for all vector fields $X, Y, Z, W \in \chi(M)$.

Then on using (7), we get

$$-(\widetilde{\nabla}_X\widetilde{R})(Y,Z)W + \eta((\widetilde{\nabla}_X\widetilde{R}(Y,Z)W)\xi = \alpha(X)\widetilde{R}(Y,Z)W. \tag{41}$$

By replacing Y and W with ξ in (41), we have

$$-(\widetilde{\nabla}_X\widetilde{R})(\xi,Z)\xi + \eta((\widetilde{\nabla}_X\widetilde{R}(\xi,Z)\xi)\xi = \alpha(X)\widetilde{R}(\xi,Z)\xi. \tag{42}$$

On the other hand, from (38) we have

$$\left(\widetilde{\nabla}_X\widetilde{R}\right)(\xi,Z)\xi = 0.$$

Therefore the above equation (42) turns into

$$\alpha(X)\widetilde{R}(\xi,Z)\xi = 0.$$

Then by virtue of (29), it is obvious that

$$\alpha(X)(1+\epsilon)[Z - \epsilon\eta(Z)\xi] = 0.$$

This implies $\alpha(X) = 0$ for any vector field $X \in \chi(M)$.

It contradicts the condition that $\alpha \neq 0$ and hence the theorem is proved. ■

Theorem 5. There is no pseudo-symmetric indefinite Kenmotsu manifold with respect to the quarter-symmetric metric connection $\widetilde{\nabla}$.

Proof. Suppose that there exists a pseudo-symmetric indefinite Kenmotsu manifold M with respect to the quarter-symmetric metric connection $\widetilde{\nabla}$. Then from (6) we have

$$\begin{aligned}(\widetilde{\nabla}_X\widetilde{R})(Y,Z)W &= 2\alpha(X)\widetilde{R}(Y,Z)W + \alpha(Y)\widetilde{R}(X,Z)W \\ &\quad + \alpha(Z)\widetilde{R}(Y,X)W + \alpha(W)\widetilde{R}(Y,Z)X \\ &\quad + g(\widetilde{R}(Y,Z)W,X)A.\end{aligned} \tag{43}$$

By a suitable contraction of (43), we get

$$\begin{aligned}(\widetilde{\nabla}_X\widetilde{S})(Z,W) &= 2\alpha(X)\widetilde{S}(Z,W) + \alpha(\xi)\widetilde{R}(X,Z,W,\xi) \\ &\quad + \alpha(Z)\widetilde{S}(X,W) + \alpha(W)\widetilde{S}(Z,X) + g(\widetilde{S}(Z,W),X)A.\end{aligned} \tag{44}$$

Taking $W = \xi$ in (44) and on using the equations (20), (27), (28) and (31), we obtain

$$\begin{aligned}(\widetilde{\nabla}_X\widetilde{S})(Z,\xi) &= 2\alpha(X)\widetilde{S}(Z,\xi) + \alpha(\xi)\widetilde{R}(X,Z,\xi,\xi) \\ &\quad + \alpha(Z)\widetilde{S}(X,\xi) + \alpha(\xi)\widetilde{S}(Z,X) + g(\widetilde{S}(Z,\xi),X)A.\end{aligned} \tag{45}$$

On the other hand, by the covariant differentiation of the Ricci tensor $\widetilde{S}$ with respect to the quarter symmetric metric connection $\widetilde{\nabla}$, we have

$$(\widetilde{\nabla}_X\widetilde{S})(Z,W) = \widetilde{\nabla}_X\widetilde{S}(Z,W) - \widetilde{S}(\widetilde{\nabla}_XZ,W) - \widetilde{S}(Z,\widetilde{\nabla}_XW). \tag{46}$$

Now by putting $W = \xi$ in (46) and on using (31), (24) and (15), we get

$$\begin{aligned}(\widetilde{\nabla}_X\widetilde{S})(Z,\xi) &= -S(Z,X) + g(Z,X)[n\epsilon - 1 - 2\epsilon] \\ &\quad + \eta(X)\eta(Z)[(1-n)(\epsilon-1) - (1+\epsilon)(n-2\epsilon)].\end{aligned} \tag{47}$$

Then by comparing the right hand sides of the equations (45) and (47), we have

$$\begin{aligned}&-S(Z,X) + g(Z,X)[n\epsilon - 1 - 2\epsilon] + \eta(X)\eta(Z)[(1-n)(\epsilon-1) - (1+\epsilon)(n-2\epsilon)] \\ &\quad = 2(1-n)(1+\epsilon)\alpha(X)\eta(Z) + (1-n)(1+\epsilon)\eta(X)\alpha(Z) \\ &\quad + \alpha(\xi)\big[S(Z,X) + g(Z,X)\big((\epsilon+2)(\epsilon-n)+2\big) + \eta(X)\eta(Z)(1+\epsilon)(n-2\epsilon)\big] \\ &\quad + (1-n)(1+\epsilon)\eta(Z)\eta(X).\end{aligned} \tag{48}$$

Replacing X and Z with ξ in the above equation, we get (since $n > 3$)

$$3(1+\epsilon)\alpha(\xi) + (1-\epsilon) = 0, \tag{49}$$

which implies

$$\alpha(\xi) = 0. \tag{50}$$

Now we show that $\alpha = 0$ holds for any vector field on M.

By taking $Z = \xi$ in (43) and on using (50), we get

$$(\widetilde{\nabla}_X\widetilde{S})(\xi,W) = 2\alpha(X)\widetilde{S}(\xi,W) + \alpha(W)\widetilde{S}(\xi,X) + g(\widetilde{S}(\xi,W),X)A. \tag{51}$$

Using the aforementioned equation along with (20), (27), (28), (46) and (50), we now obtain

$$\begin{aligned}&-S(X,W) + g(X,W)[n\epsilon - 1 - 2\epsilon] + \eta(X)\eta(W)(n - n\epsilon + 2\epsilon) \\ &\quad = 2(1-n)(1+\epsilon)\alpha(X)\eta(W) + (1-n)(1+\epsilon)\eta(X)\alpha(W) \\ &\quad + \alpha(\xi)\big[S(X,W) + g(X,W)\big((\epsilon+2)(\epsilon-n)+2\big) + \eta(X)\eta(W)(1+\epsilon)(n-2\epsilon)\big].\end{aligned} \tag{52}$$

Further, by taking $w = \xi$ in (52), we have

$$2(1-n)(1+\epsilon)\alpha(X) + \eta(X)(1-n)(1-\epsilon) = 0, \tag{53}$$

which implies $\alpha(X) = 0$ for any vector field $X \in \chi(M)$. This is a contradiction to the definition of pseudo-symmetry. Thus the theorem is proved. ■

6. CONCLUSION

This paper defines a class of almost contact metric manifolds equipped with an indefinite metric, termed indefinite Kenmotsu manifolds. Since the index of the metric generates variety of vector fields such as space-like, time-like and light-like, the geometrical features of these manifolds are widely applied in a variety of physical and geometrical fields, including the construction of super resolution sensors in electronic and communication systems, in electrical engineering, and in the general theory of relativity.

Acknowledgements: The G. V. P. College of Engineering for Women in Visakhapatnam's Senior Assistant professor, Dr. A. Kameswara Rao, is appreciated by the authors for his insightful advice during the manuscript's preparation.

Declarations of conflicting interests: The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] A. Bejancu, A. and Duggal, K. L. (1993). Real hypersurfaces of Indefinite Kaehler manifolds. *Int. J. Math. Math. Sci.*, 16(3):545-556.
[2] Chaki, M. C. (1987). On pseudo symmetric manifolds. *An. Stiint. Univ. Al. I. Cuza Iasi Sect. I a Mat.*, 33:53-58.
[3] De, U. C. and Guha, N. (1991). On generalised recurrent manifolds. *Proc. Math. Soc.*, 7:7-11.
[4] De, U. C. and Sarkar, A. (2009). On ϵ-Kenmotsu manifold. *Hardonic J.*, 32(2): 231-242.
[5] De, U. C., Shaikh, A. A. and Biswas, S. (2003). On ϕ-recurrent Sasakian manifolds. *Novi Sad J. Math.*, 33(2):43-48.
[6] Golab, S. (1975). On semi-symmetric and quarter-symmetric linear connections. *Tensor (N.S.)*, 29:249-254.
[7] Haseeb, A., Khan, M. A. and Siddiqi, M. D. (2016). Some more results on an ϵ-Kenmotsu manifold with a semi-symmetric metric connection. *Acta Math. Univ. Comenianae*, 85(1):9-20.
[8] Kenmotsu, K. (1972). A class of almost contact Riemannian manifold. *Tohoku Math. J.*, 24:93-103.
[9] Neill, B. O. (1983). Semi-Riemannian geometry with applications to relativity. *Pure and Applied Mathematics*, Academic Press, Inc., New York, 103:1-483.
[10] Pathak, G. and De, U. C. (2002). On a semi-symmetric connection in a Kenmotsu manifold. *Bull. Calcutta Math. Soc.*, 94(4):319-324.
[11] Rajendra Prasad and Shashikant Pandey. (2017). An indefinite Kenmotsu manifold endowed with quarter symmetric metric connection. *Global Journal of Pure and Applied Math.*, 13(7):3477-3495.
[12] Siddiqi, M. D. and Chaubey, S. K. (2020). η-Einstein solitons on ϵ-Kenmotsu manifolds. *Kyungpook Math. J.*, 60:805–819.
[13] Sunitha Devi, S. and Sai Prasad, K. L. (2023). On a class of Lorentzian para-Kenmotsu manifolds admitting quarter-symmetric metric connection. *Reliability: Theory and Applications*, 18(4):525-532.
[14] Sunitha Devi, S. and Sai Prasad, K. L. (2025). A class of indefinite almost paracontact metric manifolds. *International Journal of Maps in Mathematics*, 8(1):247-257.
[15] Sunitha Devi, S., Satyanarayana, T. and Sai Prasad, K. L. (2021). Curvature tensors in SP-Kenmotsu manifolds with respect to quarter-symmetric metric connection. *Reliability: Theory and Applications*, 16(4):210-218.
[16] Takahashi, T. (1969). Sasakian manifold with pseudo-Riemannian metric. *Tohoku Math. J. Second Ser.*, 21:271–290.
[17] Tripathi, M. M. (1999). On a semi symmetric metric connection in a Kenmotsu manifold. *J. Pure Math.*, 16:67-71.
[18] Tripathi, M. M. (2001). On a semi symmetric non-metric connection in a Kenmotsu manifold. *Bull. Calcutta Math. Soc.*, 93:323-330.
[19] Venkatesha and Vishnuvardhana, S. V. (2017). ϵ-Kenmotsu manifolds admitting a semi-symmetric metric connection. *Ital. J. Pure Appl. Math.*, 38:615–623.
[20] Xufeng, X. and Xiaoli, C. (1998). Two theorem on ϵ-Sasakian manifolds. *Int. J. Math. and Math. Sci.*, 21(2):249-254.
[21] Yano, K. (1970). On semi-symmetric metric connection. *Rev. Roumaine Math. Pures Appl.*, 15:1579-1586.

PROPERTIES AND APPLICATION OF ALPHA LOGARITHM TRANSFORMED SINE WEIBULL DISTRIBUTION

A.A. Osi*,[1], O.O. Ishaq[1], B.A. Abdulsalam[1], Usman Abubakar [1,2]

[1]Department of Statistics, Aliko Dangote University of Science and Technology, wudil-Kano, Nigeria.
[2]Department of Statistics, Jigawa state Polytechnic, Dutse, Nigeria.
Email: aaosi@kustwudil.edu.ng, usmanabubakar@jigpoly.edu.ng,
babington4u@gmail.com, alhassanbilal5@gmail.com

Abstract

In this research, we introduced a new submodel of the alpha logarithm transform sine-X family of distributions, called the alpha logarithm transform sine-Weibull. We derived its statistical properties, including survival function, hazard function, moments, moment-generating function, and order statistics. We also estimated the parameters using the maximum likelihood method. The performance of the maximum likelihood estimators was evaluated in terms of mean, bias, and root mean squared errors through two simulation studies. Furthermore, the practicality of this family was demonstrated using COVID-19 mortality data from Saudi Arabia and milk output data sets, where the proposed model consistently provided better fits compared to other existing distributions.

Keywords: Alpha Logarithm Transformed Sine-X family of Distributions, Weibull Distribution, Properties, Simulation, Application

1. INTRODUCTION

Modeling lifetime data is an important concept in statistics, as many existing distributions do not fit the data well, which can lead to unreliable results in various fields, including engineering, economics, biological science, and mathematics (1). This limitation has attracted researchers to introduce more flexible and versatile distributions that can serve as substitutes for current baseline distributions (3). This approach provides a framework for introducing flexible distributions along with their characterization and properties, including entropy, quantile, moment, moment-generating function, survival function, hazard rate function, and order statistics. Estimation of parameters is a crucial aspect, involving the use of maximum likelihood and other estimation methods.

In some cases, existing distributions can handle the task of fitting real data sets; however, due to certain limitations, it may be necessary to modify or enhance these distributions to make them more versatile than in their initial form. (4) examines some existing distributions and potential improvements. For example, the Rayleigh-Pareto distribution was modified using the arcsine function to create the arcsine Rayleigh-Pareto distribution (7), while the sine-Weibull distribution was explored by (18), the gamma-extended Weibull by (9), and a new alpha-power cosine Weibull was introduced by (10), among others.

The Weibull distribution is a continuous probability distribution first discovered by (17). Although it was not widely recognized until the work of (12), it found real-life applications

through the efforts of (13). The Weibull distribution is particularly applicable in situations where a skewed distribution is necessary to model data that require a heavy-tailed distribution. In such instances, the Weibull distribution is one of the appropriate options to provide accurate and reliable results.

A generalized family of distributions known as the alpha power transformation distribution was introduced by (14). This family increases the flexibility of existing models by adding a parameter. Consequently, the alpha power family and the sine-Weibull distribution can be combined to create the alpha logarithm transform sine-Weibull distribution.

The need to introduce a new probability distribution arises in various fields, including biological science, engineering, and economics, when current probability distributions do not adequately fit lifetime data that may be positively or negatively skewed. The current sine-Weibull distribution (18), which has only two parameters, may not be flexible enough to meet the requirements of standard distributions. Therefore, there is a need to extend it by adding another parameter using logarithmic transformation to enhance the flexibility of the existing sine-Weibull distribution.

2. Methods

In this section, we will discuss the proposed model, its validity check, mathematical properties, and parameter estimation.

2.1. Alpha Logarithm Transform-G Distribution (ALT)

The alpha logarithm transformed method can be obtained using the following procedure: let (G(x)) be a continuous distribution function with probability density function (g(x)). Then, (F(x)), the cumulative distribution function (CDF) of the alpha logarithm transformed family of distributions proposed by Pappas et al. (2012), takes the following form.

$$F_{ALT}(x) = 1 - \frac{log(\alpha - (\alpha - 1)G(x))}{log(\alpha)} \quad \alpha > 0 and \alpha \neq .1 \tag{1}$$

$$f_{ALT}(x) = \frac{(\alpha - 1)g(x)}{(\alpha - (\alpha - 1)G(x))log(\alpha)} \tag{2}$$

Where, G(x) is the baseline cumulative distribution function. g(x) is the corresponding probability density function and α is the shape parameter.

2.2. Weibull distribution

The weibull distribution is a two parameters continuous probability distribution introduced by If T has a weibull distribution (17), with CDF

$$F(x) = \left(1 - e^{-(\frac{x}{\lambda})^{\beta}}\right) \tag{3}$$

and PDF

$$f(x) = \frac{x^{\beta-1}}{\lambda^{\beta}} e^{-(\frac{x}{\lambda})^{\beta}} \left(1 - e^{-(\frac{x}{\lambda})^{\beta}}\right) \tag{4}$$

respectively, where $\beta > 0$ is the scale parameter and $\lambda > 0$ is the shape parameter. Weibull distribution was used by many researchers as a baseline distribution including sine weibull distribution by (18), Topp-leone weibull distribution by (19), and alpha power topp-leone weibull distribution by (20) among others.

2.3. Alpha Logarithm Transformed Sine Weibull Distribution (ALTSW-D)

Let x follows Alpha logarithm transformed sine weibull random variable with cumulative density function given as

$$F_{ALSW}(x,\alpha,\beta,\lambda) = 1 - \frac{log\left(\alpha - \bar{\alpha}Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)\right)}{log(\alpha)} \tag{5}$$

And corresponding probability density function (pdf). Given as;

$$f_{ALSW}(x,\alpha,\beta,\lambda) = \frac{\bar{\alpha}\pi\lambda\beta x^{-\lambda x^{\beta}}Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)}{2(\alpha-\bar{\alpha})log(\alpha)Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)} \tag{6}$$

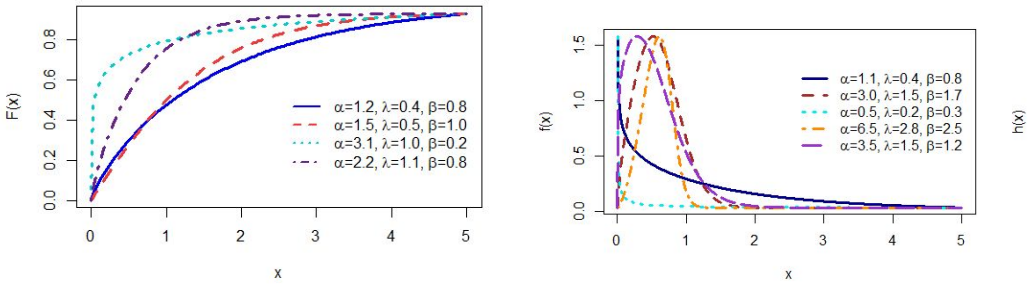


Figure 1: *CDF and PDF plot of ALTSW-D for some parameters*

Figures 1 display the PDF and CDF charts for the ALTSW-D. The risk shows growing and uni-model forms, and the PDF is right-skewed.

2.3.1 Linear representation of ALTSW-D

The probability density function of the proposed distribution is reduced as follows,

$$f_{ALSW}(x,\alpha,\beta,\lambda) = \frac{\bar{\alpha}\pi\lambda\beta x^{-\lambda x^{\beta}}e^{-\lambda x^{\beta}}Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)}{2(\alpha-\bar{\alpha})log(\alpha)Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)} \tag{7}$$

Using power series expansion

$$e^{-\lambda x^{\beta}} = \sum_{i=0}^{\infty}\frac{(-1)^i\lambda x^{\beta i}}{i!} \tag{8}$$

Substituting equation (8) into (7), we have

$$f(x) = f_{ALSW}(x) = \sum_{i=0}^{\infty}\frac{\bar{\alpha}\pi\lambda^{i+1}\beta x^{\beta i+\beta-1}(-1)^i Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)}{2i!(\alpha-\bar{\alpha})log(\alpha)Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)} \tag{9}$$

Let

$$\psi = \frac{\pi\beta\bar{\alpha}\lambda^{i+1}}{2i!log(\alpha)} \tag{10}$$

$$\Rightarrow f(x) = \sum_{i=0}^{\infty} \frac{\psi x^{\beta i+\beta-1} Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)}{\alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)} \tag{11}$$

Also let

$$\gamma = \frac{x^{\beta i+\beta-1} Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)}{\alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)} \tag{12}$$

$$\Rightarrow f(x) = \sum_{i=0}^{\infty} \gamma\psi \tag{13}$$

2.4. Mathematical Properties of ALTSW-D

In this part, some properties of the proposed distribution was derived including quantile function, survival function, hazard rate function, moment, moment generating function, order statistics and estimation of parameters using maximum likelihood method.

2.4.1 Quantile function of ALTSW-D

The quantile function is the inverse of cumulative density function use for simulation study. The quantile function of ALTSW-D were derived as follows Let

$$u = F_{ALSW}(x) = 1 - \frac{log\left(\alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)\right)}{log(\alpha)} \tag{14}$$

$$\Rightarrow u = 1 - \frac{log\left(\alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)\right)}{log(\alpha)} \tag{15}$$

$$\Rightarrow 1 - u = \frac{log\left(\alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)\right)}{log(\alpha)} \tag{16}$$

By cross multiply, we have

$$\Rightarrow (1-u)log(\alpha) = log\left(\alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)\right) \tag{17}$$

$$\Rightarrow log(\alpha)^{1-u} = log\left(\alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)\right) \tag{18}$$

Taking the exponential of both sides

$$\Rightarrow e^{log(\alpha)^{1-u}} = e^{log\left(\alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)\right)} \tag{19}$$

$$\Rightarrow (\alpha)^{1-u} = \alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right) \tag{20}$$

By collecting like terms, we have

$$\Rightarrow \alpha - (\alpha)^{1-u} = \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right) \tag{21}$$

Dividing equation (21) by $\bar{\alpha}$

$$\Rightarrow \frac{\alpha-(\alpha)^{1-u}}{\bar{\alpha}} = Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right) \tag{22}$$

$$\Rightarrow Arcsine\left(\frac{\alpha-(\alpha)^{1-u}}{\bar{\alpha}}\right) = \frac{\pi(1-e^{-\lambda x^{\beta}})}{2} \tag{23}$$

Multiply both sides by $\frac{2}{\pi}$

$$\Rightarrow 1-e^{-\lambda x^{\beta}} = \frac{2}{\pi}Arcsine\left(\frac{\alpha-(\alpha)^{1-u}}{\bar{\alpha}}\right) \tag{24}$$

$$\Rightarrow e^{-\lambda x^{\beta}} = 1-\frac{2}{\pi}Arcsine\left(\frac{\alpha-(\alpha)^{1-u}}{\bar{\alpha}}\right) \tag{25}$$

Taking the natural log of both sides of equation (25), we have

$$\Rightarrow loge^{-\lambda x^{\beta}} = log\left(1-\frac{2}{\pi}Arcsine\left(\frac{\alpha-(\alpha)^{1-u}}{\bar{\alpha}}\right)\right) \tag{26}$$

$$\Rightarrow -\lambda x^{\beta} = log\left(1-\frac{2}{\pi}Arcsine\left(\frac{\alpha-(\alpha)^{1-u}}{\bar{\alpha}}\right)\right) \tag{27}$$

Also, divide both sides by $-\lambda$

$$\Rightarrow x^{\beta} = \frac{-1}{\lambda}log\left(1-\frac{2}{\pi}Arcsine\left(\frac{\alpha-(\alpha)^{1-u}}{\bar{\alpha}}\right)\right) \tag{28}$$

Taking power of $\frac{1}{\beta}$ to both sides of equation (28)

$$\Rightarrow (x^{\beta})^{\frac{1}{\beta}} = \left\{\frac{-1}{\lambda}log\left(1-\frac{2}{\pi}Arcsine\left(\frac{\alpha-(\alpha)^{1-u}}{\bar{\alpha}}\right)\right)\right\}^{\frac{1}{\beta}} \tag{29}$$

$$\Rightarrow x_u = \left\{\frac{-1}{\lambda}log\left(1-\frac{2}{\pi}Arcsine\left(\frac{\alpha-(\alpha)^{1-u}}{\bar{\alpha}}\right)\right)\right\}^{\frac{1}{\beta}} \tag{30}$$

2.4.2 Survival and Hazard function of ALTSW-D

A random variable $X \sim ALTSW-D$ has the survival and hazard rate function given as;

$$S(x) = \frac{log\left(\alpha-\bar{\alpha}Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)\right)}{log(\alpha)} \tag{31}$$

$$H(x) = \frac{\bar{\alpha}\pi\lambda\beta x^{-\lambda x^{\beta}}e^{-\lambda x^{\beta}}Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)log(\alpha)}{2(\alpha-\bar{\alpha})log(\alpha)Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)log\left(\alpha-\bar{\alpha}Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)\right)} \tag{32}$$

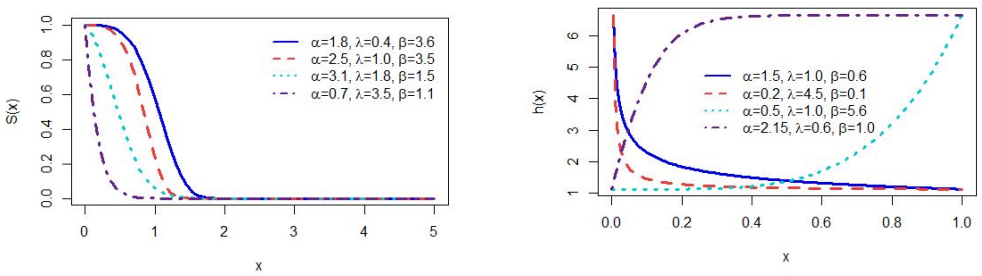


Figure 2: *Hazard rate and Survival plots of ALTSW-D for different values of parameters*

2.4.3 Renyi's Entropy of ALTSW-D

Renyi entropy forms the basis of the concept of generalized dimensions, and is given by;

$$I(x) = \frac{1}{1-\theta} log \int_{-\infty}^{\infty} f(x)^{\theta} dx \tag{33}$$

$$\Rightarrow I(x) = \frac{1}{1-\theta} log \int_{0}^{\infty} \left(\frac{\bar{\alpha}\pi\lambda\beta x^{\beta-1} e^{-\lambda x^{\beta}} Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)}{2(\alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)) log(\alpha)} \right)^{\theta} dx \tag{34}$$

$$\Rightarrow I(x) = \frac{1}{1-\theta} log \int_{0}^{\infty} (\omega\epsilon)^{\theta} dx \tag{35}$$

$$\Rightarrow I(x) = \frac{1}{1-\theta} log\omega^{\theta} + \int_{0}^{\infty} (\epsilon)^{\theta} dx \tag{36}$$

$$\Rightarrow I(x) = \frac{1}{1-\theta} \left[\theta log\omega + \int_{0}^{\infty} (\epsilon)^{\theta} dx \right] \tag{37}$$

Where

$$\omega = \frac{\bar{\alpha}\pi\lambda\beta}{2(\alpha - \bar{\alpha}) log(\alpha)} \tag{38}$$

And

$$\epsilon = \frac{x^{\beta-1} e^{-\lambda x^{\beta}} Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)}{(\alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right))} \tag{39}$$

2.4.4 Moment of ALTSW-D

The r^{th} moment about the origin of the Proposed distribution "alpha logarithm transformed sine weibull distribution" (ALSW-D) is given by.

$$\mu^{r} = E(x^{r}) = \int_{-\infty}^{\infty} x^{r} f(x) dx \tag{40}$$

$$\mu^{r} = \int_{0}^{\infty} \frac{x^{r} \bar{\alpha}\pi\lambda\beta x^{\beta-1} e^{-\lambda x^{\beta}} Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)}{2(\alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)) log(\alpha)} dx \tag{41}$$

Using power series expansion,

$$e^{-\lambda x^{\beta}} = \sum_{i=1}^{\infty} \frac{(-1)^{i} \lambda^{i} x^{\beta i}}{i!} \tag{42}$$

This implies that equation (41) become

$$\mu^{r} = \sum_{i=1}^{\infty} \int_{0}^{\infty} \frac{\bar{\alpha}\pi\lambda^{i+1}\beta x^{\beta i+\beta+r-1}(-1)^{i} Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)}{2i!(\alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)) log(\alpha)} dx \tag{43}$$

$$\mu^{r} = \sum_{i=1}^{\infty} \gamma\kappa \tag{44}$$

Where

$$\gamma = \int_{0}^{\infty} \frac{x^{\beta i+\beta+r-1} Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)}{(\alpha - \bar{\alpha} Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right))} dx \tag{45}$$

And

$$\kappa = \frac{\bar{\alpha}\pi\lambda^{i+1}\beta(-1)^i}{2i!log(\alpha)} \tag{46}$$

2.4.5 Moment Generating Function of ALTSW-D

In Statistics and Probability Theorem, the Moment Generating function is used to encodes the information about the moments of a random variable. and is given as;

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx \tag{47}$$

$$\Rightarrow M_x(t) = \int_0^{\infty} \frac{e^{tx}\bar{\alpha}\pi\lambda\beta x^{\beta-1}e^{-\lambda x^{\beta}} Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)}{2(\alpha - \bar{\alpha}Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right))log(\alpha)}dx \tag{48}$$

But

$$e^{-\lambda x^{\beta}} = \sum_{j=1}^{\infty} \frac{(-1)^j\lambda^j x^{\beta j}}{j!} \tag{49}$$

