

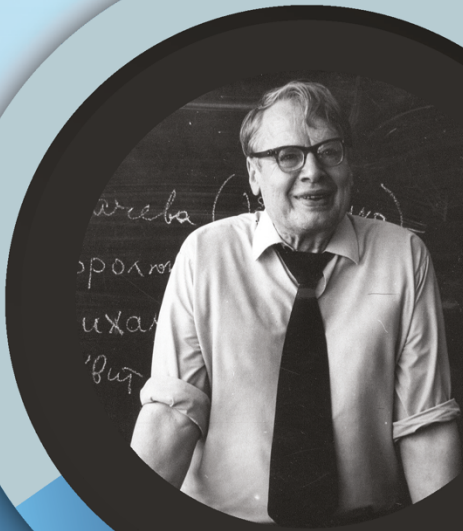
# RTA

ISSN 1932-2321

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

RELIABILITY:  
THEORY & APPLICATIONS

INTERNATIONAL  
GROUP ON  
RELIABILITY



GNEDENKO FORUM PUBLICATIONS

# 1

(77) VOL.19

MARCH

2024

SAN DIEGO

RELIABILITY

RISK ANALYSIS

MAINTENANCE

SAFETY

**ISSN 1932-2321**

© "Reliability: Theory & Applications", 2006, 2007, 2009-2024

© " Reliability & Risk Analysis: Theory & Applications", 2008

© I.A. Ushakov

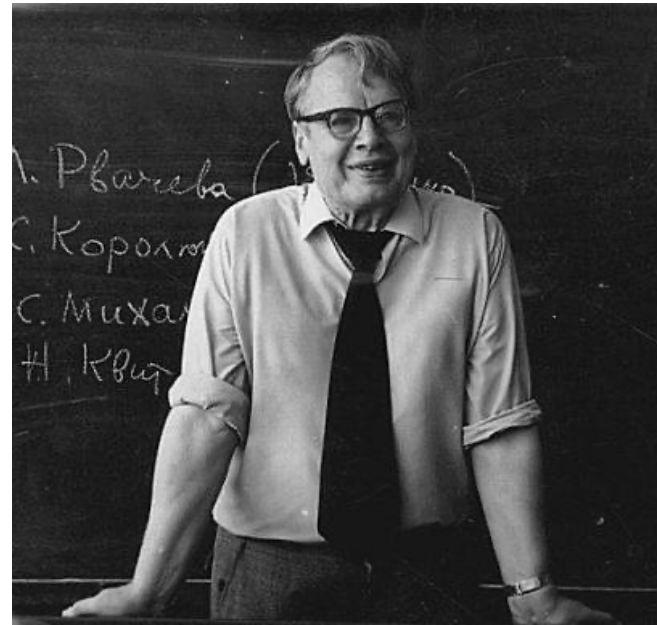
© A.V. Bochkov, 2006-2024

© Kristina Ushakov, Cover Design, 2023-2024

<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.19 No.1 (77),  
March 2024

San Diego  
2024

# 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: sazono2007@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, VNIIS (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 (SSTC NRS), 53, Chernishevskaya str., of.2, 61002, Kharkov, Ukraine  
e-mail: ma\_yastrebenetsky@sstc.com.ua

## 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, Novocheryomushkinskaya 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), Prijedor, 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](mailto:a.bochkov@gmail.com)

## Table of Contents

### ESTIMATION OF DIFFERENT ENTROPIES OF INVERSE RAYLEIGH DISTRIBUTION UNDER MULTIPLE CENSORED DATA..... 29

Hemani Sharma, Parmil Kumar

*The inverse Rayleigh distribution finds widespread applications within life testing and reliability research. Particularly, it proves invaluable in scenarios involving multiple censored data points. In this context, the Renyi, Havrda, Charvat, and Tsallis entropies of the inverse Rayleigh distribution are efficiently calculated. The maximum likelihood approach is used to get the estimators, as well as the approximate confidence interval. The mean squared errors, approximate confidence interval, and their related average length are computed. To illuminate the behavior of estimates across varying sample sizes, a comprehensive simulation study is conducted. The outcomes of the simulation study consistently reveal a downward trend in mean squared errors and average lengths as the sample size increases. Additionally, an interesting finding emerges as the censoring level diminishes. The entropy estimators progressively converge towards their true values. For practical demonstration, the effectiveness of the approach is showcased through the analysis of two real-world datasets. These applications underscore the real-world relevance of the methodology, further validating its utility in addressing complex scenarios involving censored data and inverse Rayleigh distributions.*

### A LITERATURE SURVEY ON QUEUEING MODEL WITH WORKING VACATION ..... 40

Divya K., Indhira K.

*In 2002, the Working Vacation (WV) queues were implemented as an extension of standard queueing models with vacations. During the vacation period in WV queues, the server provides service at a slower pace as opposed to the typical busy period. The objective of this survey is to provide a concise overview of the latest scholarly investigations on queueing models for WVs. The concept of a queue with WV has been implemented across various domains, encompassing computer systems, communication networks, production management, computer communication, manufacturing, and inventory systems. Additionally, it has been applied to network service, web service, file transfer service, and mail service.*

### BAYESIAN ESTIMATION OF TOPP-LEONE LINDLEY (TLL) DISTRIBUTION PARAMETERS UNDER DIFFERENT LOSS FUNCTIONS USING LINDLEY APPROXIMATION ..... 50

Nzei C. Lawrence, Adegoke M. Taiwo, Ekhosuehi N., Mbegbu I. Julian

*In this study, we present the Bayesian estimates of the unknown parameters of the Topp-Leone Lindley distribution using the maximum likelihood and Bayesian methods. In this study, the Bayes theorem was adopted for obtaining the posterior distribution of the shape parameter and scale parameter of the Topp-Leone Lindley distribution assuming the Jeffreys' (non-informative) prior for the shape parameter and the Gamma (conjugate) prior for the scale parameter under three different loss functions namely: Square Error Loss Function, Linear Exponential Loss Function and Generalized Entropy Loss Function. The posterior distribution derived for both parameters are not solvable analytically, it requires a numerical approximation techniques to obtain the solution. The Lindley approximation techniques was adopted to obtain the parameters of interest. The loss function were used to derive the estimates of both parameters with an assumption that the both parameters are unknown and independent. To ascertain the accuracy of these estimators, the proposed Bayesian estimators under different loss functions are compared with the corresponding maximum likelihood estimator using a Monte Carlo simulation on the performance of these estimators according to the mean square error and BIAS based on simulated samples*

*simulated from the Topp-Leone Lindley distribution. It was also observed for any fixed value of the parameters, as sample size increases, the mean square errors of the Bayesian Estimates and maximum likelihood estimates decrease. Also, the maximum likelihood estimates and Bayesian estimates converge to the same value as the sample gets larger except for Generalized Entropy Loss Function.*

## **SINE-WEIBULL DISTRIBUTION: MATHEMATICAL PROPERTIES AND APPLICATION TO REAL DATASETS ..... 65**

Muhammad Umar Faruk, Alhaji Modu Isa, Aishatu Kaigama

*New parameters can be added to expand families of distribution for greater flexibility or to construct covariate models in several ways. In this study, a trigonometric-type distribution called Sine-Weibull distribution was developed by adopting the Weibull distribution as the baseline distribution and Sine-G Family as the generator to generate a flexible probability distribution without the need for extra parameters. The moment, moment generating function, entropy, and order statistics are some of the mathematical aspects of this distribution that were derived. The Maximum Likelihood approach was used to estimate the new distribution's parameters. Using actual datasets, the Sine-Weibull distribution's applicability was demonstrated.*

## **RECENT DEVELOPMENTS IN THE COMPUTATION OF THE ROCOF OF MULTI-STATE SYSTEMS AND ITS APPLICATIONS ..... 73**

Guglielmo D'Amico, Fulvio Gismondi

*This paper reviews several theoretical works on the computation of the Rate of Occurrence of Failure (ROCOF) for general multi-state random systems, focusing on recent generalizations. The discussion begins by defining the ROCOF for a Markov process and discussing the main results achieved in the literature, then moves towards the richer framework represented by semi-Markov systems. The paper discusses complications that arise when extending the ROCOF to higher orders so that a measure of the association of failures in time can be obtained. The work then analyzes possible modifications in terms of a conditional version of the ROCOF, which is of special interest in applications. The findings are illustrated by a numerical example from reliability, and the broad applicability is demonstrated by a discussion of different applications in other domains.*

## **STRIP-PLOT ANALYSIS FOR THE CONSTRUCTION OF COMPLETE TRIPARTITE AND CUBIC GRAPHS..... 86**

V. Saranya, S. Kavitha

*The Strip-Plot Design (SPD) is plays an important role in the complete block designs and also using the agricultural, medical and industry fields. SPD is best suited for a two-factor experiment that has more treatments than can be accommodated by a complete block design. In a SPD, one factor is assigned to the horizontal strip plot, one factor is assigned to the vertical – strip plot and one factor is interaction plot. Also, few experimental materials may be rare while other test items may be available in altering doses of other therapeutic factors, which may be expensive or time-consuming. One of the main features of SPD involves three types of experimental errors: row - strip plot error, coloum – strip plot error and interaction plot error. Experimenting across processing steps is essential for studying the interaction of factors where certain factors come from one step and others arrive from the other. The strip-plot design is a very efficient design for investigating multiple-step processes in terms of both resources and time. Strip-plot designs are economical when the factors are hard to change and the process under research has three discrete stages. When we want to study interactions between factors where some factors are from one step and other factors from another step, it is important to conduct experiments across processing steps. The approach is flexible because it can handle experimental design problems involving factors acting at different levels, unlike the existing method. Graphs are widely used representations of both natural and human-made structures. Graph theory can be used to investigate "things that are connected to other things. "Fits nearly*



everywhere. Some tough problems become easier to solve when they are represented graphically. We reviewed the agricultural field yield of the strip-plot design and early work on the design of industrial strip-plot design in this paper. We have also described the model of strip-plot design. We, therefore, advise experimenters to ensure that their strip-plot designs contain a sufficient number of rows and columns so that valid statistical inference is possible. A bipartite graph is one in which the edges can be divided into two sets without going into sets. A complete bipartite graph is a bipartite graph that is completed. The complete tripartite graph in which the edges can be divided into three set without going into sets. The cubic graph is a graph in which all vertices have degree three. This paper describes the construction and Statistical Analysis of SPD using some particular types of graphs is discussed through numerical examples.

## **A HYBRID APPROACH TO SINGLE ACCEPTANCE SAMPLING PLANS FOR LIFETIMES MODELED WITH THE EXPONENTIAL RAYLEIGH DISTRIBUTION ..... 100**

Nandhini M, Radhika A, Jeslin J, Manigandan P

This article explores into the examination of a novel compound distribution termed the "Exponential Rayleigh distribution" in the context of truncated life testing within a sampling plan. It introduces a hybrid single acceptance sampling plan tailored for truncated life testing scenarios where the item's lifespan adheres to the Exponential Rayleigh distribution. One of the primary segments within the domain of product quality control is referred to as "sampling inspection by variables". This category encompasses procedures that involve the selection of multiple individual units based on measurements taken from a sample to assess a specific quality attribute under scrutiny. These plans, used to assess whether to accept or reject a submitted batch of items based on their observed lifetimes, are commonly known as reliability test plans. The article also outlines the development of a test plan to determine when to conclude the experiment given specific parameters like sample size, producer's risk, consumer's risk, and termination criteria. Sampling inspection, or reliability sampling, plays a pivotal role in maintaining product quality. It involves subjecting items to testing, collecting data on their lifespans, and making acceptance or rejection decisions based on the test results. When assessing an item's quality primarily based on its lifespan, which can be suitably described using a continuous probability distribution; such a plan is termed a "life test sampling plan." This article explores the application of the Exponential Rayleigh distribution within the realm of reliability sampling plans, emphasizing the utilization of hybrid censoring for life checks and median lifetime evaluations. This approach is leveraged to formulate reliability single sampling plans applicable to the Exponential Rayleigh distribution. The article utilizes binomial probabilities to compute the parameters of these sampling plans, aiming to strike a balance between protecting the interests of both the producer and the consumer while minimizing producer risks. The study involves calculating the specified median lifetime and determining design parameters like sample size and acceptance thresholds to meet predefined quality standards. The flexibility of the Exponential Rayleigh distribution in analyzing various types of lifetime data is highlighted, owing to its scale and shape parameters. To illustrate the concepts related to sampling strategies, a numerical example is provided in the sampling strategies section of the article.

## **CHARACTERIZATION OF SOME CONTINUOUS DISTRIBUTIONS BY CONDITIONAL VARIANCE OF RECORD VALUES..... 108**

Zaki Anwar, Mohd Faizan, Zakir Ali

Characterization of a probability distribution gives a unique property enjoyed by that distribution. Various approaches are available in the literature to characterize distributions through record values. Many researchers have characterized Exponential, Pareto, and Power function distributions using moments, conditional expectation, and some other characteristics of record values. In this paper, we have characterized these three distributions through conditional variance of adjacent record values. The results have been verified using numerical computation.

**SOME INFERENTIAL ASPECTS ON THREE-STATION TANDEM QUEUE ..... 113**

Ambily Jose, Agnes Jerome, M. R. Irshad

*Considered is a three-station tandem queue with service times at stations 1, 2, and 3 are exponentially distributed with customers arriving according to the Poisson process at station 1. Given that the stationary distribution is the product of three independent geometric distributions with the intensity parameters, maximum likelihood estimators and Bayes estimators of the intensity parameters based on the number of customers present at different time periods are obtained. Furthermore, the minimal posterior risk and minimum Bayes risk of the estimators are computed. Also, a simulation study is conducted to evaluate the performance of the estimators obtained.*

**RELIABILITY MODELING OF A BUTTER CHURNER AND CONTINUOUS BUTTER MAKING PRODUCTION SYSTEM..... 122**

Upasana Sharma, Drishti

*In the dairy plant, an investigation into the machine that makes butter was subjected to a reliability study in relation to the seasonal demand. In the process of expanding the butter churner into a machine that can make butter continuously, a more reliable operational model was devised. Both the models and the data acquired with MATLAB have been subjected to availability and reliability testing and analysis. In addition, the graphical analysis was carried out with the help of Code Blocks and Excel. A comparison of the two models was then covered as the final topic. It was discovered that (a) the extended model was superior to the current model, (b) the failure rate of the existing line increased, which implies that a new machine needs to be added to the line to share the load, which results in improved production, and (c) the failure rate of the extended model was lower than the failure rate of the existing model. (c) in order to maximise profits while simultaneously minimising losses The effectiveness of the system ought to be enhanced by performing routine maintenance during both the summer and the winter.*

**STREAMLINING PRODUCT DEPLOYMENT: ENHANCING EFFICIENCY THROUGH KITTING PROCESSES ..... 140**

G. Ayyappan, S. Sankeetha

*Considering a single server with two queues that is prone to unreliability. The server offers a kitting process and performs necessary checks and rectifications when required. The arrival of items follows a Markovian arrival process, while the service is distributed based on a phase type distribution. The incoming products may exhibit issues such as poor quality or defects. If either of the queues is empty, the server is unable to provide the requested service and remains inactive. Furthermore, if all queues are empty, the server goes into a vacation mode. Breakdowns, repairs, instances of customers leaving without service (reneging), and vacation periods are all modeled using an exponential distribution. To gain insights into the performance of the queueing model, various performance metrics are analyzed and represented through 2D and 3D graphs.*

**ON SOME PROPERTIES AND APPLICATIONS OF THE TYPE II HALF -LOGISTIC EXPONENTIATED FRECHET DISTRIBUTION ..... 160**

Olalekan Akanji Bello, Sani Ibrahim Doguwa, Abukakar Yahaya, Haruna Mohammed Jibril

*As the dimensions of available data for analysis continues to grow rapidly, it becomes imperative to develop new probability distributions that can more accurately represent various phenomena. In this research paper, we introduce a novel continuous probability distribution known as the Type II Half-Logistic Exponentiated Frechet Distribution, characterized by four positive parameters. This distribution expands upon the traditional Frechet distribution by introducing two additional parameters. We derive a significant density representation for this distribution. Furthermore, we delve into several statistical and mathematical properties associated with the Type*

*II Half-Logistic Exponentiated Frechet distribution. This includes explicit expressions for key metrics such as the quantile function, probability weighted moments, moments, moments generating function, reliability function, hazard function, and order statistics. To estimate the model parameters effectively, we employ a maximum likelihood estimation technique and present the results of a simulation study. Our research underscores the superiority of this new distribution by applying it to two real-world datasets. Notably, the findings demonstrate that the Type II Half-Logistic Exponentiated Frechet distribution outperforms other considered distributions in fitting the two real datasets.*

**AVAILABILITY OPTIMIZATION OF A PAINT MANUFACTURING PLANT USING GREY WOLF OPTIMIZATION: A METAHEURISTIC APPROACH ..... 175**

Ashish Kumar, Vijay Singh Maan, Monika Saini

*The primary objective of present research work is to evaluate and improve the performance and availability of the paint manufacturing plant. Paint manufacturing plant consists of five subsystem naming mixer, grinder, thinner, labelling, and filling unit. Among them labelling and filling unit have two machines in parallel configuration and both are working simultaneously. All failure and repair rates are distributed exponentially. Markov birth-death process is utilized to model the dynamic behavior of the system and its sub-components, enabling a quantitative analysis of system availability. Grey wolf optimization (GWO), a swarm-based optimization technique is used to optimize the availability of the system. Moreover, the research conducts a thorough comparison between the outcomes derived from the Markov birth-death process and the GWO technique. By harnessing the power of GWO, the study aims to further enhance the plant's overall performance.*

**DESIGNING AND EVALUATION OF SKIP-LOT SAMPLING PLAN OF TYPE SkSP-T WITH SINGLE SAMPLING PLAN AS REFERENCE PLAN UNDER THE CONDITION OF INTERVENED POISSON DISTRIBUTION ..... 183**

S. Suganya, K. Pradeepa Veerakumari

*This paper describes the scheming technique of new system of skip lot sampling plan of type SkSP-T with Single Sampling Plan as Reference plan under the condition of Intervened Poisson Distribution. The designing methodology includes the evaluation of Acceptable Quality Level, Limiting Quality Level, Operating Ratio, and Operating Characteristic curves. Tables are simulated by changing various parametric values of SkSP-T, SSP and IPD and operating characteristic curves are drawn by using R language.*

**EJAZ DISTRIBUTION A NEW TWO PARAMETRIC DISTRIBUTION FOR MODELLING DATA ..... 191**

Aijaz Ahmad, M. A. Lone, Aafaq. A. Rather

*This paper introduces a novel probability distribution known as the Ejaz distribution (ED), which is characterized by two parameters. The study offers a comprehensive analysis of this distribution, including an examination of key properties such as moments, moment-generating functions, order statistics, and reliability functions. Additionally, the paper explores the graphical representation of essential functions like the probability density function, cumulative distribution function, and hazard rate function, enhancing our visual understanding of their behavior. The distribution's parameters are estimated using the widely accepted method of maximum likelihood estimation. Through real-world examples, the paper highlights the practical applicability of the Ejaz distribution, demonstrating its performance and relevance in diverse scenarios.*

**RELIABILITY INVESTIGATION OF THE SPIRULINA PRODUCTION PLANT USING GUMBEL-HOUGAARD FAMILY COPULA ..... 202**

Priya Chaudhary, Shikha Bansal

*This study examines the consistency metrics used to evaluate the durability of a spirulina production plant, which consists of seven subsystems: cultivation pond, paddlewheel, filter unit, washing unit, spray dryer, ribbon blender, and packaging. By studying the spirulina firm, we can repair it by discovering future failures. We can increase spirulina production so that untimely failure can be prevented and production can be increased. There are two types of system failures: partial and total. While a full failure renders the system incapable of operating, a partial failure is thought to degrade the system. In contrast, repair rates follow two different types of distributions: an ordinary and an exponential distribution. The system in a partially failed or degraded condition is thought to be repaired using general time distribution. In contrast, fully failed systems are thought to be fixed using the Gumbel-Hougaard family copula distribution. Using the supplementary variable approach, the system is examined. A Chapman-Kolmogorov differential equation is created and solved by applying the Gumbel-Hougaard family Copula approach, employing the schematic representation of the system's state. supplementary variable approaches are applied to develop and resolve the differential equations related to transition diagrams, which are significant to this research. Reliability, availability, profitability, and MTTF are the critical performance metrics for the spirulina production plant. Moreover, sensitivity analysis is carried out for MTTF.*

**EFFICIENT FRAMEWORK OF SECURITY FOR INTERNET OF THINGS ..... 217**

Dr. Mihir Mehta, Dr. Kajal Patel, Dr. Komal Anadkat

*IoT security represents a highly compelling subject of research at present. The absence of a viable security solution for IoT applications could render them ineffective across various domains such as healthcare, smart homes, inventory management, smart agriculture, and more. Within the IoT architecture, security services like Confidentiality, Integrity, and Authentication play pivotal roles. In our research, we have concentrated on the Authentication service, which is fundamental for distinguishing users and devices unequivocally within a network. Authentication serves as the initial and crucial step in establishing secure communications among diverse IoT devices and users within the network. A compromised Authentication service could open the door for unauthorized users or devices to infiltrate the network, potentially leading to harmful activities like Masquerade attacks, Man-in-the-Middle (MITM) attacks, and Replay attacks. Currently, Authentication stands as a widely adopted and essential method for granting access to devices within IoT networks. Our contribution involves the development of a Multi-factor IoT Authentication Model, leveraging two key parameters: Device Context Information and Dynamic Key-based authentication. Our proposed approach begins by verifying the origin of information. If the origin is deemed valid, our model proceeds to validate the identity of the device. In the event of an intruder attempting to manipulate the device's origin from its predefined context to an alternative location, our system can swiftly detect this deviation, thereby enabling the rejection of communication requests from compromised devices. Following the verification of context information, we initiate mutual authentication between the IoT device and the server, employing the Challenge-response model. As a result of this second step, individual Session keys are generated at both the device and server sides, facilitating secure communication within a specific time window.*

**M/M/∞ QUEUE WITH IMPATIENT CUSTOMERS ..... 228**

Gulab Singh Bura

*In this paper we proposed an M/M/∞ queue with impatient customers. Generally, customers are impatient due to long waits in queue but in this work, we consider the case when customers are not impatient due to long waits but they are impatient due to the poor quality of service. We model and analyze this queueing system by using continued fraction technique and obtained the probability mass function of the customers present in the system in time dependent form. Also, we calculate the average queue size. Finally, some graphical representations are given to illustrate the model.*

**REGULARITY OF ALTERNATE QUADRA SUBMERGING POLAR FUZZY GRAPH AND ITS APPLICATION ..... 238**

Anthoni Amali A, J. Jesintha Rosline

Fuzzy soft sets and graphs are invented to solve uncertain problems in the field of Applied mathematics. It is a general mathematical tool introduced with many parameters to model the vagueness of the changing world. The insight learning of the AQSP fuzzy soft graphs paved the way to discover the extension of the AQSP fuzzy soft graph. In this research article we introduce the Regularity of AQSP fuzzy soft graph with definitions, theorems, properties, and real-life applications. The aim of this invention is mainly to obtain the parametric values in submerging level of confidence  $[-0.5, 0.5]$   $[-1, 1]$ . The scope of this new AQSP fuzzy soft graph is to solve the imprecise problems in the field of Mathematical Engineering, Bio Mathematics, Economics, Medical Science, Artificial Intelligence and Machine learning. The regularity of AQSP fuzzy soft graph is combined with the concepts of regular, totally regular, and perfectly regular. The application of this new graph is developed for governing of the women safety vehicle network in different spots with membership submerging values. The future extension can be applied in Approximate reasoning, Mathematical psychology, Decision making for medical diagnosis.

**STATISTICAL ANALYSIS OF SPLIT-PLOT DESIGN USING SPECIAL TYPE OF GRAPHS ..... 254**

V. Saranya, S. Kavitha, M. Pachamuthu, S. Vijayan

When all experimental runs cannot be done under homogeneous conditions, blocking can be utilized to increase the power for testing treatment effects. In many real-life environments, there is at least one factor that is hard to change, leading to a split-plot structure. This paper demonstrates how to generate certain graphs using main-plot and sub-plot analyses, as well as providing a catalog. As a result, during situations where the candidate set is too huge to be tractable, the design of split-plot experiments becomes computationally feasible. The designs were considered ideal because they were capable and efficient in estimating the fixed effects of the suitable statistical model given the split-plot design structure. The Split-Plot Design (SPD) is the complete block design which plays an important role in the fields of agriculture, medicine, and industries. This SPD is specifically suited for a two-factor experiment that has more treatments than can be accommodated by a complete block design. In an SPD, one factor is assigned to the main-plot. The assigned first factor is called the main-plot factor. The main-plot is then divided into subplots and the second factor is called the sub-plot factor. SPD is most used for (i) few experimental materials may be rare while the other experimental materials may be available in large quantity, (ii) the levels of one or more treatment factor or easy to change and the alteration of levels of other treatment factors are costly or time-consuming. Given the extensive study done in graph theory, it has developed to be a very broad subject in mathematics. Graphs are important because they are a visual way of expressing information. A graph shows data that is equivalent to many words. A graph can convey information that is difficult to express in words. A bipartite graph is a type of graph in which the entire graph may be divided into two bipartite sets, with edges connecting vertices in one set to vertices in the other. Vertex coloring is the procedure of assigning labels or colors to each vertex in a graph. The data set was also manually analyzed to validate the software-analyzed outcomes. R gave the same results as the manual analysis, showing that they were both correct. R is mainly command-based. The proposed approach is demonstrated using agricultural and industrial examples.

**INFERENCE ON THE TIME-DEPENDENT STRESS-STRENGTH RELIABILITY MODELS BASED ON FINITE MIXTURE MODELS ..... 268**

Krishnendu K., Annie Sabitha Paul, Drisya M., Joby K. Jose

Time-dependent stress-strength reliability engages with the chance of survival for systems with dynamic strength and/or dynamic stress. When a system is allowed to run continuously, each run will cause a change in the

strength of the system. The repeated occurrence of stress on the system over each run will affect the survival capacity of the system. In this paper, we consider the distribution of time taken for the completion of a run by the system follows gamma and the stress or strength of the system follows a finite mixture of lifetime probability models. Here we consider two cases in which the first case deals with stress and strength following a finite mixture of Weibull distribution and in the second case the stress and strength is assumed to follow a finite mixture of the power-transformed half-logistic distribution. Moreover, the strength of the system is assumed to decrease by a constant and the stress acting on the system is assumed to increase by a constant over each run. We obtained the expression of the stress-strength reliability function and explained the ML and Bayesian methods for the estimation of the reliability at various time points.

## DUAL EXPONENTIAL RATIO ESTIMATOR IN PRESENCE OF NON-RESPONSE ..... 285

Rafia Jan, T. R. Jan and Faizan Danish

The manuscript under consideration delves into a comprehensive exploration of the dual exponential ratio estimator, particularly in the context of non-response scenarios. In the following discourse, we will embark on an intricate journey through this research, emphasizing the pivotal aspects and findings that unravel the significance of this estimator in the realm of statistical estimation. The crux of this investigation revolves around evaluating the Mean Squared Error (MSE) and the Predictive Relative Efficiency (PRE) of the dual exponential ratio estimator. These two performance metrics serve as essential benchmarks for assessing the accuracy and effectiveness of the estimator. Notably, they play a crucial role in determining the estimator's suitability for practical applications, especially in situations where non-response is prevalent. To begin our exploration, it is imperative to understand the fundamental concept of the dual exponential ratio estimator. This estimator is a statistical tool employed in situations where traditional estimators may falter due to non-response, a phenomenon frequently encountered in surveys and data collection. It leverages a dual exponential model to address this challenge, making it a valuable addition to the toolkit of statisticians and researchers. The manuscript embarks on a rigorous theoretical analysis of the dual exponential ratio estimator's MSE and PRE. Through a series of mathematical derivations and proofs, the authors elucidate the underlying principles governing its performance. This theoretical foundation is crucial, as it not only establishes a solid framework for evaluating the estimator but also provides insights into its behavior under different conditions. However, theory alone can only take us so far. To validate the theoretical findings and assess the estimator's practical utility, numerical experiments are conducted. These experiments involve simulations and real-world data scenarios, allowing the authors to draw comparisons between the dual exponential ratio estimator and traditional estimators. The numerical results serve as a bridge between theory and application, offering empirical evidence of the estimator's prowess. In essence, this manuscript fills a critical gap in the field of statistical estimation by thoroughly investigating the dual exponential ratio estimator's performance in the presence of non-response. By juxtaposing its MSE and PRE with those of traditional estimators, it provides valuable insights into the potential advantages of adopting this novel approach. Moreover, the combination of rigorous theory and practical validation ensures that the findings are both intellectually sound and operationally relevant. The dual exponential ratio estimator, as explored and analyzed within these pages, emerges as a promising solution, backed by both theoretical rigor and empirical support. This research contributes not only to the theoretical foundations of statistics but also to its real-world applications, underscoring the estimator's potential to enhance the accuracy and reliability of estimation in the face of non-response complexities.

## PROCESS CAPABILITY ANALYSIS FOR NON NORMAL DATA BASED ON BOX-COX TRANSFORMATION THROUGH TESTS OF GOODNESS OF FIT ..... 297

J. Krishnan, R. Vijayaraghavan

Process capability analysis is an effective and efficient tool for quality assurance. When the distribution of the underlying quality characteristics is not normal, modifications of the basic process capability indices are required. Literature in process control provides avenues to resolve the issue of non-normality and data transformation is one of the approaches frequently applied in practice. Primarily the Box – Cox transformation (BCT) is employed to transform the non normal data into normal data which originally utilizes the method of maximum likelihood

estimation (MLE) to find the single transformation parameter  $\lambda$ . There are alternative methods to estimate the optimal parametric value  $\lambda$  using goodness of fit tests rather using MLE method. In order to bring improved estimates, this paper makes a fresh attempt to estimate process capability analysis (PCA) using transformed data through different goodness of fit tests. The simulation study uses variety of asymmetric behaviors from a Weibull distribution generating a random sample of 100 data points to find the best goodness of fit test for better process capability estimates that are compared to the standard of six sigma results for non-normal data. Final result shows that Shapiro-Wilk's (SW) and Artificial Covariate (AC) methods are performing well when compared to the method of MLE. Minitab software and R programming language were utilized for data simulation and analysis.

## **THE USE OF EXPERIMENTAL MODELLING IN THE PREDICTION OF PRODUCT RELIABILITY ..... 310**

Alena Breznická, Pavol Mikuš

When designing new systems and components, it is very important to correctly determine the degree and ability of the joint to withstand stress and load. Every new product that is intended for the market must meet the requirements for high safety and reliability during the entire life cycle. The presented article deals with the possibility of modelling the ability to withstand such a load, the principle of the interference method was used in the experimental modelling. The interference theory of reliability is based on the analysis of regularities and properties of two random variables that characterize reliability. Among these elementary properties from the point of view of reliability assessment, we can successfully use dependability and lifetime analysis. It originates from the concept of "safe life", which is deterministic, based on determining and respecting the values of reliability factors. The described approach assumes that a malfunction or a faulty function occurs when the strength limit of the object is exceeded, i.e., ability to withstand stress.

## **A STUDY ON PARTIALLY ACCELERATED LIFE TEST MODEL FOR GENERALIZED INVERSE RAYLEIGH DISTRIBUTION UNDER ADAPTIVE TYPE-II PROGRESSIVE HYBRID CENSORING ..... 320**

Intekhab Alam, Trapty Agarwal, Awakash Mishra, Aanchal Gaba

Modeling and examination of lifetime phenomena are the main aspects of statistical work in a wide variety of scientific and industrial areas. The area of lifetime information analysis has developed and extended quickly with respect to methodology, theory, and fields of applications. The point and interval maximum-likelihood estimations of generalized inverse Rayleigh distribution (GIRD) parameters and the acceleration factor are considered in this work. The estimation procedure is carried out for a partially accelerated step-stress model under adaptive Type-II progressive hybrid censored data. The biases and the mean square errors of the maximum-likelihood estimators are computed to assess their performances in the occurrence of censoring developed in this study through a Monte Carlo simulation study.

## **EFFECT OF CLASSICAL AND ROBUST REGRESSION ESTIMATORS IN THE CONTEXT OF HIGHDIMENSIONAL DATA WITH MULTICOLLINEARITY AND OUTLIERS ..... 335**

Muthukrishnan R, Karthika Ramakrishnan

Regression methods are used for the estimation and prediction in various fields of statistical study. It is a statistical method commonly used for determining the degree of relationship between a response and a number of explanatory variables. These explanatory variables may correlate each other and lead to multicollinearity. More than two predictor variables with high correlation show the existence of multicollinearity which results in the estimator having a high variance. Ordinary Least Square estimation fails to give a better regression estimator, when the model's presumptions are not met. This paper explores the various methods which can tolerate the

problems of multicollinearity and outliers. This study compares different types of regression estimators such as Ordinary Least Square, Robust, Ridge, and Liu by computing various error values such as Mean Absolute Error, Root Mean Square Error, Mean Absolute Percentage Error and R2 under real environment that has both multicollinearity and outliers. To compare the fit of the aforementioned regression models, the Akaike Information Criterion was also calculated. According to the error measures and AIC this study concludes that the Liu regression estimator performs well when compared with the other estimation methods.

## **EXPONENTIATED WEIBULL DISTRIBUTION: BAYESIAN ESTIMATION USING PROGRESSIVE TYPE I INTERVAL CENSORING ..... 342**

M. Kumar, K P Aswathi

A three-parameter distribution known as the Generalized Weibull (GW) or Exponentiated Weibull distribution is studied in this work. We construct Baye's estimators for the unknown parameters and present reliability function using progressive type I interval censoring data. Two different loss functions, namely, squared error loss and general entropy loss functions are applied to derive Baye's estimators. It is observed that there is no closed-form solution for Baye's estimators as well as for MLE. Hence, Lindley's approximation procedure is applied to obtain Bayesian estimator of unknown parameters, and Newton Rapson method is employed to obtain MLE's numerically. The corresponding reliability function is derived. Monte Carlo simulation is used to obtain MLE. Further, the performance of MLE and Bayes estimators are compared in terms of their respective MSE and Relative errors. It is noted by numerical computation that MLE's performs better than Bayes estimators. In addition to this, Bayes estimators obtained using Squared error loss function and general entropy loss function are compared. It is observed through numerical computation that general entropy loss function is better in terms of MSE.

## **METHODS FOR ENSURING AND PROVING FUNCTIONAL SAFETY OF AUTOMATIC TRAIN OPERATION SYSTEMS ..... 360**

I.B. Shubinsky, E.N. Rozenberg, H. Schäbe

The paper examines the specificity of artificial intelligence-based automatic train operation systems. Justifying the functional safety (FS) of such systems is quite difficult. The paper proposes a process for proving the functional safety of intelligent systems. A hybrid control system for a shunting locomotive was developed and analysed. It combines machine vision (MV), train protection devices and manual control by a driver. A model is presented that allows examining the functional safety of a locomotive control system layer by layer, i.e., evaluating the time to safety degradation depending on the component failure and subsequent requirement of bringing the locomotive to a complete stop. This allows to improve the FS of the shunting locomotive control system with machine vision from SIL 2 to SIL 3 and maintaining it during sufficiently long periods of time (over a quarter of the mean time to system failure). The mean time of faultless operation of a locomotive control system until it has to be brought to a complete stop for safety reasons can be increased three times. A general approach is proposed to design the functional safety of automatic train operation systems. It is based on the division of the information processing process into two subprocesses, i.e., internal intelligent information processing onboard the locomotive for the purpose of decision-making regarding track vacancy and communication of initial visual information to the operating driver for decision-making. The division of this process must be combined with redundant machine vision facilities, regular comparison of the outputs of the onboard and fixed machine vision facilities, redundant comparison outputs, smoothing of the outputs in the process of locomotive movement.



**SENSITIVITY AND PROFITABILITY ANALYSIS  
OF TWO-UNITS AMMONIA/UREA PLANT ..... 376**

Sara Salim A Oraimi, Syed Mohd Rizwan, Kajal Sachdeva

*This paper presents a reliability modelling of a two-unit ammonia/urea plant. Real maintenance data of the production plant have been used for this purpose. Four types of failure were noted: process, electrical, mechanical and instrumental failures. Both ammonia/urea formation units work in parallel and do not fail simultaneously. Various reliability indices of the plant, such as availability, busy period for repair, and expected number of repairs for each type of failure, have been obtained. Markov processes and regenerative point techniques are used for analysis. Profit analysis for the plant is also done, along with a graphical representation of various parameters. Finally, sensitivity analysis is carried out to see the impact of varied parameters on the profit function of the plant.*

**STUDIES ON A NEW MANPOWER MODEL WITH NONHOMOGENEOUS POISSON  
RECRUITMENT, PROMOTION AND LEAVING PROCESSES..... 387**

K. Suryanarayana Rao, K. Srinivasa Rao

*For proper utilization of manpower in any organization manpower modeling is needed. This paper addresses the two graded manpower model with non-stationary recruitment, promotion and leaving processes. Here it is assumed that the recruitment process in the first grade follows a NHP process which is further assumed that the promotion and leaving processes are also NHP processes. Using the difference-differential equations, the joint p.g.f of the number of employees in the organization at any time 't' is derived. The characteristics of the model such as the average number of employees in each grade, the average waiting time of an employee in each grade, the variance of the number of employees in each grade and the C.V of an employee in each grade are derived explicitly. The sensitivity analysis of the model with respect to the changes in parameter is also studied through numerical illustration. The comparative study between homogeneous Poisson recruitment and NHP recruitment is also discussed. This model also improves some of the earlier models as particular cases.*

**THE ROLE OF RECORD VALUES IN STATISTICAL INFERENCE: A REVIEW ARTICLE ..... 406**

Mahmoud A. Selim Alsanea

*The record values data have received the attention of researchers in statistics for over seven decades. Through these decades the records have played a significant and widely utilized role for statistical inference in parameter estimation, predicting future values, hypothesis tests, as well as stress-strength tests, and characterizing distributions. In this paper, the types of record values, some distributional properties, and statistical inferences of record values and their applications are reviewed. The purpose of this paper is to shed light on the role of record values in statistical inference. Therefore, we will examine this issue from two perspectives, the first perspective being estimation and the second perspective being prediction. These are through some of the most important lifetime distributions are Exponential, Weibull, Gumbel, Geometric, Pareto, Generalized exponential, Rayleigh, Lomax, and Nadarajah-Haghighi distributions. I hope that the findings of this paper will be useful for researchers in various fields and lead to further enhancement of research in record values theory and its applications.*

**APPLICATION OF THE FUZZY-SET THEORY TO ASSESS THE KNOWLEDGE OF  
ELECTRIC POWER INDUSTRY SPECIALISTS ..... 424**

V.Kh. Nasibov R.R. Alizade I.Y. Mastaliyev A.M. Ramazanli

*In most cases, the assessment of the knowledge of electric power industry workers is carried out according to a test scheme, where the correct answer is selected from the list of answers. All questions have the same difficulty*

and only the single correct answer gives a certain score. The article developed a universal model for assessing the knowledge of electric power industry workers, where using the theory of fuzzy logic and fuzzy inference, both the complexity of questions and the possibility of a partial correct answer are taken into account.

## **A CRITICAL LITERATURE REVIEW AND FUTURE PERSPECTIVE OF RAM APPROACHES FOR COMPLEX SYSTEMS IN VARIOUS PROCESS INDUSTRIES ..... 431**

Mausoof Sheikh, Dr. P.C. Tewari

*In the industrial systems there is a requirement that systems should work efficiently for long time. System performance is an important aspect for failure free operation but in real practice complete failure free operation of any production system is seldom possible. Detailed critical literature review for the past thirty-three years of Reliability, Maintainability and Availability (RAM) approaches has been carried out which can help to improve performance of Complex systems. Review of some papers provided the detailed information about past and current scenario of RAM practices in research field and industries. Different RAM tools and techniques extracted from the review may be helpful in qualitative and quantitative analysis of the complex systems. In this paper, author tried to focus on some major aspects of RAM approaches.*

## **EXACT AND CONDITIONAL BOUNDS FOR GENERALIZED CUMULATIVE ENTROPY..... 440**

Alexey V. Lebedev

*The differential entropy is a natural analog of the Shannon entropy for discrete distributions in respect to absolutely continuous distributions (with density). In modern studies, many other kinds of entropy have been introduced and analyzed, including various cumulative entropies, which are based not on the density but on the (cumulative) distribution function of random variable. Such characteristics can be used, for example, in computer vision, reliability theory, risk analysis, etc. We consider some generalizations of cumulative entropy, for a wide class of entropy generators. We use the methods of probability theory, calculus of variations and Cauchy-Bunyakovsky-Schwarz inequality. In the class of centered and normalized random variables, exact and conditional bounds are found as well as the distributions on which they are attained. By conditional bounds we understand bounds for one generalized cumulative entropy given the value of another entropy (in the class of random variables with zero mean and unit variance). This problem is analogous to the previously posed and partly solved problem on conditional bounds for expectations of sample maxima when we know the expected maximum of a sample of another size or expected maxima of two smaller samples.*

## **SOME PROPERTIES OF TSALLIS ENTROPY BASED ON A DOUBLY TRUNCATED (INTERVAL) RANDOM VARIABLE..... 448**

S. Jalayeria, G.R. Mohtashami Borzadarana, M. Khorashadizadehb

*In this paper, we first study doubly truncated (interval) Tsallis entropy and suggest doubly truncated (interval) cumulative residual Tsallis entropy (ICRT), which is an extension of cumulative residual Tsallis entropy (CRT) and the dynamic CRT defined by the aid of Sati and Gupta and of Kumar, respectively. We investigate some properties and characterization of this measure, such as its relation with doubly truncated Shannon entropy, mean residual (past) life, and hazard rate (or reversed hazard rate). Also, the twin measure, doubly truncated (interval) cumulative past Tsallis entropy, is determined, and some of its properties are studied. Moreover, their monotonicity and related aging classes of distributions are expressed, and the upper (lower) bound for them is acquired. In the end, we propose four nonparametric estimators and compare their performance by utilizing simulation data. Also, being based on the best-proposed estimator, a real data set is additionally examined.*

**GENERALIZED X-EXPONENTIAL BATHTUB SHAPED FAILURE RATE DISTRIBUTION AND ESTIMATION OF RELIABILITY OF MULTICOMPONENT STRESS-STRENGTH ..... 465**

Faryal Shabbir, Abdul Khaliq

*In an engineering setup, one is interested to know and determine the reliability of the system of different components. These components are usually subjected to different kinds of stress, and the reliability of the components needs to be estimated under stress. In this paper, we aim to estimate the reliability of a multicomponent stress-strength model assuming that the components of the system are working independently with a common life distribution. The system follows a comparatively new distribution named as; Generalized X-Exponential bathtub failure rate distribution. This paper studies the usefulness of this distribution in terms of estimating the maximum likelihood estimate of the reliability parameter and its asymptotic confidence intervals. Paper uses methods of parametric estimation and reliability estimation. Results are computed using Monte Carlo simulation for small samples. Real data set is presented to evaluate the performance of Generalized X Exponential Distribution (GXED) reliability estimator. Findings show that with the usage of proposed distribution, estimator of reliability parameter fits very well to the real-world situations*

**DEVELOPMENT OF AN INTEGRATED SAFETY SYSTEM FOR PRODUCTION FACILITIES: THE PROBLEM STATEMENT AND THE PROPOSED SOLUTION ..... 474**

Evgeny Gvozdev

*The article focuses on explosion and fire hazards at production facilities of enterprises where flammable liquids and gases, categorized by explosion and fire risks, are processed, handled, transported, and stored. The goal to be attained and the tasks to be solved towards this end are formulated in the article. Consolidated areas of knowledge, accumulating results of research into risk assessment within systems of integrated safety implemented at production facilities, are considered by the author. A model for development of a novel set of research and methodological instruments (methods, techniques, software and hardware) is presented for its further practical application. The problem of developing integrated safety systems for industrial facilities, posing explosion and fire hazards, as well as the solution, are presented by the author for the first time. The novelty of the solution lies in the computation of validity of the practical application of a novel set of research and methodological instruments. A reduction in damage from accidents and fires at production facilities is demonstrated. Ultimately, the socio-economic problem of reducing damage from accidents and fires is solved not only by Russian production facilities, but also by government agencies, including the EMERCOM of Russia (Ministry of the Russian Federation for Civil Defense, Emergencies and Elimination of Consequences of Natural Disasters), Ministry of Labor and Social Protection of Russia, and Federal Environmental, Industrial and Nuclear Supervision Service of Russia.*

**TRANSIENT AND METAHEURISTIC COST SCRUTINY OF MX/G(A, B)/1 RETRIAL QUEUE WITH RANDOM FAILURE UNDER EXTENDED BERNOULLI VACATION WITH IMPATIENT CUSTOMERS ..... 488**

Rani R, Indhira K

*The transient and metaheuristic cost analysis of a MX/G(a, b)/1 retrial queue with random failure during an extended Bernoulli vacation with impatient clients is covered in this study. Any batch that arrives and discovers the server is busy, down, or on vacation joins an orbit. In the alternative, only one new customer from the group joins the service right away, while the others join the orbit. After providing each service, the server either waits to serve the following customer with probability  $(1 - \theta)$  or goes on vacation with probability  $\theta$ . It has been found that these systems express steady-state solutions and are dependent on time probability generating functions in consideration of their Laplace transforms. We also discuss a few exceptional and particular instances. After that, the impact of different parameters on the system's effectiveness is evaluated. We are also talking about ANFIS.*

*Additional approaches employed in this study to swiftly determine the system's optimum cost include genetic algorithms (GA), artificial bee colonies (ABC), and particle swarm optimization (PSO). We also examined the graph-based convergence of several optimization algorithms.*

**REPORTING METHODOLOGY AND ALGORITHM OF MODES OF COMPLEX ENERGY SYSTEMS WITH PHASE COORDINATES ..... 510**

Huseyngulu Guliyev, Famil Ibrahimov

*A mathematical model, algorithm and program have been developed to study any types of complex asymmetric steady-state modes and transient processes of a multi-machine power system with a renewable energy source in phase coordinates, the results of which can be used in the operational control of power system operating modes with any type of emergency automation. The developed methodology and software package can also be used in industry to check the possibility of long-term operation in the considered asymmetrical mode from the point of view of the operating conditions of the system generators and electrical receivers, to determine the need to use baluns, to select their parameters and installation locations, to ensure the efficiency of asymmetrical modes, as well as for conducting various tests and analyzing accidents that have occurred.*

**BAYES ESTIMATION OF CAPABILITY INDEX USING THREE-PARAMETER WEIBULL DISTRIBUTION ..... 523**

Sonam Gubreley, Ankita Gupta, Satyanshu K. Upadhyay

*The process capability index is an important tool used in quality control and process improvement. Generally, the index is estimated under the assumption of a normal distribution, although some other distributions are also recommended in the literature. This paper instead considers a three-parameter Weibull distribution and obtains an estimate of the process capability index under the Bayesian framework. Bayesian development is based on the use of non-informative priors and the posterior sample-based inferences are drawn using an important Markov chain Monte Carlo technique, namely, the Gibbs sampler algorithm. Finally, a numerical illustration based on two real datasets is provided.*

**A NEW ALGORITHM TO SOLVE MULTI-OBJECTIVE TRANSPORTATION PROBLEM WITH GENERALIZED TRAPEZOIDAL FUZZY NUMBERS ..... 531**

Ramakant Sharma, Sohan Lal Tyagi

*Transportation Problem is a specific type of linear programming problem (LPP). Today, in the real world, the decision maker handles the multi-objectives at the same time. Fuzzy Concepts are used in LPP to handle the uncertainty and vagueness of data. This paper presents a new algorithm to solve a special type of fuzzy transportation problem (FTP) with the generalized trapezoidal fuzzy numbers (GTpFN) in which the decision maker is not certain about the exact value of transportation charge and the availabilities and requirements are the real numbers. In this Proposed Algorithm first, the fuzzy multi-objective transportation problem (FMOTP) is converted into a Crisp multi-objective transportation problem (MOTP) by the Proposed ranking function, and then the Crisp MOTP is transformed into a single objective transportation problem using the sum of objective functions values. The proposed algorithm gives an efficient compromise solution of FMOTP. To elaborate the proposed algorithm, one numerical example is solved.*

**USE OF THERMAL IMAGING METHOD OF CONTROL FOR INSPECTION OF BUILDING STRUCTURES FOR TIGHTNESS ..... 544**

Sofia Skachkova, Anton Avgutsevichs

*In any industry related to the construction of buildings and structures we have heard about the need to assess the technical condition of various objects to assess and analyze the risks associated with the possible collapse of buildings (structures), loss of life and high costs to eliminate these consequences. Since many objects fail over time, and in general to determine the wear and tear and the possible term of further safe operation, it is necessary to conduct a technical survey. The article describes the principle of operation of thermal imaging devices for determining the reliability of building structures in residential premises, and also raises problems, the solution of which can simplify the use of thermal imaging devices in the inspection of buildings and structures and reduce the economic costs of damage compensation in case of timely detection and elimination of any defects.*

**ON THE CHARACTERIZATION AND APPLICATIONS OF A THREE-PARAMETER IMPROVED WEIBULL-WEIBULL DISTRIBUTION ..... 551**

A. S. Mohammed, B. Abba, I. Abdullahi, Y. Zakari, A. I. Ishaq

*Parametric modeling of complex lifetime data characterized with nonmonotone hazard rate (NMHR) has in recent years attract the interest of many researchers and practitioners. The three-parameter improved Weibull-Weibull distribution introduced in 2022 has demonstrated a better NMHR modeling potential in the analysis of several failure times identified with bathtub hazard rate (BHR). In this study, we present the characterization, properties and two data sets' applications of the distribution. Various properties of the distribution obtained, include moment generating function, moments, skewness, kurtosis, and some types of entropy. Numerical results for mean, variance, skewness, and kurtosis are computed using simulation studies. Estimation of the distribution parameters is performed using the method of maximum likelihood, and the estimation method is assessed by Monte Carlo simulation experiments. The two illustrations further ascertain the capability of the model for modeling lifetime data from different scientific investigation areas.*

**MARSHAL-OLKIN ALPHA POWER INVERSE RAYLEIGH DISTRIBUTION: PROPERTIES, ESTIMATION AND APPLICATIONS ..... 564**

Ismaila Olawale Adegbite, Kayode Samuel Adekeye, Olubisi Lawrence Aako

*In this study, a new three-parameter distribution is introduced by extending the two-parameter Alpha Power Inverse Rayleigh distribution using Marshall-Olkin G approach. The proposed Marshall-Olkin Generalized Alpha Power Inverse Rayleigh (MOAPIR) distribution generalizes the Marshall-Olkin Inverse Rayleigh, Alpha Power Inverse Rayleigh, and Inverse Rayleigh distribution. The characterization and statistical properties of the proposed distribution such as hazard rate function, reversed hazard rate function, quantiles, moments, and order statistics were derived. The estimation of the MOAPIR distribution parameters is derived using the maximum likelihood estimation method. The performance of the proposed distribution was compared with other competing distribution using two real-life data. The goodness of fit criteria and the distribution function curve showed that the proposed distribution provides a better fit than other competing distributions of the same family of heavily positive skewed distribution.*

**ON DIFFERENT CLASSICAL ESTIMATION APPROACHES FOR TYPE I HALF LOGISTIC-  
TOPP-LEONE- EXPONENTIAL DISTRIBUTION ..... 577**

Akeem Ajibola Adepoju, Sauta S. Abdulkadir, Danjuma Jibasen

*This paper aims to propose six methods of parameter estimation in order to examine the behavior of the new Type I Half Logistic Topp-leone Exponential distribution. The methods taking into consideration are Maximum Likelihood, Anderson Darling, Least Squares, Cramer von Mises, Maximum Product of Spacing, and Weighted Least Squares Methods. The results show that all the methods are consistent, since the estimates approach the true value of the parameters for all the methods. The bias, mean square error and mean relative estimates decay as the sample size is raised. The estimates of the six methods obtained for the model, indicated that MPS estimates is the closest to the true value of the parameters across the low, moderate and high sample sizes, invariable, the MPS produces the least biasness. Buttress more, the MPS produces the least MSE all through and remain the best estimator for low, moderate and high sample size of the model. Conclusively, MPS is the most consistent among the estimators for the model.*

**ON THE Q-RAYLEIGH DISTRIBUTION AND ITS APPLICATIONS..... 588**

Ibrahim Sadok

*This paper introduces the two-parameter q-Rayleigh distribution, a powerful extension of the classical Rayleigh model for analysing real-world data. Compared to the Rayleigh, the q-Rayleigh incorporates a novel pathway parameter q, offering greater flexibility in capturing diverse data patterns. We delve into the mathematical properties of the q-Rayleigh, including its hazard rate function and quantile function, and explore parameter estimation through maximum likelihood methods. We demonstrate its superior fit compared to the widely-used Rayleigh distribution for real-world data. Moreover, we explore its application in reliability analysis. This comprehensive study makes the q-Rayleigh a compelling choice for modelling data exhibiting gradual transitions and enhanced flexibility.*

**MI-K-MEAN ALGORITHM: A NEW APPROACH FOR FINANCIAL RISK ANALYSIS WITH  
MISSING DATA IMPUTATION IN BIG DATA ..... 603**

Ravindra Kumar, Diwakar Shukla, Kamlesh Kumar Pandey, Sagar, M.P.,India , Sagar, M.P., India, Amarkantak, M.P., India