And

$$e^{tx} = \sum_{p=1}^{\infty} \frac{t^p x^p}{p!} \tag{50}$$

Substituting (50) and (49), in to (48), we have

$$M_x(t) = \sum_{j=1}^{\infty}\sum_{p=1}^{\infty} \int_0^{\infty} \frac{\bar{\alpha}\pi\lambda^{j+1}\beta(-1)^j t^p x^{\beta+\beta j+p-1} Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)}{2j!p!\left(\alpha - \bar{\alpha}Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)\right)log(\alpha)}dx \tag{51}$$

$$\Rightarrow M_x(t) = \sum_{j=1}^{\infty}\sum_{p=1}^{\infty} \tau\eta \tag{52}$$

Where

$$\tau = \int_0^{\infty} \frac{x^{\beta+\beta j+p-1} Cos\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)}{\left(\alpha - \bar{\alpha}Sin\left(\frac{\pi(1-e^{-\lambda x^{\beta}})}{2}\right)\right)}dx \tag{53}$$

And

$$\eta = \frac{\bar{\alpha}\pi\lambda^{j+1}\beta(-1)^j t^p}{2j!p!log(\alpha)} \tag{54}$$

2.4.6 Order Statistics of ALTSW-D

The order statistics of a random observation $X_{1,n} < X_{2,n} < ... < X_{n,n}$ from continues data with probability density function g(x) and cumulative distribution function G(x). The CDF and PDF of $X_{i,n}$ is defined as;

$$f_{(i,n)}(x) = \frac{n!}{(i-1)(n-i)} f(x)[F(x)]^{i-1}[1-F(x)]^{n-i} \tag{55}$$

Now using power series expansion:

$$[1-F(x)]^{n-1} = \sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}[F(x)]^j \tag{56}$$

It implies that equation (55) become:

$$f_{(i,n)}(x) = \frac{n!}{(i-1)(n-i)} \sum_{j=0}^{n-i} \binom{n-i}{j} (-1)^j f(x)[F(x)]^{i+j-1} \tag{57}$$

$$f_{(i,n)}(x) = \frac{n!(-1)^j}{(i-1)(n-i-j)} \sum_{j=0}^{n-i} \binom{n-i}{j} f(x)[F(x)]^{i+j-1} \tag{58}$$

$$f_{i,n}(x) = \sum_{j=0}^{n-i} \binom{n-i}{j} \frac{n!\bar{\alpha}\pi\lambda\beta(-1)^j x^{-\lambda x^\beta} Cos\left(\frac{\pi(1-e^{-\lambda x^\beta})}{2}\right)}{2(i-1)!(n-i-j)!j!\left(\alpha-\bar{\alpha}Sin\left(\frac{\pi(1-e^{-\lambda x^\beta})}{2}\right)\right)log(\alpha)} \times$$

$$\left[1 - \frac{log\left(\alpha - \bar{\alpha}Sin\left(\frac{\pi(1-e^{-\lambda x^\beta})}{2}\right)\right)}{log(\alpha)}\right]^{i+j-1} \tag{59}$$

2.5. Maximum Likelihood Estimation

Let $X_{1,}X_{2,}, ..., X_{n,}$ be a random sample of size "n" from ALTSW distribution then the likelihood function L is given as;

$$l = logl(x/\bar{\theta}) = log\prod_{i=1}^{n} f(x/\bar{\theta}) \tag{60}$$

$$l(\bar{\theta}) = \prod_{i=1}^{n} \frac{\bar{\alpha}\pi\lambda\beta x^{-\lambda x^\beta} Cos\left(\frac{\pi(1-e^{-\lambda x^\beta})}{2}\right)}{2(\alpha-\bar{\alpha})log(\alpha)Sin\left(\frac{\pi(1-e^{-\lambda x^\beta})}{2}\right)} \tag{61}$$

$$= nlog\pi + nlog\lambda + nlog\beta - nlog2 - nlog(log\alpha) - nlog\alpha + 2nlog\bar{\alpha} + \sum_{i=1}^{\infty} logx^{\beta-1} -$$

$$\lambda\sum_{i=1}^{\infty} x^\beta + \sum_{i=1}^{\infty} logCos\frac{\pi}{2}(1-e^{-\lambda x^\beta}) + \sum_{i=1}^{\infty} logSin\frac{\pi}{2}(1-e^{-\lambda x^\beta}) \tag{62}$$

$$\frac{\delta l}{\delta\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{\infty} \frac{\pi x^\beta e^{-\lambda x^\beta} Sin\frac{\pi}{2}(1-e^{-\lambda x^\beta})}{2Cos\frac{\pi}{2}(1-e^{-\lambda x^\beta})} - \sum_{i=1}^{\infty} \frac{\pi x^\beta e^{-\lambda x^\beta} Cos\frac{\pi}{2}(1-e^{-\lambda x^\beta})}{2Sin\frac{\pi}{2}(1-e^{-\lambda x^\beta})} \tag{63}$$

$$\frac{\delta l}{\delta\beta} = \frac{n}{\beta} + \sum_{i=1}^{\infty} logx - \lambda\beta\sum_{i=1}^{\infty} x^{\beta-1} - \sum_{i=1}^{\infty} \frac{\pi\lambda x^\beta e^{-\lambda x^\beta} e^{-\lambda x^\beta} logxSin\frac{\pi}{2}(1-e^{-\lambda x^\beta})}{2Cos\frac{\pi}{2}(1-e^{-\lambda x^\beta})} -$$

$$\sum_{i=1}^{\infty} \frac{\pi\lambda x^\beta e^{-\lambda x^\beta} e^{-\lambda x^\beta} logxCos\frac{\pi}{2}(1-e^{-\lambda x^\beta})}{2Sin\frac{\pi}{2}(1-e^{-\lambda x^\beta})} \tag{64}$$

$$\frac{\delta l}{\delta\alpha} = \frac{2n}{\bar{\alpha}} - \frac{n}{\alpha} - \frac{n}{(log\alpha)^2} \tag{65}$$

3. Results

This section will introduce the simulation study designed to examine the behavior and consistency of maximum likelihood estimation, as well as the application of real data to observe the performance of the proposed model and other competing models.

3.1. Simulation

The behavior of the maximum likelihood of ALTSW-D for certain parameter values in the first trial (i.e. $\alpha = 0.8$, $\lambda = 2.5$, $\omega = 3.4$, $\gamma = 4.5$) and second trial (i.e. $\alpha = 1.5$, $\lambda = 1.3$, $\omega = 3.8$,$\gamma = 2.8$) was investigated using a created finite sample of size n=20, 20, 100, 250, 500, and 1000. The random numbers for the ALTSW-D were generated using the quantile function. For 1000 repeats. The Means, Bias, and RMSE were then calculated. Table 1 presents the outcomes of the simulation. We concluded that the proposed model yields consistent results when predicting parameters for the mode based on the results of the Monte Carlo simulation.

Table 1: *Result of simulation for different values of parameters*

		$\alpha = 1.4$	$\beta = 1.1$	$\lambda = 0.7$	$\alpha = 0.8$	$\beta = 0.5$	$\lambda = 1.0$
n=20	Means	1.7016	1.2342	0.7151	0.9404	0.5764	1.0616
	Bias	1.7016	1.2342	0.7151	0.1404	0.0764	0.0616
	RMSE	1.7339	0.3272	0.2263	1.0495	0.1719	0.3239
n=50	Means	1.6060	1.1514	0.7073	0.9304	0.5284	1.0227
	Bias	0.2060	0.0514	0.0073	0.1304	0.0284	0.0227
	RMSE	1.2910	0.1794	0.1570	0.8533	0.0889	0.2198
n=100	Means	1.5276	1.1273	0.7015	0.8846	0.5154	1.0081
	Bias	0.1276	0.0273	0.0015	0.0846	0.0154	0.0081
	RMSE	0.9281	0.1224	0.1174	0.6152	0.0605	0.1597
n=250	Means	1.5386	1.1091	0.7044	0.9043	0.5056	1.0287
	Bias	0.1386	0.0091	0.0044	0.1043	0.0056	0.0087
	RMSE	0.7772	0.0809	0.0905	0.5350	0.0419	0.1331
n=500	Means	1.4982	1.1051	0.7013	0.8671	0.5024	1.0041
	Bias	0.0982	0.0051	0.0013	0.0671	0.0024	0.0041
	RMSE	0.6628	0.0627	0.0804	0.3994	0.0311	0.1029
n=1000	Means	1.4913	1.1000	0.7054	0.8476	0.5008	1.0051
	Bias	0.0913	0.0000	0.0054	0.0476	0.0008	0.00513
	RMSE	0.5219	0.0483	0.0640	0.3058	0.0238	0.0810

3.2. Application 1

This data set consists of the total milk output from the first calves of 107 SINDI race cows. The data was used in (15). The data set is included0.729, 0.2361, 0.514, 0.5232, 0.4438, 0.4143, 0.3821, 0.0168, 0.7261,0.4151, 0.5529, 0.535, 0.2303, 0.5744, 0.686, 0.6789, 0.675, 0.2356,0.5285, 0.6844, 0.4564, 0.5483, 0.7471, 0.48, 0.6196, 0.6012, 0.7804,0.065, 0.577, 0.4049, 0.6174, 0.3906, 0.5349, 0.4553, 0.4694, 0.3175,0.4823, 0.1479, 0.5853, 0.4371, 0.3188, 0.499, 0.3945, 0.5481,0.8147, 0.6488, 0.348, 0.4612, 0.7131, 0.5707, 0.3635, 0.5113,0.0671, 0.4741, 0.2681, 0.3627, 0.4517, 0.426, 0.6114, 0.6058,0.5878, 0.4365, 0.6907, 0.7629, 0.5912, 0.6927, 0.4576, 0.6707,0.7687, 0.4332, 0.5553, 0.1546, 0.8492, 0.3383, 0.3751, 0.3134,0.6768, 0.1525, 0.1167, 0.447, 0.5394, 0.0609, 0.515, 0.4752, 0.3323, 0.1131, 0.453, 0.3598, 0.622, 0.8781, 0.216, 0.4111, 0.5629, 0.3259,0.2747, 0.5627, 0.3406, 0.0776, 0.5447, 0.5941, 0.2605, 0.4675,0.3891, 0.6891, 0.0854, 0.3413, 0.6465

Table 2: *Information criteria measure for the fitted models using milk output data*

Distribution	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\gamma}$	$\hat{LL}$	AIC	CAIC	BIC	HQIC
WEW	100.00	10.03	0.19	71.59	82.58	173.15	174.29	179.91	175.59
BW	0.13	27.80	19.34	10.51	77.85	163.69	164.84	170.46	166.14
GEW	18.64	2.09	19.11	100.23	96.34	200.67	201.81	207.43	203.11
KwW	0.48	70.75	7.97	20.89	82.38	172.76	173.89	179.51	175.19
ALSW	1.00e+02	6.60e-03	2.81e+00	—	79.96	165.93	166.59	170.99	167.76

Table 3: *Goodness of fit measure for the fitted models using milk output data*

Distribution	KS	A	W	P-value
WEW	0.1085	0.5889	0.0793	0.7346
BW	0.0671	0.1010	0.0135	0.9937
GEW	0.1777	2.7989	0.4717	0.16
KwW	0.1078	0.5598	0.0750	0.7407
ALSW	0.0824	0.2391	0.0296	0.9485

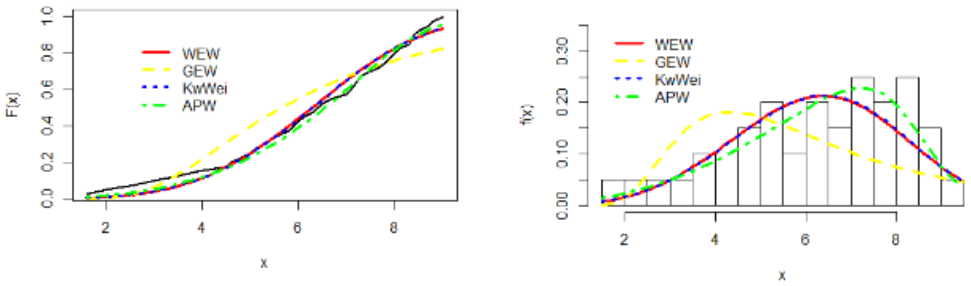


Figure 3: *Fitted PDF and CDF from first data sets*

3.3. Application 2

The first data set contains the COVID-19 mortality rate associated with Saudi Arabia during the time between July 6, 2021 and September 4, 2021. The data set contains 61 observations, and it has been used by (16). The data are as follows: 0.3086, 0.3283, 0.2865, 0.2450, 0.2852, 0.3251, 0.2636, 0.3236, 0.2824, 0.2817, 0.3012, 0.2603, 0.2997, 0.2393, 0.2785, 0.2778, 0.2375, 0.2963, 0.2167, 0.2752, 0.2353, 0.2347, 0.1951, 0.2140, 0.2329, 0.2711, 0.2126, 0.2314, 0.1924, 0.2113, 0.2683, 0.2487, 0.2674, 0.1716, 0.2666, 0.2091, 0.2278, 0.1706, 0.2271, 0.1890, 0.2077, 0.2452, 0.1319, 0.2259, 0.1504, 0.1879, 0.1689, 0.2063, 0.2249, 0.1686, 0.1310, 0.1497, 0.1309, 0.1495, 0.1121, 0.1121, 0.1307, 0.1120, 0.1306, 0.1492, 0.0932.

Table 4: *Information criteria measure for the fitted models using Covid-19 mortality data*

Distribution	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\gamma}$	$\hat{LL}$	AIC	CAIC	BIC	HQIC
WEW	100.00	9.45	0.12	79.92	-20.98	-33.95	-33.56	-23.26	-29.62
BW	-47.43	-47.04	-36.74	-43.09	-27.76	0.24	103.01	7.06	1.42
GEW	32.62	33.01	43.31	36.95	12.3076	21.42	0.06	0.61	78.93
KwW	0.58	78.31	4.45	2.84	-21.5449	-35.09	-34.69	-24.39	-30.76
ALSW	100.00	5.04	1.82	—	79.96	-48.93	-48.69	-40.91	-45.67

Table 5: *Goodness of fit measure for the fitted models using covid-19 mortality data*

Distribution	KS	A	W	P-value
WEW	0.0852	1.5891	0.2426	0.4184
BW	0.0832	0.5244	0.0829	0.4493
GEW	0.1872	7.2646	1.2791	0.0011
KwW	0.0834	1.4961	0.2276	0.4456
ALSW	0.0597	0.4988	0.0725	0.8396

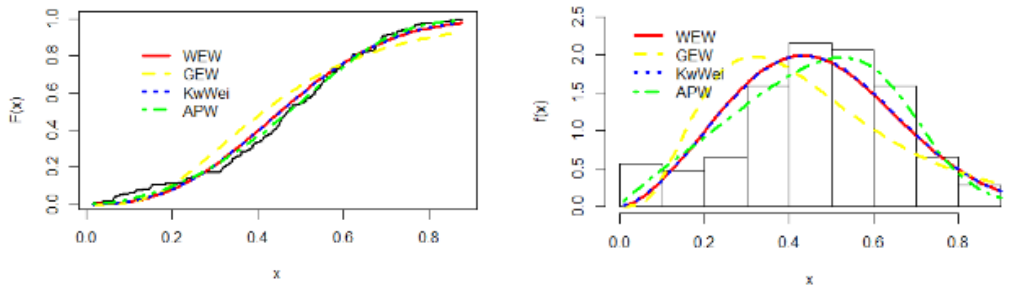


Figure 4: *Fitted PDF and CDF from second data sets*

3.4. Discussion

The analysis of the alpha logarithm transform sine Weibull distribution demonstrated its efficiency in modeling real-life data. This new distribution outperformed its competitors, as indicated by smaller values of information criteria (2), as shown in Tables 2 and 6. The simulation results in 1 indicated that the model's performance improves as the sample size increases in both the first and second trials. The fitted CDF and PDF plots from the first and second applications illustrate the flexibility of the proposed distribution compared to existing distributions.

Conclusion

This paper introduces a new distribution called the alpha power logarithm transform sine Weibull distribution. The mathematical properties were derived, and the parameter estimation of the new distribution was examined using the maximum likelihood method. The behavior of the maximum likelihood estimation was also investigated to assess consistency. In conclusion, we suggest that the proposed distribution performs better than existing distributions in terms of information criteria and goodness-of-fit tests.

Author's contribution

Abdulhameed Ado Osi: R-Coding and analysis of models performance
Olawoyin Olatunji Ishaq: Conceptualization and review
Bilal Alhassan Abdulsalam: typing and compilation
Usman Abubakar: Latex coding, manipulation, and results imputation.

Acknowledgment

The authors are grateful to the Probability and Stochastic research team and the Department of Statistics at Aliko Dangote University of Science and Technology for their contributions to improving the quality of this manuscript.

Conflict of interest:

The authors did not encounter any conflicts of interest during this work.

Ethical Consideration:

No ethical clearance is required before the commencement of this research.

Data availability statement:

The data used in this study are available in the manuscript.

References

[1] Abubakar, U., Osi, A. A., Shuaibu, A., & Salisu, L. A. (2024). SIMULATIONS AND BAYESIAN ESTIMATION OF TRUNCATED EXPONENTIAL LOG-TOPP-LEONE GENERALIZED FAMILY WITH APPLICATION TO SURVIVAL TIME DATA. *Reliability: Theory & Applications*, 19(4 (80)), 176-189.

[2] Aryal, G. R., Ortega, E. M., Hamedani, G., and Yousof, H. M. (2017). The topp- leone generated weibull distribution: regression model, characteriza- tions and applications. *International Journal of Statistics and Probability*, 6(1):126“141

[3] A. A. Suleiman, H. Daud, A. I. Ishaq, M. Othman, R. Sokkalingam, A. Usman, and A. A. Osi, The odd beta prime inverted kumaraswamy distribution with application to covid-19 mortality rate in italy,´ *Engineering Proceedings*, vol. 56, no. 1, p. 218, 2023.

[4] Z. Ahmad, E. Mahmoudi, R. Roozegarz, G. Hamedani, and N. S. Butt, Contributions towards new families of distributions: An investigation, further developments, characterizations and comparative study,´ *Pakistan Journal of Statistics and Operation Research*, pp. 99“120, 2022.

[5] Ahsanullah, M. (1991). Record values of the lomax distribution. Statistica Neer- landica, 45(1):21“29

[6] K. A. Al-Kadim and B. D. Mohammed, Rayleigh pareto distribution,´ *Journal of University of Babylon for Pure and Applied Sciences*, vol. 26, no. 1, pp. 84“95, 2018.

[7] U. Abubakar, A. A. Osi, I. A. Salisu, H. Muhammad, Y. I. Muhammad, and A. Abubakar, Arcsine rayliegh pareto distribution: Properties and application to carbon fibers data sets,´ *FUDMA JOURNAL OF SCIENCES*, vol. 8, no. 2, pp. 301“305, 2024.

[8] M. U. Faruk, A. M. Isa, and A. Kaigama, Sine-weibull distribution: Mathematical properties and application to real datasets,´ *Reliability: Theory & Applications*, vol. 19, no. 1 (77), pp. 65“72, 2024

[9] G. M. Cordeiro, M. d. C. S. Lima, A. E. Gomes, C. Q. da Silva, and E. M. Ortega,The gamma extended weibull distribution,´ *Journal of Statistical Distributions and Applications*, vol. 3, pp. 1“19, 2016.

[10] A. S. Alghamdi and M. Abd El-Raouf, A new alpha power cosine-weibull model with applications to hydrological and engineering data,´ *Mathematics*, vol. 11, no. 3, p. 673, 2023.

[11] W. Weibull, A statistical distribution function of wide applicability,´ *Journal of applied mechanics*, 1951.

[12] M. Fr©chet, Sur la loi de probabilit© de l™©cart maximum,´ *Ann. de la Soc. Polonaise de Math.*, 1927.

[13] P. Rosin, The laws governing the fineness of powdered coal,´ *J. Inst. Fuel*, vol. 7, no. 31, pp. 29“36, 1933.

[14] M. E. Mead, G. M. Cordeiro, A. Z. Afify, and H. Al Mofleh, The alpha power transformation family: properties and applications,´ *Pakistan Journal of Statistics and Operation Research*, pp. 525“545, 2019.

[15] H. Haj Ahmad, E. M. Almetwally, M. Elgarhy, and D. A. Ramadan, On unit exponential pareto distribution for modeling the recovery rate of covid-19,´ *Processes*, vol. 11, no. 1, p. 232, 2023.

[16] J. Bengalath and B. Punathumparambath, Harris extended inverted kumaraswamy distribution: Properties and applications to covid-19 data,´ *International Journal of Data Science and Analytics*, pp. 1“23, 2024

[17] Weibull, W. (1951). A statistical distribution function of wide applicability. *Jour- nal of applied mechanics*

[18] Faruk, M. U., Isa, A. M., & Kaigama, A. (2024). SINE-WEIBULL DISTRIBUTION: MATHEMATICAL PROPERTIES AND APPLICATION TO REAL DATASETS. *Reliability: Theory & Applications*, 19(1 (77)), 65-72.

[19] Tuoyo, D. O., Opone, F. C., & Ekhosuehi, N. (2021). The Topp-Leone Weibull distribution: its properties and application. *Earthline Journal of Mathematical Sciences*, 7(2), 381-401.

[20] Benkhelifa, L. (2022). Alpha power Topp-Leone Weibull distribution: Properties, characterizations, regression modeling and applications. *Journal of Statistics and Management Systems*, 25(8), 1945-1970.

A BATHTUB ADDITIVE FAILURE RATE MODEL FOR ACTUARIAL RELIABILITY

M. A. Alhassan[1]*, A. Yahaya[2], O. O. Ishaq[1,3], B. Abba[4,7], Y. Hussaini[1,5], A. Bello[1,6]

•

[1]Department of Statistics, Aliko Dangote University of Science and Technology, Wudil, Kano, Nigeria
[2]Department of Statistics, Ahmadu Bello University, Zaria, Kaduna, Nigeria
[4]Department of Mathematics, Yusuf Maitama Sule University, Kano, Nigeria
[7]School of Mathematics and Statistics, Central South University, Changsha, China
[1]*ma23038@gmail.com, [2]Abubakaryahaya@abu.edu.ng, [3]babington4u@gmail.com, [4]babba@yumsuk.edu.ng, [5]alhasan.ma@outlook.com, [6]mrwudil@gmail.com

Abstract

The actuarial approach to hardware reliability plays an important role in providing balanced avenue and fair play in dealing with multi-unit systems and dual failures occurring in such complex settings. In statistical reliability perspective, the additive failure rate methodology is believed to adequately adapt to the actuarial approach for effective system reliability. In line with this hypothetical subject, this paper introduces a lifetime model named Hybrid Weibull-Exponential Power model, based on the later phenomenon vis-à-vis the former approach. The failure rate (FR) of the proposed model displayed suitable compound trends such as increasing-decreasing and various complex forms of bathtub-curve, with and without the lingering useful life period. The consistency of its quantile function in producing effective samples was validated by simulation study, despite its nonclosed form. Other reliability properties of the model were also provided. The paper employs the maximum likelihood estimation procedure for estimating the unknown parameters of model, due to its reputable robustness in the aspect of applied statistics. The adequacy of the proposed estimator was evaluated by means of simulation study, where, negligible biases and declining MSEs were observed. The real-life performance of the proposed model in relation to three similar-structured models was weighed using bathtub failure times, where, least values of information criteria and p-values of goodness-of-fit statistics were obtained for the model. Highest loglikelihood value of the model was also observed. The additive failure rate model proposed in this paper allows for actuarial reliability analysis of compound systems, due to its outstanding performance on the dataset. The MLEs used for estimating the model parameters were proven consistent since insignificant biases and declining MSEs were obtained in all cases. Also, on the premise of the results obtained for the information criteria, the p-values for the goodness-of-fit statistics and the loglikelihood value, we infer that the new established model can better be employed over the compared models, for the dataset. Subsequently, effective system reliability with similar datasets can be achieved using the model. Succinct guide on the recent developments of the emerging family of additive FR models and their application in multi-unit system modelling, is also given.

Keywords: actuarial approach, statistical reliability, hybrid, bathtub failure rate, maximum likelihood estimation, multi-unit systems, simulation study

I. Introduction

In the industry of determining the lifetime behavior of engineered units, relevant information, such as the key phases gone through by a unit during its serviceable life, are often analyzed by reliability scientists via appropriate theoretical tools. The tools, predominantly, include classical probability models such as normal, lognormal, exponential, gamma, Rayleigh and, notably, the Weibull model [38, 40, 42]. The vital information derived, help in making essential decisions such as device warranty, optimal mode and time of replacement/or maintenance and so on.

However, with the persistent emergence of modern systems, more complex data with nonmonotonic FR behavior like the bathtub also keep recurring. Consequently, tentative decisions are believed to be made as a result of truncated information derived by employing the classical distributions for the analyses. This is because, majority of the familiar models fall flat to adequately capture non-monotone failure times (FTs), including the legend bathtub feature [39]. For example, as in the case of a single Weibull model [50], hence, leading to undesirable conclusions. The bathtub FR titled considering its form, consists of the three known consecutive stages, hypothetically shaping into a bathtub manner. First, the burn-in period, with usual high failure tendency, due to unidentified imperfections along the unit's construction process that can lead to infant mortality. Upon surviving this phase, the FR decreases and balances for some time in the second phase. This is the chance failure period or simply, useful life of the unit. As the unit is subjected to aging, the barely increasing FR begin to heighten. Thus, activates the third, wear-out period, which concludes the life cycle of the unit. Refer to Rausand and Holyland [40]. Bathtub FTs are vastly discussed in the reliability literature. These include the 36 times to first failure of 500MW generators [17]; the failure and running times of sample of 30 units [31]; 18 FTs of electronic devices [48], the early FTs of cable joints [46], the FTs of 24 DRP systems [32], the rates of failure and survivors for dielectrically insulated armature bars of a generator [18], to mention few.

To address the under-fitting problem of the classical models in representing complex FTs like the bathtub, many advanced models with desirable features were put forward by reliability statisticians to serve as better tools for which purpose classical models were used, thus, incorporating the traditional models' attributes. For instance, Lai et al. [21], proposed the Modified Weibull model for explaining bathtub FRs. It generalizes the Weibull and type I extreme value models respectively. Carrasco et al. [10], defined 4-parameter Generalized Modified Weibull model for lifetime modeling of monotone and non-monotone FRs. It extends the modified Weibull model and further having the generalized Rayleigh and extreme value distributions as special cases. Haupt and Schäbe [20], proposed a bathtub-curved FR model, however, infers not on the behavior of the infant mortality of the FR. Further, it offers not much workability to be adopted [10]. Sylwia [45] introduced a shape parameter to the Makeham's distribution which was strictly increasing. Thus, made its FRF shaping into the bathtub, hence, named it Makeham's modified distribution (MMD). Although, the author iterated on the MMD's applicability in burn-in procedure planning and preventive maintenance of non-repairable appliances, the MMD represents not well the constant region of the bathtub curve. Additionally, Alessa et al. [7] considered the bathtub FR characterization of two sets of FTs in [1] and [31] respectively, to examine the hypothetical contrast between their proposed modified exponential-Weibull (MEW) model and five other Weibull extensions. Subsequently, the MEW model was found enhanced over the rest of the models in terms of amply explaining the FTs.

Another important distribution in this regard, related to Weibull is proposed by Chen [12], with two parameters. However, inadequate in different scenarios, as it lacks a scale parameter. So, as an extension to [12], Xie *et al*. [22], proposed the Weibull extension model by adding a scale parameter to the Chen [12] model, thus, becomes flexible and plausible to practitioners. Correspondingly, Chaubey and Zang [11] extended the Chen [12] models with increasing or

bathtub FRF. Similarly, various proposals focusing on combining the Chen [12] model or its extensions, with other classical models, more so, those of the Weibull family, have been put forward for characterizing the bathtub FTs. For instance, the Additive Chen-Weibull (ACW) model introduced by Thach and Bris [47] to describe bathtub FTs, it was tested on the two bench mark FTs [1, 31], in relation to the additive Weibull (AddW) and other Weibull-modified models. The results showed that the ACW provides better fit for the FTs in [1], while the New Modified Weibull model showed better fit for the FTs in [31], seconded by the ACW and AMW models respectively. Mendez-Gonzalez *et al.* [35] proposed and related the Chen-Perks (CP) distribution to several other models including AddW, based on three case studies reported by Wang *et al.* [48], Aarset [1] and Sylwia [45] respectively. The models performed well in fitting the FTs of [45], while, the CP outperformed the rest of the models for representing the FTs in [1]. However, many AFR models fail to fit bathtub FTs well, perhaps due to the role of the Chen model in the constituting models. Although it can have bathtub FR, but may provide not a fairly good representation of the bathtub curve, this gap may be attributed to it lack of scale parameter, posed Abba *et al*. [3].