*The data mining is a tool of searching information from the data warehouse. Several mining algorithms exist in literature, one of the most common is the usual K-mean procedure. This generates centroids after every round of iteration. It is assumed that sample data is completely cleaned and noise free before the start of execution of the usual K-mean algorithm. If  $\alpha\%$  values are missing in sample data then after cleaning only  $(100-\alpha)\%$  values are available for the execution of the usual K-mean algorithm. Such bears a loss of information that affects the decision. This paper considers this problem and resolves such issue by replacing the missing data through imputed values calculated by the available values, called Mean Imputation (MI). It helps in financial risk analysis quite a lot because of risk prediction being taken on a larger sample (cleaned and imputed both). Several imputation procedures are available in literature. This paper considers the financial risk data as sample where the missing values of sample are imputed by the usual Mean-Imputation (MI) method and then on complete sample. Proposed MI-K-mean strategy is compared with no imputation usual procedure and found more efficient over the four-evaluation criterion of cluster formation while applying on risk data analysis.*

**A TYPE I HALF LOGISTIC TOPP-LEONE INVERSE LOMAX DISTRIBUTION WITH APPLICATIONS IN SKINFOLDS ANALYSIS..... 618**

Akeem Ajibola Adepoju, Sauta S. Abdulkadir, Danjuma Jibasen, Jamiu S. Olumoh

*This paper proposed a novel distribution parameterized by four parameters. This is achieved by compounding the potentials properties of the Type I half logistic topp-leone generalized distribution family with the properties of the inverse lomax distribution to form the novel Type I half logistic topp-leone inverse lomax distribution. The novel distribution is potentially capable of extending classical inverse lomax distribution. The potentiality of the shape of the probability density function of the novel distribution is worth recognizing since it produces right skewed, approximately normal, left skewed and a reverted J-shaped. Decreasing life failure shape is also observed. Distinctive features of the novel distribution such as moments, entropy, moment generating function, reliability and hazard function were derived. The estimation method explored in this study is maximum likelihood estimation. It is adopted to estimate the novel distribution unknown parameters. Real life data set was adopted to investigate the potentiality and applicability of the novel model. The type I half logistic topp-leone inverse lomax distribution outperform the recent models*

**A COMPRESSIVE STUDY ON FAULT DETECTION AND DIAGNOSIS FOR RELIABLE OPERATION OF HVAC, ENERGY BUILDINGS AND MACHINERIES ..... 631**

M. S. Patil, G.M. Malwatkar

*In Heating, Ventilation, and Air Conditioning (HVAC) systems, faults can be occurred due to various reasons such as drift deviation, valvelfan failure, water clogging, air filter obstruction, temperature sensor failure and so on. Similarly in electrical machineries faults can be occurred due to multiple causes such as phase reversal, over or under voltage, starter open/short circuit, bearing problems, insulation breakdown, overloading, thermal unbalance, environmental as well as other technical issues. The faults analysis at various stages of electrical systems are critically important for reliable operation of the system. In view of reliability and safety operations of modern sophisticated electrical systems, faults analysis and its diagnosis are necessary to avoid unaccountable losses. The faults at various stages, its causes, methods of detection and diagnosis, fault classifications are included in this work. The comment on effectiveness methods of detection of fault and diagnosis are included for electrical systems. In the industries, systems are incorporated with monitoring capacity for detection of faults at easy and early stage. This paper mainly focused on advancements in fault detection and diagnosis (FDD) methods with short review of various recent methods. This includes system information representation, methods of FDD, description of faults, fault classification, and decision actions related to maintenance, providing a systematic overview of the current state of FDD. Furthermore, the paper underscores the pivotal roles of FDD in electrical systems, emphasizing its effectiveness in identifying faulty states and taking pre-emptive actions against potential failures or accidents. The discussion extends to developments of current research in FDD approaches for electrical machineries with system monitoring, accompanied by short review of diverse and valuable FDD methodologies. The study concludes by addressing comments on recent trends, future directions, challenges, and prospective solutions in the hybrid and dynamic landscape of FDD.*

**M/M/C QUEUE WITH MULTIPLE WORKING VACATIONS AND SINGLE WORKING VACATION UNDER ENCOURAGED ARRIVAL WITH IMPATIENT CUSTOMERS ..... 650**

Prakati P, Julia Rose Mary K

*This paper demonstrates an M/M/C queuing model with Multiple working vacations and also single working vacation under encouraged arrival with impatient customers. The queuing model with the servers adopting multiple working vacation policy and single working vacation are determined separately and it is observed that the servers during working vacation(s) will be serving the customers at a slower service rate when compared during regular busy period. In addition to the above conditions, if there is a rapid increase in the customers' arrival i.e, if encouraged arrival occurs and due to this sudden growth of the queue, there may be an impatience in the behaviour of the customer. With these considerations, an M/M/C Queuing model is analysed with two*

*vacation policies separately by applying PGF method and thus the performance measures for an M/M/C Queue with Multiple Working Vacations and Single Working Vacation under Encouraged arrival with impatient customers are evaluated.*

## **BEHAVIORAL ANALYSIS AND MAINTENANCE DECISIONS OF WOOD INDUSTRIAL SUBSYSTEM USING STOCHASTIC PETRI NETS SIMULATION MODELING ..... 663**

Urvashi, Shikha Bansal

*This study aims to optimize the productivity of the plywood manufacturing system within the wood industry. A Petri nets simulation-based technique has been used to evaluate the availability analysis of the plywood manufacturing system. A Petri nets model is created to represent the modeling of the plywood system. The model is subsequently simulated using the licensed program Petri Nets (PN) GRIF 2023.7. This simulation is used to evaluate the performance of the system. In the PN simulation model, timed transitions are fired based on the failure and repair rate of the system. Immediate transitions, on the other hand, have their own guard function for firing which is coded using a logical AND-OR gate. This study also assesses the impact of the repairman on the system's availability. The system's availability is optimized by increasing the number of repairmen. However, once a specific number of repairmen is reached, the system's availability remains constant. This research is highly valuable for determining the optimal number of maintenance staff needed for the wood industrial system.*

## **PREDICTIVE MAINTENANCE SCHEME FOR PHASED MISSION SYSTEMS ..... 675**

Preeti Wanti Srivastava, Satya Rani

*In both industrial and military fields, many systems are phase mission systems (PMSs) which execute mission composed of different phases in sequence. The structure, failure behaviour, and working condition of such a system may change from phase to phase. Maintenance actions comprising corrective and preventive maintenance schemes studied in the literature are aimed at retaining the maintained system in a proper condition and improving its availability and extending its life. The present paper deals with finding optimal periodic inspection time using multi-objective criteria comprising objectives of minimizing expected maintenance cost incurred due to predictive, breakdown and periodic maintenance of a PMS, and maximizing its expected residual lifetime. The predictive maintenance is condition-based preventive maintenance that anticipates system failures in order to plan timely interventions on the system and hence improve its performance. The dependency is modelled using Gumbel-Hauggaard copula. An aircraft flight PMS comprising Taxiing phase, Take-Off phase, Cruising phase and Landing phase has been used to illustrate the method developed.*

## **PROFIT ANALYSIS OF REPAIRABLE JUICE PLANT ..... 688**

Rahul, Mohit Yadav, Hemant Kumar

*Juice is a non-fermented beverage that is obtained by squeezing fruits to increase immunity. 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 a juice plant are discussed. The juice plant has 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 reliability measures.*



**A LITERATURE REVIEW ON DEVELOPMENT OF QUEUEING NETWORKS ..... 696**

V. Narmadha, P. Rajendran

*This study conducts a quantitative research survey on the development of queueing networks over years. Development is a process of gradual change that takes place over many years, during which a theory slowly progress and attain a good state. Queueing theory has been through many developments which made its existence inevitable in every field. Queueing networks can be considered as a collection of nodes, where each node stands for a service facility. It has been proved to be a powerful and versatile tool for modelling facilities in manufacturing units and telecommunication networks. This paper presents the development in Queueing networks and its types over years. This paper's main objective is to give all the analysts and researchers the knowledge about the evolution that happened in Queueing networks over years.*

**MODELING AND ANALYSIS OF SINE POWER RAYLEIGH DISTRIBUTION: PROPERTIES AND APPLICATIONS ..... 703**

Aadil Ahmad Mir, S.P.Ahmad

*In this manuscript, a new probability model named as Sine Power Rayleigh distribution (SPRD) is proposed using a Sine-G function as generator. Various statistical properties of this new distribution were investigated, including the survival function, hazard function, reverse hazard rate, cumulative hazard function, mills ratio, quantile function, moments, moment generating function, conditional moments, entropy, and order statistics. The parameters of the proposed distribution were estimated using the method of maximum likelihood estimation. To assess the model's versatility and applicability, we conduct analyses on two real life data sets. The outcomes affirm the superior performance of the newly proposed model SPRD as compared to existing models.*

**NEW DISCRETE DISTRIBUTION FOR ZERO-INFLATED COUNT DATA ..... 717**

Peer Bilal Ahmad, Mohammad Kafeel Wani

*Over-dispersed models are commonly utilized when the variation is more than what the model actually predicts. Since one of the reasons for over-dispersion is the large number of zeros, we employ zero-inflated models instead of more traditional ones to handle this observed occurrence. We present a zero-inflated version of a discrete distribution that was developed in 2021 in our research. Significant statistical characteristics of the suggested model have been identified, such as moments, the over-dispersion feature, generating functions, and related measures, among others. We have carried the parametric estimation using the maximum likelihood estimate. Maximum likelihood estimates are checked for usefulness in a simulation exercise. We evaluated the applicability of our developed model using three real-world data sets,*

**STOCHASTIC OPTIMIZATION AND RELIABILITY ANALYSIS OF MUSHROOM PLANT .... 729**

Shakuntla Singla, Sonia, Poonam Panwar

*In the present paper the reliability model for availability analysis of mushroom plant is developed in three sub-units like water pump, winter cold standby unit A.C., and packing machine. We assume a doctor of mushroom and workers are available who examines and repairs the elements as when we need. A mathematical model of the system is developed by using all these considerations. MTSE, Availability, server of busy period and expected number of servers visit of mushroom plant are determined with the assistance of RPGT. Graphs and tables are draw to depict the behavior of various parameters such as MTSE, Availability, server of busy period and expected number of servers visits and the effect of various parameters of the plant is analyzed when repair and failure rate both are vary and also when one of them is constant*

**NUMERICAL INVESTIGATION OF RETRIAL QUEUEING INVENTORY SYSTEM WITH A CONSTANT RETRIAL RATE, WORKING VACATION, FLUSH OUT, COLLISION AND IMPATIENT CUSTOMERS ..... 744**

G. Ayyappan, N. Arulmozhi

*The retrial queueing inventory system with working vacation, flush out, balking, breakdown, and repair, as well as a constant retrial rate and orbital client collision are all examined in this study. We made the assumption that customers arrive through a Markovian arrival process and that they would get phase-type services from the server. The inventory is replenished using a (s, S) and (s, Q) strategy, and it is expected that the replenishment time will follow an exponential distribution. If there are zero inventory items, no customers in the orbit, or both, the server will go into working vacation mode. When a customer retries an orbit while the server is serving arriving customers, the orbital customer may collide with the arriving customer during that retry, in which case both of them will be shifted back into orbit; otherwise, the orbital customer may avoid colliding with the arriving customer and may rejoin the orbit for another retry. The number of customers in the orbit and the inventory level may be found in the steady state. A cost analysis is produced along with the establishment of various important performance measures. Moreover, some numerical examples are provided to clarify our mathematical notion.*

**ON  $\varphi$ -CONHARMONICALLY FLAT LORENTZIAN PARA-KENMOTSU MANIFOLDS ..... 764**

I. V. Venkateswara Rao, S. Sunitha Devi, K. L. Sai Prasad

*The present paper deals with a class of Lorentzian almost paracontact metric manifolds namely Lorentzian para-Kenmotsu (briefly LP-Kenmotsu) manifolds. We study and have shown that a quasiconformally flat Lorentzian para-Kenmotsu manifold is locally isomorphic with a unit sphere  $S_n(1)$ . Further it is shown that an LP-Kenmotsu manifold which is  $\varphi$ -conharmonically flat is an  $\eta$ -Einstein manifold with the zero scalar curvature. At the end, we have shown that a  $\varphi$ -projectively flat LPKenmotsu manifold is an Einstein manifold with the scalar curvature  $r = n(n - 1)$ .*

**DATA ANALYSIS AND CLASSICAL ESTIMATION METHODS OF THE BOUNDED POWER LOMAX DISTRIBUTION ..... 770**

Amal S. Hassan, Asma M. Khalil, Heba F. Nagy

*In this work, a novel bounded three-parameter power Lomax distribution termed the unit power Lomax (UPLoD) is presented. The UPLoD is capable of handling data with left and right skewed shapes according to its probability density function. Additionally, according to the hazard rate function, the distribution may be used to analyse data containing J-shaped hazard rates. It is possible to determine some of the distribution's mathematical characteristics like moments, probability-weighted moments, incomplete moments, residual and reversed residual life, quantile function, stress strength model, and entropy (Rényi, Havrda and Charvát, Tsallis, and Arimoto) measures. The Cramér–von Mises, weighted least squares, maximum likelihood, Anderson–Darling, maximum product of spacing, and least squares approaches are among the conventional estimating techniques that are taken into account. The performance of the resulting estimates is compared using a Monte Carlo simulation based on some precision metrics. An actual data application is presented using water capacity data, and data about the Susquehanna River's maximum flood levels to show the importance of the new distribution compared to several other known distributions.*

**LIMIT CYCLES OF LENGTH TWO IN THE RIKKER MODEL AND THEIR APPLICATION IN FISHING** ..... 790

Gurami Tsitsiashvili, Tatyana Shatilina, Marina Osipova, Tatyana Radchenkova

*The paper investigates the limit cycle of length two in the Rikker model. It is established that the dependence of the ratio of the maximum value of the cycle to the minimum depends monotonously and almost linearly on the growth coefficient of the Rikker model. Models of the parity shift of the limit cycle of length two are constructed, which is provided by a simultaneous sharp decrease/increase in the growth coefficient. On the example of the Amur salmon in 1994 It is shown that a decrease in the growth coefficient, leading to a shift in the parity of the cycle of length two, is accompanied by a low temperature during the life cycle of pink salmon, when the pink salmon population is in a state of spawn and when the young are rolling.*

**INFERENCE ON THE INVERSE POWER BURR-HATKE DISTRIBUTION UNDER TYPE II CENSORING** ..... 796

Pavitra Kumari, Vinay Kumar

*There are many real-life situations, where data require probability distribution function which have decreasing or upside-down bathtub (UBT) shaped failure rate function. The inverse power burr hatke distribution consists both decreasing and UBT shaped failure rate functions. Here, we address the different estimation methods of the parameter and reliability characteristics of the inverse Pareto distribution from both classical and Bayesian approaches. We consider classical estimation procedures to estimate the unknown parameter of inverse power burr-hatke distribution, such as maximum likelihood. Also, we consider Bayesian estimation using squared error loss function based joint priors. The Monte Carlo simulations are performed to compare the performances of the obtained estimators in mean square error sense. Finally, the flexibility of the proposed distribution is illustrated empirically using one real-life datasets. The analyzed data shows that the introduced distribution provides a superior fit than some important competing distributions such as the Weibull, inverse Pareto and Burr-Hatke distributions.*

**A COMPARATIVE STUDY OF INVENTORY MODELLING: DETERMINISTIC OVER STOCHASTIC APPROACH** ..... 804

Lalji Kumar, Pratima Singh Ghoshi, Shreyashi Saxena, Kajal Sharma

*This research study provides a comprehensive comparison of two critical approaches to inventory modelling—deterministic and stochastic. The deterministic model employs traditional optimization techniques to optimize complex systems, while the stochastic model leverages Particle Swarm Optimization (PSO) simulations to tackle the challenges posed by uncertain dynamics. This approach enables us to develop effective strategies for optimizing complex systems. After conducting sensitivity analyses, it was found that the deterministic model oversimplifies demand dynamics, whereas the stochastic model more adeptly captures market uncertainties. As a result, this study suggests that businesses adopt stochastic approaches to inventory management to better engage in adaptive decision-making, contingency planning, optimal resource allocation, risk mitigation, and realistic performance metrics. The research provides valuable insights for businesses seeking to navigate the complexities of modern supply chains.*

**SEQUENTIAL TESTING PROCEDURE FOR THE PARAMETERS OF INVERSE DISTRIBUTION FAMILY ..... 819**

K. S. Chauhan, A. Sharma

*The sequential probability ratio test is a powerful statistical tool that is frequently employed for hypothesis testing, parameter estimation, and statistical inference. The aspect of robustness is of utmost importance when employing SPRTS in practical applications. Past studies have investigated the robustness of SPRTS for specific distributions. We have developed SPRTS for a family of inverse distributions that includes eleven distinct distributions. The primary objective of this study is to investigate and evaluate the robustness of SPRTS under various conditions and distributions, focusing on the parameters of the inverse distribution family. SPRTS efficacy is measured using OC and ASN functions. This study comprehensively covers the construction and rigorous evaluation of SPRTS, particularly in testing simple null hypotheses against simple alternative hypotheses. Additionally, we investigate the robustness of SPRTS under various factors, including the presence of other parameters and specified coefficients of variation. Conclusive results, graphic representations, tables, and acceptance and rejection regions add clarity to the findings.*

**CONFIDENCE INTERVAL USING MAXIMUM LIKELIHOOD ESTIMATION FOR THE PARAMETERS OF POISSON TYPE RAYLEIGH CLASS MODEL ..... 832**

Rajesh Singh, Preeti A. Badge, Pritee Singh

*In this research paper, confidence interval using maximum likelihood estimation is obtained for Poisson type Rayleigh class for the parameters. The failure intensity function, mean time to failure function and likelihood function for the parameter is derived. Confidence interval has been obtained for the parameters using maximum likelihood estimation. To study the performance of proposed Confidence interval, average length and coverage probability are calculated by using Monte Carlo simulation technique. From the obtained intervals, it is concluded that Confidence interval for the parameters perform better for appropriate choice of execution time and certain values of parameters.*

# ESTIMATION OF DIFFERENT ENTROPIES OF INVERSE RAYLEIGH DISTRIBUTION UNDER MULTIPLE CENSORED DATA

HEMANI SHARMA AND PARMIL KUMAR



Department of Statistics, University of Jammu, J&K.  
hemanistats@gmail.com, parmil@yahoo.com

## Abstract

*The inverse Rayleigh distribution finds widespread applications within life testing and reliability research. Particularly, it proves invaluable in scenarios involving multiple censored data points. In this context, the Renyi, Havrda, Charvat, and Tsallis entropies of the inverse Rayleigh distribution are efficiently calculated. The maximum likelihood approach is used to get the estimators, as well as the approximate confidence interval. The mean squared errors, approximate confidence interval, and their related average length are computed. To illuminate the behavior of estimates across varying sample sizes, a comprehensive simulation study is conducted. The outcomes of the simulation study consistently reveal a downward trend in mean squared errors and average lengths as the sample size increases. Additionally, an interesting finding emerges as the censoring level diminishes. The entropy estimators progressively converge towards their true values. For practical demonstration, the effectiveness of the approach is showcased through the analysis of two real-world datasets. These applications underscore the real-world relevance of the methodology, further validating its utility in addressing complex scenarios involving censored data and inverse Rayleigh distributions.*

**Keywords:** inverse Rayleigh distribution, Renyi entropy, Havrda and Charvat entropy, Tsallis entropy, multiple censored.

## 1. INTRODUCTION

The concept of entropy measurement is essential in many fields, including statistics, economics, and physical, chemical, and biological phenomena. The concept of entropy was first proposed as a thermodynamic state variable in classical thermodynamics, and it is based on principles from probability theory and mathematical statistics. Although the term information theory does not have a precise meaning, it can be considered of as the study of problems involving any probabilistic system. Entropy is referred to as the amount of information found in the sample. One of the most important aspects of statistics is the study of probability distributions. Every probability distribution contains some element of uncertainty. Entropy is a phenomenon that can be utilised to provide a quantitative estimate of uncertainty. Entropy is also a measure of disorder or randomness in a probabilistic system having a large number of random states with equal probability, and is zero when the system is in a specified state with no uncertainty. In other terms, a random variable's entropy is a measure of the amount of information required to explain a random variable on average. Shannon [13] established the concept of entropy as a measure of information. Here, we focus our attention on three entropy measures- the Renyi [11], Havrad and Chavrat [7], Tsallis entropies [15]. The Renyi entropy [11] comes from information theory, whereas the Tsallis entropy [15] comes from statistical physics, and both have a wide range of applications in their respective fields. These three entropy measures are defined, accordingly, for

an arbitrary variable  $X$  with the Probability Density Function (PDF)  $f(x; \varphi)$ , where  $\varphi$  denotes the corresponding parameters.

$$R_\delta(X; \varphi) = \frac{1}{1-\delta} \log \left[ \int_{-\infty}^{\infty} f(x; \varphi)^\delta dx \right] \quad (1)$$

where  $\delta \neq 1$  and  $\delta > 0$ , and

$$HC_\delta(X; \varphi) = \frac{1}{2^{1-\delta} - 1} \left[ \int_{-\infty}^{\infty} f(x; \varphi)^\delta dx - 1 \right] \quad (2)$$

where  $\delta \neq 1$  and  $\delta > 0$ , and

$$T_\delta(X; \varphi) = \frac{1}{\delta - 1} \left[ 1 - \int_{-\infty}^{\infty} f(x; \varphi)^\delta dx \right] \quad (3)$$

where  $\delta \neq 1$  and  $\delta > 0$ ,

Lord Rayleigh [12] initially proposed the Rayleigh distribution in relation to an acoustic problem. Since then, a great deal of work has been done in numerous domains of science and technology to improve this distribution. The Rayleigh distribution's hazard function is an increasing function of time, which is an important property. If the random variable  $Y$  has a Rayleigh distribution, the random variable  $X = \frac{1}{Y}$  has an inverse Rayleigh distribution (IRD). Trayer [14] proposed the Inverse Rayleigh distribution (IRD). The IRD is used in a variety of applications, including as life tests and reliability studies. A random variable  $X$  is said to have inverse Rayleigh distribution if its PDF and CDF has the following form:

$$f(x; \sigma) = \frac{2\sigma^2}{x^3} \exp \left[ - \left( \frac{\sigma}{x} \right)^2 \right]; x > 0, \sigma > 0 \quad (4)$$

$$F(x; \sigma) = \exp \left[ - \left( \frac{\sigma}{x} \right)^2 \right]; x > 0, \sigma > 0 \quad (5)$$

Wong and Chan [17] explored the entropy of ordered sequences and the order statistic. The entropy of upper record values was studied by Baratpour et al. [4], whereas the entropy of lower record values was proposed by Morabbi and Razmkhah [?]. Abo-Eleneen [1] discussed the entropy of progressively censored samples, Cho et al. [5] estimated the entropy for the Rayleigh distribution via doubly-generalized Type II hybrid censored samples using maximum likelihood and Bayes estimators, and Hassan and Zaky [6] investigated point and interval estimation of the Shannon entropy for the for the inverse Weibull distribution under multiple censored data. Bantan et al. [3] used multiple censored data to derive the Renyi and q-entropy for the inverse Lomax distribution. To measure the Lomax distribution's dynamic cumulative residual Renyi entropy, Al-Babtain et.al [2] explored the Bayesian and non-Bayesian techniques.

However, the estimation of entropy measures for the inverse Rayleigh distribution (IR), such as the Renyi, Havrad, and Chavrat, Tsallis entropies, still an unresolved subject . The problem is examined in the context of multiple censored data in this study, which fills the gap. This is a common scenario in which many censoring levels are logically present, as it is in many situations for life assessment and survival analysis. Renyi, Havrad and Chavrat, Tsallis entropies are derived in our study after analysing the maximum likelihood estimator of  $\sigma$  . A comprehensive numerical analysis is carried out, demonstrating that the derived estimates behave well across a range of sample sizes. The mean squared errors, estimated confidence intervals, and associated average lengths are considered as benchmarks. The values of the mean squared errors and average lengths decreases as the sample size rises, according to our numerical findings. Furthermore, as the censoring level is reduced, the Renyi, Havrad and Chavrat, Tsallis entropies estimates approaches the real value. The findings are illustrated using a real-life data set.

The next is how the rest of the article is organised: Section 2 gives the Renyi, Havrad, and Chavrat,

Tsallis entropies for the inverse Rayleigh (IR) distribution. Section 3 focuses at how they can be estimated using multiple censored data. Section 4 contains the simulation and numerical results. Section 5 demonstrates how the method can be used to real-world data sets. Section 6 ends with some summing comments.

## 2. EXPRESSIONS OF THE RENYI, HAVRAD AND CHAVRAT, TSALLIS ENTROPIES

Let  $X$  be an arbitrary variable with parameter  $\sigma$  that follows the IR distribution. The Renyi entropy of  $X$  with  $\varphi = (\sigma)$  by using (1) and (3) is given as

$$R_\delta(X; \sigma) = \frac{1}{1-\delta} \log \left[ \int_0^\infty \frac{2\sigma^2}{x^3} \exp \left[ -\left(\frac{\sigma}{x}\right)^2 \right] dx \right]^\delta \quad (6)$$

Put  $\frac{\sigma}{x} = y \Rightarrow x = \frac{\sigma}{y} \Rightarrow dx = \sigma \left(-\frac{1}{y^2}\right) dy$ .

$$\begin{aligned} R_\delta(X; \sigma) &= \frac{1}{1-\delta} \log \int_0^\infty 2^\delta y^{2\delta} \left(\frac{y}{\sigma}\right)^\delta \exp(-\delta y^2) \sigma \left(\frac{-1}{y^2}\right) dy \\ &= \frac{1}{1-\delta} \log \int_0^\infty y^{3\delta-2} \frac{(-2^\delta)}{\sigma^{\delta-1}} \exp(-\delta y^2) dy \\ &= \frac{1}{1-\delta} \log \left[ \frac{-2^\delta}{\sigma^{\delta-1}} \int_0^\infty y^{3\delta-2} \exp(-\delta y^2) dy \right] \end{aligned}$$

Put  $y^2 = t \Rightarrow y = \sqrt{t} \Rightarrow dy = \frac{1}{2\sqrt{t}} dt$

$$\begin{aligned} R_\delta(X; \sigma) &= \frac{1}{1-\delta} \log \left[ \frac{-2^\delta}{\sigma^{\delta-1}} \int_0^\infty t^{\frac{3\delta-2}{2}} \exp(-\delta t) \frac{1}{2\sqrt{t}} dt \right] \\ &= \frac{1}{1-\delta} \log \left[ \frac{-2^\delta}{\sigma^{\delta-1}} \cdot \frac{1}{2} \int_0^\infty t^{\frac{3\delta}{2}-\frac{1}{2}-1} \exp(-\delta t) dt \right] \\ &= \frac{1}{1-\delta} \log \left[ \frac{-2^{\delta-1}}{\sigma^{\delta-1}} \int_0^\infty \exp(-\delta t) * t^{\frac{3\delta}{2}-\frac{1}{2}-1} dt \right] \\ R_\delta(X; \sigma) &= \frac{1}{1-\delta} \log \left[ -\left(\frac{2}{\sigma}\right)^{\delta-1} \int_0^\infty \exp(-\delta t) * t^{\frac{3\delta}{2}-\frac{1}{2}-1} dt \right] \\ R_\delta(X; \sigma) &= \frac{1}{1-\delta} \log \left[ -\left(\frac{2}{\sigma}\right)^{\delta-1} \frac{\Gamma\left(\frac{3\delta-1}{2}\right)}{\delta^{\frac{3\delta-1}{2}}} \right] \quad (7) \end{aligned}$$

with  $\delta \neq 1, \delta > 0$  and  $3\delta - 1 > 0$ .

Similarly, on using equation (7), Havrad and Chavrat entropy and Tsallis entropy of  $X$  is given by

$$\begin{aligned} HC_\delta(X; \sigma) &= \frac{1}{2^{1-\delta} - 1} \left[ \left( \int_0^\infty \frac{2\sigma^2}{x^3} \exp \left[ -\left(\frac{\sigma}{x}\right)^2 \right] dx \right)^\delta - 1 \right] \\ &= \frac{1}{2^{1-\delta} - 1} \left[ -\left(\frac{2}{\sigma}\right)^{\delta-1} \frac{\Gamma\left(\frac{3\delta-1}{2}\right)}{\delta^{\frac{3\delta-1}{2}}} - 1 \right] \quad (8) \end{aligned}$$

with  $\delta \neq 1, \delta > 0$  and  $3\delta - 1 > 0$ .

$$T_\delta(X; \sigma) = \frac{1}{\delta-1} \left[ 1 - \left( \int_0^\infty \frac{2\sigma^2}{x^3} \exp \left[ -\left(\frac{\sigma}{x}\right)^2 \right] dx \right)^\delta \right]$$

$$= \frac{1}{\delta - 1} \left[ 1 + \left( \frac{2}{\sigma} \right)^{\delta-1} \frac{\Gamma \frac{3\delta-1}{2}}{\delta^{\frac{3\delta-1}{2}}} \right] \quad (9)$$

with  $\delta \neq 1, \delta > 0$  and  $3\delta - 1 > 0$ .

The appropriate expressions of Renyi, Havrad, and Chavrat and Tsallis entropies of  $X$ , simply stated as functions of parameter  $\sigma$ , are represented by Equations (7), (8), and (9) respectively.

### 3. ENTROPY ESTIMATION

Let  $X$  be a random variable with cdf and pdf equal to  $f(x; \varphi)$  and  $F(x; \varphi)$ , respectively. We acquire  $n$  values  $x_1, x_2, \dots, x_n$  based on  $n$  units under a given test, where  $n_f$  and  $n_m$  are the number of failed and censored units, respectively. The Likelihood function for  $\varphi$  is as follows:

$$L(\varphi) = K \prod_{i=1}^n [f(x_i; \varphi)]^{\varepsilon_{i,f}} [1 - F(x_i; \varphi)]^{\varepsilon_{i,m}} \quad (10)$$

where  $K$  is a constant.

$\varepsilon_{i,f}=1$  if the  $i$ th unit failed, and 0 otherwise (so  $\sum_{i=1}^n \varepsilon_{i,f} = n_f$ )

$\varepsilon_{i,m}=1$  if the  $i$ th unit censored, and 0 otherwise (so  $\sum_{i=1}^n \varepsilon_{i,m} = n_m$ ).

By inserting (4) and (5) in (10), we can get the likelihood function of the IR distribution based on multiple censored samples is given by

$$L(\sigma) = K \prod_{i=1}^n \left[ \frac{2\sigma^2}{x^3} \exp \left[ - \left( \frac{\sigma}{x} \right)^2 \right] \right]^{\varepsilon_{i,f}} \left[ 1 - \exp \left[ - \left( \frac{\sigma}{x} \right)^2 \right] \right]^{\varepsilon_{i,m}} \quad (11)$$

The log-likelihood function is given by

$$\log l(\sigma) = \log K + 2 \sum \varepsilon_{i,f} \log(\sigma^2) - \sum_{i=1}^n \varepsilon_{i,f} \log(x_i^3) - \sum_{i=1}^n \varepsilon_{i,f} \left( \frac{\sigma}{x_i} \right)^2 + \sum_{i=1}^n \varepsilon_{i,m} \log \left[ 1 - \exp \left( - \frac{\sigma}{x_i} \right)^2 \right]$$

The MLE is obtained by maximizing  $L(\sigma)$  with respect to  $\sigma$ , and is given by

$$\begin{aligned} \frac{\partial \log l(\sigma)}{\partial \sigma} &= \frac{2n_f}{\sigma^2} \cdot 2\sigma - \sum_{i=1}^n \varepsilon_{i,f} \frac{2\sigma}{x_i^2} + \sum_{i=1}^n \varepsilon_{i,m} \left( \frac{1}{1 - \exp \left( - \frac{\sigma}{x_i} \right)^2} \right) \left[ - \exp \left( - \frac{\sigma}{x_i} \right)^2 \right] \left( \frac{-2\sigma}{x_i^2} \right) \\ &= \frac{4n_f}{\sigma} - \sum_{i=1}^n \varepsilon_{i,f} \left( \frac{2\sigma}{x_i^2} \right) + \frac{\sum_{i=1}^n \varepsilon_{i,m} \exp \left( - \frac{\sigma}{x_i} \right)^2}{\left( 1 - \exp \left( - \frac{\sigma}{x_i} \right)^2 \right)} \cdot \left( \frac{2\sigma}{x_i^2} \right) \end{aligned} \quad (12)$$

The above equation is in closed form therefore, cannot be solved manually. So the MLE estimate of  $\sigma$  is obtained with the help of matlab.

On substituting the MLE of  $\sigma$  in (7), (8) and (9), estimates for the entropies  $R_\delta(X; \sigma)$ ,  $HC_\delta(X; \sigma)$  and  $T_\delta(X; \sigma)$ , are, respectively, given by

$$R_\delta(X; \sigma) = \frac{1}{1 - \delta} \log \left[ - \left( \frac{2}{\hat{\sigma}} \right)^{\delta-1} \frac{\Gamma \frac{3\delta-1}{2}}{\delta^{\frac{3\delta-1}{2}}} \right] \quad (13)$$

with  $\delta \neq 1, \delta > 0$  and  $3\delta - 1 > 0$ .



$$HC_{\delta}(X; \sigma) = \frac{1}{2^{1-\delta} - 1} \left[ - \left( \frac{2}{\hat{\sigma}} \right)^{\delta-1} \frac{\Gamma \frac{3q-1}{2}}{\delta \frac{3q-1}{2}} - 1 \right] \quad (14)$$

with  $\delta \neq 1, \delta > 0$  and  $3\delta - 1 > 0$ .

$$T_{\delta}(X; \sigma) = \frac{1}{\delta - 1} \left[ 1 + \left( \frac{2}{\hat{\sigma}} \right)^{\delta-1} \frac{\Gamma \frac{3q-1}{2}}{\delta \frac{3q-1}{2}} \right] \quad (15)$$

with  $\delta \neq 1, \delta > 0$  and  $3\delta - 1 > 0$ .

Under sufficient regularity requirements, the MLE estimators are consistent and asymptotically normal distributed for large sample sizes. At the confidence level  $100(1-\alpha)$  with  $\alpha = (0, 1)$ , the estimated confidence interval for the Renyi entropy can be calculated as follows:

$$P \left[ -z_{\frac{\alpha}{2}} \leq \frac{\hat{R}_{\delta}(X) - R_{\delta}(X)}{\sigma_{\hat{R}_{\delta}(X)}} \leq z_{\frac{\alpha}{2}} \right] = 1 - \alpha \quad (16)$$

where  $z_{\frac{\alpha}{2}}$  is  $100(1 - \frac{\alpha}{2})$  the standard normal percentile and  $\nu$  is the significant level. As a result, approximate Renyi entropy confidence bounds can be determined, such that

$$P[\hat{R}_{\delta}(X) - z_{\frac{\alpha}{2}} \sigma_{(R_{\delta}(\hat{X}))} \leq R_{\delta}(X) \leq \hat{R}_{\delta}(X) + z_{\frac{\alpha}{2}} \sigma_{(R_{\delta}(\hat{X}))}] \cong 1 - \alpha \quad (17)$$

where  $L_H = \hat{R}_{\delta}(X) - z_{\frac{\alpha}{2}} \sigma_{(R_{\delta}(\hat{X}))}$ ,  $U_H = \hat{R}_{\delta}(X) + z_{\frac{\alpha}{2}} \sigma_{(R_{\delta}(\hat{X}))}$  are the lower and upper confidence limits for  $R_{\delta}(X)$  and  $\sigma$  is the standard deviation and  $\alpha = 0.05$ , the approximate confidence limits for Renyi entropy will be constructed with confidence levels 95%. A similar result holds for  $HC_{\delta}(X)$  and  $T_{\delta}(X)$ .

#### 4. SIMULATION STUDY

The procedure adopted to examine the performance of the Proposed estimators given by (13), (14) and (15) are as:

- 1000 random samples of sizes  $n = 50, 100, 150, 200, 300, 400$  are obtained from the IR distribution based on multiple censored samples, Using the method described in [16].
- The values of parameters are selected as  $\delta = 0.4, 1.2, 1.5$  and  $\sigma = 1.2$ . We chose  $CL = 0.5$  and  $0.7$  at random for failures at the censoring level (CL).
- The estimated value for  $\sigma$ , true values for  $R_{\delta}(X; \sigma)$ ,  $HC_{\delta}(X; \sigma)$  and  $T_{\delta}(X; \sigma)$  are obtained by (12), (7), (8) and (9), and the estimates  $\hat{R}_{\delta}(X; \sigma)$ ,  $\hat{HC}_{\delta}(X; \sigma)$  and  $\hat{T}_{\delta}(X; \sigma)$  given by (13), (14) and (15) are calculated, respectively.
- At last, the average of the derived estimates, MSEs, and ALs are computed with a threshold of 95% All the calculations are done by the use of the software Matlab and R. From the tables, the following conclusions have been made:
  - As the sample size grows, the bias and MSEs of entropy estimates fall.
  - Additionally, as the sample size grows, the ALs of estimates diminish.
  - As the sample size expands, the entropy estimations approach their true values.
  - The MSE of entropy estimates at  $CL = 0.5$  is usually less than the MSE of estimates at  $CL = 0.7$ .

These findings demonstrate the high precision of our entropy estimates.

**Table 1:** Renyi Entropy Estimates at  $CL=0.5(\sigma = 1.2, \delta = 0.4)$

n	Actual Value	Estimate	Bias	MSE	AL
50	0.9930	1.1060	0.1130	$6.33 * e^{-05}$	0.0395
100		0.9089	0.0841	$5.40 * e^{-05}$	0.0190
150		0.9465	0.0495	$4.71 * e^{-05}$	0.0121
200		0.9506	0.0424	$1.79 * e^{-05}$	0.0079
300		0.9560	0.0370	$4.56 * e^{-06}$	0.0064
400		0.9943	0.0013	$9.52 * e^{-09}$	0.0047

**Table 2:** Renyi Entropy Estimates at  $CL=0.7(\sigma = 1.2, \delta = 0.4)$

n	Actual Value	Estimates	Bias	MSE	AL
50	0.9930	0.8798	0.1132	$3.19 * e^{-04}$	0.0382
100		1.0934	0.1004	$1.01 * e^{-04}$	0.0219
150		1.0694	0.0764	$3.19 * e^{-04}$	0.0140
200		1.0533	0.0603	$2.42 * e^{-05}$	0.0099
300		0.9561	0.0369	$1.95 * e^{-05}$	0.0071
400		0.9864	0.0065	$7.26 * e^{-07}$	0.0044

**Table 3:** HC Entropy Estimates at  $CL=0.5(\sigma = 1.2, \delta = 1.5)$

n	Actual Value	Estimate	Bias	MSE	AL
50	6.5486	6.0175	0.5311	0.0056	0.2407
100		6.0733	0.4753	0.0023	0.1215
150		6.2092	0.3394	$7.67 * -04$	0.0828
200		6.3493	0.1993	$1.98 * e^{-04}$	0.0635
300		6.6832	0.1346	$6.03 * e^{-05}$	0.0446
400		6.5667	0.0181	$8.16 * e^{-07}$	0.0328

**Table 4:** HC Entropy Estimates at  $CL=0.7(\sigma = 1.2, \delta = 1.5)$

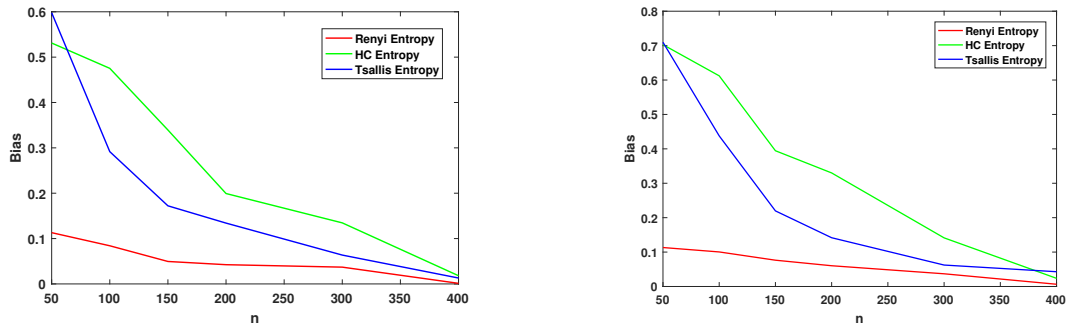
n	Actual Value	Estimate	Bias	MSE	AL
50	6.5486	5.8458	0.7028	0.0099	0.2338
100		5.9364	0.6122	0.0037	0.1187
150		6.1541	0.3945	0.0010	0.0821
200		6.2186	0.3300	$5.44 * e^{-05}$	0.0622
300		6.4073	0.1413	$6.65 * e^{-05}$	0.0427
400		6.5249	0.0237	$1.40 * e^{-07}$	0.0326

**Table 5:** Tsallis Entropy Estimates at  $CL=0.5(\sigma = 1.2, \delta = 1.2)$

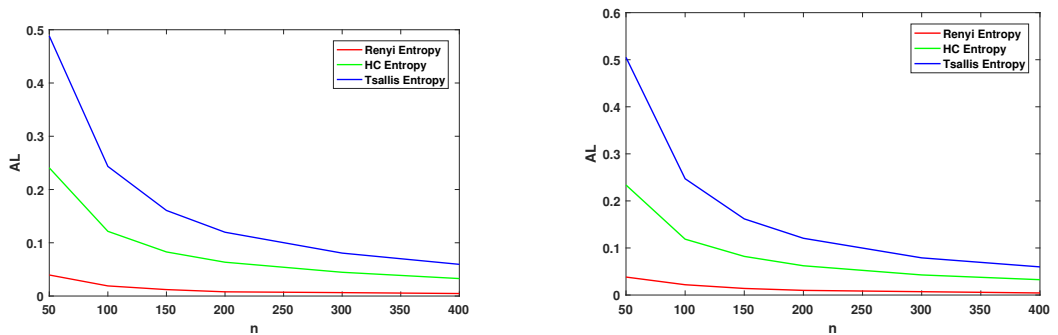
n	Actual Value	Estimate	Bias	MSE	AL
50	11.9156	12.5135	0.5979	0.0036	0.4883
100		12.2072	0.2916	0.0017	0.2434
150		12.0878	0.1722	$9.88 * e^{-05}$	0.1607
200		12.0497	0.1341	$1.19 * e^{-05}$	0.1198
300		11.9792	0.0636	$2.02 * e^{-05}$	0.0806
400		11.9027	0.0130	$4.19 * e^{-07}$	0.0595

**Table 6:** Tsallis Entropy Estimates at  $CL=0.7(\sigma = 1.2, \delta = 1.2)$

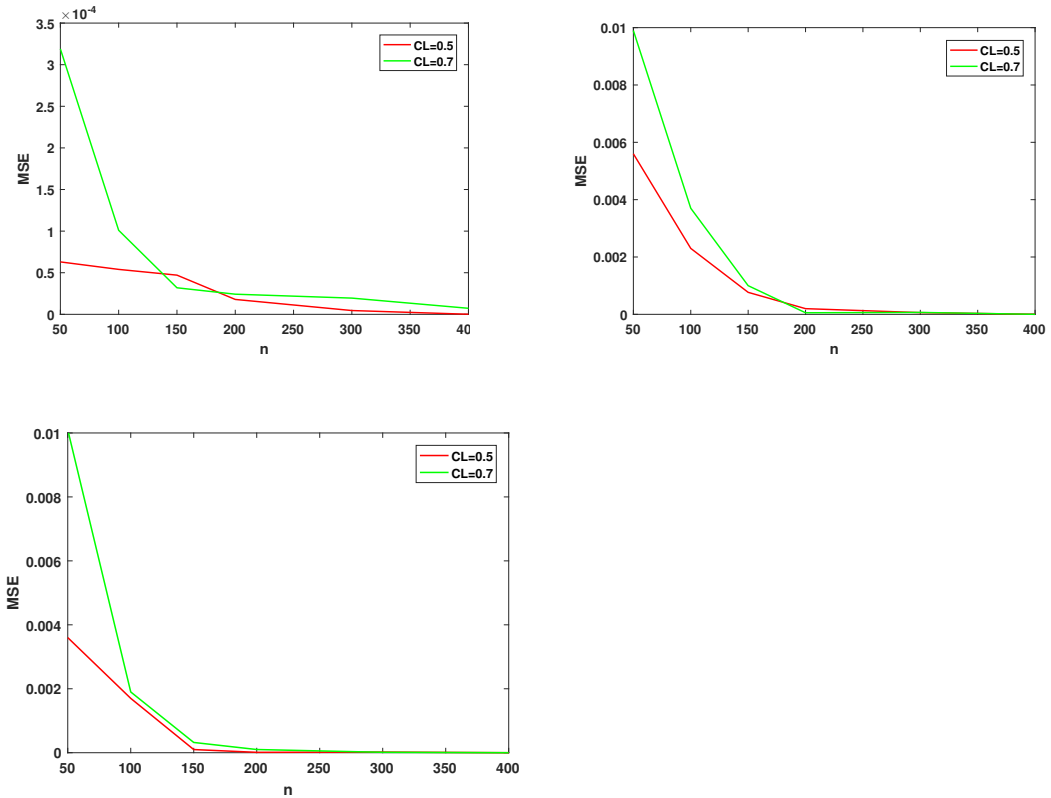
n	Actual Value	Estimate	Bias	MSE	AL
50	11.9156	12.6254	0.7098	0.0101	0.5050
100		12.3534	0.4377	0.0019	0.2471
150		12.1353	0.2197	$3.21 * e^{-04}$	0.1618
200		12.0573	0.1417	$1.00 * e^{-04}$	0.1206
300		11.8533	0.0623	$1.29 * e^{-05}$	0.0790
400		11.9585	0.0429	$4.60 * e^{-07}$	0.0598



**Figure 1:** (a) Bias of Renyi, Havrda and Charvat, Tsallis entropy at  $CL=0.5$  and (b) Bias of Renyi, Havrda and Charvat, Tsallis entropy at  $CL=0.7$



**Figure 2:** (a) Average Length of Renyi, Havrda and Charvat, Tsallis entropy at  $CL=0.5$  and (b) Average Length of Renyi, Havrda and Charvat, Tsallis entropy at  $CL=0.7$



**Figure 3:** (a) MSE of Renyi entropy at CL=0.5 and CL=0.7, (b) MSE of Havrda and Charvat entropy at CL=0.5 and CL=0.7 and (c) MSE of Tsallis entropy at CL=0.5 and CL=0.7

## 5. DATA ANALYSIS

To demonstrate the effectiveness of our estimation methods, we utilize the dataset pertaining to fatigue failure times of twenty-three ball bearings as documented in [8]. This dataset has been extensively employed in various research investigations.

**Dataset I:** 0.1788, 0.2892, 0.3300, 0.4152, 0.4212, 0.4560, 0.4840, 0.5184, 0.5196, 0.5412, 0.5556, 0.6780, 0.6864, 0.6888, 0.8412, 0.9312, 0.9864, 1.0512, 1.0584, 1.2792, 1.2804, 1.7340. The Kolmogorov-Smirnov (K-S) distance and its corresponding p-value for the actual dataset are calculated as 0.1440 and 0.6988 respectively. These values suggest that the observed dataset aligns well with the inverse Rayleigh distribution. This assertion gains further validation through the visualization of the empirical Cumulative Distribution Function (ECDF) plot, the quantile-quantile (Q-Q) plot, and the Histogram, showcased in figures 4 and 5. Derived from the complete sample, the maximum likelihood estimate of the parameter sigma is determined as 0.4681, with a standard error of 0.0499.

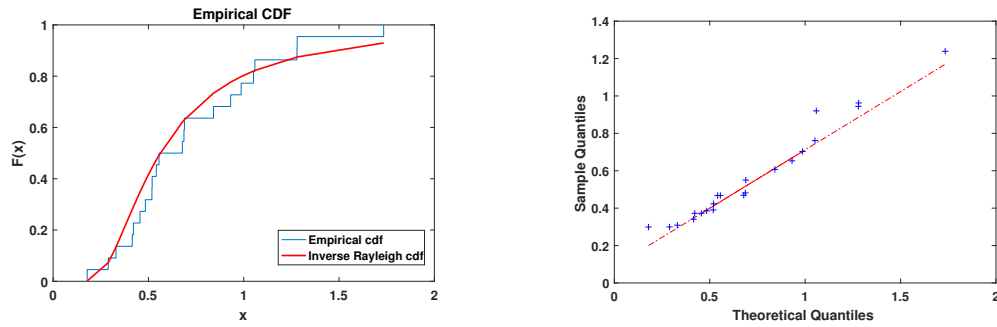


Figure 4: (a) Ecdf plot for the dataset I (b) Q-Q plot for the dataset I

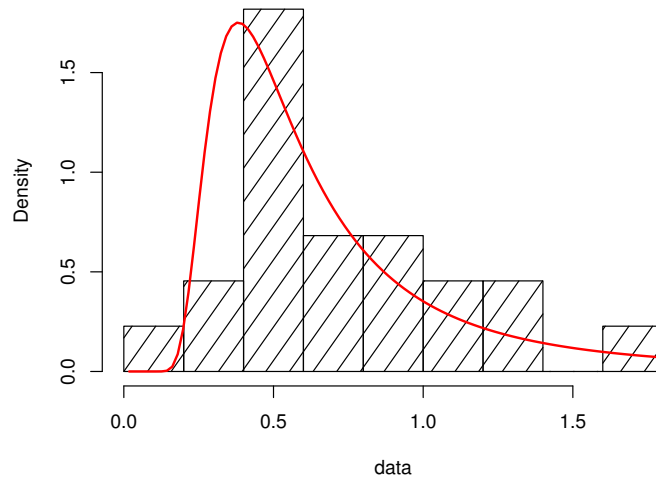


Figure 5: Plot of the fitted density for dataset I

**Dataset II:** The second dataset, sourced from [9], encompasses monthly actual tax revenues in Egypt spanning from January 2006 to November 2010. These data points are measured in 1000 million Egyptian pounds and exhibit the following sequence: 5.9, 20.4, 14.9, 16.2, 17.2, 7.8, 6.1, 9.2, 10.2, 9.6, 13.3, 8.5, 21.6, 18.5, 5.1, 6.7, 17, 8.6, 9.7, 39.2, 35.7, 15.7, 9.7, 10, 4.1, 36, 8.5, 8, 9.2, 26.2, 21.9, 16.7, 21.3, 35.4, 14.3, 8.5, 10.6, 19.1, 20.5, 7.1, 7.7, 18.1, 16.5, 11.9, 7, 8.6, 12.5, 10.3, 11.2, 6.1, 8.4, 11, 11.6, 11.9, 5.2, 6.8, 8.9, 7.1, 10.8. The Kolmogorov-Smirnov (K-S) distance and its corresponding p-value for this dataset stand at 0.08219 and 0.8203, respectively. These results suggest a fitting match with the inverse Rayleigh distribution. This assertion gains further support from the visual analyses, including the Empirical CDF plot, Quantile-Quantile (Q-Q) plot, and Histogram are depicted in figures 6 and 7. The maximum likelihood estimate for the parameter sigma, obtained from the complete dataset is 9.3595, with a standard error of 0.6092. Table 7 and 8 present estimates for different entropy measures in both datasets. These tables reveal as the parameter  $\delta$  increases, Renyi entropy demonstrates an ascending trend, whereas Tsallis and HC entropies exhibit a descending trend with the increase of  $\delta$ . Additionally, the estimates are notably influenced by the level of censoring.

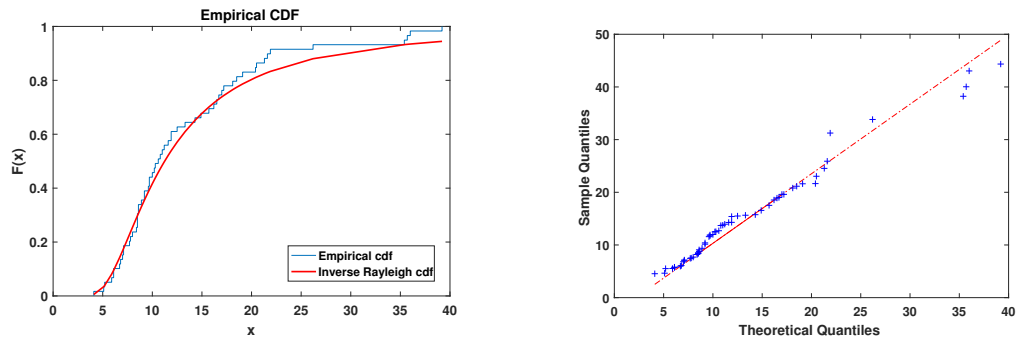


Figure 6: (a) Ecdf plot for the dataset II (b) Q-Q plot for the dataset II

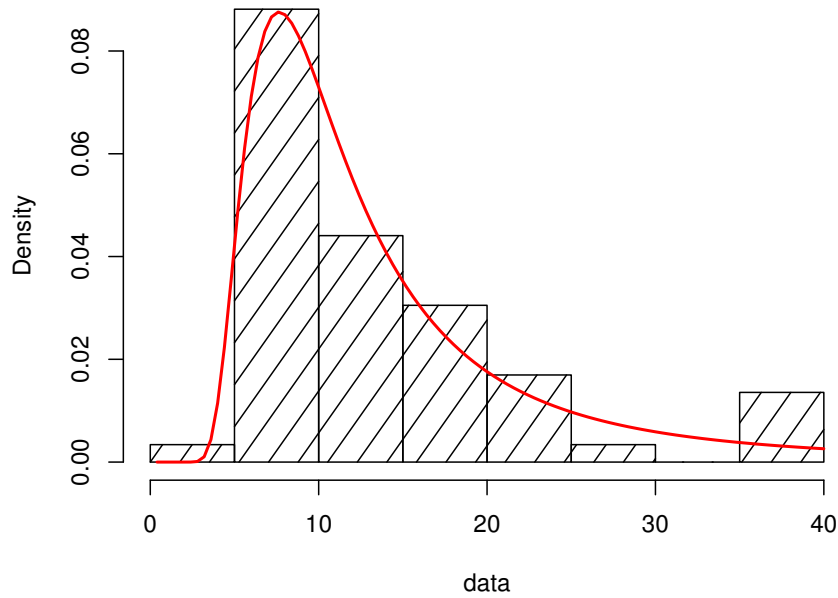


Figure 7: Plot of the fitted density for dataset II

Table 7: Estimated of Renyi entropy, Tsallis entropy and HC entropy at CL=0.5, 0.7 for Dataset I.