One of the very rare two parameter lifetime models with possible bathtub FRs for complex system reliability, is the Exponential-power model. It was introduced, with Weibull type parameterization, by Smith and Bain [44], for modelling wear-out failures. It has U-shaped form of bathtub FR [26, 29]. Moreover, Choy and Walker [13] extended the family while Barriga *et al*. [9] annexed a shape parameter to the model and proposed a complementary version, to handle variety of non-monotone FRs. Abba *et al.* [4] came-up with flexible exponential power-Gompertz (FEPG4) by methodizing the two component models for complex reliability modelling.

- Relevance of Additive Failure Rate models in Actuarial reliability

The actuarial approach to hardware reliability entails the analysis of systems of several units (say $\mathcal{u}_i, i = 1, \dots, n$). Whereby, the strength of the units and the entire data regarding the loads of operation, all combined, are described by the distributional probability, $\mathbb{G}(t)$, of the time-to-failure, T [40]. Although modelling the strength and the load is implicit, however, reliability function, FR and other reliability properties of the component units can be determined via the distributional probability, $\mathbb{G}(t)$. Where appropriate, the relations, $R(t) = 1 - \mathbb{G}(t)$ and $\hbar(t) = \frac{\mathbb{g}(t)}{R(t)}$ can simply be adopted, where, $R(t)$, $\mathbb{g}(t)$ and $\hbar(t)$ are defined in Table 1.

In recent times, statistical reliabilists offer considerable score of advanced models which are largely hybrid in structure, for modelling the bathtub and other complex FTs. These models are mostly formulated based on multi-unit systems assumption [2, 3, 41, 47, 50, etc.], i.e several components assembled in a system (mostly in series), in order to function in one direction for achieving one goal. Thus, points to their potential appropriateness for effective actuarial reliability analysis of hardware systems. In this direction, the additive-wise methodology of linear combination of two or more FRs of certain probability models, with the intent of reliability studies is gaining more attention in reliability theory. Admittedly, researchers devoted greatly towards introducing more distributions based on the additive phenomenon, particularly in the last couple of years, see for instance, [3], [5], [6], [47] and the references within. These hybrid models are usually termed as additive FR (AFR) models (for the compositional manner of their FRFs). They are formulated through their FRFs, which is obtained as a summation of two or more FRFs or reliability functions (RFn) of desirable classical probability distributions. The resultant compound models form another family of probability distributions, with members such as, the AddW of Xie and Lai [50], exponential-log logistic additive FR model of Rosaiah *et al.* [41], the AFR models derived by Abid and Hassan [6];; additive Perks (AddP) of Méndez-González *et al*. [33];; flexible exponential-power Gompertz of Abba *et al*. [4]; additive Gompertz-Weibull of Abba and Wang [2]; additive Chen (AddC) of Méndez-González *et al*. [34]; flexible Dhillon-Weibull Competing Risk model of Abba *et al.* [5], etc. These models are also known as multi-risk or competing risk distributions, due to their multi-risk representation ability. The distributions have lately become increasingly appealing to studies, for their versatility in modelling bathtub and other complex FTs.

Albeit, the AFR methodology can be regarded as old approach, see Refs. [16], [36] and [19]. However, its resurgence with more applications in reliability and survival analysis can be traced back nearly three decades after the establishment of Additive Weibull model [50]. In motivation of this additive methodology, Rosaiah *et al.* [41] and, Abid and Hassan [6] explained the fact that the reliability probabilities of such compound systems are either evaluated by the system's failure mode as a whole, such as

$$R(t) = \exp\left\{-\int_0^t \hbar(u)du\right\}, \tag{1}$$

or via the independent reliability probabilities of the units that define the whole system using their respective FRs. Formally described in section II. The AFR distributions are mostly applied to the modelling of complicated FTs with two or more failure mechanisms, or to the system reliability with serially independent subunits. A series system consists of serially installed units (say, 1, 2, …, r). Thus, subsequent units are considered a failure by the failure of the first unit in the installation line. Its brief illustration is given below.



Figure 1: *An instance of a series system*

Notwithstanding, great portion of the modern hybrid models presented as sensitive to the bathtub FTs, fail at some point, to represent effectively such FTs. Refer, for instance, to Abba and Wang [2], on the inability of notable bathtub distributions in representing the bathtub-curved times of early failures of cable joints [46]. The more realistic ones were able to fairly characterize the bathtub curve, however, overlook the worthwhile lingering life stage. For instance, the flexible additive Weibull model [25] and Xgamma-Burr XII [37] have not been able to sufficiently contain the constant and/or end-of-life regions of some bathtub FTs. Thus, the process of estimating and representing devices' FTs through these models may be further from the reality. Notwithstanding, a handful of distributions [43], [33] and [5], put forward in this regard perform satisfactorily well. However, these models may perhaps be outnumbered by the unprecedented overflow of the FR problems involving some variants of the bathtub and other non-monotone trends.

Thus, given above importunate concern, particularly, regarding the AFR models pointed out earlier constituting of Chen as submodel, specifically, the ACW that motivates this study. For the nonexistence of scale parameter in the Chen distribution, these models may not be well fit for properly monitoring the spread of bathtub datasets - where, the value of the scale parameter controls the dispersion a distribution. Hitherto, the quest for a much more satisfying lifetime model in this regard remains incessant. Also, given the importance of the actuarial approach to system reliability, there is need for simplified methodologies in its regard.

Accordingly, considering the AFR criterion, this study aims to establish a FT model for describing monotone and non-monotone FTs of non-dependent dual failures. It is also, purposeful, giving regard to the multi-unit assumption of the actuarial reliability analysis of series systems of independent sub-units. The proposed Hybrid Weibull Exponential-Power (HWEP) model can be deemed an alternate modification to the ACW model by swapping the Exponential-power for the Chen model accordingly.

For achieving this study, we adhered to these key motivations, (i) proposing an AFR model with the ample capability to explain the complex FTs of technical units based on actuarial phenomenon, (ii) running Monte Carlo trials to assess the consistency of the quantile function of the model, (iv) estimating the parameters of the model using MLE technique, and also, checking the consistency of the MLEs (v) testing the applicability of the model in real-life on bathtub FT data and evaluating its competitive performance using the information criteria.

Subsequent to the overview and pertinence put forward in section I, the research hierarchy goes as this: section II explains the AFR method for compounding statistical models for system reliability. Swift discussion on the parent models used as the baseline for the constructing the new model. Then, the proposed model was introduced with some of its reliability attributes, including its quantile function. The symbolizations utilized in the whole text are interpreted in this part as well. Results of the simulation studies and reliability relevance testing for the model were presented in section III. Section IV contains general discussions and section V concludes the study.

II. Methods

Succinct explanation on the methodology adopted, the underlying statistical models utilized as the baseline and the proposed model are all given in this section.

I. Additive Failure Rate Approach

Suppose that, random variables $T_{i;\,2\le i\le k}$ are the FTs of finite independent subunits of a system. Alternatively, set the random variables $T_{i;\,2\le i\le k}$ as FTs of multiple independent failure modes (FMs) suffered by a system. In both scenarios, suppose also that each of the FTs follow a specific probability model with FRF, $h_i(t)$ and RFn, $R_i(t)$. Then, by the AFR approach discussed earlier, we build a hybrid statistical model for system reliability via the sum of baseline models' FRF as:

$$h(t) = \sum_i h_i(t)\,, i = 2,..,k. \tag{2}$$

Then again, the product of the underlying models' RFn can be considered as:

$$R(t) = \prod_i R_i(t) = \exp\left\{-\sum_i \int_0^t h_i(u)du\right\}, i = 2,..,k. \tag{3}$$

Refer also to Murthy *et al.* [38], Thach and Bris [47], Mohammad *et al.* [37], Abba *et al.* [5], etc.

II. The Baseline Models

Concise introductions based on this study's portion of interest regarding the Weibull as well as exponential-power models are given accordingly as follows.

- The Weibull model

The well-known Weibull distribution by Weibull [48] can have its FRF and RFn respectively as,

$$h(t;\alpha,\gamma) = \alpha\gamma^{\alpha}\,t^{\alpha-1}, t > 0. \tag{4}$$

$$R(t;\alpha,\gamma) = e^{-(\gamma t)^{\alpha}}, t \ge 0;\ \alpha,\gamma > 0. \tag{5}$$

However, its FRF can only discretely assume increasing, decreasing or constant behavior [23, 47]. In consequence, numerous studies considered remodeling the distribution. Refer to Johnson *et al*. [23] for more elaboration. Also, Murthy *et al*. [38] gave wider, whereas, Lai and Xie [27] offered brief details on the notable developments of the Weibull models due then.

- The Exponential-power model

The exponential-power model is a life-casting model first studied by Smith and Bain [44] with the possibility of a bathtub or U-shaped FRF at its shape's value, $\tau < 1$ [26, 29]. Thus, making it one of the very rare natural two-parameter models with such attribute. The model is useful in case of reliability modeling of a unit that fails quickly during its wear-out phase, after being stable for some period [44]. It has its FRF and RFn respectively as:

$$h(t;\tau,\nu) = \tau\nu^{\tau}t^{\tau-1}e^{(\nu t)^{\tau}}, \tag{6}$$

$$R(t;\tau,\nu) = e^{1-e^{(\nu t)^{\tau}}}, t > 0;\ \tau\nu > 0. \tag{7}$$

Notable developments of the exponential-power model were given in section I.

- Symbolizations

Across this research work, we adhered to these symbolizations below.

Table 1: *Symbolizations utilized in the study*

$\hbar$	Number of units subjected to life test, also, number of FMs
T_i	i^{th} unit uptime until failure, $1 \leq i \leq \hbar$
$\mathbb{G}(\cdot)$	Cumulative distribution function (CDF) of T_i
$\mathbb{g}(\cdot)$	Probability density function (PDF) of T_i
$\hbar(\cdot)$	$\frac{\mathbb{g}(\cdot)}{R(\cdot)}$ FRF of T_i
$R(\cdot)$	$= 1 - \mathbb{G}(\cdot)$, reliability function (RFn) of T_i
$W(\alpha, \gamma)$	Weibull distribution with α, γ as parameters
$EP(\nu, \tau)$	Exponential power model with ν, τ as parameters

III. The Proposed HWEP Model and its Properties

This subsection describes the proposed model, its reliability and other distributional properties.

- Model specification

Define $T_i, (i = 1,2)$ as FTs resulting from $(\hbar = 2)$ independent forms of failure suffered by either of two sub-units of a system at a time. Next, set T as the overall FT of the system which can be interpreted as the first unit-specific time of failure arrived by the system. Otherwise, T may be taken as the first mode-specific failure arrived by a unit. Thus, $T = min\{T_1, T_2\}$. Further, assume that $T_1 \sim W(\alpha, \gamma)$ with FRF and RFn in Eq. (4) and Eq. (5) while, $T_2 \sim EP(\nu, \tau)$ with FRF and RFn in Eq. (6) and Eq. (7), respectively. Where W and EP stand as corresponding acronyms for Weibull and exponential-power models. That is to say, the proposed model, termed as Hybrid Weibull-Exponential Power (HWEP), describes the FTs or FMs of the subunits. In either case, the components' FRs are assumed to abide by Eq. (4) and Eq. (6) respectively. Thus, reference to Eq. (2), if $T \sim HWEP(\alpha, \gamma, \nu, \tau)$, then, we obtain the corresponding FRF, $\hbar(t; \alpha, \gamma, \nu, \tau)$ of T as:

$$\begin{aligned} \hbar(t; \alpha, \gamma, \nu, \tau) &= \hbar_W(t; \alpha, \gamma) + \hbar_{EP}(t; \tau, \nu), \\ &= \alpha\gamma^{\alpha}\, t^{\alpha-1} + \tau\nu^{\tau} t^{\tau-1} e^{(\nu t)^{\tau}}, \quad t > 0, \end{aligned} \tag{8}$$

with $\boldsymbol{\xi} = (\alpha, \gamma, \nu, \tau)'$ denoting the HWEP's vector of parameters, among which $\alpha, \tau > 0$ and also $\gamma, \nu \geq 0$, stand dually as corresponding shape and scale.

The probability of a unit to execute its set function for a quoted time interval under usual (or preset) settings, determines its reliability [38]. The corresponding RFn, $R(t)$, is defined as:

$$R(t) = e^{-H(t)}, \tag{9}$$

where, $H(t)$, is the cumulative FR. It is obtained from Eq. (8) as:

$$H(t) = \int_0^t \hbar(u)du = (\gamma t)^{\alpha} + e^{(\nu t)^{\tau}} - 1. \tag{10}$$

Thus, putting Eq. (10) into Eq. (9) yields the following:

$$R(t) = \exp\left\{-(\gamma t)^{\alpha} - e^{(\nu t)^{\tau}} + 1\right\}, t > 0. \tag{11}$$

Equally, Eq. (11) can easily be attained via Eq. (3) using Eq. (5) and Eq. (7) respectively. Moreover, the associated CDF, $\mathbb{G}(t)$, of T is obtained from Eq. (11) as:

$$\mathbb{G}(t) = 1 - \exp\left\{-(\gamma t)^{\alpha} - e^{(\nu t)^{\tau}} + 1\right\}, \quad t > 0. \tag{12}$$

While, by differentiating the CDF in Eq. (12), with respect to t, we obtain the PDF, $\mathbb{g}(t)$, as:

$$\mathbb{g}(t) = \left[\alpha\gamma^{\alpha}\, t^{\alpha-1} + \tau\nu(\nu t)^{\tau-1} e^{(\nu\, t)^{\tau}}\right] \exp\left\{-\left[(\gamma t)^{\alpha} + e^{(\nu t)^{\tau}} - 1\right]\right\}, t > 0. \tag{13}$$

- FR attributes of the HWEP model

In reliability, the FR of a unit (say T), is the probability of failure at some time t, conditional on the unit's survival until then [28, 8]. Expressed mathematically as $Pr(t < T \leq t + \Delta t | T > t)$. It is quite significant; it explicitly interprets the quantity of system's failure within some time frame.

The FR of the proposed HWEP model Figure 2, exhibits monotonously decreasing and increasing trends, as shown in panel (a) and (b) respectively. It also possesses variety of bathtub shapes as in panel (c) and (d). Notably, the much-needed shape of the bathtub with the extended functional life span is uniquely observed in (c), for relatively higher-ordered values of α, and τ between 0.35 and 0.75. The FR sometimes reveal an increasing-decreasing, V-pattern, as seen in (b).

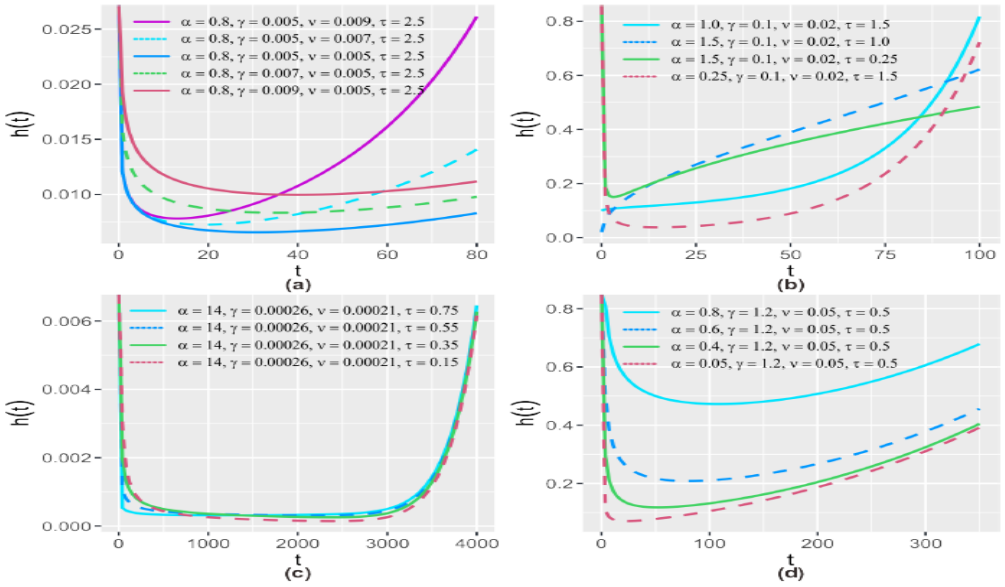


Figure 2: *Different FR Curves of the proposed HWEP model.*

- Mean Residual Life (MRL) and Mean Time-to-Failure (MTTF)

Let $X \sim HWEP$, then, at time t, the conditional statement $X - t|X > t$, gives the $X's$ residual life. Thus, MRL is the FT expected, of a system which has lived through some time t, expressed as:

$$M(t) = \mathbb{E}[X - t|X > t] = \frac{1}{R(t)} \int_0^{+\infty} R(x+t)dx,$$

$$= \frac{1}{R(t)} \int_0^{+\infty} e^{-(\gamma(x+t))^{\alpha} - e^{(\nu(x+t))^{\tau}} + 1} dx,$$

By series expansion of $e^{-e^{(\nu(x+t))^{\tau}}}$, we have

$$= \frac{1}{R(t)} \sum_{k,\ell=0}^{+\infty} \frac{(-1)^k k^{\ell} (\nu)^{\ell\tau} e^1}{k!\,\ell!} \int_0^{+\infty} (x+t)^{\ell\tau} e^{-(\gamma(x+t))^{\alpha}} dx. \tag{14}$$

We define $y = (\gamma(x+t))^{\alpha}$, so that, $(x+t)^{\ell\tau+1} = \gamma^{-(\ell\tau+1)}((\gamma(x+t))^{\alpha})^{\frac{\ell\tau+1}{\alpha}}$. Thus,

$$M(t) = \frac{1}{\alpha R(t)} \sum_{k,\ell=0}^{+\infty} \frac{(-1)^k (k)^{\ell} (\nu)^{\ell\tau} e^1}{k!\,\ell!\,\gamma^{\ell\tau+1}} \Gamma\left(\frac{\ell\tau+1}{\alpha}\right). \tag{15}$$

Where, $\Gamma(.)$, represents the gamma incomplete function, $\int_0^{+\infty} y^{\frac{\ell\tau+1}{\alpha}-1} e^{-y} dy$.

The reciprocal relationship between the FR and MRL functions can be observed in the curves in Figure 3 (a) and (b) respectively, under the same parameter settings.

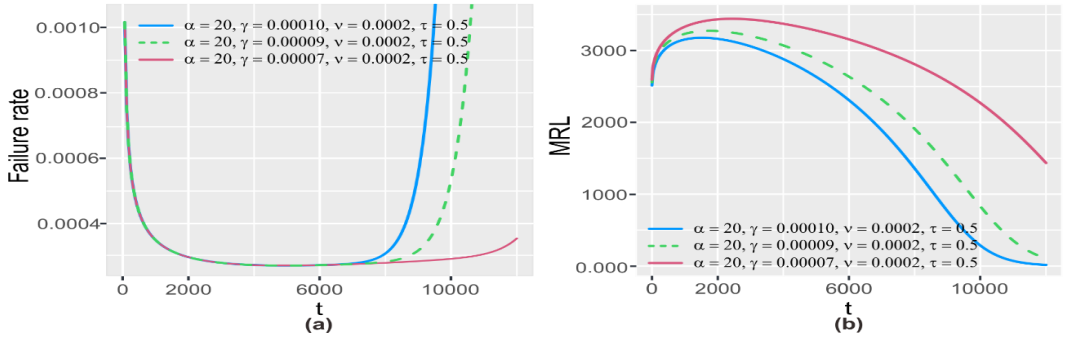


Figure 3: *FR and MRL curves of the proposed HWEP model*

Next, MTTF refers to the mean operational time of a system ahead of its failure. For HWEP unit, it is given as:

$$MTTF = \mathbb{E}[T] = \int_0^{+\infty} R(t)dt = \int_0^{+\infty} e^{-(\gamma t)^{\alpha} - e^{(\nu t)^{\tau}} + 1} dt.$$

Reference to Eq. (14)-(15), take $y = (\gamma t)^{\alpha}$ and $t^{\ell\tau+1} = \gamma^{-(\ell\tau+1)}((\gamma t)^{\alpha})^{\frac{\ell\tau+1}{\alpha}}$. Thus,

$$= \frac{1}{\alpha} \sum_{\hbar,\ell=0}^{+\infty} \frac{(-1)^{\hbar}(\hbar)^{\ell}(\nu)^{\ell\tau} e^{1}}{\hbar!\, \ell!\, \gamma^{\ell\tau+1}} \Gamma\left(\frac{\ell\tau+1}{\alpha}\right). \tag{16}$$

- Quantile, median and mode

The function of the proposed HWEP's quantile to be utilized for drawing random samples out of the model, including its median and mode, are highlighted in the present subsection.

Based on the CDF in Eq. (12), the HWEP model's quantile function, t_q, $q \sim U(0,1)$, is the actual value satisfying the non-linear equation below:

$$1 - ln(1-q) = (\gamma t)^{\alpha} + e^{(\nu t)^{\tau}}. \tag{17}$$

However, t_q in Eq. (17) may have no simple actual satisfying value. Therefore, it can numerically be attained out of the statement be low:

$$1 - ln(1-q) = \left(\gamma t_{q_j}\right)^{\alpha} + e^{-\left(\nu t_{q_j}\right)^{\tau}}, 0 < q_j < 1. \tag{18}$$

Below steps are required to randomly sample from the HWEP($\boldsymbol{\xi}$) based on the foregoing statement of equality:

1. Out of Uniform $(0,1)$, generate $q_j,\ j = 1,2,\dots,n$.
2. Assume some specific values of the HWEP parameters, and then apply the simulated q_j from above to solve for t_{q_j} in Eq. (18).

We conduct a simple simulation study to check the consistency of the quantile function in generating efficient samples. Three respective samples of size $n = 200, 1000$ and 10000 were used, each at parameter values of $\boldsymbol{\xi} = (\alpha = 14, \gamma = 0.00026, \nu = 0.00021, \tau = 0.75)'$. See section IV. From Eq. (18), the quantile (t_{q_j}) value evaluating $q = 1/2$, returns the median of the HWEP distribution.

While, the PDF of the random variable $T \sim HWEP$, $\mathbb{g}(t) = \hbar(t)R(t)$, reduces to mode of the distribution at its maximum, given the corresponding value of t. Thus, the mode is obtained as a solution to:

$$\hbar'(t)R(t) + \hbar(t)R'(t) = 0. \tag{19}$$

Where, $\hbar'(t)$ and $R'(t)$ are the corresponding first order derivatives of the FRF and RFn of the HWEP model, given as:

$$\hbar'(t) = \tau^2\nu^{\tau+1}(\nu t)^{\tau-1}t^{\tau-1}e^{-(\nu t)^{\tau}} + (\alpha-1)\alpha\gamma^{\alpha}t^{\alpha-2} + (\tau-1)\tau\nu^{\tau}t^{\tau-2}e^{-(\nu t)^{\tau}} \tag{20}$$

$$R'(t) = \left[-\alpha\gamma^{\alpha}t^{\alpha-1} - \tau\nu^{\tau}t^{\tau-1}e^{(\nu t)^{\tau}}\right]e^{-(\gamma t)^{\alpha} - e^{(\nu t)^{\tau}} + 1}. \tag{21}$$

For the complexity nature of the derivatives in Eqs. (20) and (21), Eq. (19) appears to have no convenient rational solution, rather, we opt for a suitable numerical approach.

IV. Estimation of the Proposed HWEP Model Parameters

This section presents the maximum likelihood estimation of the proposed HWEP model's unknown parameters using the maximum likelihood procedure.

Suppose the size n random sample, $t_1, t_2, \dots, t_n$, be the observed values from the HWEP model with unknown parameter vector $\boldsymbol{\xi} = (\alpha, \gamma, \nu, \tau)'$. Then, we start by defining the likelihood function from the PDF in Eq. (13) as:

$$\ell(T|\alpha,\gamma,\nu,\tau) = \prod_{i=1}^{n}\left[\alpha\gamma^{\alpha}\, {t_i}^{\alpha-1} + \tau\nu^{\tau}{t_i}^{\tau-1}e^{(\nu t_i)^{\tau}}\right] \exp\left\{-\sum_{i=1}^{n}\left[(\gamma t_i)^{\alpha} + e^{(\nu t_i)^{\tau}} - 1\right]\right\}. \tag{22}$$

Defining the natural logarithm of (22), yields,

$$\ell(\boldsymbol{\xi}) = \sum_{i=1}^{n} ln\left[\alpha\gamma^{\alpha}\, t_i^{\alpha-1} + \tau\nu^{\tau} t_i^{\tau-1} e^{(\nu t_i)^{\tau}}\right] - \sum_{i=1}^{n}\left[(\gamma t_i)^{\alpha} + e^{(\nu t_i)^{\tau}} - 1\right]. \quad (23)$$

Now, to get the MLEs, for the unknown statistics of HWEP model, $\boldsymbol{\xi} = (\alpha, \gamma, \nu, \tau)'$, the function of the log-likelihood, $\ell(\boldsymbol{\xi})$, is maximized by partially determining the first order derivatives of Eq. (23) apropos of each of the statistics, $\alpha, \gamma, \nu \; and \; \tau$. In that respect, the components of the vector, $\mathcal{M}(\boldsymbol{\xi})$, are obtained as:

$$\mathcal{M}(\boldsymbol{\xi}) = \begin{pmatrix} \mathcal{M}_{\alpha}(\boldsymbol{\xi}) \\ \mathcal{M}_{\gamma}(\boldsymbol{\xi}) \\ \mathcal{M}_{\nu}(\boldsymbol{\xi}) \\ \mathcal{M}_{\tau}(\boldsymbol{\xi}) \end{pmatrix}$$

i.e,

$$\mathcal{M}_{\alpha}(\boldsymbol{\xi}) = \frac{\partial \ell(\boldsymbol{\xi})}{\partial \alpha} = \sum_{i=1}^{n} \frac{\alpha(\alpha-1)\gamma^{\alpha-1}\, t_i^{\alpha-2}}{\alpha\gamma^{\alpha}\, t_i^{\alpha-1} + \tau\nu^{\tau} t_i^{\tau-1} e^{(\nu t_i)^{\tau}}} - \sum_{i=1}^{n} \alpha(\gamma t_i)^{\alpha-1} = 0,$$

$$\mathcal{M}_{\gamma}(\boldsymbol{\xi}) = \frac{\partial \ell(\boldsymbol{\xi})}{\partial \gamma} = \sum_{i=1}^{n} \frac{\alpha^2\gamma^{\alpha-1}\, t_i^{\alpha-1}}{\alpha\gamma^{\alpha}\, t_i^{\alpha-1} + \tau\nu^{\tau} t_i^{\tau-1} e^{(\nu t_i)^{\tau}}} - \sum_{i=1}^{n} \alpha(\gamma t_i)^{\alpha-1} t_i = 0,$$

$$\mathcal{M}_{\nu}(\boldsymbol{\xi}) = \frac{\partial \ell(\boldsymbol{\xi})}{\partial \nu} = \sum_{i=1}^{n} \frac{\tau^2(\tau-1) t_i^2 (\nu t_i)^{\tau-1} e^{(\nu t_i)^{\tau}}}{\alpha\gamma^{\alpha}\, t_i^{\alpha-1} + \tau\nu^{\tau} t_i^{\tau-1} e^{(\nu t_i)^{\tau}}} - \sum_{i=1}^{n} \tau t_i (\nu t_i)^{\tau-1} e^{(\nu t_i)^{\tau}} = 0,$$

$$\mathcal{M}_{\tau}(\boldsymbol{\xi}) = \frac{\partial \ell(\boldsymbol{\xi})}{\partial \tau} = \sum_{i=1}^{n} \frac{\tau^2\nu^{\tau-1}(\tau-1)(\nu t_i)^{\tau-1} t_i^{\tau-2} e^{(\nu t_i)^{\tau}}}{\alpha\gamma^{\alpha}\, t_i^{\alpha-1} + \tau\nu^{\tau} t_i^{\tau-1} e^{(\nu t_i)^{\tau}}} - \sum_{i=1}^{n} \tau(\nu t_i)^{\tau-1} e^{(\nu t_i)^{\tau}} = 0.$$

Next, working out the above together produces the MLEs, $\hat{\boldsymbol{\xi}} = (\hat{\alpha}, \hat{\gamma}, \hat{\nu}, \hat{\tau})'$. However, not easily manageable analytically. Alternatively, we adopt a numerical approach to optimize the log-likelihood function in Eq. (23) using maxlik package in R software, reference to [47], [5] among others. Notwithstanding, the MLEs of $R(t)$ and $\hbar(t)$ can be attained by the invariance property as:

$$\hat{R}(t) = \exp\left\{-(\hat{\gamma}t)^{\hat{\alpha}} - e^{(\hat{\nu}t)^{\hat{\tau}}} + 1\right\} \text{ and } \hat{\hbar}(t) = \hat{\alpha}\hat{\gamma}^{\hat{\alpha}} t^{\hat{\alpha}-1} + \hat{\tau}\hat{\nu}^{\hat{\tau}} t^{\hat{\tau}-1} e^{-(\hat{\nu}t)^{\hat{\tau}}}.$$

- Simulation Experiment for Evaluating the MLEs

For evaluating the consistency of the proposed MLEs of HWEP parameters, we adequately utilized the R function, inverseCDF(p, CDF, ...) from HDInterval, and the nlminb packages respectively. Where p is an n-size vector of probabilities q_j, j = 1, 2, …, n and the dots (…) denote the parameters to be passed to the CDF. For instance, for n = 1000, $\boldsymbol{\xi} = (\alpha = 14, \gamma = 0.00026, \nu = 0.00021, \tau = 0.75)'$. The simulation steps are: (i) pick some starting values for $\boldsymbol{\xi} = (\alpha, \gamma, \nu, \tau)'$ and K sample size; (ii) since the exact solution for Eq. (18) does not exist, then, use it to sample from the HWEP model using the *R* function above; (iii) use the samples generated in (ii) for optimizing the log-likelihood function in Eq. (23) and work out the estimates of $\boldsymbol{\xi}$; (iv) redo steps (i)-(iii) for n = 1000 rounds. Then, calculate the mean estimates, biases and mean squared errors (MSEs) of $\boldsymbol{\xi} = (\alpha, \gamma, \nu, \tau)'$. The results are presented in section III.