$\delta$	CL=0.5			CL=0.7		
	$R_\delta(X)$	$T_\delta(X)$	$HC_\delta(X)$	$R_\delta(X)$	$T_\delta(X)$	$HC_\delta(X)$
1.2	-2.6986	13.5778	14.7472	-2.2308	12.8115	15.4360
2	-1.7333	6.6593	2.4727	-1.2654	4.5447	2.7547

Table 8: Estimated of Renyi entropy, Tsallis entropy and HC entropy at CL=0.5, 0.7 for Dataset II.

$\delta$	CL=0.5			CL=0.7		
	$R_\delta(X)$	$HC_\delta(X)$	$T_\delta(X)$	$R_\delta(X)$	$HC_\delta(X)$	$T_\delta(X)$
1.2	0.3961	9.6191	20.7654	0.4187	9.5983	20.8244
2	1.3615	1.2562	12.4407	1.3841	1.2505	12.6790

## 6. CONCLUSION

In this article, the Renyi, Havrda and Charvat, Tsallis entropies of the inverse Rayleigh distribution are estimated using multiple censored data. Using maximum likelihood and plugging approach, we present an efficient estimation strategy. The Renyi, Havrda and Charvat, and Tsallis entropies estimates' behaviour is measured in terms of mean squared errors and average lengths. According to numerical results, the bias and mean squared errors of our estimators decreases as the sample size grows. It's also worth noting that as the sample size grows, the average length of our estimators shrinks. As a result, the proposed estimates show to be efficient, giving new valuable tools with potential relevance in a wide range of applications involving the inverse Rayleigh distribution's entropy. The paper concludes with an applications to a real-world data sets. In upcoming research endeavors, one could explore the assessment of entropies using both Bayesian and E-Bayesian methodologies across various censoring scenarios.

**Acknowledgments:** This work is supported by the Department of Science and Technology (DST).

## REFERENCES

- [1] Abo-Eleneen, Z. A. (2011). The entropy of progressively censored samples. *Entropy*, 13(2):437–449.
- [2] Al-Babtain, A. A., Hassan, A. S., Zaky, A. N., Elbatal, I. and Elgarhy, M. (2021). Dynamic cumulative residual renyi entropy for lomax distribution: Bayesian and non-bayesian methods. *J. AIMS Mathematics*, 6(4):3889–3914.
- [3] Bantan, R. A. R., Elgarhy, M., Chesneau, C. and Jamal, F.(2020). Estimation of entropy for inverse lomax distribution under multiple censored data. *Entropy*, 22(6).
- [4] Baratpour, S., Ahmadi, J. and Arghami, N.R. (2007). Entropy properties of record statistics. *Statistical Papers*, 48:197–213.
- [5] Cho, Y., Sun, H. and Lee, K. (2014). An estimation of the entropy for a rayleigh distribution based on doubly-generalized type-ii hybrid censored samples. *Entropy*, 16(7):3655–3669.
- [6] Hassan, A. S. and Zaky, A. N. (2019). Estimation of entropy for inverse weibull distribution under multiple censored data. *Journal of Taibah University for Science*, 13(1):331–337.
- [7] Havrda, J. and Charvat, F. (1967). Quantification method in classification processes: concept of structural  $\alpha$ -entropy. *Kybernetika*, 3:30–35.
- [8] Lawless, J. F. (2011). *Statistical models and methods for lifetime data*. John Wiley & Sons, New York, NY, USA.
- [9] Mead, M.E. (2016). On five-parameter lomax distribution: Properties and applications. *Pak. J. Stat. Oper. Res.*, 1:185–199.
- [10] Morabbi, H. and Razmkhah, M. (2010). Entropy of hybrid censoring schemes. *Journal of Statistical Research of Iran*, 6(2).
- [11] Renyi, A. (1961). On the measure of entropy and information. *Proceedings of the fourth Berkely symposium on mathematical statistics and probability*, 1:547–561.
- [12] Rayleigh S J. W. S. (1880). On the resultant of a large number of vibrations of the some pitch and of arbitrary phase. *Philosophical Magazine*, 10:73–78.
- [13] Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423.
- [14] Trayer, VN. (1964). Inverse rayleigh (ir) model. *Proceedings of the Academy of Science, Doklady Akad, Nauk Belarus, USSR*.
- [15] Tsallis, C. (1968). Possible generalization of boltzmann–gibbs statistics. *Journal of Statistical Physics*, 52:479–487.
- [16] Wang, F. K. and Cheng, Y. F. (2010). Em algorithm for estimating the burr xii parameters with multiple censored data. *Quality and Reliability Engineering International*, 26.
- [17] Wong, K.M. and Chan, S. (1990). The entropy of ordered sequences and order statistics. *IEEE Transactions on Information Theory*, 36(2):276–284.

# A LITERATURE SURVEY ON QUEUEING MODEL WITH WORKING VACATION

DIVYA K<sup>1</sup> AND INDHIRA K\*



Department of Mathematics, School of Advanced Sciences,  
Vellore Institute of Technology, Vellore-632014, Tamil Nadu, India.  
divya.k2020@vitstudent.ac.in  
Correspondence email: kindhira@vit.ac.in.

## Abstract

*In 2002, the Working Vacation (WV) queues were implemented as an extension of standard queueing models with vacations. During the vacation period in WV queues, the server provides service at a slower pace as opposed to the typical busy period. The objective of this survey is to provide a concise overview of the latest scholarly investigations on queueing models for WVs. The concept of a queue with WV has been implemented across various domains, encompassing computer systems, communication networks, production management, computer communication, manufacturing, and inventory systems. Additionally, it has been applied to network service, web service, file transfer service, and mail service.*

## Keywords:

Working Vacation Queue(WVQ),  $M/M/1$  and  $M/G/1$  queue,  $GI/M/1$  and  $GI/G/1$  queue, Retrial queue, Discrete time  $Geo/G/1$  queue, Multi-server queue,  $M^{[X]}/M/1$  -Batch arrival queue, MAP queue.

## 1. Introduction

In the realm of service industries like healthcare and manufacturing, as well as computer systems, the queueing model plays a vital role. This mathematical concept, known as queuing theory, finds applications in predicting queue lengths and waiting durations when different types of customers are served by distinct servers following various queue disciplines.

One interesting aspect of queueing systems is the idea of a "working vacation" (WV). Traditionally, when there are no customers or the server experiences a failure, the system goes on vacation, and the server stops serving customers entirely. However, a WV introduces a more efficient approach where the server continues working with different service rates during vacation times, rather than coming to a complete halt. This way, the server can make better use of its idle time. Model for WV is shown in 1.

Our focus in this review paper is on the literature surrounding WV models. The idea of vacation, which involves utilising the idle time of a server for additional work in a secondary system, was first introduced by Levy and Yechiali in 1975 [36]. Subsequently, the concept of a WV was afterwards introduced by Servi and Finn [63]. Over the last three decades, WV queueing models have emerged as a prominent subject of interest within the field of queuing theory.

The objective of this paper is to present a comprehensive overview of the progress achieved



in the examination of arrival and service operations in diverse WV models. We'll explore the application of WVs in the  $M/M/1$  and  $M/G/1$  queueing models in Section 2, while Section 3 will delve into the models for the  $GI/M/1$  and  $GI/G/1$  queues with WVs. Furthermore, Section 4 will cover recent research on retrial queueing models incorporating WVs. Finally, in Section 5, we'll discuss some of the most recent developments in WV models. The paper will conclude in Section 6, summarizing the key findings and offering concluding remarks to aid readers in understanding the field of WVQ.

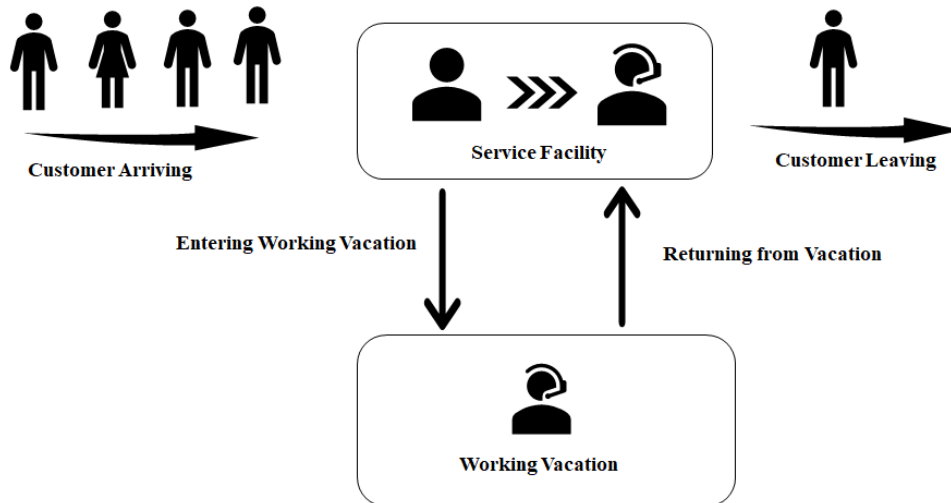


Figure 1: Queueing System with Working Vacation

## 2. An $M/M/1$ and $M/G/1$ Queue Models with Working vacations

The concept of vacations in queueing models was first explored by Levy and Yechiali [36]. They utilized decomposition results to derive the optimal vacation size. Servi and Finn [63] introduced a semi-vacation policy and derived an  $M/M/1$  queue with multiple WVs (MWV). They also provided explicit formulas for average, variance, and distribution of time and number of customers in the system. Wu and Takagi [78] extended Servi and Finn's [63]  $M/M/1$  model to an  $M/G/1/WV$  model, considering general distributions for both service times and WVs. They further obtained the Laplace-Stieltjes Transform (LST) for the distribution of vacation sizes.

Numerous studies followed, exploring different aspects of WV models. Liu et al. [50] analyzed the stochastic decomposition structures of the number of customers and sojourn time in  $M/M/1/WV$  queues. Zhang and Xu [89] investigated an  $M/M/1$  queue with MWV and N-policy. Li et al. [39] studied an  $M/G/1$  queue with exponential WVs using matrix analytic methods. Xu et al. [80] examined  $M/M/1$  queue with SWV, utilizing quasi birth and death (QBD) process and matrix-geometric solution (MGS) method.

The research expanded to consider various scenarios, such as server breakdowns and disasters. Kim et al. [30] explored the  $M/G/1$  queue with disasters and working breakdown services. Additionally, WV models were studied with different impatient behaviors, multiple types of WVs, and variant service interruptions [83, 66, 76].

Vacation interruption (VI) models emerged, where vacation and VI are interconnected, and the server may interrupt vacation based on specific system indices. Jihong Li and Naishuo Tian [41] introduced VI, analyzing the  $M/M/1$  queue using QBD process and MGS method. Zhang and Hou [84] extended this to an  $M/G/1$  queue with WV and VI, obtaining queue length

distribution and service status.

The integration of WVs and service interruption due to server breakdowns added strength to queueing models. Various analytical methods, such as generating functions, were employed [24, 14, 35, 88]. Imbalanced behavior of servers was also considered [51, 40, 21, 17].

Overall, extensive research has been conducted to understand the dynamics of queueing models with WVs and vacation interruptions, offering valuable insights into optimizing system performance and resource utilization.

### 3. An $GI/M/1$ and $GI/G/1$ Queue Models with Working vacations

In the context of general input (GI) queue models with WVs (WV), several studies have been conducted. Baba [5] explored a  $GI/M/1$  queue with WV, extending Servi and Finn's  $M/M/1/WV$  system to a  $GI/M/1/WV$  model. Building on this, Banik et al. [7] analyzed the  $GI/M/1/N$  queue with a MWV policy. Li and Tian [42] delved into the details of a  $GI/M/1$  queue with SWV, where the server can continue working at a reduced rate during the vacation period.

Zhang and Hou [86] studied the  $GI/M/1/N$  queue with a variant of MWV and obtained the queue length distribution at different time periods using the supplementary variable technique (SVT) and embedded Markov chain (EMC) method. Goswami et al. [19] developed the  $GI/M(n)/1$  queue model with finite buffer, considering state-dependent services and state-dependent MWV. Vijayalaxmi et al. [34] focused on a limited buffer come-back arrival single server queueing system with multiple state-dependent exponential WV.

Ye and Liu [82] presented the  $GI/M/1$  queue with SWV and derived the stationary distribution of the system size at arrival time using the matrix-geometric solution (MGS) method. They also found the stationary distribution of the system size at arbitrary time using the semi-Markov process (SMP) method. Panda et al. [56] explored an infinite buffer come-back arrival queue with MWV policy, considering general bulk service (a,b)-rule.

In the context of general input and vacation interruption models, where the server goes on vacation when there are no customers, several studies have been conducted. Li and Tian [38] presented WV and VI in a discrete-time  $GI/Geo/1$  queue using the MGS approach. Ji-hong et al. [25] studied a  $GI/M/1$  queue with WVs and vacation interruptions. Zhao et al. [90] introduced setup time with VI policy and investigated a single server general input queue with set-up period, WV, and VI, obtaining the distribution of the number of customers in the system and waiting time.

Chen et al. [11] analyzed PH (Phase-type) WVs and vacation interruptions in  $GI/M/1$  queues. They obtained steady-state distributions for the queue length and waiting time of customers and revealed stochastic decomposition structures of the queue length and waiting time using the method of matrix analytic method (MAM).

Li et al. [68] considered Bernoulli schedule rule and studied the start-up period, SWV, and vacation interruption in the  $GI/M/1$  queue. Goswami and Mund [18] dealt with impatient customers in a single server renewal arrival batch service queue with MWV and balking. They determined the probability distribution of queue length at pre-arrival epoch using the EMC method.

### 4. Retrial Queue Models with Working Vacations

Retrial queues are mathematical models used in queueing theory to describe systems with finite capacity where arriving jobs that find the system busy will wait for a while before attempting to enter again. Which is shown in Fig. 2. Such systems can be found in various real-world scenarios like restaurant reservations, telecommunication networks, and packet switching networks. Recently, the combination of retrial queues with WVs (WV) has become a subject of thorough investigation.

Studies have been conducted on different types of retrial queues incorporating WVs. For instance, T. Van Do [74] analyzed the stability of the  $M/M/1$  retrial queue with WV. Tao et al. [69] used the matrix analytic method to propose conditions for stability in the  $M/M/1$  retrial queue with WV interruption under N-policy. Several researchers, such as Li et al. [45], Gao et al. [16], and Aissani et al. [2], explored various aspects of single server retrial queues with WVs and vacation interruptions. Further research delved into specific aspects of retrial queues with

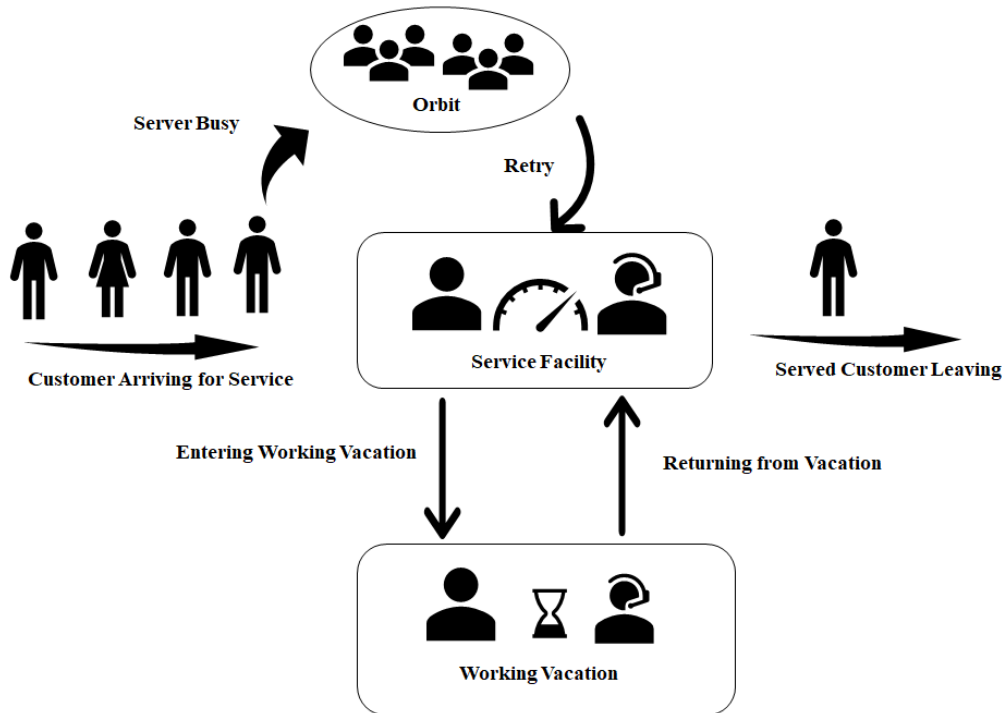


Figure 2: Retrial Queueing Model with Working Vacation

WV. For example, Upadhyaya [73] examined a discrete-time  $Geo^{[X]}/Geo/1$  retrial queue with WV and derived various performance measures using the matrix-geometric method. Rajadurai et al. [61, 60] addressed RQ systems with general retrial times, feedback, balking, multiple WVs, and vacation interruptions using the supplementary variable technique.

Other studies considered specific features of retrial queues, such as starting failure, preemptive priority, balking customers, and Bernoulli feedback, in the presence of WVs and vacation interruptions [20, 59, 43, 46]. The effects of bulk arrivals, constant retrial rates, and socially optimal balking strategies were also investigated [53, 54, 12].

In conclusion, the combination of retrial queues with WVs has attracted significant attention in recent research, leading to a better understanding of system behaviors and performance measures in various queueing scenarios.

## 5. Other Working Vacation Models

### 5.1. Discrete time Queue Models with Working Vacations

Discrete-time (DT) queues with vacations have been extensively investigated by various researchers, owing to their wide range of applications in digital communication systems and telecommunication networks, such as B-ISDN, ATM, and related technologies.

Li [37] studied a discrete-time  $Geo/G/1$  queueing system with multiple WVs, where the server operates at a reduced rate during vacation periods. Li and Tian [38] introduced a discrete-time queue model, where customer arrivals and service completions occur at discrete-time instants, in the  $GI/Geo/1$  framework. Li and Zhang [44] examined a discrete-time  $Geo/Geo/1$  queue with server breakdowns and repairs. Yang et al. [81] investigated the equilibrium joining/balking behavior in the discrete-time  $Geo/Geo/1$  queueing model with multiple WVs.

## 5.2. Multi Server Queue Models with Working Vacations

Krishnamoorthy and Shreenivasan [31] investigated a two-server  $M/M/2$  queueing system, where one server remains idle while the other goes on vacation if there are no customers waiting for service. Vijayashree and Janani [77] conducted a transient analysis of an  $M/M/c$  queue subjected to multiple exponential WVs.

Bouchentouf et al. [9] studied a heterogeneous two-server queueing system with Bernoulli feedback and multiple WVs, considering impatient customers. They obtained performance measures and the steady-state probability of the queueing model. Sharma and Kumar [64] analyzed a multi-server queueing system with essential two-phase repair and multiple WVs. They employed the Runge-Kutta method to find the time-dependent probability.

## 5.3. Batch Arrival Queue Models with Working Vacations

Xu et al. [79] examined a batch arrival  $M^{[X]}/M/1$  queue with single working vacation (SWV), using the matrix analytic method (MAM) to derive the probability generating function (PGF) of the stationary system length. Baba [6] investigated a batch arrival  $M^{[X]}/M/1$  queue with multiple working vacations (MWV) and obtained the exact Laplace-Stieltjes Transform (LST) of the stationary waiting time distribution.

Gao and Yao [15] demonstrated a batch arrival  $M^{[X]}/G/1$  queue with randomized WVs, allowing for at most  $J$  vacations. Laxmi and Rajesh [32] extended Baba's work [6] by incorporating the concept of variant WVs. They analyzed a single-server batch arrival infinite-buffer queueing system with various types of WVs. Laxmi and Rajesh [33] further expanded on their previous research and explored the effects of different WVs on a batch arrival queue with reneging and server breakdowns.

Thangaraj and Rajendran [70] discussed a batch arrival queueing system with two types of service and vacations. Niranjan et al. [55] analyzed a bulk arrival queueing model with batch size-dependent service and WVs.

## 5.4. Markovian Arrival Process Queue Models With Working Vacations

The Markov Arrival Process (MAP) system represents another significant advancement in the research of WV models. Zhang and Hou [85] conducted a study on a  $MAP/G/1$  queue with  $N$ -policy WVs and vacation interruptions. They successfully determined the distribution of the system size at the pre-arrival epoch and the Laplace-Stieltjes Transform (LST) of waiting time using the supplementary variable technique (SVT) and matrix analytic method (MAM).

Sreenivasan et al. [65] expanded on the work of Li and Tian [41] by incorporating MAP arrivals, Phase-type (PH) services, and  $N$ -policy vacation queue models. Liu et al. [49] examined a cold standby repairable system with WVs and interruptions, utilizing the MAP arrival queueing model. Chakravarthy and Kulshrestha [10] investigated the  $MAP/PH/1$  type queueing model with WVs, server breakdowns, and repairs.

## 6. Conclusion

In conclusion, this survey provides an in-depth exploration of the development of working vacation (WV) queueing models from their early stages to the present. The pioneering researchers who have contributed to the field of WV queueing policies are presented. Readers gain a comprehensive understanding of the current state of WV queueing models through this survey. A wide array of research papers have been reviewed, and proper citations have been included.

This survey offers readers a holistic view of the diverse applications of WV queueing models in various scenarios. It highlights the significance of WV models in predicting queue lengths, waiting durations, and other essential performance measures in queueing systems.

## REFERENCES

- [1] P.K. Agrawal, A. Jain and M. Jain, *M/M/1 Queueing Model with Working Vacation and Two Type of Server Breakdown*, Journal of Physics: Conference Series Vol.1849(1), pp. 012021(2021).
- [2] A. Aissani, S. Taleb, T. Kernane, G. Saidi and D. Hamadouche, *An M/G/1 retrial queue with working vacation*, Advances in Systems Science,,pp.443-452 (2014).
- [3] S.I. Ammar, *Transient solution of an M/M/1 vacation queue with a waiting server and impatient customers*, Journal of the Egyptian Mathematical Society, Vol 25(3),pp.337-342(2017).
- [4] D. Arivudainambi, P. Godhandaraman, and P. Rajadurai. *Performance analysis of a single server retrial queue with working vacation* Opsearch, Vol 51(3),pp. 434-462 (2014 ).
- [5] Y. Baba, *Analysis of a GI/M/1 queue with multiple working vacations*, Operations Research Letters, Vol 33(2), pp.201-209 (2005).
- [6] Y. Baba, *The  $M^{[X]}/M/1$  queue with multiple working vacation*,(2012).
- [7] A.D. Banik, U.C. Gupta, and S.S Pathak, *On the GI/M/1/N queue with multiple working vacations—analytic analysis and computation*, Applied Mathematical Modelling, Vol 31(9), pp.1701-1710 (2007).
- [8] A.A. Bouchentouf, M. Cherfaoui, and M. Boualem, *Performance and economic analysis of a single server feedback queueing model with vacation and impatient customers*, Opsearch, Vol 56(1),pp.300-323 (2019).
- [9] A.A. Bouchentouf, A. Guendouzib, and A. Kandoucib, *Performance and economic study of heterogeneous M/M/2/N feedback queue with working vacation and impatient customers*, ProbStat Forum, Vol 12,pp. 15–35 (2019).
- [10] S.R. Chakravarthy, and R. Kulshrestha, *A queueing model with server breakdowns, repairs, vacations, and backup server*, Operations Research Perspectives, Vol 7, pp.100131 (2020).
- [11] H.Y. Chen, J.H. Li, and N.S. Tian, *The GI/M/1 queue with phase-type working vacations and vacation interruption*, Journal of Applied Mathematics and Computing, Vol 30(1),pp. 121-141(2009).
- [12] N.H. Do, T. Van Do, and A. Melikov, *Equilibrium customer behavior in the M/M/1 retrial queue with working vacations and a constant retrial rate*. Operational Research, Vol 20(2), pp.627-646 (2020).
- [13] B.T. Doshi, *Queueing systems with vacations—a survey*, Queueing systems, Vol 1(1), pp.29-66 (1986).
- [14] S. Gao, and Z. Liu, *An M/G/1 queue with single working vacation and vacation interruption under Bernoulli schedule*, Applied Mathematical Modelling, Vol 37(3), pp.1564-1579 (2013).
- [15] S. Gao, and Y. Yao, *An  $M^{[X]}/G/1$  queue with randomized working vacations and at most J vacations*, International Journal of Computer Mathematics, Vol 91(3),pp. 368-383 (2014).
- [16] S. Gao, J. Wang, and W.W. Li. *An M/G/1 retrial queue with general retrial times, working vacations and vacation interruption*, Asia-Pacific Journal of Operational Research, Vol 31(02),pp. 1440006 (2014).

- [17] C. Goswami, and N. Selvaraju. *A working vacation queue with priority customers and vacation interruptions* International Journal of Operational Research, Vol 17(3), pp. 311-332 (2013).
- [18] V. Goswami, and G.B. Mund. *Analysis of renewal input batch service queue with impatient customers and multiple working vacations*, International Journal of Management Science and Engineering Management, Vol 15(2), pp.96-105 (2020).
- [19] V. Goswami, P.V. Laxmi, and K. Jyothsna, *Analysis of GI/M(n)/1 queue with state-dependent multiple working vacations*, Opsearch, Vol 50(1), pp. 106-124 (2013).
- [20] M. Gowsalya, and D. Arivudainambi, *Stochastic analysis of an M/G/1 retrial queue subject to working vacation and starting failure* AIP Conference Proceedings, Vol 2095(1), pp. 030009 (2019).
- [21] E. Hertini, C. Harisbaya, and J. Nahar. *Queueing model using sojourn time distribution with single working vacation and vacation interruption*. IOP Conference Series: Materials Science and Engineering, Vol. 567(1), pp. 012036 (2019).
- [22] M. Jain, and A. Jain. *Working vacations queueing model with multiple types of server breakdowns*, Applied Mathematical Modelling, Vol 34(1), pp. 1-13 (2010).
- [23] M. Jain, S. Dhibar, and S.S. Sanga, *Markovian working vacation queue with imperfect service, balking and retrial*, Journal of Ambient Intelligence and Humanized Computing, pp. 1-17 (2021).
- [24] M. Jain, G.C. Sharma, and R. Sharma. *Working vacation queue with service interruption and multi optional repair*, International Journal of Information and Management Sciences, Vol 22(2), pp 157-175 (2011).
- [25] L. Jihong, T. Naishuo, and M. Zhanyou. *Performance Analysis of GI/M/1 Queue with Working Vacations and Vacation Interruption*, Applied Mathematical Modeling, Vol 32(12), pp. 2715-2730 (2008).
- [26] P.K. Joshi, S. Gupta, and K.N. Rajeshwari. *An M/G/1 Model with Multiple Vacation Queueing System*, South East Asian J. of Mathematics and Mathematical Sciences, Vol 16(1), pp.37-50 (2020).
- [27] E. Kasim, and G. Gupur. *Functional analysis method for the M/G/1 queueing model with single working vacation*, Open Mathematics, Vol 16(1), pp. 767-791 (2018).
- [28] E. Kasim, and G. Gupur. *Point spectra of the operator corresponding to the M/M/1 queueing model with working vacation and vacation interruption* Journal of Mathematical Research with Applications, Vol 39(1), pp. 75-88 (2019).
- [29] J.C. Ke, C.H. Wu, and Z.G. Zhang. *Recent developments in vacation queueing models: a short survey*, International Journal of Operations Research, Vol 7(4), pp. 3-8 (2010).
- [30] B.K. Kim, and D.H. Lee. *The M/G/1 queue with disasters and working breakdowns*, Applied Mathematical Modelling, 38(5-6), 1788-1798 (2014).
- [31] A. Krishnamoorthy, and C. Sreenivasan. *An M/M/2 queueing system with heterogeneous servers including one with working vacation*, International Journal of Stochastic Analysis, (2012).
- [32] P.V. Laxmi, and P. Rajesh. *Analysis of variant working vacations on batch arrival queues*, Opsearch, Vol 53(2), pp. 303-316 (2016).
- [33] P.V. Laxmi, and P. Rajesh. *Variant working vacations on batch arrival queue with renegeing and server breakdowns*, East Afr. School. Multidiscip. bull, Vol 1, pp. 2617-4413 (2018).
- [34] P.V. Laxmi, V. Goswami, and V. Suchitra. *Analysis of GI/M(n)/1/N queue with single working vacation and vacation interruption*, International Journal of Computational and Mathematical Sciences, Vol 7, pp. 58-64 (2013).
- [35] D.H. Lee, and B.K. Kim. *A note on the sojourn time distribution of an M/G/1 queue with a single working vacation and vacation interruption*, Operations Research Perspectives, Vol 2, pp. 57-61 (2015).
- [36] Y. Levy, and U. Yechiali. *Utilization of idle time in an M/G/1 queueing system*, Management Science, Vol 22(2), pp. 202-211 (1975).

- [37] J.H. Li. *Analysis of the discrete-time Geo/G/1 working vacation queue and its application to network scheduling*, Computers and Industrial Engineering, Vol 65(4), pp. 594-604 (2013).
- [38] J.H. Li, and N.S. Tian. *The discrete-time GI/Geo/1 queue with working vacations and vacation interruption*, Applied Mathematics and Computation, Vol 185(1), pp. 1-10 (2007).
- [39] J.H. Li, N.S. Tian, Z.G. Zhang, and H.P. Luh. *Analysis of the M/G/1 queue with exponentially working vacations—a matrix analytic approach*, Queueing systems, Vol 61(2), pp.139-166 (2009).
- [40] J. Li, and T. Li. *An M/G/1 G-queue with Server Breakdown, Working Vacations and Bernoulli Vacation Interruption*, IAENG International Journal of Applied Mathematics, Vol 50(2), 2020.
- [41] J. Li, and N. Tian. *The M/M/1 queue with working vacations and vacation interruptions*, Journal of Systems Science and Systems Engineering, Vol 16(1), pp. 121-127 (2007).
- [42] J. Li, and N. Tian. *Performance analysis of a GI/M/1 queue with single working vacation*, Applied Mathematics and Computation, Vol 217(10), pp. 4960-4971 (2011).
- [43] J. Li, T. Li, and J. Xu. *An M/M/1 Retrial Queue with Working Vacation and Feedback*, 2019.
- [44] T. Li, and L. Zhang. *Discrete-time Geo/Geo/1 Queue with Negative Customers and Working Breakdowns*, International Journal of Applied Mathematics, Vol 47(4), 2017.
- [45] T. Li, Z. Wang, and Z. Liu. *Geo/Geo/1 retrial queue with working vacations and vacation interruption* Journal of Applied Mathematics and Computing, Vol 39(1), pp. 131-143 (2012).
- [46] T. Li, L. Zhang, and S. Gao. *An M/G/1 retrial queue with balking customers and Bernoulli working vacation interruption*, Quality Technology and Quantitative Management, Vol 16(5), pp. 511-530 (2019).
- [47] T. Li, L. Zhang, and S. Gao. *An M/G/1 retrial queue with single working vacation under Bernoulli schedule*, RAIRO-Operations Research, Vol 54(2), pp. 471-488 (2020).
- [48] C.H. Lin, and J.C. Ke. *Multi-server system with single working vacation*. Applied Mathematical Modelling, Vol 33(7), pp.2967-2977 (2009)
- [49] B. Liu, L. Cui, Y. Wen, and J. Shen. *A cold standby repairable system with working vacations and vacation interruption following Markovian arrival process* Reliability Engineering and System Safety, Vol 142, pp. 1-8 (2015).
- [50] W.Y. Liu, X.L. Xu, and N.S. Tian. *Stochastic decompositions in the M/M/1 queue with working vacations*, Operations Research Letters, Vol 35(5), PP.595-600 (2007).
- [51] S. Majid, and P. Manoharan. *Analysis of the M/M/1 queue with single working vacation and vacation interruption*, International Journal of Mathematics Trends and Technology, Vol 47(1), pp.32-40 (2017)
- [52] P. Manoharan, and T. Jeeva. *Impatient Customers in an M/M/1 Working Vacation Queue with a Waiting Server and Setup Time*, Journal of Computer and Mathematical Sciences, Vol 10(5), pp. 1189-1196 (2019).
- [53] S.P.B. Murugan, and R. Vijaykrishnaraj. *A bulk arrival retrial queue with feedback and exponentially distributed multiple working vacation*, J Comput Math Sci, Vol 10, pp.81-91 (2019).
- [54] S.P.B. Murugan, and R. Vijaykrishnaraj. *A bulk arrival retrial queue with non-Persistent customers and exponentially distributed multiple working vacation*, AIP Conference Proceedings, Vol.2177(1) , pp. 020064 (2019), AIP Publishing LLC.
- [55] S.P. Niranjana, K. Indhira, and V.M. Chandrasekaran. *Analysis of bulk arrival queueing system with batch size dependent service and working vacation* AIP Conference Proceedings, Vol.1952(1), pp. 020061 (2018). AIP Publishing LLC.
- [56] G. Panda, A.D. Banik, and D. Guha. *Stationary analysis and optimal control under multiple working vacation policy in a GI/M(a,b)/1 queue*, Journal of Systems Science and Complexity, Vol 31(4), pp.1003-1023 (2018).
- [57] J. Patterson, and A. Korzeniowski. *M/M/1 Model With Unreliable Service and a Working Vacation*, International Journal of Statistics and Probability, Vol 8(2), (2019).

- [58] P. Rajadurai . *A study on M/G/1 retrial queueing system with three different types of customers under working vacation policy*, International Journal of Mathematical Modelling and Numerical Optimisation, Vol 8(4), pp. 393-417 (2018).
- [59] P. Rajadurai. *A study on M/G/1 preemptive priority retrial queue with Bernoulli working vacations and vacation interruption*, International Journal of Process Management and Benchmarking, Vol 9(2), pp. 193-215 (2019).
- [60] P. Rajadurai, M.C. Saravanarajan, and V.M. Chandrasekaran. *A study on M/G/1 feedback retrial queue with subject to server breakdown and repair under multiple working vacation policy*, Alexandria Engineering Journal, Vol 57(2), pp.947-962 (2018).
- [61] P. Rajadurai, M. Sundararaman, S.I. Ammar, and D. Narasimhan. *Analysis of M/G/1 priority retrial G-queue with bernoulli working vacations*, Advances in Algebra and Analysis, pp. 383-391, Birkh?user, Cham (2018).
- [62] N. Selvaraju, and C. Goswami. *Impatient customers in an M/M/1 queue with single and multiple working vacations*. Computers and Industrial Engineering, Vol 65(2), pp.207-215 (2013).
- [63] L.D. Servi, and S.G. Finn. *M/M/1 queues with working vacations (m/m/1/wv)*, Performance Evaluation, Vol 50(1), pp.41-52 (2002).
- [64] R. Sharma, and G. Kumar. *Multi-Server M/M/c Queue and Multiple Working Vacation under Phase Repair*, In 2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE), pp. 181-185 (2020). IEEE.
- [65] C. Sreenivasan, S.R. Chakravarthy, and A. Krishnamoorthy. *MAP/PH/1 queue with working vacations, vacation interruptions and N policy*, Applied Mathematical Modelling, Vol 37(6), pp.3879-3893 (2013).
- [66] R. Sudhesh, and A. Azhagappan. *RETRACTED ARTICLE: Transient analysis of an M/M/1 queue with variant impatient behavior and working vacations*, Opsearch, Vol 55(3), pp. 787-806 (2018).
- [67] M. Sundararaman, P. Rajadurai, and D. Narasimhan. *An M/G/1 retrial queueing system with atmost J number of working vacataions*, International Journal of Pure and Applied Mathematics, Vol 119(6), pp. 151-159 (2018).
- [68] L. Tao, Z. Liu, and Z. Wang. *The GI/M/1 queue with start-up period and single working vacation and Bernoulli vacation interruption*, Applied mathematics and computation, Vol 218(8), pp. 4401-4413 (2011).
- [69] L. Tao, Z. Liu, and Z. Wang. *M/M/1 retrial queue with collisions and working vacation interruption under N-policy*, RAIRO-Operations Research, Vol 46(4), pp. 355-371 (2012).
- [70] M. Thangaraj, and P. Rajendran. *Analysis of batch arrival queueing system with two types of service and two types of vacation*, International Journal of Pure and Applied Mathematics, Vol 117(11), pp.263-272 (2017).
- [71] N.S. Tian, J.H. Li, and Z.G. Zhang. *Matrix analytic method and working vacation queues—a survey*, International Journal of Information and Management Sciences, Vol 20(4), pp. 603-633 (2009).
- [72] N. Tian, X. Zhao, and K. Wang. *The M/M/1 queue with single working vacation*, International Journal of Information and Management Sciences, Vol 19(4), pp. 621-634 (2008).
- [73] S. Upadhyaya. *Working vacation policy for a discrete-time Geo<sup>[X]</sup>/Geo/1 retrial queue*, Opsearch, Vol 52(4), pp. 650-669 (2015).
- [74] T. Van Do. *M/M/1 retrial queue with working vacations*, Acta Informatica, Vol 47(1), pp.67-75 (2015).
- [75] M. Varalakshmi, V.M. Chandrasekaran, and M.C. Saravanarajan. *A study on M/G/1 retrial G-queue with two phases of service, immediate feedback and working vacations*, In IOP conference series: materials science and engineering, Vol.263(4), pp. 042156, IOP Publishing (2017).
- [76] K.V. Vijayashree, and A. Anjuka. *Stationary analysis of a fluid queue driven by an M/M/1 queue with working vacation* Quality Technology and Quantitative Management, Vol 15(2), pp. 187-208 (2018).



- [77] K.V. Vijayashree, and B. Janani. *Transient Analysis of an M/M/c Queue Subject to Multiple Exponential Working Vacation*, Applied Mathematical Sciences, Vol 9(74), pp.3669-3677 (2015).
- [78] D.A. Wu, and H. Takagi. *M/G/1 queue with multiple working vacations*, Performance Evaluation, Vol 63(7), pp. 654-681 (2006).
- [79] X.L. Xu, Z.J. Zhang, and N.S. Tian. *Analysis for the  $M^{[X]}/M/1$  working vacation queue*, International journal of information and management sciences, Vol 20(3), pp.379-394 (2009).
- [80] X. Xu, Z. Zhang, and N. Tian. *The M/M/1 queue with single working vacation and set-up times*, International Journal of Operational Research, Vol 6(3), pp.420-434 (2009).
- [81] B. Yang , Z. Hou , and J. Wu. *Analysis of the equilibrium strategies in the Geo/Geo/1 queue with multiple working vacations*, Quality Technology and Quantitative Management, Vol 15(6), pp. 663-685 (2018).
- [82] Q. Ye, and L. Liu. *Performance analysis of the GI/M/1 queue with single working vacation and vacations*, Methodology and computing in Applied Probability, Vol 19(3), pp.685-714 (2017).
- [83] H. Zhang, and G. Zhou. *M/M/1 queue with m kinds of differentiated working vacations*, Journal of Applied Mathematics and Computing, Vol 54(1-2), pp.213 (2017).
- [84] M. Zhang, and Z. Hou. *Performance analysis of M/G/1 queue with working vacations and vacation interruption*, Journal of Computational and Applied Mathematics, Vol 234(10), pp.2977-2985 (2010).
- [85] M. Zhang, and Z. Hou. *Performance analysis of MAP/G/1 queue with working vacations and vacation interruption*, Applied Mathematical Modelling, Vol 35(4), pp.1551-1560 (2011).
- [86] M. Zhang, and Z. Hou.. *Steady state analysis of the GI/M/1/N queue with a variant of multiple working vacations*, Computers and Industrial Engineering, Vol 61(4), pp. 1296-1301 (2011).
- [87] M. Zhang, and Z. Hou.. *M/G/1 queue with single working vacation*, Journal of Applied Mathematics and Computing, Vol 39(1), pp.221-234 (2012).
- [88] M. Zhang, and Q. Liu. *An M/G/1 G-queue with server breakdown, working vacations and vacation interruption*, Opsearch, Vol 52(2), pp. 256-270 (2015).
- [89] Z.J. Zhang, and X.L. Xu. *Analysis for the M/M/1 queue with multiple working vacations and N-policy*, International Journal of Information and Management Sciences, Vol 19(3), pp.495-506 (2008).
- [90] G.H Zhao, X.X Du, and N.S. Tian. *GI/M/1 queue with set-up period and working vacation and vacation interruption*, Int. J. Inform. Manage. Sci, Vol 20, pp. 351-363 (2009).

# BAYESIAN ESTIMATION OF TOPP-LEONE LINDLEY (TLL) DISTRIBUTION PARAMETERS UNDER DIFFERENT LOSS FUNCTIONS USING LINDLEY APPROXIMATION

\*<sup>1</sup>Nzei C. Lawrence; <sup>2</sup>Adegoke M. Taiwo; <sup>3</sup>Ekhosuehi N.; <sup>4</sup>Mbegbu I. Julian

•

\*<sup>1, 3, 4</sup>Department of Statistics, University of Benin, Benin City, Nigeria

<sup>2</sup>Department of Mathematics and Statistics, First Technical University, Ibadan., Nigeria

\*<sup>1</sup>lawrencedumebi@gmail.com, <sup>2</sup>adegoketaiwo@gmail.com,

<sup>3</sup>nosakhare.ekhosuehi@uniben.edu, <sup>4</sup>julian.mbegbu@uniben.edu

## Abstract

*In this study, we present the Bayesian estimates of the unknown parameters of the Topp-Leone Lindley distribution using the maximum likelihood and Bayesian methods. In this study, the Bayes theorem was adopted for obtaining the posterior distribution of the shape parameter and scale parameter of the Topp-Leone Lindley distribution assuming the Jeffreys' (non-informative) prior for the shape parameter and the Gamma (conjugate) prior for the scale parameter under three different loss functions namely: Square Error Loss Function, Linear Exponential Loss Function and Generalized Entropy Loss Function. The posterior distribution derived for both parameters are not solvable analytically, it requires a numerical approximation techniques to obtain the solution. The Lindley approximation techniques was adopted to obtain the parameters of interest. The loss function were used to derive the estimates of both parameters with an assumption that the both parameters are unknown and independent. To ascertain the accuracy of these estimators, the proposed Bayesian estimators under different loss functions are compared with the corresponding maximum likelihood estimator using a Monte Carlo simulation on the performance of these estimators according to the mean square error and BIAS based on simulated samples simulated from the Topp-Leone Lindley distribution. . It was also observed for any fixed value of the parameters, as sample size increases, the mean square errors of the Bayesian Estimates and maximum likelihood estimates decrease. Also, the maximum likelihood estimates and Bayesian estimates converge to the same value as the sample gets larger except for Generalized Entropy Loss Function.*

**Keywords:** Bayesian estimation, Prior Distribution, Loss Functions, Lindley's Approximation, Topp-Leone Lindley distribution

## 1. INTRODUCTION

Topp and Leone [1] introduced a distribution with finite support whose cumulative distribution function (cdf) has a closed form-expression called the Topp-Leone (the J-Shaped) distribution. This distribution has been used to model several phenomenon representing the time until the occurrence of a particular event. Data from such studies are called the survival data or lifetime data. Nadarajah and Kotz [2] studied and disclosed the usefulness of the Topp-Leone distribution in the analysis of interval-bounded data. In their study of the mathematical properties, it was observed that the Topp-Leone distribution exhibit bathtub failure rate functions and the closed form of the moments werederived, which disclosed the wide range of its applications in reliability study. The disclosure

of the important properties of the Topp-Leone distribution by Nadarajah and Kotz [2] has attracted the interest of authors which is evident in statistical literature. For instance, see the work of Ghitany et al [3], Zhou et al [4], Kotz and Seier [5], Nadarajah [6], Zghoul [7], amongst others. The cumulative frequency distribution (cdf) and probability density function (pdf) of the Topp-Leone (TL) distribution are respectively given as

$$G(t) = t^\alpha(2-t)^\alpha = [t(2-t)]^\alpha = [1-(1-t)^2]^\alpha, \quad 0 < t < 1, \quad \alpha > 0 \quad (1)$$

and

$$g(t) = 2\alpha(2-t)[1-(1-t)^2]^{\alpha-1} \quad 0 < t < 1, \quad \alpha > 0 \quad (2)$$

The TL distribution is on a unit interval support (0,1); this means that it cannot be used in the analysis of survival data, which are not on a unit interval support. To overcome this setback, Al-Shomrani et al [8] presented the Topp-Leone generated family of distribution with cdf and pdf given as

$$G(x; \alpha, \Phi) = F(x; \Phi)^\alpha (2 - F(x; \Phi))^\alpha = \left[1 - (\bar{F}(x; \Phi))^2\right]^\alpha, \quad x > 0, \quad \alpha > 0 \quad (3)$$

and

$$g(x; \alpha, \Phi) = 2\alpha f(x) \bar{F}(x; \Phi) \left[1 - (\bar{F}(x; \Phi))^2\right]^{\alpha-1}, \quad x > 0, \quad \alpha > 0 \quad (4)$$

Where  $f(x; \Phi)$ ,  $F(x; \Phi)$  and  $\bar{F}(x; \Phi)$  are respectively the pdf, cdf and survival functions of the baseline distribution and  $\Phi$  is the vector of parameters of the baseline distribution. Nzei and Ekhosuehi [9] used the logit of the TL-G family to presented the Topp-Leone Lindley (TL-L) distribution with the probability density function (pdf) and the cumulative distribution function (cdf) for the Topp-Leone Lindley (TL-L) distribution respectively expressed as;

$$g(x) = \frac{2\alpha\theta^2}{\theta+1} (1+x) \left(\frac{\theta+1+\theta x}{\theta+1}\right) e^{-2\theta x} \left\{1 - \left[\frac{\theta+1+\theta x}{\theta+1} e^{-\theta x}\right]^2\right\}^{\alpha-1}, \quad x > 0, \quad \alpha, \theta > 0 \quad (5)$$

and

$$G(x) = \left\{1 - \left[\frac{\theta+1+\theta x}{\theta+1} e^{-\theta x}\right]^2\right\}^\alpha, \quad x > 0, \quad \alpha, \theta > 0 \quad (6)$$

The Reliability (survival) function of the TL-L distribution is given as

$$R(x) = 1 - \left\{1 - \left[\frac{\theta+1+\theta x}{\theta+1} e^{-\theta x}\right]^2\right\}^\alpha \quad (7)$$

In addition, the corresponding hazard rate function of the TL-L distribution is expressed as

$$h(x) = \frac{\frac{2\alpha\theta^2}{\theta+1} (1+x) \left(\frac{\theta+1+\theta x}{\theta+1}\right) e^{-2\theta x} \left\{1 - \left[\frac{\theta+1+\theta x}{\theta+1} e^{-\theta x}\right]^2\right\}^{\alpha-1}}{1 - \left\{1 - \left[\frac{\theta+1+\theta x}{\theta+1} e^{-\theta x}\right]^2\right\}^\alpha} \quad (8)$$

The CDF, pdf and hazard rate function of the TL-L distribution are shown in Figure (1), (2) and (3) respectively for different values of the parameters  $\alpha$  and  $\theta$ .

The aim of this study is to obtain the Bayesian estimates of the parameters  $\alpha$  and  $\theta$  for TL-L distribution under different loss functions. The Bayesian framework is considered under the square error loss function (SELF) presented by Legendre [10] and Gauss [11], linear exponential (LINEX) loss function presented by Varian [12] and general entropy loss function (GELF) presented by Calabria and Pulcini [13] to obtain the Bayes estimators of the unknown parameters  $\alpha$  and  $\theta$ .

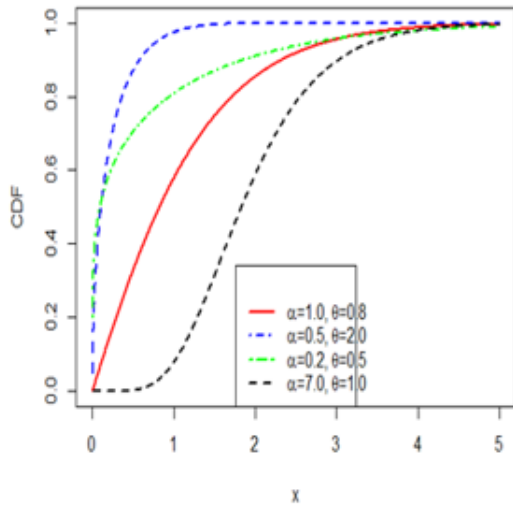
## 2. THE MAXIMUM LIKELIHOOD ESTIMATION (MLE)

Let  $x_i, i=1,2,3,\dots,n$  be a random sample from the TL-L distribution, then the maximum likelihood function of (5) denoted by  $L(x, \delta)$  is defined as:

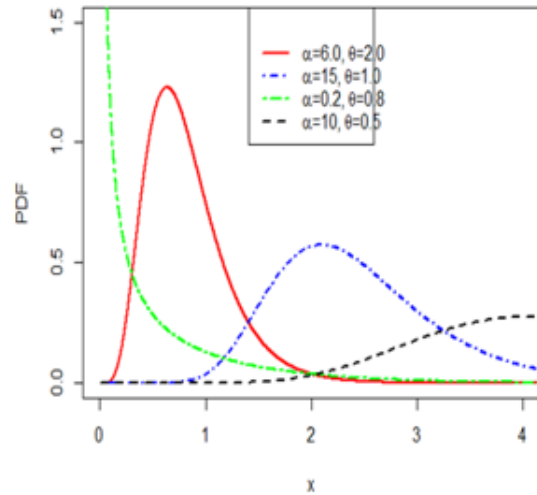
$$L(x; \alpha, \theta) = \prod_{i=1}^n \left\{ \frac{2\alpha\theta^2}{\theta+1} (1+x) \left( \frac{\theta+1+\theta x}{\theta+1} \right) e^{-2\theta x} \left[ 1 - \left( \frac{\theta+1+\theta x}{\theta+1} e^{-\theta x} \right)^2 \right]^{\alpha-1} \right\} \quad (9)$$

and the log-likelihood function denoted by  $\ell_n(x, \alpha, \theta)$  is given as

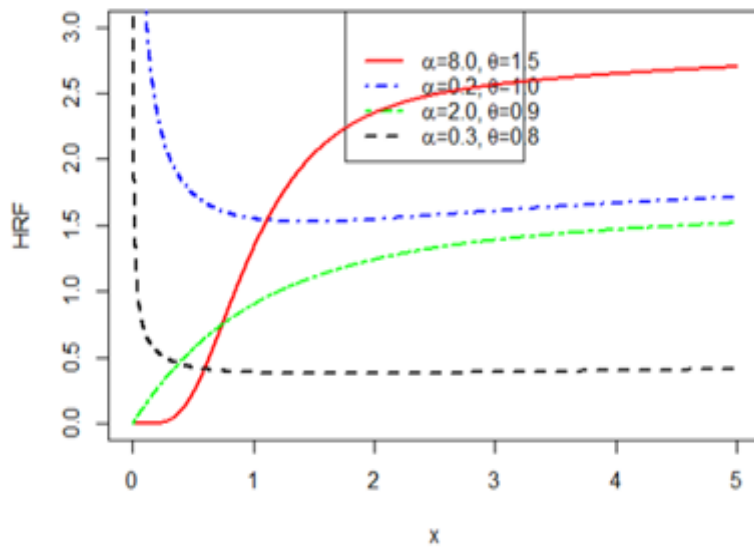
$$\begin{aligned} \ell_n(x, \alpha, \theta) = & n \ln(2\alpha\theta^2) - n \ln(\theta+1) - 2\theta \sum_{i=1}^n x_i + \sum_{i=1}^n \ln(1+x_i) + \sum_{i=1}^n \ln \left( \frac{\theta+1+\theta x_i}{\theta+1} \right) \\ & + (\alpha-1) \sum_{i=1}^n \ln \left[ 1 - \left( \frac{\theta+1+\theta x_i}{\theta+1} e^{-\theta x_i} \right)^2 \right] \end{aligned} \quad (10)$$



**Figure 1:** The CDF of TL-L Distribution



**Figure 2:** The PDF of TL-L Distribution



**Figure 3:** The HRF of TL-L Distribution

To obtain the MLEs of the TL-L, we solve the equations of the partial derivatives of the log-likelihood function with respect to the parameters  $\frac{\partial \ell_n}{\partial \alpha} = 0$  and  $\frac{\partial \ell_n}{\partial \theta} = 0$ . These partial derivatives with respect to the parameters  $\alpha$  and  $\theta$  are:

$$\frac{\partial \ell_n}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \ln \left[ 1 - \left( \frac{\theta + 1 + \theta x}{\theta + 1} e^{-\theta x} \right)^2 \right] \quad (11)$$

$$\frac{\partial \ell_n}{\partial \theta} = \frac{2n}{\theta} - \frac{n}{\theta + 1} - 2 \sum_{i=1}^n x + \frac{1}{\theta + 1} \sum_{i=1}^n \frac{x}{\theta + 1 + \theta x} + 2\theta(\alpha - 1) \sum_{i=1}^n \frac{(\theta + 1 + \theta x)e^{-2\theta x}}{(\theta + 1) \left[ (\theta + 1)^2 - (\theta + 1 + \theta x)e^{-2\theta x} \right]} \quad (12)$$

The solution of  $\frac{\partial \ell_n}{\partial \alpha} = 0$ , is the MLE of  $\hat{\alpha}$  which given as

$$\hat{\alpha} = \frac{-n}{\sum_{i=1}^n \ln \left[ 1 - \left( \frac{\theta + 1 + \theta x}{\theta + 1} e^{-\theta x} \right)^2 \right]} \quad (13)$$

By replacing  $\alpha$  (12) in  $\frac{\partial \ell_n}{\partial \theta} = 0$  with estimate in (13), we have expression in terms of the parameter  $\theta$  as

$$\frac{2n}{\theta} - \frac{n}{\theta + 1} - 2 \sum_{i=1}^n x + \frac{1}{\theta + 1} \sum_{i=1}^n \frac{x}{\theta + 1 + \theta x} + 2\theta \left( \frac{-n}{\sum_{i=1}^n \ln \left[ 1 - \left( \frac{\theta + 1 + \theta x}{\theta + 1} e^{-\theta x} \right)^2 \right]} - 1 \right) \sum_{i=1}^n \frac{(\theta + 1 + \theta x)e^{-2\theta x}}{(\theta + 1) \left[ (\theta + 1)^2 - (\theta + 1 + \theta x)e^{-2\theta x} \right]} = 0 \quad (14)$$

Obviously, (14) is a complex equation, which cannot be solved analytically. Hence, solving (12) and (13) simultaneously to obtain the maximum likelihood estimates of  $\alpha$  and  $\theta$  requires iterative approach such as Newton-Raphson iterative scheme as presented by Obisesan et al [14] and Bakari et al [15] amongst others. This Newton-Raphson method can be performed with R-Software package.