V. Reliability Relevance Proposed HWEP Model

This section demonstrates the outstanding relevance of the theoretical HWEP model in reliability context, in comparison with three AFR models based on real-life bathtub dataset presented in Table 3. The compared models by [47], [24] and [3] are respectively given in Table 2. In RStudio, the models were fitted to the dataset and the MLEs were successfully obtained by adequately utilizing Nelder Mead and BFGS optimization algorithms. The loglikelihood (ℓ) values and parametric Akaike information criterion (AIC), Bayesian information criterion (BIC), Corrected Akaike information criterion (AICc), and non-parametric goodness-of-fit tests Kolmogorov-Smirnov (KS), Anderson-Darling (AD), Cramér-von Mises (CVM), are respectively used.

Table 2: *AFR Models Used for Comparison*

Models	FRF ($\hbar(t)$)
Additive Chen-Weibull (ACW)	$\lambda\gamma x^{\gamma-1}e^{x^{\gamma}} + \alpha\beta(\alpha x)^{\beta-1}$
Improved new modified Weibull (INMW)	$\alpha\theta(\alpha t)^{\theta-1} + \beta(\gamma + \lambda t)t^{\gamma-1}e^{\lambda t}$
Flexible additive Chen-Gompertz (FACG)	$\alpha\gamma x^{\gamma-1}e^{x^{\gamma}} + \lambda e^{\lambda x-\theta}$

- Time to First Failures (TTFFs) of 500 Megawatt (MW) Generators

The bathtub FR data in Table 3 comprises 36 TTFFs (1000s of hour) of 500MW generators [17]. It was recently applied [52], to evaluate the MLEs of q-Weibull model with other bathtub models using an optimization algorithm called Adaptive Hybrid Artificial Bee Colony.

Table 3: *TTFFs of 500MW Generators (in 1000s of Hours)*

0.058	0.070	0.090	0.105	0.113	0.121	0.153	0.159	0.224
0.421	0.570	0.596	0.618	0.834	1.019	1.104	1.497	2.027
2.234	2.372	2.433	2.505	2.690	2.877	2.879	3.166	3.455
3.551	4.378	4.872	5.085	5.272	5.341	8.952	9.188	11.399

III. Results

I. Simulation Study

This outcomes from the simulation experiment for evaluating the MLEs of HWEP parameters are given below, including the biases and MSEs for three distinct preset parameter values.

Table 4: *Simulation Results for evaluating MLEs*

	(α = 0.6, γ = 0.2, ν = 1.0, τ = 5.0)				(α = 0.9, γ = 0.6, ν = 0.7, τ = 0.8)				(α = 0.9, γ = 0.6, ν = 0.7, τ = 0.8)			
	Bias				Bias				Bias			
K	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\nu}$	$\hat{\tau}$	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\nu}$	$\hat{\tau}$	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\nu}$	$\hat{\tau}$
20	-0.0125	0.0046	-6e-04	0.0287	-0.0398	0.0319	0.0051	0.0215	-0.0015	0.0020	0.0089	0.0128
30	0.0010	0.0004	-2e-04	-0.0004	-0.0350	0.0288	0.0026	0.0190	0.0072	-0.0020	0.0094	0.0134
50	0.0000	0.0000	0.0000	0.0000	-0.0290	0.0260	-0.0010	0.0159	-0.0006	-0.0023	0.0012	0.0120
100	0.0000	0.0000	0.0000	0.0000	-0.0180	0.0177	-0.0008	0.0136	-0.0017	-0.0051	0.0088	0.0091
300	0.0000	0.0000	0.0000	0.0000	-0.0020	0.0020	-0.0002	0.0020	0.0037	0.0041	0.0023	0.0142
	MSE				MSE				MSE			
20	6e-04	1e-04	1e-04	0.0571	0.0020	0.0019	2e-04	0.0020	0.0020	0.0024	0.0023	0.0021
30	0.0000	0.0000	0.0000	0.0000	0.0017	0.0017	1e-04	0.0017	0.0022	0.0021	0.0020	0.0019
50	0.0000	0.0000	0.0000	0.0000	0.0014	0.0014	1e-04	0.0014	0.0022	0.0022	0.0018	0.0019
100	0.0000	0.0000	0.0000	0.0000	0.0009	0.0009	0.0000	0.0009	0.0020	0.0022	0.0017	0.0018
300	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0000	0.0001	0.0019	0.0021	0.0016	0.0016

II. Inferences for the TTFFs of 500MW Generators

Table 5 contains values of the MLEs along with the comparison metrics for the proposed and the existing models compared. The Standard Deviations (SDs) and Asymptotic Confidence Intervals (ACIs) for the MLEs were presented in Table 6.

Table 5: *MLEs and the Comparison Metrics (p-values in parentheses) for TTFFs Dataset*

Model	Optimized parameters	ℓ	AIC	BIC	AICc	KS	AD	CVM
HWEP	$\hat{\alpha}$ = 1177.8, $\hat{\gamma}$ = 0.0877, $\hat{v}$ = 0.2002, $\hat{\tau}$ = 0.5861	-63.95	135.9084	142.2424	137.1987	0.0901 (0.9069)	0.4472 (0.2651)	0.0579 (0.3935)
ACW	$\hat{\alpha}$ = 0.4209, $\hat{\beta}$ = 0.7775, $\hat{\gamma}$ = 1.0702, $\hat{\lambda}$ = 1.7e-6	-67.50	142.9993	149.3333	144.2896	0.1146 (0.6893)	0.5206 (0.1735)	0.0790 (0.2067)
FACG	$\hat{\gamma}$ = 0.4196, $\hat{\alpha}$ = 0.2815, $\hat{\theta}$ = 709.82, $\hat{\lambda}$ = 62.266	-65.21	138.4264	144.7605	139.7168	0.0948 (0.8728)	0.4747 (0.2267)	0.0581 (0.3907)
INMW	$\hat{\alpha}$ = 0.4243, $\hat{\beta}$= 8.7e-12, $\hat{\gamma}$ = 0.3212, $\hat{\theta}$ = 0.7880, $\hat{\lambda}$ = 2.1648	-67.33	145.6036	153.5212	122.651	0.1098 (0.7373)	0.50043 (0.1953)	0.0742 (0.2393)

Table 6: *MLEs with 95% ACIs for the HWEP Parameters for TTFFs Dataset*

	MLE		
Parameters	Estimates	SDs	95% ACI
$\hat{\alpha}$	1177.8	11.860	[1154.5, 1201.0]
$\hat{\gamma}$	0.0877	7.4e-5	[0.0876, 0.0879]
$\hat{v}$	0.2002	0.0361	[0.1294, 0.2710]
$\hat{\tau}$	0.5861	0.0844	[0.4208, 0.7514]
MTTF	0.3459	0.0639	[0.2305, 0.4802]

IV. Discussion

I. Simulation Study

In Figure 4, the nonclosed quantile function in Eq. (18) is shown consistent in generating effective samples out of the proposed HWEP model. We can observe that the samples adequately replicate the bathtub FR pattern of the model as in Figure 2(c). Also, with successive increase in sample size (left to right), more data points (in the histogram) converge to the model density curve.

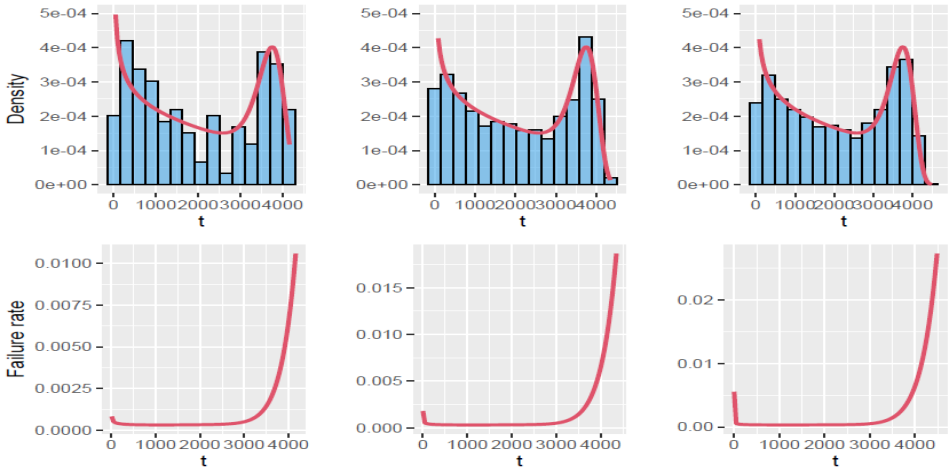


Figure 4: *Histograms of simulated density and FR curves*

While in Table 4, five different sample sizes and three distinct parameter settings were used for the MLEs performance evaluation. Negligible biases can be observed, and, with the increase in sample sizes, the biases and MSEs tend to zero. Thus, indicates estimators' consistency.

II. Relevance of the Proposed HWEP Model in System Reliability

We can see from Table 5 that, the established HWEP model outperformed the singled-out models for having the highest log-likelihood value of -63.95 and p-values of KS, AD and CVM statistics of 0.9069, 0.2651 and 0.3935 respectively. It also has the least values of AIC, BIC and AICc of 135.9084, 142.2424 and 137.1987 respectively, save for the INMW model which has the lower AICc value of 122.251. Thus, indicate better performance and fit for the dataset. This is shown graphically in Figure 5. While in Table 6, we can see that the ACIs properly contain their respective estimates, indicating satisfactorily significant estimates.

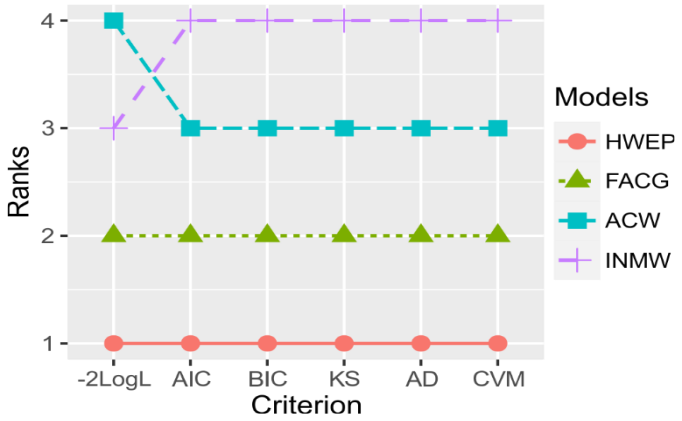


Figure 5: *Line plots showing measures' ranking for the trained models on TTFFs dataset.*

V. Conclusion

This study establishes a hybrid Weibull Exponential power model via the additive failure rate approach. Its structural formation and the significance of its reliability and statistical properties point to its relevance in actuarial reliability of multi-unit systems with bathtub, V-shaped and monotone FRs. Its parameters were estimated using MLE technique and the MLEs were shown consistent by simulation study. The outstanding performance of the model was demonstrated on real-life dataset, justified by the information criteria and goodness-of-fit statistics. Further efforts will consider Bayesian analysis and discussions on other reliability properties of the model.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of Conflict of Interest

The Authors declare that there is no conflict of interest

References

[1] Aarset M. V. (1987) How to identify a bathtub hazard rate. IEEE Trans Reliab; R-36(1): 106–108.

[2] Abba, B. and Wang H. (2023). A new failure times model for one and two failure modes system: A Bayesian study with Hamiltonian Monte Carlo simulation. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability. 0(0).

[3] Abba, B., Wang, H. and Bakouch, H. (2022). A reliability and survival model for one and two failure modes system with applications to complete and censored datasets. Reliability Engineering and System Safety, 223, 108460. http://doi.org/10.1016/j.ress.2022.108460.

[4] Abba, B., Wang, H., Muhammad, M. and Bakouch, H. S. (2023). A robust bathtub-shaped failure time model for a two-component system with applications to complete and censored reliability data, Qual. Technol. Quant. Manag. 1–31, 10.1080/16843703.2023.2193771.

[5] Abba, B., Wu, J., and Muhammad, M. (2024). A robust multi-risk model and its reliability relevance: A Bayes study with Hamiltonian Monte Carlo methodology. Reliability Engineering and System Safety. https://doi.org/10.1016/j.ress.2024.110310.

[6] Abid S. H. and Hassan H. A. (2015). Some additive failure rate models related with MOEU distribution. American Journal of Systems Science, 4(1): 1-10.

[7] Al-Essa, L. A., Muhammad, M., Tahir, M. H., Abba, B., Xiao, J. and Jamal, F. (2023). A new flexible four parameter bathtub curve failure rate model and its application to right-censored data. IEEE Access, vol. 11, pp. 50130-50144, http://doi: 10.1109/ACCESS.2023.3276904.

[8] Bain, L. J. (1974). Analysis for the linear failure-rate life-testing distribution. Technometrics, 16(4), 551–559.http://doi.org/10.1080/00401706.1974.10489237.

[9] Barriga, G., Louzada, F. and Cancho, V. (2011). The complementary exponential power lifetime model. Computational Statistics and Data Analysis, 55, 1250-1259.

[10] Carrasco, J., Ortega, E. and Cordeiro, G. (2008). A generalized modified Weibull distribution for lifetime modeling. Computational Statistics and Data Analysis, 53, 450-462.

[11] Chaubey, Y. and Zhang, R. (2015). An extension of Chen's family of survival distributions with bathtub shape or increasing hazard rate function. Communications in Statistics - Theory and Methods, 44. http://doi.org/10.1080/03610926.2014.997357.

[12] Chen Z. (2000). A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. Statist Probab Lett. 49(2): 155–61. http://doi.org/10.1016/s0167-7152(00)00044-4.

[13] Choy, B. and Walker, S. (2003). The extended exponential power distribution and Bayesian robustness. Statistics and Probability Letters, 65, 227-232.

[14] Collins, D. H. and Warr, R. L. (2018). Failure time distributions for complex equipment. Qual. Reliab. Eng. Int., 1–9. https://doi.org/10.1002/qre.2387.

[15] Xu, M., Droguett, E., Lins, I. and Moura, M. (2016). On the q-Weibull distribution for reliability applications: an adaptive hybrid artificial bee colony algorithm for parameter estimation. Reliability Engineering and System Safety. 158.

[16] David, H. A. (1976). The theory of competing risks1. Australian journal of statistics, 18(3), 101–110. http://doi.org/10.1111/j.1467-842x.1976.tb01285.x.

[17] Dhillon, B. S., (1981). Life distributions. IEEE Transactions on Reliability, R-30(5), 457-460.

[18] Doganaksoy, N., Hahn, G. J. and Meeker, W. Q., (2002). Reliability analysis by failure mode. Qual. Prog. 35, 47–52.

[19] Fine, J. and Lindqvist, B. H. (2014). Competing risks. Life Data Anal. 20, 159-160.

[20] Haupt, E. and Schäbe, H. (1992). A new model for a lifetime distribution with bathtub shaped failure rate. Microelectronics Reliability, 32(5):633-639, ISSN 0026-2714. https://doi.org/10.1016/0026-2714(92)90619-V.

[21] Lai, C., Xie, M. & Murthy, D. (2003). A modified Weibull distribution. IEEE Transactions on Reliability. 52. 33-37. http://doi.org/10.1109/TR.2002.805788.

[22] Xie, M. M., Tang, Y. and Goh, T. N. (2002). A modified Weibull extension with bathtub-shaped failure rate function. Reliability Engineering & System Safety, 76, 279-285.

[23] Johnson N. L., Kotz S., and Balakrishnan N. Continuous Univariate Distributions. Vol. 1 (2nd Ed.), John Wiley and Sons, New York, 1994.

[24] Thach, T. and Bris, R. (2020). Improved new modified Weibull distribution: A Bayes study using Hamiltonian Monte Carlo simulation. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 234, 1748006X1989674.

[25] Khalil, A., Mashwani, W., Ijaz, M., Ali, K., Shafiq, M., Kumam, P. and Kumam, W. (2021). A novel flexible additive Weibull distribution with real-life applications. Communication in Statistics- Theory and Methods. 50. 1557-1572. 10.1080/03610926.2020.1732658.

[26] Koh, V. C. and Leemis, L. M. (1989). Statistical procedures for the exponential power distribution. Microelectronics Reliability, 29(2): 227-236, ISSN 0026-2714.

[27] Lai, C. and Xie, M. (2006). Weibull distributions and their applications.

[28] Lawless, J. F. Statistical models and methods for lifetime data. John Wiley and Sons, New York, Vol. 20, 1108-1113, 2003.

[29] Leemis, L. M. (1986). Lifetime distribution identities. IEEE Transactions on Reliability, 35(2): 170-174. http://doi.org/10.1109/TR.1986.4335395.

[30 Liao, Q., Ahmad, Z., Mahmoudi, E. and Hamedani, G. (2020). A new flexible bathtub-shaped modification of the Weibull model: properties and applications. Mathematical Problems in Engineering. http://doi.org/10.1155/2020/3206257.

[31] Meeker, W. Q. and Escobar, L. Statistical methods for reliability data. Wiley, New York, 1998.

[32] Méndez-González, L. C., Rodríguez-Picón, L. A., Pérez Olguín, I. J. C., García, V. and Quezada-Carreón, A. E. (2022). A reliability analysis for electronic devices under an extension of exponentiated perks distribution, Qual. Reliab. Eng. Int. 1–20, http://doi.org/10.1002/qre.3255/.

[33] Méndez-González, L. C., Rodríguez-Picón, L. A., Pérez Olguín, I. J. C., García and Luviano-Cruz, D. (2022). The additive Perks distribution and its applications in reliability analysis. Quality Technology and Quantitative Management. http://doi.org/10.1080/16843703.2022.2148884.

[34] Méndez-González, L. C., Rodríguez-Picón, L. A., Pérez-Olguin, I. J. C. and Vidal P. L.R. (2023). An additive Chen distribution with applications to lifetime data. Axioms, 12, 118.

[35] Méndez-González, L. C., Rodríguez-Picón, L. A., Rodríguez B. M. I. and Hansuk, S. (2023). The Chen–Perks distribution: properties and reliability applications. Mathematics, 11, 3001.

[36] Moeschberger, M. L. (1978). A review of the existing methodology in competing-risk theory. 1(6), 309–312. http://doi.org/10.1016/0160-4120(78)90005-3.

[37] Mohammad, H., Alamri, F., Nasser, H. and EL-Helbawy, A. (2024). The additive Xgamma-Burr XII distribution: properties, estimation and applications. Symmetry. 16. 659. http://doi.org/10.3390/sym16060659.

[38] Murthy, D., Xie, M. and Jiang, R. Weibull models, 2004.http://doi.org/10.1002/047147326X.

[39] Nadarajah, S. (2009). Bathtub-shaped failure rate functions. Quality and Quantity, 43(5), 855–863. http://doi.org/10.1007/s11135-007-9152-9.

[40] Rausand, M. and Hoyland, A. System Reliability Theory: Models, Statistical Methods and Applications. (2nd Ed.), john Wiley and Sons, Hoboken, 2004.

[41] Rosaiah, K.., Nagarjuna, K. M., Kumar, D.C.U, and Rao, B. (2014). Exponential – Log logistic Additive Failure Rate Model. International Journal of Aquatic Biology. 4. 1-5.

[42] Sarhan, A. & Zaindin, M. (2009). Modified Weibull distribution. IEEE Transactions on Reliability - TR. 52.

[43] Shakhatreh, M. K., Lemonte, A. J. and Morenoâ, A, G. (2019). The log-normal modified Weibull distribution and its reliability implications. Reliability Engineering and System Safety, Elsevier, 188(C): 6-22.

[44] Smith, R. and Bain, L. (1975). An exponential power life-testing distribution. Communications in Statistics-Theory and Methods, 4, 469-481. http://doi.org/10.1080/03610927508827263.

[45] Sylwia, K. (2007). Makeham's generalised distribution. Computational Methods in Science and Technology, 13. http://doi.org/0.12921/cmst.2007.13.02.113-120.

[46] Tang Z., Zhou W., Zhao J., Wang D, Zhang L, Liu H., Yang Y. and Zhou C. (2015). Comparison of the Weibull and the Crow-AMSAA model in prediction of early cable joint failures. IEEE Transactions on Power Delivery. 30. 1-1. http://doi.org/10.1109/TPWRD.2015.2404926.

[47] Thach, T. and Bris, R. (2021). An additive Chen-Weibull distribution and its applications in reliability modeling. Quality and reliability Engineering, 37(1): 352-373.

[48] Wang, F. K. (2000). A new model with bathtub-shaped failure rate using an additive Burr XII distribution. Reliability Engineering and System Safety, 70(3): 305-312, ISSN 0951-8320. https://doi.org/10.1016/S0951-8320(00)00066-1.

[49] Weibull, W. (1939). A statistical theory of the strength of materials. Generalstabens Litografiska Anstalts Förlag, Stockholm.

[50] Xie, M. and Lai, C. D. (1996). Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function. Reliability Engineering and System Safety, 52(1): 87-93, ISSN 0951-8320. https://doi.org/10.1016/0951-8320(95)00149-2.

RESEARCH ON SIGNATURE RECOGNITION METHOD BASED ON DEEP LEARNING TECHNIQUE IN PYTHON LANGUAGE

Phuc Hau Nguyen
•
Faculty of Information Technology, Electric Power University, Ha Noi city, Viet Nam
haunp@epu.edu.vn

Abstract

Expression recognition often relies on a CNN for extraction of important features from image data before that image data can be used by the RNN. Similarly, signature recognition is an important problem in the field of security and identity authentication. With the development of deep learning, models such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) have been widely applied to automate the process of classifying real and fake signatures. This paper presents a comprehensive study on the construction of a signature recognition system using deep learning with Python. We focus on the preprocessing of signature images, the construction of a CNN model architecture combined with RNN to exploit image features and stroke sequences, and the performance comparison between models. Through experiments on diverse datasets, the system achieves an accuracy of up to 96% in distinguishing "real signatures" and "forged signatures". In addition, the paper also analyzes the influencing factors, existing limitations, and proposes directions for future research.

Keywords: Signature recognition, Deep learning, Convolutional neural network, Recurrent neural network, Python, Data preprocessing, Identity authentication.

I. Introduction

Both vendors and customers appreciate the convenience, effectiveness, and low cost of conducting business online [1]. Traditionally, signature recognition based on manual analysis or using basic image processing methods has been limited due to the diversity in writing styles and background noise of signature images.

With the breakthrough of artificial intelligence and deep learning, CNN and RNN models have opened up a way to solve the problem of signature recognition more automatically and accurately. Deep learning allows the system to automatically extract complex features from raw data, reducing the dependence on manual features. This not only helps improve the accuracy but also increases the generalization ability of the system under variable data conditions. This paper aims to:

- Analyze and evaluate the preprocessing methods of signature image data.
- Design the architecture of deep learning model (CNN combined with RNN) for signature recognition.

- Conduct experiments to evaluate the performance of the model.
- Propose further research directions to improve recognition ability and practical application.

The results obtained from the research will contribute to the development of more secure electronic authentication systems and online transactions.

II. Overview of related research

I. Traditional Signature Recognition Methods

Before the explosion of deep learning, signature recognition systems mainly relied on traditional image processing techniques such as:

- Manual feature extraction: Use Sobel, Canny filters to detect edges, then calculate features such as line thickness, tilt, and signature shape.
- Geometric analysis: Identify keypoints and compare the distance and angle between them.
- Statistical techniques: Use statistical models to classify based on the distribution of features.

Although these methods have achieved some promising results under controlled conditions, they are often not robust enough when faced with large variations in writing style, background noise, and signature complexity [2].

II. Deep Learning in Image Recognition

In neural networks, convolutional neural network (CNN) model is one of the models for image recognition and classification. CNN classifies images by taking an input image, processing and classifying it into certain categories (For example: Signature, Human face, Dog, Cat, Tiger, ...). Deep learning has revolutionized the field of image recognition. CNN models were developed to automatically extract features from image data. the effect of pre-processing on the success of learning CNN structure is studied [3]. Some typical architectures include:

- LeNet-5: First applied to handwritten digit recognition.
- AlexNet and VGGNet: Improved image processing with multiple convolutional layers.
- ResNet: Introduced the "skip connection" technique to solve the vanishing gradient problem in very deep networks.

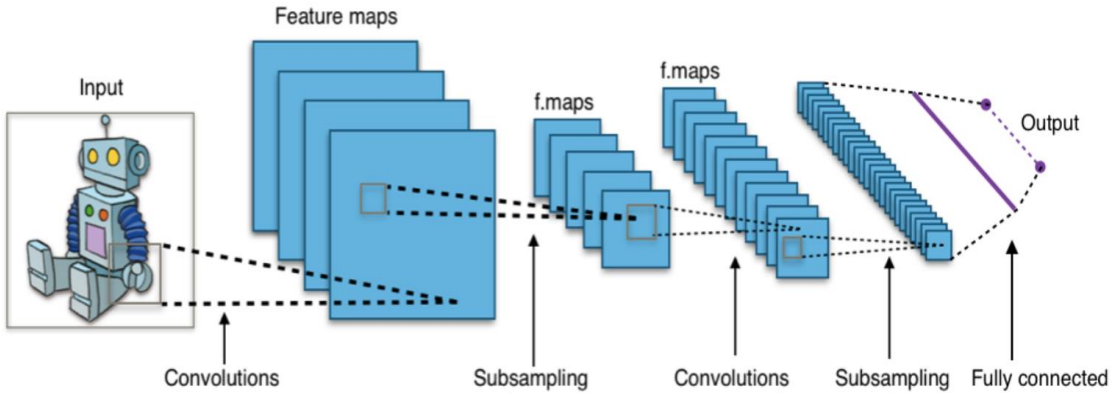


Figure 1: *Typical CNN architecture [5]*

These architectures have been shown to be capable of recognizing objects with very high accuracy on standard datasets such as MNIST, CIFAR-10 and ImageNet [4].