### 3. BAYESIAN ESTIMATION (BE)

The main belief of Bayesian statistics that distinguishes it from the classical statistics is that it consider the parameter(s) of the given model to be random variables with prior distribution denoted by  $\pi(\Phi)$ . In this Section, we discuss the Bayesian estimates for the parameters of the TL-L distribution using the Jeffreys' (non-informative) prior for  $\alpha$  and the Gamma (conjugate) prior for  $\theta$  under some loss functions namely; squared error loss function (SELF), linear exponential loss function (LINEX) and general entropy loss function (GELF). We discuss these loss functions and the priors briefly as follows:

#### 3.1 The Square Error Loss Function (SELF)

The square error loss function, which is the simplest and the most commonly used symmetric loss function in the literature by authors, see Rastogi and Merovci [16] and Sangeeta et al [17] amongst others. It is defined as

$$L_{SELF}(\hat{\Phi}, \Phi) = (\hat{\Phi} - \Phi)^2 \quad (15)$$

The Bayesian estimate under SELF is  $\hat{\Phi}_{BSELF} = E_{\Phi}(\Phi | \underline{x})$ .

This is the expectation considered with regard to the posterior density. SELF assigns the same magnitude of error to both over estimation and under estimation because of its symmetric nature, which is not always true in many practical scenario Kaur et al. [18].

### 3.2 The Linear Exponential Loss Function (LINEX)

Varian [23] presented an asymmetric loss function defined as

$$L_{LINEX}(\hat{\Phi}, \Phi) = e^{m(\hat{\Phi}-\Phi)} - m(\hat{\Phi}-\Phi) - 1 \quad (16)$$

Where  $m \neq 0$  is the shape parameter of the LINEX loss function. Zellner[19] studied the properties of this loss function and showed that for  $m > 0$ , over estimation is more costly than under estimation. When  $m < 0$ , the loss function increases almost exponentially for  $d < 0$  and almost linearly for  $d > 0$ , where  $d = \hat{\Phi} - \Phi$ . The Bayesian estimate under the LINEX loss function is given as

$$\hat{\Phi}_{BLINEX} = -\frac{1}{m} \ln \left[ E_{\Phi} \left( e^{-m\Phi} | \underline{x} \right) \right] \quad (17)$$

### 3.3 The General Entropy Loss Function (GELF)

The general entropy loss function (GELF) was proposed by Calabria and Pulcini [13] as an alternative to the modified LINEX loss function and it is defined as

$$L_{GELF}(\hat{\Phi}, \Phi) = \left( \frac{\hat{\Phi}}{\Phi} \right)^k - k \ln \left( \frac{\hat{\Phi}}{\Phi} \right) - 1 \quad (18)$$

Where  $k \neq 0$  and it determines the shape of the loss function. When  $k < 0$ , it shows there is more of under estimation than over estimation. On the other hand, when  $k > 0$  shows more of over estimation than under estimation. The Bayes estimate of  $\Phi$  under the general entropy loss function is given as

$$\hat{\Phi}_{BGELF} = \left[ E_{\Phi} \left( \Phi^{-k} | \underline{x} \right) \right]^{-\frac{1}{k}} \quad (19)$$

It is important to note that for  $k = -1$ ,  $\hat{\Phi}_{GELF} = \hat{\Phi}_{SELF}$  i.e. the general entropy loss function reduces to the square error loss function at  $k = -1$ .

### 3.4 Prior Distributions:

The choice of prior distribution for an unknown parameter(s) is an important part of Bayesian statistics. For the Bayes estimate of the parameters  $\alpha$  and  $\theta$ , we consider the Jeffreys' (non-informative) prior for  $\alpha$  and the Gamma (conjugate) prior for  $\theta$ . Then the prior distributions are defined below as:

$$\pi_1(\Phi) \propto \sqrt{I(\Phi)} \quad (20)$$

Where  $I(\Phi) = -E \left\{ \frac{\partial^2 \ell_n}{\partial \Phi^2} \right\}$  which is the Fisher's Information. For the TL-L distribution, the Jeffreys' prior of  $\alpha$  is defined as

$$\pi_1(\alpha) = \frac{1}{\alpha} \quad \alpha > 0 \quad (21)$$

and

$$\pi_2(\theta) = \frac{q^p}{\Gamma(p)} \theta^{p-1} e^{-q\theta} \quad \theta > 0, p > 0, q > 0 \quad (22)$$

The joint prior distribution of the parameters  $\alpha$  and  $\theta$  is defined as a combination of the priors as

$$\pi(\alpha, \theta) = \frac{q^p}{\alpha \Gamma(p)} \theta^{p-1} e^{-q\theta} \quad (23)$$

### 3.5 Posterior Distribution

The posterior distribution function of an unknown probability distribution parameter  $\Phi$  is the formula used to compute the conditional probability density of the distribution parameter  $\Phi$  given the data  $X = x$  through the Bayes formula defined as

$$P(\Phi | \underline{x}) = \frac{L(x | \Phi)\pi(\Phi)}{\int_{\Theta} L(x | \Phi)\pi(\Phi)d\Phi} \quad (24)$$

Where the prior distribution of the unknown parameter is  $\pi(\Phi)$ ,  $L(x | \Phi)$  is the likelihood function of the density of  $X$  and  $\Phi$  is vector of the unknown parameter. Then the posterior distribution of the TL-L distribution parameters  $\alpha$  and  $\theta$  is obtained by substituting (9) and (23) into (24) to be

$$P(\alpha, \theta | \underline{x}) = \frac{\frac{2^n \alpha^{n-1} \theta^{2n}}{(\theta+1)^n \Gamma(p)} \prod_{i=1}^n (1+x) \left( \frac{\theta+1+\theta x}{\theta+1} \right) \left\{ 1 - \left[ \frac{\theta+1+\theta x}{\theta+1} e^{-\theta x} \right]^2 \right\}^{\alpha-1} e^{-\theta(2x+q)}}{\int_0^\infty \int_0^\infty \frac{2^n \alpha^{n-1} \theta^{2n}}{(\theta+1)^n \Gamma(p)} \prod_{i=1}^n (1+x) \left( \frac{\theta+1+\theta x}{\theta+1} \right) \left\{ 1 - \left[ \frac{\theta+1+\theta x}{\theta+1} e^{-\theta x} \right]^2 \right\}^{\alpha-1} e^{-\theta(2x+q)} d\alpha d\theta} \quad (25)$$

Obviously, the posterior distribution in (25) for the estimation of TL-L parameters,  $\alpha$  and  $\theta$  is in a rational form which cannot be reduced to a closed form, making tedious to evaluate the posterior distribution in order to obtain the Baye's estimators. However, one can use the approach developed by Lindley [20], to approximate these Bayes estimators.

### 3.6 Lindley's Approximation

Lindley [20] developed a method for reducing the posterior distribution in Bayesian estimation, which involves integral that can't be expressed in closed form. This method provides a simplified form of Bayesian estimator, which makes it easier to apply in practice. Several authors have used the Lindley approximation to obtain the Bayes estimate for some lifetime distribution in the literature; amongst whom are Hummara and Ahmad [21], Adegoke et al ([22], [23]), Kamran et al [24], Bashiru et al [25], etc. Lindley developed an asymptotic approximation to the ratio

$$I(X) = \frac{\int_{(\alpha, \theta)} Z(\alpha, \theta) e^{L(\alpha, \theta) + U(\alpha, \theta)} \partial(\alpha, \theta)}{\int_{(\alpha, \theta)} e^{L(\alpha, \theta) + U(\alpha, \theta)} \partial(\alpha, \theta)} \quad (26)$$

Where  $Z(\alpha, \theta)$  is a function of the distribution parameter  $\alpha$  and  $\theta$ ,  $L(\alpha, \theta)$  is the log-likelihood function and  $U(\alpha, \theta)$  is the log of the prior distribution function  $\pi(\alpha, \theta)$ . Therefore,  $I(X)$  is evaluated as

$$I(X) = Z(\alpha, \theta) + \frac{1}{2} [Z_{11} \sigma_{11} + Z_{22} \sigma_{22}] + (U_1 Z_1 \sigma_{11} + U_2 Z_2 \sigma_{22}) + \frac{1}{2} [L_{111} Z_1 \sigma_{11}^2 + L_{222} Z_2 \sigma_{22}^2] + \frac{1}{2} [L_{122} Z_1 \sigma_{11} \sigma_{22} + L_{112} Z_2 \sigma_{11} \sigma_{22}] \quad (27)$$

Therefore, for an unknown parameter  $\alpha$ , the Lindley approximation is can be expressed as

$$E[Z(\alpha | \underline{x})] = Z(\hat{\alpha}, \hat{\theta}) + \frac{1}{2} [Z_{11} \sigma_{11}] + (U_1 Z_1 \sigma_{11}) + \frac{1}{2} [L_{111} Z_1 \sigma_{11}^2 + L_{122} Z_1 \sigma_{11} \sigma_{22}] \quad (28)$$

Similarly, for an unknown parameter  $\theta$ , the Lindley approximation is can be expressed as

$$E[Z(\theta|x)] = Z(\hat{\alpha}, \hat{\theta}) + \frac{1}{2} [Z_{22}\sigma_{22}] + (U_2 Z_2 \sigma_{22}) + \frac{1}{2} [L_{222} Z_2 \sigma_{22}^2 + L_{221} Z_2 \sigma_{11} \sigma_{22}] \quad (29)$$

Where the elements of the Lindley approximation in (27 - 29) are as given below

$$Z_1 = \frac{\partial Z(\alpha, \theta)}{\partial \alpha}; \quad Z_2 = \frac{\partial Z(\alpha, \theta)}{\partial \theta}; \quad Z_{11} = \frac{\partial^2 Z(\alpha, \theta)}{\partial \alpha^2}; \quad Z_{22} = \frac{\partial^2 Z(\alpha, \theta)}{\partial \theta^2}; \quad Z_{12} = \frac{\partial^2 Z(\alpha, \theta)}{\partial \alpha \partial \theta}$$

$$\text{and } Z_{21} = \frac{\partial^2 Z(\alpha, \theta)}{\partial \theta^2 \partial \alpha}.$$

$$P(\alpha, \theta) = \ln \pi(\alpha, \theta) = p \ln q - \ln \Gamma(p) - \ln \alpha + (p-1) \ln \theta - q\theta;$$

$$U_1 = \frac{\partial P(\alpha, \theta)}{\partial \alpha} = -\frac{1}{\alpha} \quad \text{and} \quad U_2 = \frac{\partial P(\alpha, \theta)}{\partial \theta} = \frac{p-1}{\theta} - q.$$

$$L_{111} = \frac{\partial^3 \ln L(\alpha, \theta)}{\partial \alpha^3} = \frac{2n}{\alpha^3}$$

$$L_{222} = \frac{\partial^3 \ln L(\alpha, \theta)}{\partial \theta^3} = \frac{2n}{\theta^3} - \frac{2n}{(\theta+1)^3} + \frac{1}{\theta+1} \sum_{i=1}^n \frac{x(1+x)^2}{(\theta+1+\theta x)^3} - \frac{2}{(\theta+1)^2} \sum_{i=1}^n \frac{x(1+x)}{(\theta+1+\theta x)^2}$$

$$+ \frac{2}{(\theta+1)^3} \sum_{i=1}^n \frac{x}{(\theta+1+\theta x)} + \frac{2(\alpha+1)}{\theta+1} \frac{\sum_{i=1}^n \frac{CBA'' - CAB'' - C'A'B + C'B'A}{C^2}}{\sum_{i=1}^n \frac{BA' - AB'}{B^2}} - \frac{2(\alpha-1)}{(\theta+1)^2} \frac{\sum_{i=1}^n \frac{BA' - AB'}{B^2}}{\sum_{i=1}^n \frac{BA' - AB'}{B^2}}$$

$$L_{122} = \frac{\partial^3 \ln L(\alpha, \theta)}{\partial \alpha \partial \theta^2} = \frac{2}{\theta+1} \sum_{i=1}^n \frac{BA' - AB'}{B^2} - \frac{2}{(\theta+1)^2} \sum_{i=1}^n \frac{A}{B}$$

$$L_{112} = \frac{\partial^3 \ln L(\alpha, \theta)}{\partial \alpha^2 \partial \theta} = 0$$

$$\sigma_{11} = -\frac{1}{L_{11}} = \frac{\alpha^2}{n} \quad \text{and}$$

$$\sigma_{22} = -\frac{1}{L_{22}} = \left[ \frac{n}{\theta^2} - \frac{n}{(\theta+1)^2} - \frac{1}{\theta+1} \sum_{i=1}^n \frac{x(1+x)}{(\theta+1+\theta x)^2} + \frac{2}{(\theta+1)^2} \sum_{i=1}^n \frac{x}{(\theta+1+\theta x)} - \frac{2(\alpha+1)}{(\theta+1)} \right. \\ \left. + \frac{\sum_{i=1}^n \frac{BA' - AB'}{B^2}}{\sum_{i=1}^n \frac{BA' - AB'}{B^2}} \frac{2(\alpha-1)}{(\theta+1)^2} \frac{\sum_{i=1}^n \frac{A}{B}}{\sum_{i=1}^n \frac{A}{B}} \right]^{-1}$$

Where

$$A = x(\theta+1+\theta x)[(\theta+1)(\theta+1+\theta x)-1]e^{-2\theta x}$$

$$A' = x \left\{ [(\theta+1)(\theta+1+\theta x)-1] [1-x-2\theta x-2\theta x^2] + (\theta+1+\theta x)[(\theta+1+\theta x) + (\theta+1)(1+x)] \right\} e^{-2\theta x}$$

$$B = (\theta+1)^2 - (\theta+1+\theta x)^2 e^{-2\theta x}$$

$$B' = 2(\theta+1) + 2(\theta+1+\theta x)(\theta x^2 + \theta x - 1) e^{-2\theta x}$$

$$C = \left\{ (\theta+1)^2 - (\theta+1+\theta x)^2 e^{-2\theta x} \right\}^2$$



$$C' = 2 \left( (\theta + 1)^2 - (\theta + 1 + \theta x)^2 e^{-2\theta x} \right) \left\{ 2(\theta + 1) - 2(\theta + 1 + \theta x) (\theta x^2 \theta x - 1) e^{-2\theta x} \right\}$$

### 3.7 Lindley Approximation under the Different Loss Functions

In this section, we consider the Bayes estimators of the TL-L parameters  $\alpha$  and  $\theta$  are obtained assuming that both  $\alpha$  and  $\theta$  are unknown, using the prior in (23) under three different loss functions:

#### 3.7.1 Under Squared Error Entropy Loss Function

a) For the parameter  $\alpha$ , it can be seen from the SELF estimator that  $Z(\alpha, \theta) = \alpha$ , then  $Z_1 = 1$ , and  $Z_2 = Z_{11} = Z_{22} = 0$ , we have

$$\hat{\alpha}_{SELF} = E(\alpha | x) = \alpha \left[ 1 + \frac{\alpha}{2n} H_1 H_2 \right] \quad (30)$$

Where  $H_1 = \frac{2}{\theta + 1} \sum_{i=1}^n \frac{BA' - AB'}{B^2} - \frac{2}{(\theta + 1)^2} \sum_{i=1}^n \frac{A}{B}$  and

$$H_2 = \left[ \frac{n}{\theta^2} - \frac{n}{(\theta + 1)^2} - \frac{1}{\theta + 1} \sum_{i=1}^n \frac{x(1+x)}{(\theta + 1 + \theta x)^2} + \frac{2}{(\theta + 1)^2} \sum_{i=1}^n \frac{x}{(\theta + 1 + \theta x)} - \frac{2(\alpha + 1)}{(\theta + 1)} + \sum_{i=1}^n \frac{BA' - AB'}{B^2} \frac{2(\alpha - 1)}{(\theta + 1)^2} \sum_{i=1}^n \frac{A}{B} \right]^{-1}$$

b) For the parameter  $\theta$ , it can be seen from the SELF estimator that  $Z(\alpha, \theta) = \theta$ , then  $Z_2 = 1$ , and  $Z_1 = Z_{11} = Z_{22} = 0$ , we have

$$\hat{\theta}_{SELF} = E[Z(\theta | x)] = \theta + H_2 \left( \frac{p-1}{\theta} - q + \frac{1}{2} H_2 H_3 \right) \quad (31)$$

Where

$$H_2 = \left[ \frac{n}{\theta^2} - \frac{n}{(\theta + 1)^2} - \frac{1}{\theta + 1} \sum_{i=1}^n \frac{x(1+x)}{(\theta + 1 + \theta x)^2} + \frac{2}{(\theta + 1)^2} \sum_{i=1}^n \frac{x}{(\theta + 1 + \theta x)} - \frac{2(\alpha + 1)}{(\theta + 1)} + \sum_{i=1}^n \frac{BA' - AB'}{B^2} \frac{2(\alpha - 1)}{(\theta + 1)^2} \sum_{i=1}^n \frac{A}{B} \right]^{-1}$$

and

$$H_3 = \frac{2n}{\theta^3} - \frac{2n}{(\theta + 1)^3} + \frac{1}{\theta + 1} \sum_{i=1}^n \frac{x(1+x)^2}{(\theta + 1 + \theta x)^3} - \frac{2}{(\theta + 1)^2} \sum_{i=1}^n \frac{x(1+x)}{(\theta + 1 + \theta x)^2} + \frac{2}{(\theta + 1)^3} \sum_{i=1}^n \frac{x}{(\theta + 1 + \theta x)} + \frac{2(\alpha + 1)}{\theta + 1} \sum_{i=1}^n \frac{CBA'' - CAB'' - C'A'B + C'B'A}{C^2} - \frac{2(\alpha - 1)}{(\theta + 1)^2} \sum_{i=1}^n \frac{BA' - AB'}{B^2}$$

### 3.7.2 Under LINEX Loss Function

a) For the parameter  $\alpha$ , it can be seen from the LINEX estimator that  $Z(\alpha, \theta) = e^{-m\alpha}$ , then  $Z_1 = -me^{-m\alpha}$ ,  $Z_{11} = m^2e^{-m\alpha}$  and  $Z_2 = Z_{22} = 0$  we have

$$\hat{\alpha}_{LINEX} = E\left(e^{-m\alpha} | \underline{x}\right) = e^{-m\alpha} \left[ 1 + \frac{m\alpha}{2} (\alpha - H_1 H_2) \right] \quad (32)$$

Where

$$H_1 = \frac{2}{\theta+1} \sum_{i=1}^n \frac{BA' - AB'}{B^2} - \frac{2}{(\theta+1)^2} \sum_{i=1}^n \frac{A}{B} \quad \text{and}$$

$$H_2 = \left[ \frac{n}{\theta^2} - \frac{n}{(\theta+1)^2} - \frac{1}{\theta+1} \sum_{i=1}^n \frac{x(1+x)}{(\theta+1+\theta x)^2} + \frac{2}{(\theta+1)^2} \sum_{i=1}^n \frac{x}{(\theta+1+\theta x)} - \frac{2(\alpha+1)}{(\theta+1)} \right. \\ \left. + \sum_{i=1}^n \frac{BA' - AB'}{B^2} \frac{2(\alpha-1)}{(\theta+1)^2} \sum_{i=1}^n \frac{A}{B} \right]^{-1}$$

b) For the parameter  $\theta$ , it can be seen from the LINEX estimator that  $Z(\alpha, \theta) = e^{-m\theta}$ , then  $Z_2 = -me^{-m\theta}$ ,  $Z_{22} = m^2e^{-m\theta}$  and  $Z_1 = Z_{11} = 0$ , we have

$$\hat{\theta}_{LINEX} = E\left(e^{-m\theta} | \underline{x}\right) = e^{-m\theta} \left\{ 1 + mH_2 \left[ \frac{1}{2} (m - H_2 H_3) + q - \frac{p-1}{\theta} \right] \right\} \quad (33)$$

Where

$$H_2 = \left[ \frac{n}{\theta^2} - \frac{n}{(\theta+1)^2} - \frac{1}{\theta+1} \sum_{i=1}^n \frac{x(1+x)}{(\theta+1+\theta x)^2} + \frac{2}{(\theta+1)^2} \sum_{i=1}^n \frac{x}{(\theta+1+\theta x)} - \frac{2(\alpha+1)}{(\theta+1)} \right. \\ \left. + \sum_{i=1}^n \frac{BA' - AB'}{B^2} \frac{2(\alpha-1)}{(\theta+1)^2} \sum_{i=1}^n \frac{A}{B} \right]^{-1}$$

and

$$H_3 = \frac{2n}{\theta^3} - \frac{2n}{(\theta+1)^3} + \frac{1}{\theta+1} \sum_{i=1}^n \frac{x(1+x)^2}{(\theta+1+\theta x)^3} - \frac{2}{(\theta+1)^2} \sum_{i=1}^n \frac{x(1+x)}{(\theta+1+\theta x)^2} + \frac{2}{(\theta+1)^3} \sum_{i=1}^n \frac{x}{(\theta+1+\theta x)} \\ + \frac{2(\alpha+1)}{\theta+1} \sum_{i=1}^n \frac{CBA'' - CAB'' - C'A'B + C'B'A}{C^2} - \frac{2(\alpha-1)}{(\theta+1)^2} \sum_{i=1}^n \frac{BA' - AB'}{B^2}$$

### 3.7.3 Under GELF Loss Function

a) For the parameter  $\alpha$ , it can be seen from the GELF estimator that  $Z(\alpha, \theta) = \alpha^{-k}$ , then  $Z_1 = -k\alpha^{-(k+1)}$ ,  $Z_{11} = k(k+1)\alpha^{-(k+2)}$  and  $Z_2 = Z_{22} = 0$  we have

$$\hat{\alpha}_{GELF} = E\left(\alpha^{-k} | \underline{x}\right) = \alpha^{-k} \left[ 1 + \frac{k\alpha}{2n} (k+1 - H_1 H_2) \right] \quad (34)$$

Where

$$H_1 = \frac{2}{\theta+1} \sum_{i=1}^n \frac{BA' - AB'}{B^2} - \frac{2}{(\theta+1)^2} \sum_{i=1}^n \frac{A}{B} \quad \text{and}$$

$$H_2 = \left[ \frac{n}{\theta^2} - \frac{n}{(\theta+1)^2} - \frac{1}{\theta+1} \sum_{i=1}^n \frac{x(1+x)}{(\theta+1+\theta x)^2} + \frac{2}{(\theta+1)^2} \sum_{i=1}^n \frac{x}{(\theta+1+\theta x)} - \frac{2(\alpha+1)}{(\theta+1)} + \sum_{i=1}^n \frac{BA' - AB'}{B^2} \frac{2(\alpha-1)}{(\theta+1)^2} \sum_{i=1}^n \frac{A}{B} \right]^{-1}$$

b) For the parameter  $\theta$ , it can be seen from the GELF estimator that  $Z(\alpha, \theta) = \alpha^{-k}$ , then  $Z_2 = -k \alpha^{-(k+1)}$ ,  $Z_{22} = k(k+1) \alpha^{-(k+2)}$  and  $Z_1 = Z_{11} = 0$ , we have

$$\hat{\theta}_{GELF} = E(\theta^{-k} | \underline{x}) = \theta^{-k} \left\{ 1 + k\theta H_2 \left[ \frac{1}{2} ((k+1)\theta - H_2 H_3) + q - \frac{p-1}{\theta} \right] \right\} \quad (35)$$

Where

$$H_2 = \left[ \frac{n}{\theta^2} - \frac{n}{(\theta+1)^2} - \frac{1}{\theta+1} \sum_{i=1}^n \frac{x(1+x)}{(\theta+1+\theta x)^2} + \frac{2}{(\theta+1)^2} \sum_{i=1}^n \frac{x}{(\theta+1+\theta x)} - \frac{2(\alpha+1)}{(\theta+1)} + \sum_{i=1}^n \frac{BA' - AB'}{B^2} \frac{2(\alpha-1)}{(\theta+1)^2} \sum_{i=1}^n \frac{A}{B} \right]^{-1}$$

and

$$H_3 = \frac{2n}{\theta^3} - \frac{2n}{(\theta+1)^3} + \frac{1}{\theta+1} \sum_{i=1}^n \frac{x(1+x)^2}{(\theta+1+\theta x)^3} - \frac{2}{(\theta+1)^2} \sum_{i=1}^n \frac{x(1+x)}{(\theta+1+\theta x)^2} + \frac{2}{(\theta+1)^3} \sum_{i=1}^n \frac{x}{(\theta+1+\theta x)} + \frac{2(\alpha+1)}{\theta+1} \sum_{i=1}^n \frac{CBA'' - CAB'' - C'A'B + C'B'A}{C^2} - \frac{2(\alpha-1)}{(\theta+1)^2} \sum_{i=1}^n \frac{BA' - AB'}{B^2}$$

## 4. NUMERICAL ANALYSIS

### 4.1 Monte Carlo Simulation Study

In this section, a Monte Carlo simulation study was carried out with R Statistical software to compare the performance and accuracy of the proposed Bayesian estimators and their maximum likelihood estimates counterpart of TL-L distribution parameters  $\alpha$  and  $\theta$  by using mean square Errors (MSE) and the BIAS given as:

$$MSE = \frac{1}{N} \sum_{i=1}^N (\hat{\Phi} - \Phi)^2$$

and

$$BIAS = \frac{1}{N} \sum_{i=1}^N |\hat{\Phi} - \Phi|$$

Where N is the number of samples. In each simulation, we generate N=10,000 samples of size  $n = 30, 50, 100, 200, 500, 1000$  from TL-L distribution for some sets of parameter values  $\alpha = 0.84, 1.6, 2, 2.5$  and  $\theta = 0.5, 2, 2.5$ . We assume that  $p$  takes the values  $p = 2, 5, 8, 10$ ;  $q$  takes the values  $q = 1, 2, 5, 12$ ;  $m$  takes the values  $m = 1, 6, 8, 15$  and  $c$  takes the values

$c = -0.25, -0.5, -0.65 - 0.75$ . These results presented in Tables 1- 4 below showed the mean, MSE's and bias for estimating the parameters  $\alpha$  and  $\theta$ .