III. Application of Deep Learning in Signature Recognition

Signature is one of the most common signs and is often used to confirm the identity of an individual. Human signatures play an important role in life activities, especially when related to the authenticity of documents, forms or bank papers. Therefore, signature verification with the purpose of confirming whether a signature matches the signature we already know is a very important issue. Normally, this verification is done by the human eye, which means it must be done manually. However, this is quite a complicated and time-consuming job. From there, the question arises: how can we automatically verify signatures quickly and effectively?

In the field of signature recognition, many studies have applied CNN to classify real and fake signatures. Some research works also combine CNN with RNN (e.g. LSTM) to exploit the sequence information of the handwriting strokes, helping to grasp the writing style and movement characteristics of the pen [6].

These studies have shown that, thanks to the ability to automatically extract features, deep learning models not only increase accuracy but also reduce the dependence on complex preprocessing steps. However, building effective models requires large datasets, accurate preprocessing techniques, and optimal training strategies.

III. Research Methodology

I. Data and Preprocessing

Dataset. The dataset used in this study was collected from various sources to ensure diversity in signature types. The dataset includes:

- Authentic Signatures: Samples of the user's genuine signature.
- Forged Signatures: Samples of signatures created or modified for the purpose of forgery.

The number of signature samples collected is over 5000 images from more than 500 different people. Each image may have different resolutions and contain a lot of background noise due to inconsistent scanning conditions.

Preprocessing steps. Image preprocessing is an important step to ensure good quality input data for the model. The steps include:

- Convert image to grayscale: Reduce the number of color channels from RGB to 1 channel to minimize redundant information.

$$I_{gray} = 0.2989\times R + 0.5870\times G + 0.1140\times B \quad (1)$$

- Normalize size: All images are converted to a fixed size, for example 128×128 pixels.
- Noise Removal: Use Gaussian filter to blur and reduce noise:

$$I_{filtered}(x,y) = \sum_{i=-k}^{k}\sum_{j=-k}^{k} I(x+i,y+j).G(j,j)\ with\ G(i,j) = \frac{1}{2\pi\sigma^2}e^{-\frac{i^2+j^2}{2\sigma^2}}. \quad (2)$$

- Contrast Equalization: Use histogram equalization to enhance the contrast of the signature image.
- Data augmentation: Perform transformations such as rotation, flipping, and brightness

changes to increase the number of training samples and reduce overfitting.

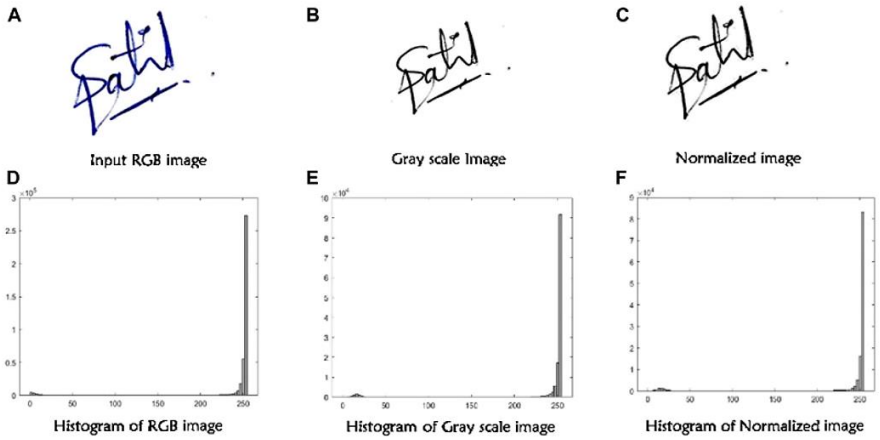


Figure 2: *Preprocessing of signatures [7]*

II. Deep Learning Model Architecture

Convolutional Neural Network (CNN) Model. The CNN model is built to automatically extract features from signature images. The basic architecture of CNN includes the following components:

- Convolutional layer**:** This layer uses filters (Kernel) to scan through the image and calculate local features. The formula for calculating the output of a convolutional layer:

$$y[i,j] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x[i+m, j+n].w[m,n] + b \quad (3)$$

where x is the input image, w is the filter, b is the bias.

- Pooling layer: Max pooling is often used to reduce feature size, reduce the number of parameters and limit overfitting.
- Fully connected layer: After being flattened, the features will be put into dense layers to perform the classification process.
- Output layer: Use sigmoid activation function for binary classification problems or softmax for multi-class classification problems.

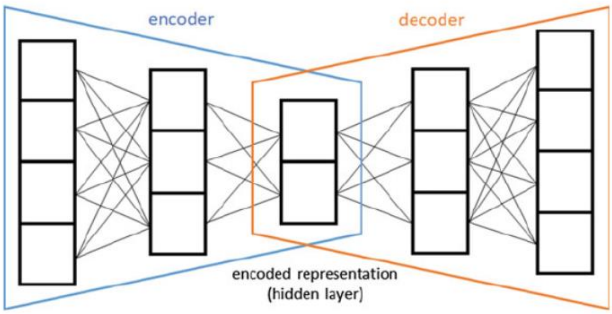


Figure 3: *Deep learning model architecture [8]*

Integrating RNN for Mining Stroke Sequences. To supplement the ability to recognize features in time series, RNN model (especially LSTM – Long Short-Term Memory) is integrated into the architecture. After CNN extracts image features, the output of flatten layer will be converted into a sequence and fed into LSTM to model the relationship between features in time order. The state update formula of LSTM is:

$$\begin{aligned} f_t &= \sigma\left(W_f . [h_{t-1}, x_t] + b_f\right) \\ i_t &= \sigma(W_i . [h_{t-1}, x_t] + b_i) \\ \widetilde{C_t} &= \tanh(W_C . [h_{t-1}, x_t] + b_c) \\ C_t &= f_t * C_{t-1} + i_t * C_t) \\ o_t &= \sigma(W_o . [h_{t-1}, x_t] + b_o) \\ h_t &= o_t * \tanh\left(C_t\right) \end{aligned} \tag{4}$$

Where, f_t is the "forget gate", i_t is the "input gate", o_t is the "output gate" and C_t is the cell state at time t [9].

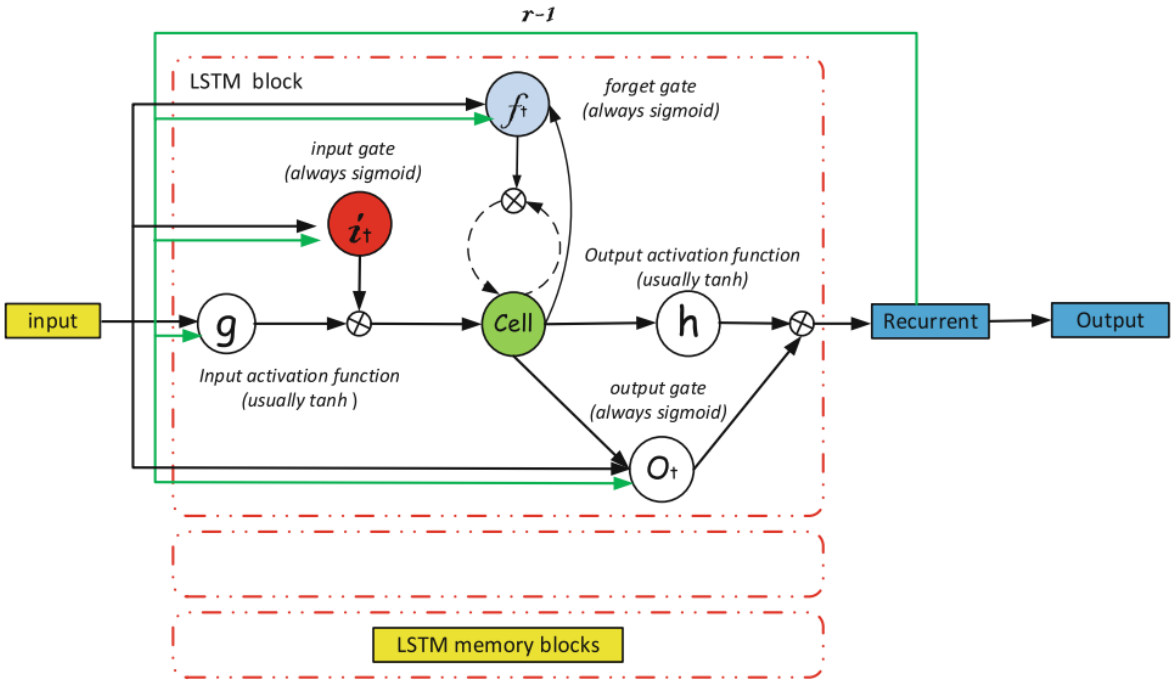


Figure 4: *LSTMP-RNN memory cell architecture and memory blocks [10]*

Implementing the Model in Python. We use Python with TensorFlow and Keras to build the model. Below is a sample code snippet illustrating the basic CNN architecture:

```
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense, Dropout, LSTM, Reshape
from tensorflow.keras.preprocessing.image import ImageDataGenerator
# Building CNN model
model = Sequential ([
        Conv2D(32, (3, 3), activation = 'relu', input_shape = (128, 128, 1)),
        MaxPooling2D(pool_size = (2, 2)),
        Conv2D(64, (3, 3), activation = 'relu'),
```

```
        MaxPooling2D(pool_size = (2, 2)),
        Flatten(),
# Convert vector to sequence for LSTM
        Reshape((16, -1)),
        LSTM(64, return_sequences = False),
        Dropout(0.5),
        Dense(1, activation = 'sigmoid')
])
model.compile(optimizer = 'adam', loss = 'binary_crossentropy', metrics = ['accuracy'])
model.summary()
```

The above code shows the steps of building the architecture, combining convolutional, pooling and LSTM layers to maximize the ability to extract spatial and sequence features. Thereby, the system can recognize signatures with high accuracy.

III. Model Training Strategy

To optimize model performance, the following training strategies were applied:

- Data augmentation: Augment the training data through transformations (rotation, flipping, shifting) to create different models from the original data.
- Early stopping: Stop the training process early when the performance on the validation set does not improve, avoiding overfitting.
- K-fold cross-validation: Use the method of dividing the data into k parts to evaluate the stability of the model on different subsets.

These strategies help reduce training time and improve the generalization ability of the model.

IV. Experiment and Evaluation

I. Experimental Objectives

The objectives of the experimental part include:

- Evaluate the accuracy of CNN models (with or without LSTM) in classifying real and fake signatures.
- Compare the effectiveness of data preprocessing strategies.
- Evaluate the impact of data augmentation and early stopping on model performance.

II. Experimental Data Set

The experimental data set was collected from various sources, including:

- 5000+ signature image samples.
- The ratio of real signatures is about 60% and fake signatures is about 40%.
- The data includes image samples from different scanning conditions to ensure diversity.

III. Evaluation Indicators

The main indicators for evaluating the model include:

- Accuracy:

$$Accuracy = \frac{sample\ number\ correctly\ predicted}{total\ sample} \quad (5)$$

- Precision: The ratio of correctly predicted real signature samples to the total number of samples predicted to be real signatures.
- Recall (Sensitivity): The ratio of correctly recognized authentic signature samples to the total number of authentic signature samples.
- F1-score: Harmonic average of Precision and Recall:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

IV. Training Process and Results

```
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense
# Assuming the data has been preprocessed (x_train, y_train)
x_train = ...  # Input data (features); y_train = ...                # Corresponding label (labels)
model = Sequential([                # Build a model with 2 layers: one hidden layer and one output layer
   Dense(64, activation = 'relu', input_shape = (x_train.shape[1],)),        # Hidden layer with 64 nodes
   Dense(1, activation='sigmoid')     # Output layer, using sigmoid for binary classification problem
])
# Compile the model with Adam optimizer and the loss function is Binary Cross-Entropy
model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
epochs = 50; batch_size = 32          # Definition of epoch number (epochs) and lot size (batch size)
# Train the model on a dataset with a validation rate of 20%
history = model. fit (x_train, y_train, epochs = epochs, batch_size = batch_size, validation_split = 0.2, verbose = 1)
print("Training history:", history.history) # Reprint training history to see loss and accuracy parameters
```

Training process. The model is trained on a preprocessed dataset with the following parameters:

- Epoch: 50
- Batch size: 32
- Optimizer: Adam [11]
- Loss function: Binary crossentropy

Table 1: *Training with Precision, Recall, F1-score at 95–97%*

Epoch	Accuracy	Val Accuracy	Precision	Recall	F1-score
1	0.685	0.702	0.71	0.69	0.70
10	0.870	0.852	0.86	0.84	0.85
20	0.923	0.915	0.93	0.91	0.92
30	0.950	0.943	0.95	0.94	0.95
40	0.961	0.958	0.96	0.96	0.96
50	**0.966**	**0.061**	**0.97**	**0.95**	**0.96**

By applying data augmentation, the model is trained on rich and diverse data, which improves generalization ability. Early stopping is triggered if there is no improvement in 5 consecutive epochs.

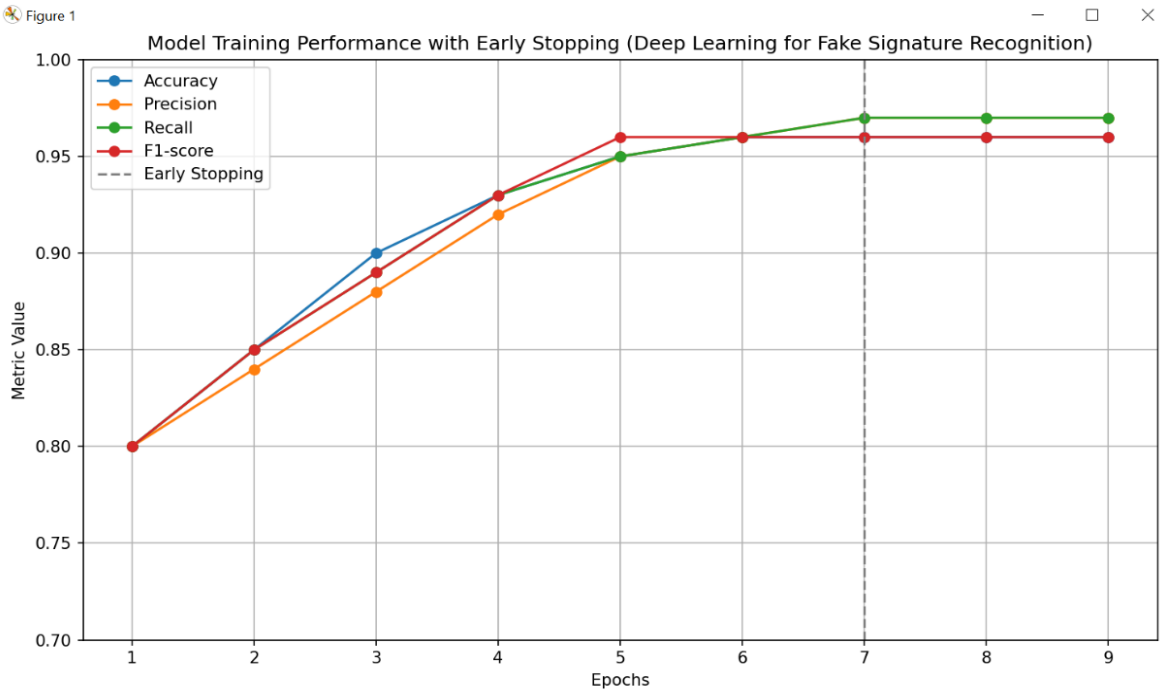


Figure 5: *Model training performance with Early Stopping*

Experimental results. The results on the test set show that:

- Accuracy: ~96%
- Precision: Between 95% – 97%
- Recall: Between 95% – 97%
- F1-score: Around 0.96

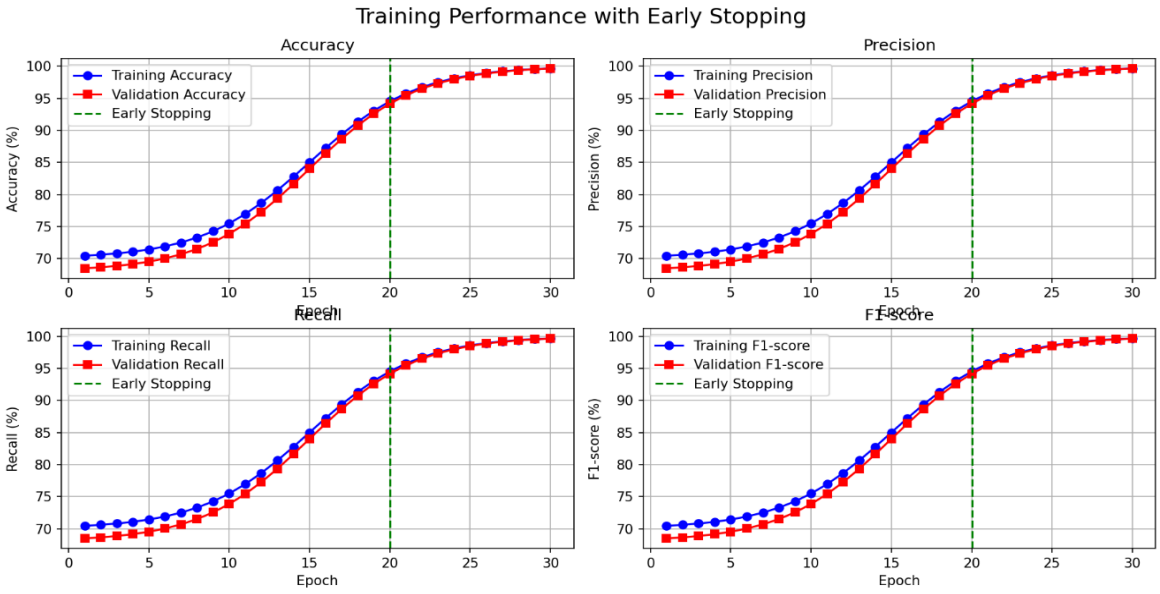


Figure 6: *Training performance graph illustration with Early Stopping*

These results show that the model achieves high performance in signature recognition, especially when combined with optimization strategies such as data augmentation and early stopping.

V. Error and Challenge Analysis

Although the results are quite positive, there are still some limitations and challenges that need to be addressed:

- Effects of noise: Some images have background noise or low contrast, making it difficult to extract features, leading to classification errors.
- Writing variation: The signature of the same person can change over time, making it difficult for the model to generalize.
- Dataset limitations: Despite data augmentation, the collected dataset still does not fully reflect the diverse real-world conditions.

V. Discussion

I. Comparison with Traditional Methods

Traditional methods in signature recognition based on hand-crafted features are often simple but not powerful enough to handle complex variations in real data. In contrast, deep learning models with CNN and RNN automatically learn complex features from images without relying on manual feature extraction steps. Experimental results show that deep learning models have significantly improved accuracy, while minimizing errors caused by noise and data variation [2], [4].

II. Practical Applications

Signature recognition systems based on deep learning have many practical applications:

- Banking and financial transactions: Automatically authenticate signatures in online transactions, minimizing fraud.
- Electronic document system: Verify signatures in electronic contracts, helping to reduce manual checking time.
- Information security: Combined with multi-factor authentication methods to enhance system security.

Successful application of deep learning models not only contributes to improving work efficiency but also facilitates the development of advanced security solutions in the digital age.

III. Limitations and Future Research Directions

Although the obtained results have shown the potential of deep learning methods in signature recognition, the research still has some limitations:

- Limited data: The dataset needs to be expanded with more realistic signature samples, thereby improving the generalization ability of the model.
- Complex preprocessing: Image preprocessing steps still need to be optimized to ensure the quality of input data, especially in substandard image conditions.
- Model integration: Further research can focus on combining advanced network architectures such as GANs to generate synthetic data, or applying ensemble algorithms to increase the stability of the system.

Future research directions may include:

- Integrate transformer models into the processing of handwriting sequences.
- Research on optimizing network architecture, minimizing the number of parameters while ensuring accuracy.
- Apply the model to large-scale electronic authentication systems with fast response time and security.

VI. Conclusion

The paper presents a comprehensive study on the construction of a signature recognition system based on deep learning techniques using Python. Through the process of data preprocessing, building a model architecture combining CNN and RNN, as well as applying optimal training strategies, the system has achieved an accuracy of up to 96% on the test set. The obtained results confirm the potential of deep learning in the signature recognition problem and open up a new research direction to improve the performance as well as the applicability in electronic authentication systems and online transactions.

The integration of mathematical formulas illustrating the calculation process in convolutional layers and LSTM shows the accuracy and transparency of the model, helping the research community to evaluate and reproduce experimental results. At the same time, the illustration of the preprocessing process and model architecture provides an intuitive view for readers.

Future research will focus on expanding the data, optimizing preprocessing steps, and integrating advanced models to address the remaining challenges, especially in real-world conditions with diverse signature patterns and complex noise.

Acknowledgements: This research is funded by the Electric Power University, Hanoi, Vietnam under the 2025 research project (DTKHCN.29/2025).

References

[1] Edward H. Freeman. (2004). Digital Signatures and Electronic Contracts. *September 2004, EDPACS 13(3):8-12*. DOI:10.1201/1086/44312.13.2.20040501/81647.2

[2] LeCun, Y., Bottou, L., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521(7553)*, 436-444.

[3] Şaban Öztürk, Bayram Akdemir. (2018). Effects of Histopathological Image Pre-processing on Convolutional Neural Networks. *International Conference on Computational Intelligence and Data Science (ICCIDS 2018). Procedia Computer Science 132 (2018)* 396–403

[4] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv:*1409.1556.

[5] https://en.wikipedia.org/wiki/Convolutional_neural_network (date accessed: 14.02.25).

[6] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780.

[7] Shivanand S. Gornale, Sathish Kumar, Abhijit Patil and Prakash S. Hiremath. (2021). Behavioral Biometric Data Analysis for Gender Classification Using Feature Fusion and Machine Learning. *Original research. May 2021 | Volume 8 | Article 685966*. www.frontiersin.org. doi: 10.3389/frobt.2021.685966

[8] Mohammad-Parsa Hosseini, Senbao Lu, Kavin Kamaraj, Alexander Slowikowski and Haygreev C. Venkatesh. (2020). Deep Learning Architectures. *Springer Nature Switzerland AG 2020. Deep Learning: Conceptsand Architectures, Studies in Computational Intelligence* 866. https://doi.org/10.1007/978-3- 030-31756-0_1

[9] Chollet, F. (2017). Deep Learning with Python. *Manning Publications*.

[10] Pattanapong Chantamit-o-pas, Madhu Lata Goyal. (2018). Long Short-Term Memory Recurrent Neural Network for Stroke Prediction. *Chapter in Lecture Notes in Computer Science*. DOI: 10.1007/978-3-319-96136-1_25

[11] Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization

[12] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. *MIT Press*.

[13] Pham, H., et al. (2019). Signature Verification using Deep Convolutional Neural Networks. *International Journal of Computer Vision, 127(3)*, 567-584.

[14] O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv*:1511.08458.

[15] Hung, T.B.; Tien, M.L. (2021). Facial expression recognition with CNN-LSTM. *In Research in Intelligent and Computing in Engineering. Advances in Intelligent Systems and Computing*. Volume 1254.

BI STAGE FSSP WITH TRIANGULAR INTUITIONISTIC FUZZY NUMBER AND COMPARISON BETWEEN HEURISTICS

Pooja Kaushik[1*], Sonia[2], Deepak Gupta[3], Sonia Goel[4*]

•

[1,2,3,4]Department of Mathematics, Maharishi Markandeshwar Engineering College, Maharishi Markandeshwar (Deemed to be University), Mullana (Ambala), Haryana, India
[1*]drtobekaushik2024@gmail.com, [2]dhimansonia13@gmail.com, [3]guptadeepak@yahoo.co.in, [4*]sonia.mangla14@gmail.com

Abstract

Scheduling theory focuses on the best way to distribute limited resources among tasks throughout time and allocating limited resources to tasks in the best possible way. This field's history begins when two people first fought over a resource and devised a strategy to divide it without resorting to violence. In real world applications, scheduling is a vital yet highly intricate issue. In industrial processes, managing production with the goal of reducing makespan is an essential duty. The processing time of each job on each machine has been taken to be a real number for the majority of scheduling problems studied already. However, processing times in real-world applications are frequently uncertain, meaning that human error or operational flaws may cause them to fluctuate easily. This study explores a two-stage FSSP, where parallel machines execute jobs in the both stages. The processing times in the flow shop scheduling are represented by fuzzy numbers in this work. This paper gives comparative study to address the FSSP under an uncertain environment based on triangular intuitionistic fuzzy numbers (TrIFNs). To accurately model uncertainty and imprecision in processing times, we employ TIFNs, which provide a more comprehensive representation of vagueness compared to traditional fuzzy sets. To find the best work sequence with the goal of minimizing the makespan, fuzzy and scheduling principles are applied to the flow shop scheduling challenges. The main goal is to optimize the flow time by comparing the efficacy of three distinct optimization approaches. A rigorous computational analysis is conducted to evaluate their performance, leveraging fuzzy arithmetic operations and defuzzification techniques. The three heuristics' performances are compared and their performance is analyzed. An analysis of the performance of three heuristics i.e. B&B, Dannenbring exhaustive search and the division algorithm proposed by santhi in previous paper has been conducted. The findings offer valuable insights into the applicability of fuzzy scheduling methodologies in complex manufacturing environments, contributing to the broader discourse on uncertainty modelling in production systems. The comparison is shown by a numerical illustration. This study provides a comparison and evaluation of the scheduling problem which also helps the reader comprehend various environmental assumptions, system limitations, and useful heuristics for further research.

Keywords: flow time (makespan), triangular intuitionistic fuzzy number, flow shop scheduling, transportation technique, B&B, Dannenbring exhaustive search

I. Introduction

Scheduling is a fundamental decision-making process in operation research and industrial management, where tasks or jobs are assigned to resources such as machines, workers, or computers over a specified time horizon. The primary goal of scheduling is to optimize performance measures such as minimizing makespan, reducing job tardiness, improving resource utilization, and balancing workloads. Scheduling problems are encountered in various domains, including manufacturing, logistics, healthcare, and computing, making it a critical area of study. Depending on the nature of the system, scheduling problems can be categorized into single-machine, parallel-machine, job shop, open shop, and flow shop scheduling, each presenting unique challenges and optimization criteria. Among these, flow shop scheduling is a widely studied class of scheduling problems where jobs follow the same sequence of machines. In a typical flow shop, a set of jobs is processed across multiple machines in the same order, ensuring a structured workflow. The classical permutation FSSP restricts job sequences to remain unchanged across all machines, making it a constrained but well-defined problem. The objective in flow shop scheduling often revolves around minimizing total flow time, reducing machine idle time, or improving energy efficiency.

The two-machine FSSP was first introduced [1], which provided an optimal sequencing rule to minimize makespan. However, as the number of machines and complexity increase, the problem becomes NP-hard, requiring heuristic and metaheuristic approaches for efficient solutions. This flexibility increases practical applicability in real-world industries, such as automobile manufacturing, semiconductor production, and textile industries. Traditional deterministic models assumed fixed processing times, but uncertainties in real-world systems led to the adoption of fuzzy set theory [2]. This study considered uncertainties in real-life scheduling scenarios which led to the development of fuzzy FSSP, where processing times are represented as fuzzy numbers rather than deterministic values. This approach allows for handling imprecision and variability in job execution times, making scheduling more robust against disruptions.

FSSPs have been widely studied due to their applications in manufacturing and production systems. Work on two-stage FSSP was done[3] and later optimized[4] providing an optimal sequencing rule for makespan minimization. This work on sequencing and scheduling laid the foundation for understanding how to efficiently allocate tasks over time while optimizing various performance measures and focused on job sequencing (determining the order of tasks) and scheduling (assigning tasks to resources over time) in manufacturing and service systems. The study [5] provides an early evaluation of flow shop sequencing heuristics. This foundational work introduced and assessed heuristic techniques, which have influenced modern scheduling methods. The study's insights remain relevant as a benchmark for contemporary heuristic development.

The work [6] introduced a method to rank fuzzy subsets which is useful in scenarios where decision making involves imprecise or uncertain information and presented a method for ranking fuzzy subsets of the unit interval. Since fuzzy sets lack a strict ordering due to their membership functions, traditional ranking methods are inadequate. Yager introduced an approach based on integral and expectation-like computations to assign a numerical index to fuzzy subsets, facilitating comparison. However, the study [7] proposed intuitionistic fuzzy sets (IFSs) incorporating both membership and non-membership degrees, making them more effective in handling hesitation in scheduling. The study [8] addressed scheduling challenges in two machine flow shop where job processing times are uncertain and represented them as triangular fuzzy number.

The work [9] explored the problem of FSSP with two machines and a single transport facility under a fuzzy environment and focused on optimizing job scheduling while considering uncertain setup times, which are modeled using fuzzy logic. Unlike traditional approaches that assume fixed setup times, this method incorporates fuzzy numbers to handle variability, making the model more applicable to real-world industrial settings. The extension of the Dannenbring method to

optimize flow shop-type production while considering job waiting times and weights has been considered [10]. Their study suggests modifications or enhancements to the original heuristic, making it more applicable to modern production environments. A novel method incorporating triangular intuitionistic fuzzy numbers to address scheduling uncertainties was proposed [11]. This approach aligns with recent trends in fuzzy logic applications, aiming to enhance decision-making in complex, uncertain scheduling environments.

Also, the study [12] proposed an optimization model for a two-stage FSSP with equipotential machines at the first stage. Their approach likely incorporates mathematical optimization or heuristic techniques to enhance scheduling efficiency. The researchers [13] presented B&B technique for solving two-stage FSSPs. The B&B method is a well-known exact optimization approach, and this study extends its application to scenarios with equipotential machines, ensuring efficient job allocation across stages. The work [14] compared heuristic methods, particularly NEH (Nawaz-Enscore-Ham) and CDS (Campbell-Dudek-Smith), within a fuzzy environment and assessed their performance in bi-stage flow shop scheduling, highlighting the trade-offs between computational efficiency and optimality in uncertain conditions.

II.Preliminaries

I.Fuzzy Set

A fuzzy set, a mathematical representation of a set that allows for vagueness and uncertainty, is defined in equation (1).

$$\chi = \{(y, \mu_{\chi}(y)) \mid y \in Y\} \tag{1}$$

where χ is the fuzzy set, Y is the set of all possible elements, $\mu_{\chi}(y)$ is the membership function, which assigns a degree of membership to each element y in Y.

II.Fuzzy Number

A fuzzy number is a mathematical representation of a numerical value that is uncertain or imprecise, is defined as in equation (2):

$$\chi = \{(y, \mu_{\chi}(y)) \mid y \in \mathbb{R}\} \tag{2}$$

where, χ is the fuzzy number, $\mathbb{R}$ is the set of real numbers, $\mu_{\chi}(y)$is membership function, which assigns a degree of membership to each real number y.

III. Triangular Fuzzy Number

A fuzzy number Å = (à, ß, ç) is called triangular fuzzy number (graphically shown in figure 1) if its membership function satisfy the following conditions

- à to ß is an increasing function
- ß to ç is decreasing function
- à≤ ß≤ ç

and is given by equation (3).

$$\mu_{Å}(x)=\begin{cases} 0 & for x < à \\ \frac{x-à}{ß-à} for\ à \le x \le ß \\ \frac{ç-x}{ç-ß} for\ ß \le x \le ç \\ 0 & for x > ç \end{cases} \tag{3}$$

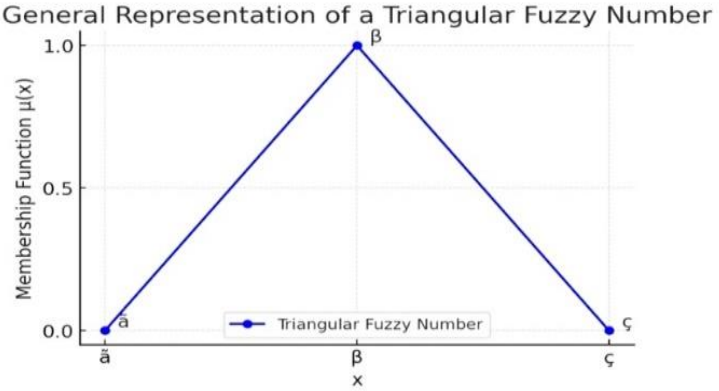


Figure 1: *Graph of triangular fuzzy number*

IV. Intuitionistic Fuzzy Number (IFN)

An IFN is a mathematical representation of a numerical value that is uncertain or imprecise, and also takes into account the degree of hesitation or uncertainty. It is defined as:

$$\chi = \{(y, \mu_\chi(y), \nu_\chi(y)) \mid y \in \mathbb{R}\} \quad (4)$$

where χ is the intuitionistic fuzzy number, $\mathbb{R}$ is the set of real numbers, $\mu_\chi(y)$ is membership and $\nu_\chi(y)$is non-membership function assigning a degree of membership and non-membership to each real number y respectively.

V. Triangular Intuitionistic Fuzzy Number (TrIFN)

An IFN G = $(g_1, g_2, g_3)(\acute{g}_1, g_2, \acute{g}_3)$ is said to be TrIFN if itsmembership and the non-membership function are given in equation (5) and graphically presented in figure 2.

$$\mu_G(y) = \begin{cases} \frac{y-g_1}{g_2-g_1}, & if\, g_1 \le y \le g_2 \\ \frac{g_3-y}{g_3-g_2}, & if\, g_2 \le y \le g_3 \\ 0, & otherwise \end{cases}$$

$$\nu_G(y) = \begin{cases} \frac{g_2-y}{g_2-\acute{g}_1}, & if\, \acute{g}_1 \le y \le g_2 \\ \frac{y-g_2}{\acute{g}_3-g_2}, & if\, g_2 \le y \le \acute{g}_3 \\ 1, & otherwise \end{cases} \quad (5)$$

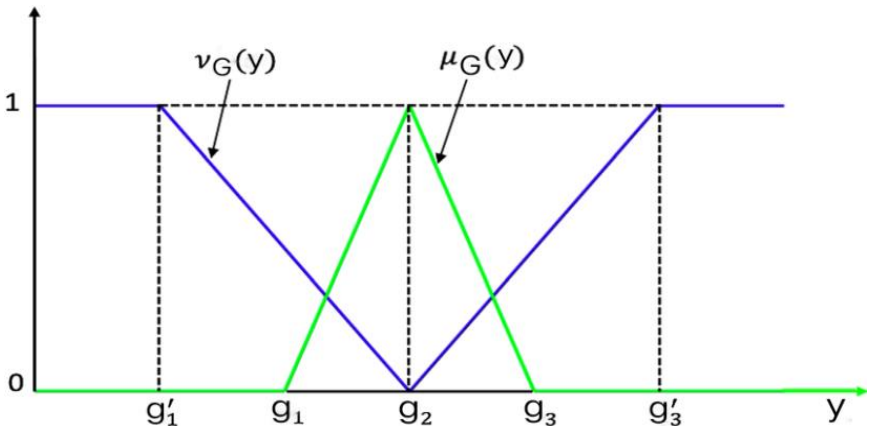


Figure 2: *Graph of triangular Intuitionistic fuzzy number*

VI.Arithmetic operators

Let G = (g_1,g_2,g_3);$(\acute{g}_1,g_2,\acute{g}_3)$ and H= (h_1,h_2,h_3);$(\acute{h}_1,h_2,\acute{h}_3)$ be two TrIFNs then

- Addition Let G+H = $(g_1+h_1,g_2+h_2,g_3+h_3)$ $(\acute{g}_1+\acute{h}_1,g_2+h_2,\acute{g}_3+\acute{h}_3)$
- Subtraction Let G-H = $(g_1-h_1,g_2-h_2,g_3-h_3)$ $(\acute{g}_1-\acute{h}_1,g_2-h_2,\acute{g}_3-\acute{h}_3)$

VII.Ranking of TrIFN

The ranking of TrIFN is defined as in equation (6)

$$R(A)=\frac{1}{3}\left[\frac{(\acute{g}_3-g_1)(g_2-2\acute{g}_3-2\acute{g}_1)+(g_3-g_1)(g_1+g_2+g_3)+3(\acute{g}_3{}^2-\acute{g}_1{}^2)}{(\acute{g}_3-\acute{g}_1+g_3-g_1)}\right] \quad (6)$$

If $R.G.\leq R.H.$, then $G\leq H$.

III. Problem Formulation

Assume that three parallel equipotential machines are to be used for both stages (i=1,2,3……, n; j=1,2,3……, m). Every job's processing time is taken as TrIFN. The available time and unit operating cost of each parallel machine are also provided. The following is the problem's model in matrix form, represented in table 1.

Table 1: *Mathematical model of the problem*

Job	Processor $\widehat{U}$					Processor $\widehat{V}$				
i	$\widehat{U}_1$	$\widehat{U}_2$	$\widehat{U}_3$	$\widehat{U}_n$	Processing time of$\widehat{U}$	$\widehat{V}_1$	$\widehat{V}_2$	$\widehat{V}_3$	$\widehat{V}_m$	Processing time of $\widehat{V}$
1	u_{11}	u_{12}	u_{13}	u_{1n}	(c_1,d_1,e_1);$(\hat{c}_1,d_1,\hat{e}_1)$	v_{11}	v_{12}	v_{13}	v_{1n}	(g_1,h_1,k_1);$(\hat{g}_1,h_1,\hat{k}_1)$
2	u_{21}	u_{22}	u_{23}	u_{2n}	(c_2,d_2,e_2);$(\hat{c}_2,d_2,\hat{e}_2)$	v_{21}	v_{22}	v_{23}	v_{2n}	(g_2,h_2,k_2); $(\hat{g}_2,h_2,\hat{k}_2)$
3	u_{31}	u_{32}	u_{33}	u_{3n}	(c_3,d_3,e_3); $(\hat{c}_3,d_3,\hat{e}_3)$	v_{31}	v_{32}	v_{33}	v_{3n}	(g_3,h_3,k_3); $(\hat{g}_3,h_3,\hat{k}_3)$
4	u_{41}	u_{42}	u_{43}	u_{4n}	(c_4,d_4,e_4); $(\hat{c}_4,\hat{d}_4,\hat{e}_4)$	v_{41}	v_{42}	v_{43}	v_{4n}	(g_4,h_4,k_4); $(\hat{g}_4,h_4,\hat{k}_4)$
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
n	u_{n1}	u_{n2}	u_{n3}	u_{nn}	(c_n,d_n,e_n); $(\hat{c}_n,d_n,\hat{e}_n)$	v_{n1}	v_{n1}	v_{n1}	v_{nm}	(g_n,h_n,k_n); $(\hat{g}_n,h_n,\hat{k}_n)$
t_{pj}	t_{11}	t_{12}	t_{13}	t_{1n}		t_{21}	t_{22}	t_{23}	t_{2m}	

I. Notations

(c_n,d_n,e_n);$(\hat{c}_n,d_n,\hat{e}_n)$: Processing time of n^{th} job on processor $\widehat{U}$
(g_n,h_n,k_n);$(\hat{g}_n,h_n,\hat{k}_n)$: Processing time of n^{th} job on processor $\widehat{V}$

$\widehat{U}_n$: n equipotential machines of processor $\widehat{U}$
$\widehat{V}_m$: n equipotential machines of processor $\widehat{V}$
t_{pj}: Transportation time of moving jobs from machine A to B
u_{ij}: unit cost of i^{th}job on $\widehat{U}_j{}^{th}$parallel machine
v_{ij}: unit cost of i^{th}job on $\widehat{V}_j{}^{th}$parallel machine

II. Assumptions

- Each device operates the same set of tasks.
- Each job is processed on machine B once it has been finished on machine A.
- All machines can start simultaneously.
- The operational costs of each type A and type B machine vary.
- Machines are always operational and never experience downtime or failures, ensuring continuous availability for scheduling.
- Once a task begins, it must be finished without interruption.

III. Algorithm

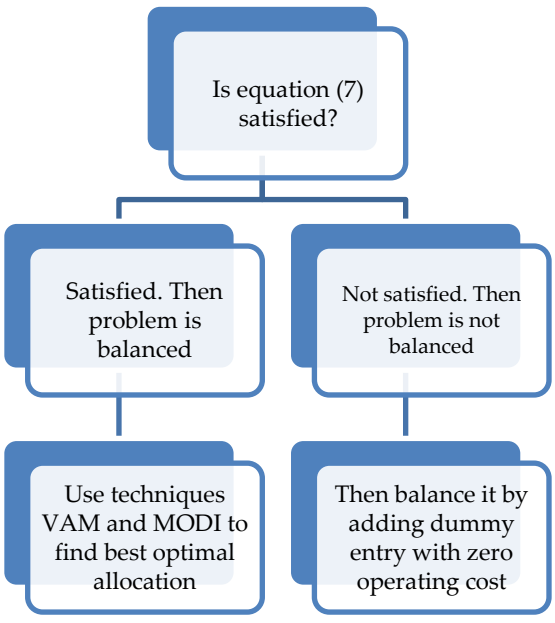
The following course of action can be used to solve the FSSP given in table 1. Using the ranking formula, we convert the triangular intuitionistic fuzzy utilisation time into a crisp value in both phases. Next, in order to determine the best allocation, we use transportation techniques such as VAM. The MODI method can then be used to optimize the working time of each job on the parallel equipotential machines and to verify that the solution reached is optimal. The best sequence is then determined using B&B.

Step 1: Use ranking rule given in equation (6) to defuzzify the prosessing time taken as TrIFNs into crisp values.

Step 2: Apply transportation technique like VAM and MODI. To apply any transportation technique, we examine the constraint

$$\sum_{j=1}^{3} t_{1j} = \sum_{i=1}^{n} d_i \quad (7)$$

To check that the problem is balanced or not. Here come two situations:



Step 3: Use branch and bound method by applying the rule

$$g^{'}=\max(\sum_{i=1}^{n} U_{ij}\}_{j=1,2..m}+\min_{i\in J_k^{'}}\{\max_{j=1,2..p} V_{ij})\&g^{''}=\max_{i\in J_k}\{U_{ij}\}_{j=1,2..m}+\max\{\sum_{i}^{n} V_{ij}\}_{j=1,2..p}$$

Then evaluate $G = \max(g^{'}, g^{''})$ for all the jobs.

Find the minimum of all values of G and start from that vertex as the first job in the optimal subsequence. Repeat the above process till we reach termination point of branch which will give us the most ideal optimal sequence of jobs.

IV.Numerical problem

Three machines $\hat{U}_1, \hat{U}_2$ and $\hat{U}_3$ of type $\hat{U}$ and three machines $\hat{V}_1, \hat{V}_2$ and $\hat{V}_3$ of type $\hat{V}$ are accessible for various amounts of time and have differing running expenses, as presented in table 2.

Table 2: *Numerical problem*

Job	Processor $\hat{U}$				Processor $\hat{V}$			
i	$\hat{U}_1$	$\hat{U}_2$	$\hat{U}_3$	Processing time of $\hat{U}$	$\hat{V}_1$	$\hat{V}_2$	$\hat{V}_3$	Processing time of $\hat{V}$
1	12	15	13	(10,11,12) ;(8,11,13)	11	12	8	(6,8,9); (5,7,10)
2	9	11	10	(9,10,11) ;(8,10,12)	9	11	7	(10,11,12);(8,11,13)
3	9	7	11	(11,12,13);(10,12,15)	7	9	10	(8,9,11); (7,9,12)
4	13	8	8	(11,12,13);(10,12,15)	15	12	13	(6,7,8); (5,7,9)
5	10	12	9	(12,14,15);(11,14,16)	10	9	12	(8,9,10); (7,9,11)
Available time	19	21.5	18.1		15	12.7	16	

Solution:

Step 1: Changing the TrIFN into crisp value using the ranking formula given in equation (6)

(10,11,12) ;(8,11,13) changed into 10.7 In similar fashion (6,8,9); (5,7,10) changed into 7.5

(9,10,11) ;(8,10,12) changed into 10 (10,11,12);(8,11,13) changed into 10.8

(11,12,13);(10,12,15) changed into 12 (8,9,11); (7,9,12) changed into 9.4

(11,12,13);(10,12,15) changed into 12.2 (6,7,8); (5,7,9) changed into 7

(12,14,15);(11,14,16) changed into 13.7 (8,9,10); (7,9,11) changed into 9

and the changed crisp values are given in table 3.

Table 3: *Reduced problem after defuzzifying TrIFN into crisp value*

Job	Processor $\hat{U}$				Processor $\hat{V}$			
i	$\hat{U}_1$	$\hat{U}_2$	$\hat{U}_3$	Processing time of $\hat{U}$	$\hat{V}_1$	$\hat{V}_2$	$\hat{V}_3$	Processing time of $\hat{V}$
1	12	15	13	10.7	11	12	8	7.5
2	9	11	10	10	9	11	7	10.8
3	9	7	11	12	7	9	10	9.4
4	13	8	8	12.2	15	12	13	7
5	10	12	9	13.7	10	9	12	9
Available time	19	21.5	18.1		15	12.7	16	

Step 2: We determine the best way to distribute processing time by applying transportation technique VAM first and then MODI which is given in table 4&5.

Table 4 :*Optimal distribution of processing time on machine* $\widehat{U}_1, \widehat{U}_2$ *and* $\widehat{U}_3$

i	$\widehat{U}_1$	$\widehat{U}_2$	$\widehat{U}_3$
1	9	0	1.7
2	10	0	0
3	0	12	0
4	0	9.5	2.7
5	0	0	13.7

Table 5: *Optimal distribution of processing time on machine* $\widehat{V}_1, \widehat{V}_2$*and* $\widehat{V}_3$

i	$\widehat{V}_1$	$\widehat{V}_2$	$\widehat{V}_3$
1	0	0	7.5
2	2.3	0	8.5
3	9.4	0	0
4	0	7	0
5	3.3	5.7	0

After applying transportation techniques VAM and MODI, the time is allocated on machines and the reduced problem is given in table 6.

Table 6: *Reduced problem's tabular representation*

Job	Processor $\widehat{U}$				Processor $\widehat{V}$			
i	$\widehat{U}_1$	$\widehat{U}_2$	$\widehat{U}_3$	Processing time of $\widehat{U}$	$\widehat{V}_1$	$\widehat{V}_2$	$\widehat{V}_3$	Processing time of $\widehat{V}$
1	9	0	1.7	10.7	0	0	7.5	7.5
2	10	0	0	10	2.3	0	8.5	10.8
3	0	12	0	12	9.4	0	0	9.4
4	0	9.5	2.7	12.2	0	7	0	7
5	0	0	13.7	13.7	3.3	5.7	0	9

Now we apply Branch and bound (B&B) algorithm and calculate the lower bound of jobs to obtain the sequence of jobs for optimal solution. Branch and bound algorithm is applied to further optimize the allocation obtained after the application of VAM and MODI.

Table 7: *Depicting lower bound of jobs*

(i)	$g^{'}$ =max $(\sum_{i=1}^{n} K_{ij}) + \min_{i\in J_k}(p_i)$	$g^{''} = \max_{i\in J_k} K_{ij}+\sum_{i}^{n} p_i$	G=max $\{g^{'}, g^{''}\}$
1	max $(\sum_{i=1}^{n} K_{ij}) + \min_{i\in\{2,3,4\}}(p_i)$	$\max_{i\in\{2,3,4\}} K_{ij}+\sum_{i}^{n} p_i$	27.2
	=21.5+5.7=27.2	=9+16=25	
2	max $(\sum_{i=1}^{n} K_{ij}) + \min_{i\in\{1,3,4\}}(p_i)$	$\max_{i\in\{1,3,4\}} K_{ij}+\sum_{i}^{n} p_i$	27.2
	=21.5+5.7=27.2	=10+16=25	
3	max $(\sum_{i=1}^{n} K_{ij})+ \min_{i\in\{1,2,4\}}(p_i)$	$\max_{i\in\{1,2,4\}} K_{ij}+\sum_{i}^{n} p_i$	28
	=21.5+5.7=27.2	=12+16=25	
4	max $(\sum_{i=1}^{n} K_{ij})+ \min_{i\in\{1,2,3\}}(p_i)$	$\max_{i\in\{1,2,3\}} K_{ij}+\sum_{i}^{n} p_i$	27.2
	=21.5+5.7=27.2	=9.5+16=25	
5	max $(\sum_{i=1}^{n} K_{ij})+ \min_{i\in\{1,2,3\}}(p_i)$	$\max_{i\in\{1,2,3\}} K_{ij}+\sum_{i}^{n} p_i$	29.7
	=21.5+5.7=27.2	=13.7+16=25	

Hence from table 7, lower bound is 27.2 corresponding to job 1,2 and 4.Without loss of generality let us fix job 1 at 1^{st} place. Now we continue the same process and determine the nodes for the second branch.This process is continued till all jobs gets place in the optimal order.

Table 8: *Depicting lower bound of jobs when job 1 is fixed at first place*

(i)	$g^{'}$ =max $(\sum_{i=1}^{n} K_{ij}) + \min_{i\in J_k}(p_i)$	$g^{''} = \max_{i\in J_k} K_{ij}+\sum_{i}^{n} p_i$	G=max $\{g^{'}, g^{''}\}$
12	max $(\sum_{i=1}^{n} K_{ij}) + \min_{i\in\{2,3,4\}}(p_i)$	$\max_{i\in\{2,3,4\}} K_{ij}+\sum_{i}^{n} p_i$	43.5
	=21.5+5.7=27.2	=27.5+16=43.5	
13	max $(\sum_{i=1}^{n} K_{ij}) + \min_{i\in\{1,3,4\}}(p_i)$	$\max_{i\in\{1,3,4\}} K_{ij}+\sum_{i}^{n} p_i$	37.4
	=21.5+5.7=27.2	=21.4+16=37.4	
14	max $(\sum_{i=1}^{n} K_{ij})+ \min_{i\in\{1,2,4\}}(p_i)$	$\max_{i\in\{1,2,4\}} K_{ij}+\sum_{i}^{n} p_i$	32.5
	=21.5+5.7=27.2	=16.5+16=32.5	
15	max $(\sum_{i=1}^{n} K_{ij})+ \min_{i\in\{1,2,3\}}(p_i)$	$\max_{i\in\{1,2,3\}} K_{ij}+\sum_{i}^{n} p_i$	37.1
	=21.5+7=28.5	=21.1+16=37.1	

Lower bound comes to be 32.5 as shown in table 8 which is corresponding to subsequence 14. Therefore job 4 is fixed at2^{nd} place. We continue searching for the next task in the ideal order.

Table 9: *Depicting lower bound of jobs when job 1 is fixed at first place and 4 at second place*

(i)	$g^{'}$ =max $(\sum_{i=1}^{n} K_{ij}) + \min_{i\in J_k}(p_i)$	$g^{''} = \max_{i\in J_k} K_{ij}+\sum_{i}^{n} p_i$	G=max $\{g^{'}, g^{''}\}$
142	max $(\sum_{i=1}^{n} K_{ij}) + \min_{i\in\{2,3,4\}}(p_i)$	$\max_{i\in\{2,3,4\}} K_{ij}+\sum_{i}^{n} p_i$	34.5
	=21.5+5.7=27.2	=18.5+16=34.5	
143	max $(\sum_{i=1}^{n} K_{ij}) + \min_{i\in\{1,3,4\}}(p_i)$	$\max_{i\in\{1,3,4\}} K_{ij}+\sum_{i}^{n} p_i$	46.9
	=21.5+5.7=27.2	=30.9+16=46.9	
145	max $(\sum_{i=1}^{n} K_{ij})+ \min_{i\in\{1,2,4\}}(p_i)$	$\max_{i\in\{1,2,4\}} K_{ij}+\sum_{i}^{n} p_i$	39.8
	=21.5+7=28.5	=23.8+16=39.8	

Lower bound comes to be 34.5 which is corresponding to subsequence 142 given in table 9. Therefore, job 2 is fixed at third place. Now proceeding in the similar fashion to find the next job.

Table 10:*Depicting lower bound of jobs when job 1,4 and 2 is fixed at first, second and third placerespectively.*

(i)	$g^{'}$ =max $(\sum_{i=1}^{n} K_{ij}) + \min_{i\in J_k}(p_i)$	$g^{''} = \max_{i\in J_k} K_{ij}+\sum_{i}^{n} p_i$	G=max $\{g^{'}, g^{''}\}$
1423	max $(\sum_{i=1}^{n} K_{ij}) + \min_{i\in\{2,3,4\}}(p_i)$	$\max_{i\in\{2,3,4\}} K_{ij}+\sum_{i}^{n} p_i$	37.4
	=21.5+5.7=27.2	=21.4+16=37.4	
1425	max $(\sum_{i=1}^{n} K_{ij}) + \min_{i\in\{1,3,4\}}(p_i)$	$\max_{i\in\{1,3,4\}} K_{ij}+\sum_{i}^{n} p_i$	35.4
	=21.5+9.4=30.9	=19.4+16=35.4	

From table 10, min.{G}=35.4 which is related to the subsequence 1425. So first, second, third and fourth place in the required optimal sequence is acquired by job 1,4,2,5 respectively. Also, by default job 3 will take the last i.e. fifth place. Therefore {1,4,2,5,3} is the required optimal sequence. Best feasible solution is given in table 11 .

Table 11: *In-out table*

Job	Processor $\hat{U}$			Processor $\hat{V}$		
i	$\hat{U}_1$	$\hat{U}_2$	$\hat{U}_3$	$\hat{V}_1$	$\hat{V}_2$	$\hat{V}_3$
1	0-9	0	0-1.7	0	0	9-16.5
4	0	0-9.5	1.7-4.4	0	9.5-16.5	0
2	9-19	0	0	19-21.3	0	19-27.5
5	0	0	4.4-18.1	21.3-24.6	18.1-23.8	0
3	0	9.5-21.5	0	24.6-34	0	0

V. Gantt Chart

The solution obtained in table 11 is shown by Gantt chart in figure 3.

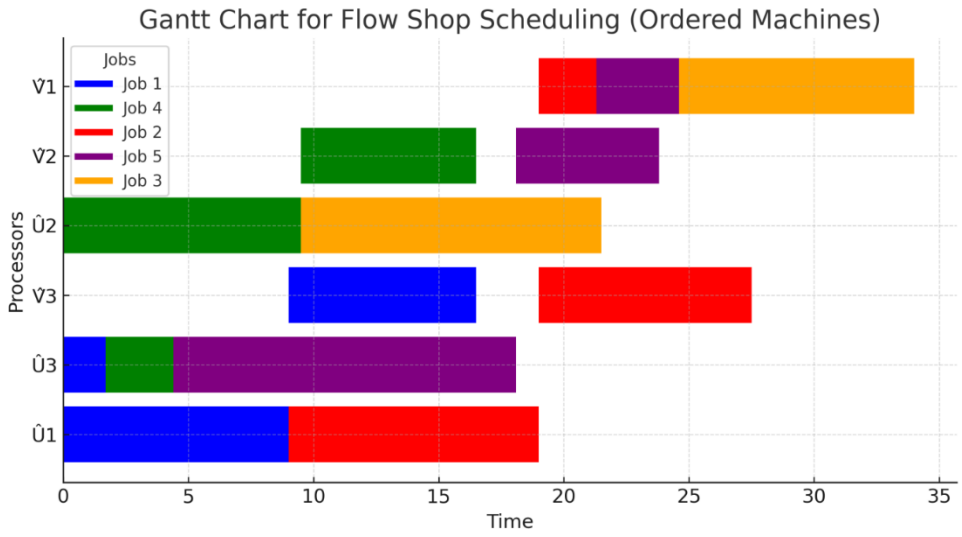


Figure 3: *Gantt chart*

The Gantt Chart as shown in figure 3, helps in scheduling, tracking progress, and ensuring efficient resource. It helps to define tasks, deadlines, resource allocation, organizes work in a logical sequence and tracks completed, ongoing, and pending tasks allocation.
Here

- Each row represents a task in the project.
- The horizontal axis shows time.
- The length of each bar indicates the task duration.
- Colour coding shows task completion and different colours indicating different jobs.

VI. The Dannenbring Method

Dannenbring Method is a heuristic approach for solving the FSSP. It was proposed by D.G. Dannenbring in 1977 as an improvement over traditional heuristics like Johnson's rule, aiming to minimize total time to complete all jobs. Johnson method (1954) served as the basis for the Dannenbring approach. Scheduling issues with more than two machines can be resolved using this approach [Panneerselvan 1999]. The method prioritizes job sequencing based on mean processing times and positional weight calculations. It is particularly useful for FSSPsincluding machines more than two where Johnson'srule may not be optimal.
The Dannenbring method (1977) gives the following formula in equation (8)

$$Þ_{i1} = \sum_{i=1}^{m} (m - i + 1) * t_{ij}$$
$$Þ_{i2} = \sum_{i=1}^{m} i * t_{ij} \quad (8)$$

Steps of the Dannenbring Method:

- Compute the average processing time of each job across all machines.
- Compute Positional Weights: Positional weight is calculated to balance job placement in the sequence, considering how a job affects the overall makespan.
- Rank Jobs Based on Weight & Priority Rules: Jobs with higher priority (based on mean time and weight) are placed in the sequence to optimize flow.
- Sequence Jobs Accordingly: The method assigns jobs in a way that balances bottleneck and reduces idle time between machines.

Solution:

First step in the Dannenbring approach is obtaining the expected values for every job given in table 12.

Table12: *Expected values of each job*

Job	Processor $\widehat{U}$	Processor $\widehat{V}$
1	10.7	7.5
2	10	8.5
3	12	9.4
4	12.2	7
5	13.7	5.7

Now, we find the values of $Þ_{i1}$ and $Þ_{i2}$.

Table13: *Positional weight of jobs*

Job	Processor $\widehat{U}$	Processor $\widehat{V}$
1	36.4	25.7
2	37	27
3	42.8	30.8
4	38.4	26.2
5	38.8	25.1

From table 13, the optimal order is 3-2-4-1-5.
Table 14 gives in-out table and optimal flow time.

Table14: *Makespan Value*

Job	Processor $\widehat{U}$			Processor $\widehat{V}$		
i	$\widehat{U}_1$	$\widehat{U}_2$	$\widehat{U}_3$	$\widehat{V}_1$	$\widehat{V}_2$	$\widehat{V}_3$
3	0	0-12	0	12-21.4	0	0
2	0-10	0	0	21.4-23.7	0	10-18.5
4	0	12-21.5	0-2.7	0	21.5-28.5	0
1	10-19	0	2.7-4.4	0	0	19-26.5
5	0	0	4.4-18.1	23.7-27	28.5-34.2	0

Therefore, makespan from the optimal sequence comes out to be 34.2

VII. Division Algorithm (Method proposed by Selvakumari and S. Santhi [11])

Step 1: Convert TrIFN into crisp value.
Step 2: Divide all entries of rows by row's highest value and place that value in the top right corner of each cell.
Step 3: Divide all the entries of the columns by column's highest value and place that value in the bottom left corner of each cell.
Step 4: Add both the numbers obtained in step 2&3 and write in a table.
Step 5: Choose the row corresponding to minimum value, mark that job as first job in optimal order and remove the row.
Step 6: Repeat the process and obtain the optimal order.

Solution: Defuzzified values are given the table 15.

Table15: *Problem after defuzzification*

Job	Processor $\widehat{U}$	Processor $\widehat{V}$
1	10.7	7.5
2	10	8.5
3	12	9.4
4	12.2	7
5	13.7	5.7

In table 15, divide all the entries of each row by the row's highest value and place that value in the top right corner of each cell.
Similarly, divide all the entries of the columns by column's highest value and place that value in

the bottom left corner of each cell.

Allocate these calculated effective values to each cell and place them as shown in table 16.

Table16: *Effective value assigned to jobs*

Job	Processor $\hat{U}$		Processor $\hat{V}$	
		1		0.7
1	10.7		7.5	
	0.78		0.79	
		1		0.85
2	10		8.5	
	0.73		0.90	
		1		0.78
3	12		9.4	
	0.87		1	
		1		0.57
4	12.2		7	
	0.89		0.74	
		1		0.42
5	13.7		5.7	
	1		0.61	

Add both the numbers obtained in the table 16. Then choose the row corresponding to minimum value, marking that job as first job in optimal order and remove the row. Repeat the process and obtain the optimal order.

Table17: *Total effective values of each job*

Job	Processor $\hat{U}$	Processor $\hat{V}$
1	1.78	1.49
2	1.73	1.75
3	1.87	1.78
4	1.89	1.31
5	2	1.03

In table 17, value in the fifth row is lowest. Therefore, delete the 5^{th} row and place 5th job at first place in optimal sequence as S_5. Proceeding in this manner, optimal order obtained is S_5>S_4>S_1>S_2>S_3 given in table 18.

Table 18: *In-Out table*

Job	Processor $\hat{U}$			Processor $\hat{V}$		
i	$\hat{U}_1$	$\hat{U}_2$	$\hat{U}_3$	$\hat{V}_1$	$\hat{V}_2$	$\hat{V}_3$
5	0	0	0-13.7	13.7-17	13.7-19.4	0
4	0	0-9.5	13.7-16.4	0	16.4-23.4	0
1	0-9	0	16.4-18.1	0	0	18.1-25.6
2	9-19	0	0	19-21.3	0	25.6-34.1
3	0	9.5-21.5	0	21.5-30.9	0	0

Therefore, minimum flow time is 34.1

VIII. Practical Situation

A pharmaceutical company is developing three new drugs (D1, D2, D3) for different medical conditions.Since processing times in both stages involve uncertainty (due to variations in formulation success rates, lab conditions, patient recruitment, and regulatory delays), they are modeled using Triangular Intuitionistic Fuzzy Numbers (TrIFNs). The development process consists of two main stages:

Stage 1: Drug Formulation & Initial Testing
The drug is formulated in the lab, and its chemical composition is tested. Processing times for different formulations are uncertain due to variability in raw materials, laboratory conditions, and expert evaluations. These uncertain processing times are modeled using Triangular Intuitionistic Fuzzy Numbers (TIFNs) to capture both hesitation and imprecision in expert assessments.

Stage 2: Clinical Trials & Regulatory Approval
Once initial tests are passed, the drug enters clinical trials, followed by regulatory approval. The duration of trials depends on patient responses, ethical approvals, and unexpected side effects, introducing further uncertainty in scheduling.
The aim is to optimizetotal time for all drugs to pass through both stages while considering uncertainty in processing times. Since Stage 2 depends on Stage 1 completion, delays in Stage 1 affect the overall timeline.

IX. Comparison between B&B, Dannenbring Method and Divison Algorithm

B&B is a systematic search method that explores solution spaces using a tree structure, pruning branches that cannot yield optimal solutions. This significantly reduces computational complexity and makes it suitable for large-scale problems like the Traveling Salesman Problem (TSP), knapsack, and job scheduling. In contrast, Dannenbring's Exhaustive Search evaluates all possible solutions without pruning, ensuring that the optimal solution is found but at the cost of high computational expense. This method becomes impractical for large problems due to its exponential time complexity and high memory requirements. Division algorithm offers a pragmatic solution by incorporating fuzzy logic to handle imprecise data,near optimal solutions suitable for complex and real world applications.

X. Future Scope

Future studies should integrate deep learning, reinforcement learning, and Type-2 fuzzy systems to improve large-scale scheduling under extreme uncertainty. The growing adoption of metaheuristic techniques, fuzzy ranking models, and hybrid optimization approaches indicates that TrIFN-based scheduling will continue to evolve, addressing more complex industrial challenges.

Conflicts of interests: The authors declare that there are no conflicts of interest.

References

[1] Johnson, S. M. (1954). Optimal two- and three-stage production schedules with setup times included. *Naval Research Logistics Quarterly*, 1:61–68.

[2] Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8:338–353.

[3] Gupta, J. N. D. (1971). A functional heuristic algorithm for the flow shop scheduling problem. *Operations Research*, 19:343–351.

[4] Panwalkar, S. S., Dudek R. A., and Smith M. L. (1973). Sequencing research and the industrial scheduling problem. *Symposium on the Theory of Scheduling and its Application, 86:29-38.*

[5] Dannenbring, D. G. (1977). An evaluation of flow shop sequencing heuristics. *Management Sci*., 23: 1174-1182.

[6] Yager R. R. (1981). A procedure for ordering fuzzy subsets of the unit interval. *Information Sciences*, 24:143–61.

[7] Atanassov, K.T. and Atanassov, K.T. (1999). Intuitionistic fuzzy sets. *Physica-Verlag HD*, 35:1-137.

[8] Sanuja, P., Xueyan, S. (2006). A new approach to two machine flow shop problem with uncertain processing time. *Optimization and Engineering*, 7:329–43.

[9] Gupta, D., Sharma, S., and Aggarwal, S. (2013). Flow shop scheduling on 2-machines with setup time and single transport facility under fuzzy environment. *Opsearch*, 50: 14-24.

[10] Rahmani, U., Wulan, E.R., Huda, A.F., Supian, S. and Subiyanto, S. (2019). Scheduling of flowshop Type Production with Dannenbring Method to Obtain Optimal Results with Observing Waiting Time and Weight of Jobs. *An International Scientific Journal*, 127:212-224.

[11] Selvakumari, K., and Santhi, S. (2020). A new approach for solving flow shop scheduling problem with triangular intuitionistic fuzzy number. *Advances in Mathematics: Scientific Journal*, 9:5815-5825.

[12] Gupta, D.,Goel, S., and Mangla, N. (2022). Optimization of production scheduling in two stage flow shop scheduling problem with m equipotential machines at first stage. *International Journal of System Assurance Engineering and Management*, 13:1162-1169.

[13] Gupta, D., and Goel, S. (2022). Branch and bound technique for two stage flow shop scheduling model with equipotential machines at every stage. *International Journal of Operational Research*, 44: 462-472.

[14] Kaushik, P., Gupta D. And Goel, S. (2025). Comparative Study Of B&B With Heuristics NEH And CDS For Bi Stage Flow Shop Scheduling Problem Under Fuzzy Environment. *Journal of Mechanics of Continua and Mathematical Sciences*, 20: 73-89.

A NEW DISCRETE ENTROPIC MODEL AND ITS APPLICATION TO CONTINGENCY TABLE INFERENCE

Poonam Kumari[1], Nishi Kumari[2]

•

[1] Department of Mathematics, Magadh Mahila College, Patna University, Patna, India
poonamkumari1865@gmail.com
[2]Department of Mathematics, Patna University, Patna, India
nishiraj1510@gmail.com

Abstract

A variety of parametric and non-parametric information models, both discrete and continuous, are well-established. However, there remains a continued need to develop new parametric models that offer greater flexibility in analyzing complex systems. Information entropy is closely linked to the field of statistics, providing valuable insights. This paper introduces a novel discrete information entropic model and explores its applications in statistical analysis. We have demonstrated the efficacy of this new model by applying it to the inference of contingency tables, a cornerstone of statistical analysis. Using the maximum entropy principle and Lagrange method of multiplier, we have derived an important relation between the chi-square and the entropy measure. By incorporating this novel entropy measure, we have expanded the scope and applicability of the maximum entropy principle in this domain, enabling more robust and informative statistical inferences.

Keywords: Discrete probability distribution, entropy, entropic model, maximum entropy principle, Lagrange method of multiplier, contingency table.

I. Introduction

Entropic models have significantly advanced fields like Physics and Engineering, and now play a fundamental role across all areas of Mathematical Sciences. This concept underwent a remarkable transformation in the 20th century with the development of Shannon's [1] mathematical theory of communication. He proposed entropy as a measure of information content, which also facilitates decision-making in problems related to message transmission over noisy or noiseless channels. The non-parametric nature of this entropic model has limited its applicability in various domains. This necessitates the development of more flexible parametric entropic models. Recognizing the inherent strengths and limitations of this entropy, numerous novel generalized entropic models have been formulated, finding successful applications across diverse disciplines.

One particular area where entropic models are of paramount importance is in the analysis of contingency tables, which are frequently used in statistics to analyze the relationship between two or more categorical variables. Learning from contingency tables involves understanding the

dependencies and associations between different variables, often relying on methods such as the maximum entropy principle to make inferences about the underlying data distribution. By imposing constraints based on observed data, the maximum entropy principle provides a way to estimate distributions in situations where limited information is available.

This paper proposes a novel discrete entropic model designed to enhance the flexibility and applicability of the maximum entropy principle in the context of contingency table inference. By extending the range of possible entropy distributions, our model allows for more accurate estimations, even in cases of limited or incomplete data. This new model is based on an innovative formulation that better reflects the underlying structure of the data and can adapt to various types of relationships between variables. The paper is organized as follows: Section II outlines the existing literature review, while Section III describes the proposed entropic model. Section IV demonstrates the application of this entropic model to contingency table inference. Finally, Section V discusses the conclusion and suggests directions for future research.

II. Literature Review

In 1948, Shannon [1] established the following expression for the influential and fundamental additive discrete entropic model:

$$H(P) = -\sum_{i=1}^{n} p_i \log p_i, \tag{1}$$

where $P = (p_1, p_2, \ldots, p_n)$ is the probability measure on a finite set $\{x_1, x_2, \ldots, x_n\}$ such that $p_i \geq 0$ for all $i = 1,2,\ldots,n$ and $p_1 + p_2 + \ldots + p_n = 1$. The logarithm in equation (1) is taken to the base 2, and by convention, $0 \log 0 = 0$.

Arimoto [2] proposed the subsequent generalized entropic model represented by the expression

$$H_t(P) = \frac{t}{t-1}\left[1 - \left(\sum_{i=1}^{n} p_i^t\right)^{1/t}\right]; \quad t > 0, t \neq 1. \tag{2}$$

Behara and Chawla [3] derived the generalized entropic model recognized as τ - entropy and defined as

$$H_\tau(P) = \frac{1 - \left(\sum_{i=1}^{n} p_i^{1/\tau}\right)^{\tau}}{1 - 2^{\tau-1}}; \ \tau > 0, \tau \neq 1. \tag{3}$$

Sharma and Mittal [4] introduced a two-parameter entropy measure in the following form:

$$H_{\alpha,\beta}(P) = \frac{1}{1-\beta}\left[\left(\sum_{i=1}^{n} p_i^{\alpha}\right)^{\frac{1-\beta}{1-\alpha}} - 1\right]; \ \alpha, \beta > 0, \alpha \neq 1, \beta \neq 1, \alpha \neq \beta. \tag{4}$$

Kapur [5] developed several probabilistic models of entropy, one of which is

$$H_b^a(P) = \frac{1}{a+b-2}\left[\sum_{i=1}^{n} p_i^a + \sum_{i=1}^{n} p_i^b - 2\right]; \ a \geq 1, b \leq 1 \text{ or } a \leq 1, b \geq 1. \tag{5}$$

Recognizing the broad scope of entropic models in probability spaces, Parkash and Kakkar [6] examined the following model:

$$H^a(P) = \frac{\sum_{i=1}^{n} p_i^a \log\left(\frac{p_i^a}{\sum_{i=1}^{n} p_i^a}\right)}{\sum_{i=1}^{n} p_i^a}; \ a > 0. \tag{6}$$

The applications of the above model were extended to the field of coding theory for the development of innovative codeword lengths. To advance the literature in the area of coding theory, Parkash and Kakkar [6], [7] extended existing information theoretic entropic models and derived the following expressions:

$$S_a(P)=\frac{1}{1-a}\left(\prod_{i=1}^{n} p_i^{\,p_i(a-1)}-1\right);\ a>0, a\neq 1, \tag{7}$$

$$_aS(P)=\frac{\sum_{i=1}^{n} p_i a^{\log p_i}-1}{1-a};\ a>1, \tag{8}$$

$$^aS(P)=\frac{a^{\sum_{i=1}^{n} p_i \ln p_i}-1}{1-a};\ a>1, \tag{9}$$

$$S_a^b(P)=\frac{a^{\frac{1}{b-1}\ln\left(\sum_{i=1}^{n} p_i^b\right)}-1}{1-a};\ a>1, 0<b<1. \tag{10}$$

Bulinski and Kozhevin [8] investigated the applications of entropy functions to obtain features that can be effectively used for feature selection problems. Lu et al. [9] observed that measures of predictability in physiological indicators, based on entropy metrics, have been widely utilized in the field of medical diagnosis and clinical documentation. The authors proposed a novel entropy-based pattern recognition approach by integrating singular spectrum analysis with entropy measures for the assessment of physiological signals.

Sholehkerdar et al. [10] pointed out that entropy-based metrics are frequently proposed for evaluating the effectiveness of image fusion, owing to their single-layer execution of entropy and minimal parameter set. In their study, the authors presented a theoretical exploration of an image fusion quality measure based on Tsallis [11] entropy. The aim of their research was to assess whether the proposed quality measure could meet the expected performance standards of an ideal information-based image fusion distinction metric. To evaluate the Tsallis [11] quality measure, the authors employed an image realization model to derive a closed-form expression for distinction, while weighted averaging was used as the fusion technique. Their findings showed that the Tsallis-based quality measure affected the anticipated performance, particularly in relation to variations in signal-to-noise ratio and the influence of entropy order on the measured distinction. Additionally, the authors conducted experiments with real images, and the results demonstrated consistency with their theoretical predictions.

Lenormand et al. [12] investigated the applications of entropy-based models in urban environments. They observed that characterizing and quantifying spatial inequalities across urban landscapes remains a complex and challenging task, exacerbated by the growing availability of vast geo-located datasets. Their research findings demonstrated that the attractiveness of a location, as quantified by entropy, is a significant predictor of the socioeconomic status of the surrounding area and can thus be utilized as an indicator for various socioeconomic factors.

Manzoor et al. [13] presented entropic models in the field of chemistry, noting that, inspired by Shannon's [1] entropy, graph entropies with topological indices have effectively quantified the information content of chemical graphs and complex systems.

Recently, Parkash and Mukesh [14] developed the following parametric information model to support the continuation of their findings:

$$H_{a,b}(P)=\frac{\sum_{i=1}^{n} p_i^a-\sum_{i=1}^{n} p_i^b-n^{1-a}+n^{1-b}}{(b-a)(b+a-1)};\quad a\geq 1, b\leq 1 \text{ or } a\leq 1, b\geq 1. \tag{11}$$

Siddiqui et al. [15] computed and analyzed several distinct entropy measures in conjunction with diverse crystallographic structures. They worked on the Zagreb entropies, hyper and augmented Zagreb entropies, and forgotten and Balaban entropies for the crystallographic structures of the cuprite and titanium difluoride.

Teixeira et al. [16] remarked that there is no universally established definition for conditional Tsallis [11] entropy. The traditional definition of Tsallis entropy relies on a parameter α, which converges to Shannon's [1] entropy as α approaches 1. In their research, the authors presented three proposed classifications of conditional Tsallis entropy as outlined in the literature. Furthermore, they introduced an alternative natural approach for conditional Tsallis entropy and compared it with the existing ones. Finally, the authors provided insightful suggestions, emphasizing that the application of entropies across various fields, particularly in information theory and its connection to communication, holds significant potential for the development of numerous valuable information measures, such as mutual information, information symmetry, and information distances.

Segura et al. [17] emphasized that entropy plays a crucial role as a physical concept in any configuration or environment, particularly within the business context. They pointed out that within organizations; a common issue is inefficiency, disruption of relationships, and widespread misunderstandings, which ultimately lead to dysfunction. In this context, a system is defined as a group of individuals characterized by specific attributes, which interact with one another and exist within a particular environment to achieve certain objectives. Entropy, in this sense, refers to a measure of the constraints present within a system that influence its operations. The authors further noted that, in the face of market changes and uncertainty, entropy remains a persistent factor in companies. However, it can also be part of a progression, as new business decisions are made. In this regard, entropy allows us to understand the underlying causes of disorder within organizations and provides tools to monitor changes or disruptions in information.

Saraiva [18] provided a concise and unexpected perspective on Shannon's entropy, encompassing specific properties and demonstrating the model's applicability in two distinct domains from its original context: biological diversity and an innovative study on student migration.

Zhang and Shi [19] emphasized that Shannon's entropy is a fundamental building block of information theory and an essential component of Machine Learning (ML) algorithms. The authors highlighted asymptotic properties that require no assumptions about the underlying data distribution. These properties enable interval estimation and statistical testing with general Shannon's entropy. Friesner et al. [20] investigated the applications of entropy in the field of survey research. Stoyanov et al. [21] applied the maximum entropy approach to derive a novel principle for the *M*-indeterminacy of distributions on the positive half-line (Stieltjes case). Additionally, the authors illustrated how maximum entropy enhances the understanding of symmetry properties and *M*-indeterminacy. Some noteworthy related work on entropy measures and their applications has been carried out by Jizba and Korbel [22], Gao and Deng [23], Elgawad et al. [24], Bajic et al. [25], Fowler and Heckman [26], Li et al. [27], Bisikalo et al. [28], Gregorio et al. [29], among others.

III. A New Discrete Entropic Model

We introduce a new parametric entropic model of order a and type b defined by the expression

$$H_{a,b}(P) = -\sum_{i=1}^{n} p_i \log p_i + \frac{1}{b-a}\sum_{i=1}^{n}\left[1 - p_i^{(b-a)p_i}\right];\ a \neq b, b-a > 0. \tag{12}$$

Obviously, for $b = 1$, we have

$$H_{a,1}(P) = H_a(P) = -\sum_{i=1}^{n} p_i \log p_i + \frac{1}{1-a}\sum_{i=1}^{n}\left[1 - p_i^{(1-a)p_i}\right];\ a > 0. \tag{13}$$

It can be proved that $\lim_{a\to 0} \frac{1}{1-a}\sum_{i=1}^{n}\left(1-p_i^{(1-a)p_i}\right) = -\sum_{i=1}^{n} p_i \log p_i$. Therefore the entropy measure (13) reduces to $\lim_{a\to 0} H_a(P) = -2\sum_{i=1}^{n} p_i \log p_i$, which is similar to Shannon's entropy. Thus, equation (12) represents a generalized measure of entropy.

To prove the validity of the proposed entropy measure, we verify its fundamental properties.

(i) The entropy is non-negative.

For all $b-a>0$, we have $p_i^{(b-a)p_i} \le 1$. Therefore, $\frac{1}{b-a}\sum_{i=1}^{n}\left[1-p_i^{(b-a)p_i}\right] \ge 0$ for all $b-a>0$.

Also, we have $-\sum_{i=1}^{n} p_i \log p_i \ge 0$.

Thus, we get $$-\sum_{i=1}^{n} p_i \log p_i + \frac{1}{b-a}\sum_{i=1}^{n}\left[1-p_i^{(b-a)p_i}\right] \ge 0$$

i.e., $$H_{a,b}(P) \ge 0.$$

(ii) A degenerate distribution concentrated at a single element has entropy zero.

If we take $(p_1, p_2, p_3, \ldots, p_n) = \delta_1 = (1, 0, 0, \ldots, 0)$ or $(p_1, p_2, p_3, \ldots, p_n) = \delta_2 = (0, 1, 0, \ldots, 0)$ or $(p_1, p_2, p_3, \ldots, p_n) = \delta_n = (0, 0, 0, \ldots, 1)$, then, obviously, $H_{a,b}(P) = 0$. Therefore the entropy defined by equation (12) vanishes for a degenerate distribution.

(iii) The entropy does not change on inserting an impossible event.

We have

$$H_{a,b}(p_1, p_2, \ldots, p_n, 0) = -\sum_{i=1}^{n} p_i \log p_i + \frac{1}{b-a}\sum_{i=1}^{n}\left[1-p_i^{(b-a)p_i}\right] + \frac{1}{b-a}\left[1-(0)^{(b-a)0}\right]$$

$$= -\sum_{i=1}^{n} p_i \log p_i + \frac{1}{b-a}\sum_{i=1}^{n}\left[1-p_i^{(b-a)p_i}\right]$$

$$= H_{a,b}(p_1, p_2, \ldots, p_n).$$

Thus, the entropy is unaffected by inserting an impossible event.

(iv) The entropy is a symmetric function of $p_1, p_2, \ldots, p_n$.

Obviously, if $p_1, p_2, \ldots, p_n$ are reordered, then $H_{a,b}(P)$ remains unchanged. Hence, $H_{a,b}(P)$ is a symmetric function of $p_1, p_2, \ldots, p_n$.

(v) The entropy is a concave function of $p_1, p_2, \ldots, p_n$.

Differentiating $H_{a,b}(P)$ w.r.to p_i, we get

$$\frac{dH_{a,b}(P)}{dp_i} = -1 - \log p_i - p_i^{(b-a)p_i}(1 + \log p_i).$$

$$\therefore \quad \frac{d^2 H_{a,b}(P)}{dp_i^2} = -\frac{1}{p_i} - p_i^{(b-a)p_i}\left\{(b-a)(1+\log p_i)^2 + \frac{1}{p_i}\right\} < 0.$$

Hence $H_{a,b}(P)$ is a concave function of $p_1, p_2, \ldots, p_n$.

The concavity of $H_{a,b}(P)$ can also be seen from the graph of $H_{a,b}(P)$ plotted against p_i for $n = 2$ and $a = 2, b = 4$.

Table 1: $H_{a,b}(P)$ *against* p_i *for* $n=2$ *and* $a=2, b=4$

p_i	$H_{a,b}(P)$
0	0
0.1	0.739892
0.2	1.109399
0.3	1.335030
0.4	1.459862
0.5	1.50000
0.6	1.459862
0.7	1.335030
0.8	1.109399
0.9	0.739892
1	0

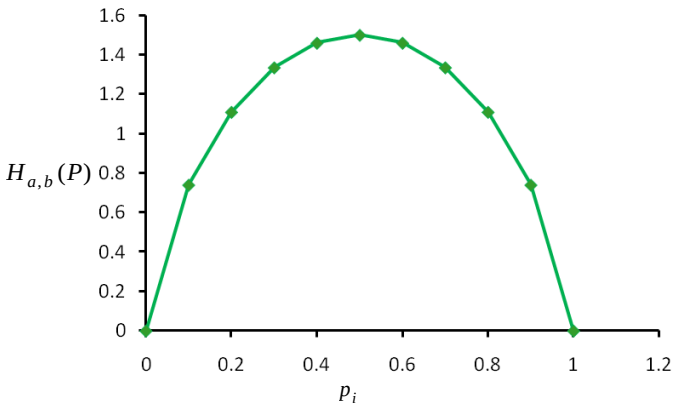


Figure 1: Concavity of $H_{a,b}(P)$ with respect to p_i

(vi) The entropy is a continuous function of $p_1, p_2, \ldots, p_n$.

The continuity of $H_{a,b}(P)$ is obvious from the graph of $H_{a,b}(P)$ plotted against p_i.

(vii) The entropy attains its maximum value at the uniform distribution.

Consider the Lagrange's function

$$L = -\sum_{i=1}^{n} p_i \log p_i + \frac{1}{b-a}\sum_{i=1}^{n}\left(1 - p_i^{(b-a)p_i}\right) - \lambda\left(\sum_{i=1}^{n} p_i - 1\right).$$

Differentiating the above equation w.r.to p_i and equating the derivative to zero, we get

$$\log p_i + p_i^{(b-a)p_i}(1+\log p_i) = -(1+\lambda).$$

The above relation is true only if $p_1 = p_2 = \ldots = p_n$. Further, since $\sum_{i=1}^{n} p_i = 1$, we have $p_i = \frac{1}{n}$ for all $i = 1,2,\ldots,n$. Thus, the maximum value of $H_{a,b}(P)$ is attained at the uniform distribution.

(viii) When entropy is at its maximum, it becomes an increasing function of n.

If $\phi(n)$ denotes the maximum value of the entropy, then

$$\phi(n) = \log n + \frac{n}{b-a}\left[1 - \left(\frac{1}{n}\right)^{\frac{b-a}{n}}\right].$$

Therefore $$\phi'(n) = \frac{1}{n} + \left(\frac{1}{n}\right)^{\frac{b-a}{n}+1}\left[1 + \log\frac{1}{n}\right] + \frac{1}{b-a}\left[1 - \left(\frac{1}{n}\right)^{\frac{b-a}{n}}\right] > 0.$$

This shows that $\phi(n)$ is an increasing function of n.

IV. Application of the Proposed Entropy Measure to Contingency Table Inference

Let ω_{ij} denote the observed frequency in the i^{th} row and j^{th} column of a $p \times q$ contingency table. Let $r_1, r_2, ..., r_p$ denote the sum of all the elements in $1^{\text{st}}, 2^{\text{nd}}, ..., p^{\text{th}}$ rows and let $c_1, c_2, ..., c_q$ denote the sum of all the elements in $1^{\text{st}}, 2^{\text{nd}}, ..., q^{\text{th}}$ columns respectively. Then, we have

$$\sum_{j=1}^{q} \omega_{ij} = r_i \; ; \quad i = 1, 2, ..., p, \tag{14}$$

$$\sum_{i=1}^{p} \omega_{ij} = c_j \; ; \quad j = 1, 2, ..., q, \tag{15}$$

$$\sum_{j=1}^{q}\sum_{i=1}^{p} \omega_{ij} = \sum_{i=1}^{p} r_i = \sum_{j=1}^{q} c_j = X. \tag{16}$$

Now, we suppose that the frequencies are unknown and we estimate these values from the available information of row totals and column totals. Since row totals and column totals are equal, there are $p + q - 1$ independent constraints. With $p + q - 1$ constraints and pq cells, our problem is to find the frequency in each of the pq cells.

Let ν_{ij} denote the estimated frequency in the i^{th} row and j^{th} column. To find ν_{ij}, we use the entropy function introduced in equation (12) and apply the Maximum Entropy Principle. Since the probability of the $(i, j)^{\text{th}}$ cell frequency is $\frac{\nu_{ij}}{X}$, the corresponding entropy function is

$${}_1H^{a,b} = -\sum_{j=1}^{q}\sum_{i=1}^{p} \frac{\nu_{ij}}{X} \log \frac{\nu_{ij}}{X} + \frac{1}{b-a}\sum_{j=1}^{q}\sum_{i=1}^{p}\left[1 - \left(\frac{\nu_{ij}}{X}\right)^{(b-a)\frac{\nu_{ij}}{X}}\right]. \tag{17}$$

Next, we maximize (17) subject to the constraints

$$\sum_{j=1}^{q} \nu_{ij} = r_i \; ; \quad i = 1, 2, ..., p, \tag{18}$$

$$\sum_{i=1}^{p} \nu_{ij} = c_j \; ; \quad j = 1, 2, ..., q, \tag{19}$$

$$\sum_{j=1}^{q}\sum_{i=1}^{p} \nu_{ij} = \sum_{i=1}^{p} r_i = \sum_{j=1}^{q} c_j = X. \tag{20}$$

Now, we apply Lagrange's method to find the maximum value of ${}_1H^{a,b}$ subject to the constraints (18), (19) and (20). The corresponding Lagrangian is

$$L \cong -\sum_{j=1}^{q}\sum_{i=1}^{p} \frac{\nu_{ij}}{X} \log \frac{\nu_{ij}}{X} + \frac{1}{b-a}\sum_{j=1}^{q}\sum_{i=1}^{p}\left[1 - \left(\frac{\nu_{ij}}{X}\right)^{(b-a)\frac{\nu_{ij}}{X}}\right] - \lambda_i\left(\sum_{j=1}^{q} \nu_{ij} - r_i\right) - \mu_j\left(\sum_{i=1}^{p} \nu_{ij} - c_j\right) - \xi\left(\sum_{j=1}^{q}\sum_{i=1}^{p} \nu_{ij} - X\right)$$

or $$L \cong -\sum_{j=1}^{q}\sum_{i=1}^{p}\frac{\nu_{ij}}{X}\log\frac{\nu_{ij}}{X}+\frac{1}{b-a}\sum_{j=1}^{q}\sum_{i=1}^{p}\left[1-e^{(b-a)\frac{\nu_{ij}}{X}\log\left(\frac{\nu_{ij}}{X}\right)}\right]$$
$$-\lambda_i\left(\sum_{j=1}^{q}\nu_{ij}-r_i\right)-\mu_j\left(\sum_{i=1}^{p}\nu_{ij}-c_j\right)-\xi\left(\sum_{j=1}^{q}\sum_{i=1}^{p}\nu_{ij}-X\right)$$

or $$L \cong -\sum_{j=1}^{q}\sum_{i=1}^{p}\frac{\nu_{ij}}{X}\log\frac{\nu_{ij}}{X}+\frac{1}{b-a}\sum_{j=1}^{q}\sum_{i=1}^{p}\left[1-\left(1+(b-a)\frac{\nu_{ij}}{X}\log\left(\frac{\nu_{ij}}{X}\right)\right)\right]$$
$$-\lambda_i\left(\sum_{j=1}^{q}\nu_{ij}-r_i\right)-\mu_j\left(\sum_{i=1}^{p}\nu_{ij}-c_j\right)-\xi\left(\sum_{j=1}^{q}\sum_{i=1}^{p}\nu_{ij}-X\right)$$

or $$L \cong -2\sum_{j=1}^{q}\sum_{i=1}^{p}\frac{\nu_{ij}}{X}\log\frac{\nu_{ij}}{X}-\lambda_i\left(\sum_{j=1}^{q}\nu_{ij}-r_i\right)-\mu_j\left(\sum_{i=1}^{p}\nu_{ij}-c_j\right)-\xi\left(\sum_{j=1}^{q}\sum_{i=1}^{p}\nu_{ij}-X\right)$$

or $$L \cong -\frac{2}{X}\sum_{j=1}^{q}\sum_{i=1}^{p}\nu_{ij}(\log\nu_{ij}-\log X)-\lambda_i\left(\sum_{j=1}^{q}\nu_{ij}-r_i\right)-\mu_j\left(\sum_{i=1}^{p}\nu_{ij}-c_j\right)-\xi\left(\sum_{j=1}^{q}\sum_{i=1}^{p}\nu_{ij}-X\right)$$

or $$L \cong -\frac{2}{X}\left[\sum_{j=1}^{q}\sum_{i=1}^{p}\nu_{ij}\log\nu_{ij}-\log X\sum_{j=1}^{q}\sum_{i=1}^{p}\nu_{ij}\right]-\lambda_i\left(\sum_{j=1}^{q}\nu_{ij}-r_i\right)-\mu_j\left(\sum_{i=1}^{p}\nu_{ij}-c_j\right)-\xi\left(\sum_{j=1}^{q}\sum_{i=1}^{p}\nu_{ij}-X\right).$$

Differentiating the above equation with respect to ν_{ij} and equating the derivative to zero, we get

$$-\frac{2}{X}[(1+\log\nu_{ij})-\log X]-\lambda_i-\mu_j-\xi=0$$

or $$\log\nu_{ij}=-\frac{1}{2}(X\lambda_i-2\log X+2)-\frac{X}{2}(\mu_j+\xi)$$

or $$\nu_{ij}=e^{-\frac{1}{2}(X\lambda_i-2\log X+2)}e^{-\frac{X}{2}(\mu_j+\xi)}$$

or $$\nu_{ij}=\theta_i\psi_j, \tag{21}$$

where $\theta_i=e^{-\frac{1}{2}(X\lambda_i-2\log X+2)}$ and $\psi_j=e^{-\frac{X}{2}(\mu_j+\xi)}$.

Substituting the value of ν_{ij} from equation (21) in equations (18), (19) and (20), we get

$$\sum_{j=1}^{q}\theta_i\psi_j=r_i;\ \ i=1,2,\ldots,p,$$

$$\sum_{i=1}^{p}\theta_i\psi_j=c_j;\ \ j=1,2,\ldots,q,$$

$$\sum_{j=1}^{q}\sum_{i=1}^{p}\theta_i\psi_j=\sum_{i=1}^{p}r_i=\sum_{j=1}^{q}c_j=X$$

or $$\theta_i\sum_{j=1}^{q}\psi_j=r_i;\ \ i=1,2,\ldots,p, \tag{22}$$

$$\psi_j\sum_{i=1}^{p}\theta_i=c_j;\ \ j=1,2,\ldots,q, \tag{23}$$

$$\left(\sum_{j=1}^{q}\psi_j\right)\left(\sum_{i=1}^{p}\theta_i\right)=X. \tag{24}$$

Multiplying equations (22) and (23), we get

$$\theta_i\psi_j\left(\sum_{j=1}^{q}\psi_j\right)\left(\sum_{i=1}^{p}\theta_i\right)=r_ic_j$$

or $$\theta_i\psi_jX=r_ic_j \qquad \text{[Using equation (24)]}$$

or $$\theta_i \psi_j = \frac{r_i c_j}{X}$$

or $$\nu_{ij} = \frac{r_i c_j}{X}. \tag{25}$$

This ν_{ij} gives the maximum frequency estimate of $(i, j)^{\text{th}}$ cell. We shall represent this maximum frequency estimate by ρ_{ij}. Then, from equation (20), we have

$$\sum_{j=1}^{q}\sum_{i=1}^{p}\rho_{ij} = X. \tag{26}$$

The corresponding maximum entropy $H_{\max}^{a,b}$ is given by

$$H_{\max}^{a,b} = -\sum_{j=1}^{q}\sum_{i=1}^{p}\frac{\rho_{ij}}{X}\log\frac{\rho_{ij}}{X} + \frac{1}{b-a}\sum_{j=1}^{q}\sum_{i=1}^{p}\left[1-\left(\frac{\rho_{ij}}{X}\right)^{(b-a)\frac{\rho_{ij}}{X}}\right]$$

$$= -\sum_{j=1}^{q}\sum_{i=1}^{p}\frac{r_i c_j}{X^2}\log\frac{r_i c_j}{X^2} + \frac{1}{b-a}\sum_{j=1}^{q}\sum_{i=1}^{p}\left[1-\left(\frac{r_i c_j}{X^2}\right)^{(b-a)\frac{r_i c_j}{X^2}}\right]$$

$$= -\sum_{i=1}^{p}\frac{r_i}{X}\log\frac{r_i}{X} - \sum_{j=1}^{q}\frac{c_j}{X}\log\frac{c_j}{X} + \frac{1}{b-a}\sum_{j=1}^{q}\sum_{i=1}^{p}\left[1-e^{(b-a)\frac{r_i c_j}{X^2}\log\left(\frac{r_i c_j}{X^2}\right)}\right]$$

$$= -\sum_{i=1}^{p}\frac{r_i}{X}\log\frac{r_i}{X} - \sum_{j=1}^{q}\frac{c_j}{X}\log\frac{c_j}{X} - \frac{1}{b-a}\sum_{j=1}^{q}\sum_{i=1}^{p}(b-a)\frac{r_i c_j}{X^2}\log\left(\frac{r_i c_j}{X^2}\right)$$

$$= -\sum_{i=1}^{p}\frac{r_i}{X}\log\frac{r_i}{X} - \sum_{j=1}^{q}\frac{c_j}{X}\log\frac{c_j}{X} - \frac{1}{b-a}\sum_{j=1}^{q}\sum_{i=1}^{p}(b-a)\frac{r_i c_j}{X^2}\left(\log\frac{r_i}{X} + \log\frac{c_j}{X}\right)$$

$$= -\sum_{i=1}^{p}\frac{r_i}{X}\log\frac{r_i}{X} - \sum_{j=1}^{q}\frac{c_j}{X}\log\frac{c_j}{X} - \frac{1}{b-a}\left(\sum_{j=1}^{q}\frac{c_j}{X}\right)\left[\sum_{i=1}^{p}(b-a)\frac{r_i}{X}\log\frac{r_i}{X}\right]$$
$$- \frac{1}{b-a}\left(\sum_{i=1}^{p}\frac{r_i}{X}\right)\left[\sum_{j=1}^{q}(b-a)\frac{c_j}{X}\log\frac{c_j}{X}\right]$$

$$= -\sum_{i=1}^{p}\frac{r_i}{X}\log\frac{r_i}{X} - \sum_{j=1}^{q}\frac{c_j}{X}\log\frac{c_j}{X} - \frac{1}{b-a}\sum_{i=1}^{p}\log\left(\frac{r_i}{X}\right)^{(b-a)\frac{r_i}{X}} - \frac{1}{b-a}\sum_{j=1}^{q}\log\left(\frac{c_j}{X}\right)^{(b-a)\frac{c_j}{X}}$$

i.e., $$H_{\max}^{a,b} = H_1^{a,b} + H_2^{a,b}, \tag{27}$$

where $H_1^{a,b}$ represents the entropy of the row totals, i.e.,

$$H_1^{a,b} = -\sum_{i=1}^{p}\frac{r_i}{X}\log\frac{r_i}{X} + \frac{1}{b-a}\sum_{i=1}^{p}\left[1-\left(\frac{r_i}{X}\right)^{(b-a)\frac{r_i}{X}}\right]$$

and $H_2^{a,b}$ represents the entropy of the column totals, i.e.,

$$H_2^{a,b} = -\sum_{j=1}^{q} \frac{c_j}{X} \log \frac{c_j}{X} + \frac{1}{b-a} \sum_{j=1}^{q} \left[1 - \left(\frac{c_j}{X} \right)^{(b-a)\frac{c_j}{X}} \right].$$

Let $H^{a,b}$ denote the entropy of the observed cell frequencies ω_{ij}. Then, we have

$$H^{a,b} = -\sum_{j=1}^{q} \sum_{i=1}^{p} \frac{\omega_{ij}}{X} \log \frac{\omega_{ij}}{X} + \frac{1}{b-a} \sum_{j=1}^{q} \sum_{i=1}^{p} \left[1 - \left(\frac{\omega_{ij}}{X} \right)^{(b-a)\frac{\omega_{ij}}{X}} \right].$$

Obviously, $H_{\max}^{a,b} \geq H^{a,b}$

i.e., $H_1^{a,b} + H_2^{a,b} \geq H^{a,b}$.

Thus the gap between the two frequencies is

$$G = H_1^{a,b} + H_2^{a,b} - H^{a,b} \geq 0.$$

Now, $G = H_1^{a,b} + H_2^{a,b} - H^{a,b}$

$$= -\sum_{i=1}^{p} \frac{r_i}{X} \log \frac{r_i}{X} + \frac{1}{b-a} \sum_{i=1}^{p} \left[1 - \left(\frac{r_i}{X} \right)^{(b-a)\frac{r_i}{X}} \right] - \sum_{j=1}^{q} \frac{c_j}{X} \log \frac{c_j}{X}$$

$$+ \frac{1}{b-a} \sum_{j=1}^{q} \left[1 - \left(\frac{c_j}{X} \right)^{(b-a)\frac{c_j}{X}} \right] + \sum_{j=1}^{q} \sum_{i=1}^{p} \frac{\omega_{ij}}{X} \log \frac{\omega_{ij}}{X} - \frac{1}{b-a} \sum_{j=1}^{q} \sum_{i=1}^{p} \left[1 - \left(\frac{\omega_{ij}}{X} \right)^{(b-a)\frac{\omega_{ij}}{X}} \right]$$

$$= -\sum_{i=1}^{p} \frac{r_i}{X} \log \frac{r_i}{X} + \frac{1}{b-a} \sum_{i=1}^{p} \left[1 - e^{(b-a)\frac{r_i}{X}\log\frac{r_i}{X}} \right] - \sum_{j=1}^{q} \frac{c_j}{X} \log \frac{c_j}{X}$$

$$+ \frac{1}{b-a} \sum_{j=1}^{q} \left[1 - e^{(b-a)\frac{c_j}{X}\log\frac{c_j}{X}} \right] + \sum_{j=1}^{q} \sum_{i=1}^{p} \frac{\omega_{ij}}{X} \log \frac{\omega_{ij}}{X} - \frac{1}{b-a} \sum_{j=1}^{q} \sum_{i=1}^{p} \left[1 - e^{(b-a)\frac{\omega_{ij}}{X}\log\frac{\omega_{ij}}{X}} \right]$$

$$= -\sum_{i=1}^{p} \frac{r_i}{X} \log \frac{r_i}{X} - \sum_{i=1}^{p} \frac{r_i}{X} \log \frac{r_i}{X} - \sum_{j=1}^{q} \frac{c_j}{X} \log \frac{c_j}{X} - \sum_{j=1}^{q} \frac{c_j}{X} \log \frac{c_j}{X}$$

$$+ \sum_{j=1}^{q} \sum_{i=1}^{p} \frac{\omega_{ij}}{X} \log \frac{\omega_{ij}}{X} + \sum_{j=1}^{q} \sum_{i=1}^{p} \frac{\omega_{ij}}{X} \log \frac{\omega_{ij}}{X}$$

i.e., $$G = 2 \left[-\sum_{i=1}^{p} \frac{r_i}{X} \log \frac{r_i}{X} - \sum_{j=1}^{q} \frac{c_j}{X} \log \frac{c_j}{X} + \sum_{j=1}^{q} \sum_{i=1}^{p} \frac{\omega_{ij}}{X} \log \frac{\omega_{ij}}{X} \right]. \quad (28)$$

From equation (14), we have

$$\sum_{j=1}^{q} \omega_{ij} = r_i$$

$$\Rightarrow \quad \frac{r_i}{X} = \sum_{j=1}^{q} \frac{\omega_{ij}}{X}$$

$$\Rightarrow \quad \frac{r_i}{X} \log \frac{r_i}{X} = \sum_{j=1}^{q} \frac{\omega_{ij}}{X} \log \frac{r_i}{X}$$

$$\Rightarrow \quad \sum_{i=1}^{p} \frac{r_i}{X} \log \frac{r_i}{X} = \sum_{i=1}^{p} \sum_{j=1}^{q} \frac{\omega_{ij}}{X} \log \frac{r_i}{X}. \quad (29)$$

Similarly, from equation (15), we have

$$\sum_{i=1}^{p} \omega_{ij} = c_j$$

$$\Rightarrow \quad \frac{c_j}{X} = \sum_{i=1}^{p} \frac{\omega_{ij}}{X}$$

$$\Rightarrow \quad \frac{c_j}{X} \log \frac{c_j}{X} = \sum_{i=1}^{p} \frac{\omega_{ij}}{X} \log \frac{c_j}{X}$$

$$\Rightarrow \quad \sum_{j=1}^{q} \frac{c_j}{X} \log \frac{c_j}{X} = \sum_{j=1}^{q} \sum_{i=1}^{p} \frac{\omega_{ij}}{X} \log \frac{c_j}{X}. \tag{30}$$

Using equations (29) and (30), equation (28) reduces to

$$G = 2 \sum_{j=1}^{q} \sum_{i=1}^{p} \frac{\omega_{ij}}{X} \log \frac{\frac{\omega_{ij}}{X}}{\frac{r_i}{X} \frac{c_j}{X}}$$

or

$$G = 2 \sum_{j=1}^{q} \sum_{i=1}^{p} \frac{\omega_{ij}}{X} \log \frac{\omega_{ij}}{\rho_{ij}}. \tag{31}$$

Obviously, $G = 0$ if and only if $\frac{\omega_{ij}}{\rho_{ij}} = 1$, i.e., $\omega_{ij} = \rho_{ij}$, where ω_{ij} and ρ_{ij} are the observed frequencies and the maximum frequency estimates respectively..

Thus we find that frequency estimates determined from Maximum Entropy Principle give the same values as provided by the hypothesis of independence.

Now, we consider the case when $G > 0$. In this case, let

$$\omega_{ij} = \rho_{ij} + \varepsilon_{ij}, \tag{32}$$

where $-1 < \varepsilon_{ij} < 1$ and ε_{ij} are very small quantities. Then

$$\sum_{j=1}^{q} \sum_{i=1}^{p} \omega_{ij} = \sum_{j=1}^{q} \sum_{i=1}^{p} \rho_{ij} + \sum_{j=1}^{q} \sum_{i=1}^{p} \varepsilon_{ij}.$$

But, from equations (16) and (26), $\sum_{j=1}^{q} \sum_{i=1}^{p} \omega_{ij} = \sum_{j=1}^{q} \sum_{i=1}^{p} \rho_{ij} = X$. Therefore

$$\sum_{j=1}^{q} \sum_{i=1}^{p} \varepsilon_{ij} = 0. \tag{33}$$

Now,

$$G = 2 \sum_{j=1}^{q} \sum_{i=1}^{p} \frac{\rho_{ij} + \varepsilon_{ij}}{X} \log \frac{\rho_{ij} + \varepsilon_{ij}}{\rho_{ij}}$$

$$= \frac{2}{X} \sum_{j=1}^{q} \sum_{i=1}^{p} (\rho_{ij} + \varepsilon_{ij}) \log \left(1 + \frac{\varepsilon_{ij}}{\rho_{ij}} \right)$$

$$= \frac{2}{X} \sum_{j=1}^{q} \sum_{i=1}^{p} (\rho_{ij} + \varepsilon_{ij}) \left[\frac{\varepsilon_{ij}}{\rho_{ij}} - \frac{(\varepsilon_{ij} / \rho_{ij})^2}{2} + \ldots \right]$$

$$= \frac{2}{X} \sum_{j=1}^{q} \sum_{i=1}^{p} \varepsilon_{ij} + \frac{1}{X} \sum_{j=1}^{q} \sum_{i=1}^{p} \frac{\varepsilon_{ij}^2}{\rho_{ij}}.$$

But, from equation (33), we have $\sum_{j=1}^{q}\sum_{i=1}^{p}\varepsilon_{ij} = 0$. Therefore

$$G = \frac{1}{X}\sum_{j=1}^{q}\sum_{i=1}^{p}\frac{\varepsilon_{ij}^{2}}{\rho_{ij}} = \frac{1}{X}\sum_{j=1}^{q}\sum_{i=1}^{p}\frac{(\omega_{ij}-\rho_{ij})^{2}}{\rho_{ij}} = \frac{1}{X}\chi^{2},$$

where χ^2 denotes the chi-square. Thus, we have

$$\chi^{2} = XG \tag{34}$$

i.e., chi-square is X times the additional information provided by the observed frequencies over the information previously obtained by the row totals and column totals. This interpretation aligns with the fundamental purpose of the chi-square test in contingency tables: assessing how well the observed data fits the expected distribution.

V. Conclusion

Through the present work, we have tried to contribute to the ongoing development of more flexible and powerful tools for statistical analysis, offering a new perspective on contingency table inference and providing a foundation for future advancements in entropy-based methods. The proposed model leverages entropy as a central concept to capture the underlying dependencies and associations between variables in a more flexible and computationally efficient manner than traditional methods. The significance of the work lies not only in the development of a new entropy model but also in its practical applications. By expanding the applicability of the maximum entropy principle, the proposed model provides a powerful tool for statistical inference in a wide range of domains, including social sciences, biology, economics, and any field where contingency tables are commonly used. Additionally, the enhanced flexibility of the model opens up new possibilities for exploring complex dependencies and interactions within datasets, ultimately leading to more accurate and meaningful insights.

References

[1] Shannon, C. E. (1948). A mathematical theory of communication. Bell. System Technical Journal, 27: 379–423, 623–659.

[2] Arimoto, S. (1971). Information theoretical consideration on estimation problems. *Information and Control,* 19: 181–194.

[3] Behara, M. and Chawla, J. S. (1974). Generalized gamma-entropy. *Selecta statistica canadiana,* 2: 15–38.

[4] Sharma, B. D. and Mittal, D. P. (1975). New non-additive measures of entropy for discrete probability distributions. Journal of Mathematical Sciences, 10: 28–40.

[5] Kapur, J.N. (1986). Four families of measures of entropy. *Indian Journal of Pure and Applied Mathematics,* 17: 429–449.

[6] Parkash, O. and Kakkar, P. (2012). Development of two new mean codeword lengths. Information Sciences, 207: 90–97.

[7] Parkash, O. and Kakkar, P. (2014). New information theoretic models, their detailed properties and new inequalities. Canadian Journal of Pure and Applied Sciences, 8(3): 3115-3123.

[8] Bulinski, A. and Kozhevin, A. (2019). Statistical estimation of conditional Shannon entropy. *ESAIM Probabability and Statistics,* 23: 350–386.

[9] Lu,Y., Wang, M., Wu, W., Zhang, Q., Han, Y., Kausar, T., Chen, S., Liu, M., and Wang, B. (2020). Entropy-Based Pattern Learning Based on Singular Spectrum Analysis Components for Assessment of Physiological Signals. Complexity, Article ID 4625218.

[10] Sholehkerdar, A., Tavakoli, J. and Liu, Z. (2020). Theoretical analysis of Tsallis Entropy-Based quality measure for weighted averaging Image Fusion. Information Fusion, 58: 69-81.

[11] Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, 52: 479–487.

[12] Lenormand, M., Samaniego, H., Chaves, J.C., Vieira, V.D.F., Barbosa da Silva, M.A.H., Evsukoff, A.G. (2020). Entropy as a measure of attractiveness and socioeconomic complexity in Rio de Janeiro metropolitan area. *Entropy*, 22: 368.

[13] Manzoor, S., Siddiqui, M. K. and Ahmad, S. (2020). On entropy measures of molecular graphs using topological indices. Arabian Journal of Chemistry, 13(8):6285–6298.

[14] Parkash, O. and Mukesh (2021). Two new parametric entropic models for discrete probability distributions. Turkish Journal of Computer and Mathematics Education, 12(6): 2949–2954.

[15] Siddiqui, M. K., Manzoor, S., Ahmad, S. and K. A. Kaabar, M. K.A. (2021). On Computation and Analysis of Entropy Measures for Crystal Structures. Graph Invariants and Their Applications.

[16] Teixeira, A. Souto, A. and Antunes, L. (2021). On Conditional Tsallis Entropy. *Entropy*, *23*(11): 1427.

[17] Segura, E. A., Zamar, M. D. G. and Squillante, M. (2021). Examining the Research on Business Information-Entropy Correlation in the Accounting Process of Organizations. *Entropy*, *23*(11):1493.

[18] Saraiva, P. (2023). On Shannon entropy and its applications. Kuwait Journal of Science, 50(3): 194–199.

[19] Zhang, J. and Shi, J. (2022). Asymptotic normality for plug-in estimators of generalized Shannon's entropy. *Entropy*, 24: 683. 62.

[20] Friesner, D., Bozman, CS, McPherson, M., Valente, F. and Zhang, A. (2023). Information entropy as a quality control tool in survey research. *Journal of Marketing Development and Competitiveness, 17(1)*.

[21] Stoyanov, J.M., Tagliani, A. and Inverardi, P.L.N. Maximum (2024). Entropy Criterion for Moment Indeterminacy of Probability Densities. *Entropy*, *26*(2), 121.

[22] Jizba, P. and Korbel, J. (2019). Maximum entropy principle in statistical inference: Case for Non-Shannonian entropies. *Physical Review Letters,* 122, 120601.

[23] Gao, X. and Deng, Y. (2020). The Pseudo-Pascal Triangle of Maximum Deng Entropy. *International Journal of Computers, Communications & Control,* 15: 1–10.

[24] Elgawad M. A. A., Barakat H. M., Xiong S. and Alyami S. A. (2021). Information Measures for Generalized Order Statistics and Their Concomitants under General Framework from Huang-Kotz FGM Bivariate Distribution. Entropy 23(3): 335(1–17).

[25] Bajic D., Đaji´c V. and Milovanovi´c B. (2021). Entropy analysis of COVID-19 cardiovascular signals. *Entropy,* 23(1): 87 (1–15).

[26] Fowler, R. and Heckman, J.J. (2022). Misanthropic entropy and renormalization as a communication channel. *International Journal of Modern Physics, A* 37(16): 2250109, 18.

[27] Li, Y., Geng, B. and Jiao, S. (2022). Dispersion entropy-based Lempel-Ziv complexity: A new metric for signal analysis. Chaos Solitons & Fractals 161: 112400.

[28] Bisikalo, O., Kharchenko,V. Kovtun, V., Krak, L. and Pavlov, S. (2023). Parameterization of the Stochastic Model for Evaluating Variable Small Data in the Shannon Entropy Basis. *Entropy*, *25* (2), 184.

[29] Gregorio, J. D., Sánchez, D. and Toral, R. (2024). Entropy Estimators for Markovian Sequences: A Comparative Analysis, *Entropy*, 26(1), 79.