From the results of the simulation study in Table 1 – 4, we summarize our observations as follows:

- i. For any fixed values of the parameters  $\alpha$  and  $\theta$ , as sample size increases, the MSEs of all the estimators, both MLEs and Bayesian Estimates decrease.
- ii. The values of the hyper parameters from the prior distribution have minimal effect on the posterior estimates.
- iii. Generally, the terms of MSEs of the MLEs and Bayesian estimates converge to the same value as for the large sample except for GELF.

## 4.2 Real Data Analysis

This section present the application of TLL distribution to real data set. This data set represent 66 breaking stress of carbon fibers (in Gba) which was reported in Nicholas and Padgett [26].

3.70, 2.74, 2.73, 2.50, 3.60, 3.11, 3.27, 2.87, 1.47, 3.11, 3.56, 4.42,  
 2.41, 3.19, 3.22, 1.69, 3.28, 3.09, 1.87, 3.15, 4.90, 1.57, 2.67, 2.93,  
 3.22, 3.39, 2.81, 4.20, 3.33, 2.55, 3.31, 3.31, 2.85, 1.25, 4.38, 1.84,  
 0.39, 3.68, 2.48, 0.85, 1.61, 2.79, 4.70, 2.03, 1.89, 2.88, 2.82, 2.05,  
 3.65, 3.75, 2.43, 2.95, 2.97, 3.39, 2.96, 2.35, 2.55, 2.59, 2.03, 1.61,  
 2.12, 3.15, 1.08, 2.56, 1.80, 2.53

**Table 1:** Showing mean of ML and Bayesian estimates with corresponding MSEs and Bias for  $\theta = \alpha = 2$  (while  $p = 2, q = 1, m = 1, c = -0.75$ )

n	Method	$\theta$			$\alpha$		
		MEAN	MSE	BIAS	MEAN	MSE	BIAS
30	ML	2.563265	0.317267	0.563265	2.155780	0.024267	0.155780
	LINEX	2.563265	0.317267	0.563265	2.156686	0.024550	0.156686
	GELF	2.019669	0.000386	0.019669	1.656376	0.118120	0.343623
	SELF	2.559342	0.312864	0.559342	2.156988	0.024645	0.156988
50	ML	2.485253	0.235470	0.485253	2.125277	0.015694	0.125277
	LINEX	2.485253	0.235470	0.485253	2.130306	0.017000	0.130306
	GELF	1.965990	0.001156	0.034009	1.686560	0.098244	0.313439
	SELF	2.473202	0.223021	0.473202	2.131028	0.017197	0.131028
100	ML	2.219903	0.048357	0.219903	2.067709	0.004585	0.0677091
	LINEX	2.219903	0.048357	0.219903	2.066873	0.004473	0.066873
	GELF	1.817202	0.033414	0.182797	1.692596	0.094496	0.307403
	SELF	2.218925	0.047928	0.218925	2.066875	0.004473	0.066875
200	ML	2.194559	0.037853	0.194559	1.940733	0.003512	0.059266
	LINEX	2.194559	0.037853	0.194559	1.959422	0.001751	0.040577
	GELF	1.798195	0.040554	0.201380	1.723795	0.076289	0.276204
	SELF	2.190094	0.036136	0.190094	1.959858	0.001720	0.040141
500	ML	1.926363	0.005422	0.073636	2.017662	0.000311	0.017662
	LINEX	1.926363	0.005422	0.073636	2.017107	0.000292	0.017107
	GELF	1.634335	0.133710	0.365665	1.763728	0.055834	0.236271
	SELF	1.925588	0.005537	0.074411	2.017155	0.000294	0.017155
1000	ML	1.975739	0.000588	0.024260	2.010489	0.000110	0.010489
	LINEX	1.975739	0.000588	0.024260	2.007554	0.000057	0.007553
	GELF	1.666054	0.111519	0.333945	1.779841	0.048469	0.220158
	SELF	1.975331	0.000608	.024668	2.007565	0.000057	0.007565

**Table 2:** Showing mean of ML and Bayesian estimates with corresponding MSEs and Bias for  $\theta = \alpha = 2.5$  (while  $p = 5, q = 2, m = 8, k = -0.5$ )

N	Method	$\theta$			$\alpha$		
		MEAN	MSE	BIAS	MEAN	MSE	BIAS
30	ML	3.200958	0.491343	0.700958	2.746747	0.060884	0.246747
	LINEX	3.200958	0.491343	0.700958	2.746587	0.060805	0.246587
	GELF	0.318750	4.757852	2.181250	0.364048	4.562259	2.135951
	SELF	3.199980	0.489972	0.699980	2.746916	0.060967	0.246961
50	ML	3.103670	0.364418	0.603670	2.733993	0.054753	0.233993
	LINEX	3.103670	0.364418	0.603670	2.734001	0.054756	0.234001
	GELF	1.753758	0.556876	0.746241	0.365735	4.555087	2.134265
	SELF	3.101468	0.361764	0.601468	2.734241	0.054486	0.234241
100	ML	2.775719	0.076021	0.275719	2.659853	0.025531	0.159853
	LINEX	2.775719	0.076021	0.275719	2.659853	0.025531	0.159853
	GELF	0.362098	4.570621	2.137901	1.630771	0.755558	0.869228
	SELF	2.775484	0.075891	0.275484	2.659435	0.025420	0.159435
200	ML	2.749392	0.062196	0.249392	2.463737	0.001314	0.036262
	LINEX	2.749392	0.062196	0.249392	2.463329	0.001344	0.036670
	GELF	0.367475	4.547659	2.132524	0.405834	4.385529	2.094165
	SELF	2.748470	0.061737	0.248470	2.464147	0.001285	0.035858
500	ML	2.411848	0.007770	0.088151	2.516045	0.000257	0.016045
	LINEX	2.411848	0.007770	0.088151	2.515688	0.000246	0.015688
	GELF	0.415480	4.345220	2.084519	0.397501	4.420501	2.102498
	SELF	2.411668	0.007803	0.088338	2.515719	0.000247	0.015719
1000	ML	2.472058	0.000780	0.027941	2.512703	0.000161	0.012703
	LINEX	2.472058	0.000780	0.027941	2.512889	0.000166	0.012889
	GELF	0.404940	4.389274	2.095059	0.397943	4.418641	2.102056
	SELF	2.471967	0.000785	0.028032	2.512922	0.000166	0.012922

**Table 3:** Showing mean of ML and Bayesian estimates with corresponding MSEs and Bias for  $\theta=0.5$  and  $\alpha = 1.6$  (while  $p = 10, q = 5, m = 15, c = -0.25$ )

n	Method	$\theta$			$\alpha$		
		MEAN	MSE	BIAS	MEAN	MSE	BIAS
30	ML	0.482010	1.249900	1.117989	1.725412	1.501634	1.225412
	LINEX	0.482010	1.249900	1.117989	1.725519	1.501897	1.225519
	GELF	0.833072	0.588178	0.766927	1.146106	0.417453	0.646106
	SELF	0.482010	1.249900	1.117989	1.602093	1.221410	1.102093
50	ML	0.494062	1.223098	1.105937	1.700421	1.441010	1.200421
	LINEX	0.494062	1.223098	1.105937	1.701528	1.443671	1.201528
	GELF	0.838310	0.580171	0.761689	1.141972	0.412128	0.641972
	SELF	0.494062	1.223098	1.105937	1.700645	1.441547	1.200645
100	ML	0.547370	1.108029	1.052629	1.654412	1.332668	1.154412
	LINEX	0.547370	1.108029	1.052629	1.624175	1.264541	1.124175
	GELF	0.859336	0.548582	0.74663	1.127087	0.393238	0.627087
	SELF	0.547371	1.108026	1.052628	1.553157	1.109139	1.053157
200	ML	0.553372	1.095429	1.046627	1.613726	1.240386	1.113726
	LINEX	0.553372	1.095429	1.046627	1.613730	1.240395	1.113730
	GELF	0.862086	0.544517	0.737914	1.126783	0.391974	0.626078
	SELF	0.553371	1.095430	1.046628	1.725438	1.501699	1.225438

500	ML	0.616878	0.966528	0.983122	1.607956	1.227566	1.107956
	LINEX	0.616878	0.966528	0.983122	1.607959	1.227574	1.107959
	GELF	0.888384	0.506396	0.711615	1.125163	0.391029	0.625163
	SELF	0.630569	0.939795	0.969430	1.613727	1.240389	1.113727
1000	ML	0.635364	0.903520	0.964635	1.553046	1.108906	1.053046
	LINEX	0.635364	0.903520	0.964635	1.553455	1.109767	1.053455
	GELF	0.891125	0.502502	0.708874	1.116361	0.3799015	0.616361
	SELF	0.635354	0.930540	0.964645	1.607957	1.227568	1.107957

**Table 4:** Showing mean of ML and Bayesian estimates with corresponding MSEs and Bias for  $\theta=2$  and  $\alpha=0.84$  (while  $p=5, q=12, m=8, c=-0.65$ )

n	Method	$\theta$			$\alpha$		
		MEAN	MSE	BIAS	MEAN	MSE	BIAS
30	ML	2.631936	3.211034	1.791936	0.767969	1.517899	1.232030
	LINEX	2.631936	3.211034	1.791936	0.738588	1.591409	1.261411
	GELF	1.764271	0.854281	0.924271	0.820932	1.390366	1.179068
	SELF	2.408270	2.459486	1.568270	0.738978	1.592766	1.261921
50	ML	2.553476	2.936002	1.713476	0.781702	1.484247	1.218297
	LINEX	2.553476	2.936002	1.713476	0.766525	1.521766	1.233474
	GELF	1.826392	0.972970	0.986392	0.841222	1.342981	1.158778
	SELF	2.543565	2.871552	1.694565	0.766439	1.522096	1.233560
100	ML	2.231311	1.935746	1.391311	0.787585	0.469048	1.212414
	LINEX	2.231311	1.935746	1.391311	0.804772	1.429149	1.195227
	GELF	1.634268	0.630865	0.794268	0.885795	1.341462	1.114206
	SELF	2.132726	1.671155	1.292726	0.829791	1.369402	1.170208
200	ML	2.170852	1.771167	1.330852	0.8445506	1.335063	1.155449
	LINEX	2.170852	1.771167	1.330852	0.829822	1.369314	1.170177
	GELF	1.676593	0.699902	0.83659	0.866981	1.283888	1.133019
	SELF	2.216419	1.894586	1.376419	0.802809	1.433578	1.197190
500	ML	1.895829	1.114774	1.055829	0.854625	1.311883	1.145374
	LINEX	1.895829	1.114774	1.055829	0.826858	1.376269	1.173141
	GELF	1.546722	0.499456	0.706722	0.887273	1.238159	1.112726
	SELF	1.956499	1.246570	1.116499	0.831958	1.364320	1.168041
1000	ML	1.958634	1.251342	1.118634	0.855785	1.309226	1.144214
	LINEX	1.958634	1.251342	1.118634	0.902505	1.206101	1.097494
	GELF	1.512338	0.452302	0.672533	0.942452	1.118742	1.057547
	SELF	1.890693	1.103958	1.050693	0.912619	1.183110	1.087380

**Table 5:** The Point Estimates of Topp-Leone Lindley distribution parameters through MLE, LINEX, GELF and SELF  $p=6, q=4, m=5, k=-0.75$

Parameters	MLE	LINEX	GELF	SELF
$\theta$	0.7128402	0.7128402	0.7768919	0.7128402
$\alpha$	6.339166	6.339166	3.995067	6.339166

## 5. CONCLUSION

In estimating the parameters of probability distribution in survival analysis, Bayesian mechanism examines the nature uncertainty and provide a judicious framework for studying such problems. In this study, we considered the Bayesian Estimation (BE) for the Topp-Leone distribution parameters. The BEs were obtained using Lindley's approximation under three different loss functions, which includes Square Error Loss Function (SELF), Linear Exponential Loss Function (LINEX) and Generalized Entropy Loss Function (GELF). Monte Carlo simulation was carried out to examine the behavior of the maximum likelihood (ML) and Bayesian Estimators, which was investigated through the mean square error (MSE) and bias of the estimators. It was also observed for any fixed value of the parameters, as sample size increases, the MSEs of the Bayesian Estimates and MLEs decrease. Also, the MLEs and Bayesian estimates converge to the same value as the sample gets larger except for GELF. Generally, it was observed that the results obtained from the MLE, SELF and LINEX are more consistent than that of GELG.

### Conflicts of Interest

The authors declared that there is no conflict of interest in this work.

## REFERENCE

- [1] Topp, C.W and Leone, F.C. (1955). A Family of J-Shaped Frequency Functions. *American Journal of Statistical Association*, 50 (269): 209-219.
- [2] Nadarajah, S. and Kotz, S. (2003). Moments of some J-shaped distributions. *Journal of Applied Statistics*, 30(3):311-317.
- [3] Ghitany, M.E., Kotz, S. and Xie, M. (2005). On some reliability measures and their stochastic orderings for the Topp-Leone distribution. *Journal of Applied Statistics*, 32: 715-722.
- [4] Zhou, M., Yang, D.W., Wang, Y. and Nadarajah, S. (2006). Some J-shaped distributions: Sums, products and ratios. *Proceedings of the Annual Reliability and Maintainability Symposium*, 175-181.
- [5] Kotz, S. and Seie, r E. (2007). Kurtosis of the Topp Leone distributions. *International Journal Statistics*, 1-15.
- [6] Nadarajah, S. (2009). Bathtub-shaped failure rate functions. *Quality and Quantity*, (43): 855-863.
- [7] Zghoul, A. A. (2011) Record values from a family of J-shaped distributions. *Statistica*, 71: 355-365.
- [8] Al-Shomrani, A., Arif, O., Shawky, K., Hanif, S. and Shahbaz, M. Q. (2016) Topp-Leone family of distributions: some properties and application. *Pakistan Journal of Statistics and Operation Research*, 12(3): 443-451.
- [9] Nzei, L. C. and Ekhosuehi, N. (2020). Topp-Lindley Distribution. *Journal of Mathematic Association of Nigeria (ABACUS)*, 47 (1): 20-34.
- [10] Legendre, A. (1805). New Method for the Dermination of Orbits of Comets. *Courcier, Paris, France*.
- [11] Gauss, C. F. (1810). Least Squares Method for the Combinations of Observation (Translated by J. Bertrand, 1955). *Mallet-Bach. Paris, France*.
- [12] Varian, H. R. (1975). A Bayesian approach to real estate assessment. *North Holland, Amsterdam*, 195 -208.
- [13] Calabria, R. and Pulcini, G. (1996). Point estimation under asymmetric loss function for left truncated exponential samples. *Communication in Statistics-Theory and Methods*, 25(3):585-600.
- [14] Obisesan, K. O., Adegoke, T. M., Adekanmbi, D. B. and Lawal M. (2015). Numerical approximation to intractable likelihood functions. *Perspectives and Developments in*

- Mathematics*, 301-324.
- [15] Bakari, H. R., Adegoke, T. M., and Yahya, A.M. (2016). Application of Newton Raphson method to non-linear models. *International Journal of Mathematics and Statistics*, 4(4):21-31.
- [16] Rastogi, M. K., and Merovci, F. (2018). Bayesian estimation for parameters and reliability characteristic of the Weibull Rayleigh distribution. *Journal of King Saud University of Science*, 30:472–478
- [17] Sangeeta, A., Kalpana, K. M. and Ritu, K. (2019). Bayes estimators for the reliability and hazard rate functions of Topp-Leone distribution using Type-II censored data. *Communication in Statistics-Simulation and Computation*, DOI: 10.1080/03610918.2019.1602646
- [18] Kaur, K., Arora, S. and Mahajan, K. K. (2015). Bayesian Estimation of Inequality and Poverty Indices in Case of Pareto Distribution Using Different Priors under LINEX Loss Function. *Advances in Statistics. Article ID 964824*. DOI:10.1155/2015/964824.
- [19] Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *American Journal of Statistical Association*, 81(394):446–51. DOI:10.1080/01621459.1986.10478289.
- [20] Lindley, D. V. (1980). Approximate Bayesian Method. *Trabajos de Estadística y de Investigación Operativa*, 31:223–45.
- [21] Hummara, S. and Ahmad, S. P. (2015). Bayesian Approximation Techniques for Kumaraswamy Distribution. *Mathematical Theory and Modeling*, 5(5): 49 – 60
- [22] Adegoke, T. M., Yahya, W. B. and Adegoke, G. K. (2018). Inverted generalized exponential. *Annals of Statistical Theory and Application*, 1: 1-10.
- [23] Adegoke, T. M., Nasiri, P., Yahya, W. B., Adegoke, G. K., Afolayan, R. B. and Yahaya, A. M. (2019). Bayesian estimation of Kumaraswamy distribution under different loss functions. *Annals of Statistical Theory and Application* 2:90-102.
- [24] Kamran, A., Zamir, H., Noreen, R., Amjad, A., Muhammad, T., Sajjad. A. K., Sadaf, M., Umair, K., and Dost, M. K. (2020). Bayesian Estimation of Gumbel Type-II Distribution under Type-II Censoring with Medical Applications. *Hindawi-Computational and Mathematical Methods in Medicine*, ID 1876073, <https://DOI.org/10.1155/2020/1876073>
- [25] Sule, B. O., Adegoke, T. M. and Uthman, K. T. (2021). Bayes Estimators of Exponentiated Inverse Rayleigh Distribution using Lindleys Approximation. *Asian Research Journal of Mathematics*, 17(2): 60-71.
- [26] Nichols, M. D and Padgett, W. J. (2006). A bootstrap control chart for Weibull percentiles. *Quality and Reliability Engineering International*, 22(2): 141-151.



# SINE-WEIBULL DISTRIBUTION: MATHEMATICAL PROPERTIES AND APPLICATION TO REAL DATASETS

<sup>1</sup>Muhammad Umar Faruk, <sup>2</sup>Alhaji Modu Isa, <sup>3</sup>Aishatu Kaigama

•

<sup>1,2,3</sup>Department of Mathematics and Computer Science, Borno State University, Nigeria  
alhajimoduisa@gmail.com

## Abstract

*New parameters can be added to expand families of distribution for greater flexibility or to construct covariate models in several ways. In this study, a trigonometric-type distribution called Sine-Weibull distribution was developed by adopting the Weibull distribution as the baseline distribution and Sine-G Family as the generator to generate a flexible probability distribution without the need for extra parameters. The moment, moment generating function, entropy, and order statistics are some of the mathematical aspects of this distribution that were derived. The Maximum Likelihood approach was used to estimate the new distribution's parameters. Using actual datasets, the Sine-Weibull distribution's applicability was demonstrated.*

**Keywords:** Sine-G Family, Weibull Distribution, Probability Distribution, Maximum Likelihood Estimator

## I. Introduction

Distribution functions, their properties and interrelationships play a significant role in modeling naturally occurring phenomena. For this reason, a large number of distribution functions, which were found applicable to many events in real life, have been proposed and defined in literature. Various methods exist in defining statistical distributions. Many of these arose from the need to model naturally occurring events. For example, the Normal distribution addresses real-valued variables that tend to cluster at a single mean value, while the Poisson distribution models discrete rare events. Yet few other distributions are functions of one or more distributions.

To explain real world phenomena, statistical distributions are widely applied. Their theory is widely studied due to the utility of statistical distributions, and new distributions are developed. In the field of probability theory and statistics, the search for creating a more effective and scalable distribution of probability remains high [1]. Numerous standard distributions have been extensively used over the past decades for modeling data in several fields such as Engineering, Economics, Finance, Biological, Environmental and Medical Sciences etc. However, generalizing these standard distributions has produced several compound distributions that are more flexible compared to the baseline distributions. For this reason, several methods for generating new families of distributions have been studied.

Weibull distribution is a continuous probability distribution. It is one of different distributions used to describe particle size with major application in survival analysis, weather

forecast and reliability engineering. The Weibull distribution is a continuous probability distribution. It was named after Swedish mathematician Waloddi Weibull, who describe it in detail in 1951, although it was first recognized by [2] and first applied by [3] to describe a unit size of distribution. Weibull distribution exist with scale and shape parameters. This distribution has become very popular in analyzing lifetime data and for many applications where a skewed distribution is required. Inducing of a new shape parameter(s) introduces a model into greater family of distributions and can give significantly skewed and heavy-tailed distributions and also provides greater flexibility in the form of new distribution.

Even when there is uncertainty about the future in real life, decisions still need to be taken. Thus, uncertainty issues must be dealt with by decision-making processes. Probability is one of the frequently employed strategies for addressing uncertainty in planning and management. In order to create a family of hybrid distributions that are more effective than their parent distributions, many researchers have focused on the idea of combining two or more probability distributions. By adding one or more parameters, these distributions become more flexible and can track a variety of random phenomena that are difficult to model using their parent distributions. The laws of generality, which state that when a particular distribution has more than four parameters, it undermines the performance of the model, can sometimes be breached by such compounding or extended distributions.

Many researchers have come up with new families of trigonometric in recent times. Some of these families include: exponentiated sine-generated family of distributions by [4], Sin-G class of distributions by [5], Sec-G Class by [7], Sine Square distribution by [8], Sine Inverse Lomax Generated Family by [9], Sine Burr XII by [10], Sine Kumaraswamy-G family of distributions by [11], Sine Topp-Leone family by [12], Sine-Exponential Distribution by [13] and Sine Power Lomax distribution by [14] (2021).

The quest for developing more efficient and flexible probability distribution remains strong in the field of probability theory and statistics. However, there is no single probability distribution that is suitable for different data sets. Therefore, there is a need to come up with their extended forms to give substitutive adaptable models or as to form a better representation of the data. Thus, this has triggered the need to extend the existing classical Weibull distributions. Therefore, this gives a gap of coming up with a distribution (Sine-Weibull Distribution) capable of handling a dataset that behaved negatively or positively skewed. Hence, this research is aimed at developing a new probability distribution function called Sine-Weibull Distribution.

## II. Methods

### 2.1 The Weibull Distribution

A continuous random variable  $X$  is said to have followed a Weibull distribution if its cdf is expressed as;

$$H(x, k, \lambda) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k}, \quad x > 0 \tag{1}$$

and the pdf is also expressed as;

$$h(x, k, \lambda) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{(k-1)} e^{-\left(\frac{x}{\lambda}\right)^k} \quad x > 0 \tag{2}$$

### 2.2 Sine G Family of Probability Distribution

Let  $H(x)$  be the cumulative distribution function (cdf) of a univariate continuous distribution and  $h(x)$  be the corresponding probability density function (pdf), then, the Sine-G family of probability distribution according to [5] Kumar *et al.*, (2015) is given by:

$$F(x, \xi) = \int_0^{\frac{\pi}{2}H(x, \xi)} \cos t \, dt = \sin \left\{ \frac{\pi}{2}H(x, \xi) \right\} \quad (3)$$

and its corresponding pdf is given by:

$$f(x, \xi) = \frac{\pi}{2}h(x, \xi) \cos \left\{ \frac{\pi}{2}H(x, \xi) \right\} \quad (4)$$

where  $H(x, \xi)$  and  $h(x, \xi)$  are the cdf and the pdf of any baseline distribution with vector parameter  $\xi$ .

### 2.3 The New Sine Weibull Distribution

The pdf and cdf of the new sine Weibull distribution are given in equation (5) and (6):

$$f(x, k, \lambda) = \frac{\pi}{2} \left( \frac{x}{\lambda} \right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \cos \left\{ \frac{\pi}{2} \left[ 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \right] \right\} \quad (5)$$

And

$$F(x, k, \lambda) = \sin \left\{ \frac{\pi}{2} \left[ 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \right] \right\} \quad (6)$$

The survival function  $S(x)$ , hazard function  $h(x)$ , reverse hazard function  $r(x)$  and the quantile function  $Q(u)$  are given below:

$$S(x) = 1 - F(x) = 1 - \sin \left\{ \frac{\pi}{2} \left[ 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \right] \right\} \quad (7)$$

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{\frac{\pi}{2} \left( \frac{x}{\lambda} \right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \cos \left\{ \frac{\pi}{2} \left[ 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \right] \right\}}{1 - \sin \left\{ \frac{\pi}{2} \left[ 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \right] \right\}} \quad (8)$$

$$r(x) = \frac{f(x)}{F(x)} = \frac{\pi}{2} \left( \frac{x}{\lambda} \right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \cot \left\{ \frac{\pi}{2} \left[ 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \right] \right\} \quad (9)$$

$$Q(U) = F^{-1} \lambda \left\{ -\log \left( 1 - \frac{2 \sin^{-1} U}{\pi} \right) \right\}^{\frac{1}{k}} \quad (10)$$

### 2.4. Parameter Estimation

The parameters of the newly developed Sine-Weibull distribution will be estimated using the method of maximum likelihood (MLE). Moment and moment generating function (mgf) will be used in determine the mean, variance, skewness and kurtosis, among other properties, of the proposed distribution.

#### 2.4.1 Method of Maximum Likelihood

Let  $Y_1, Y_2, \dots, Y_n$  independent, identically distributed (*iid*) random sample of a random variable  $Y$  with *pdf* given by  $f(y/\delta)$ , then the likelihood function  $L(\delta: y)$  of  $Y_1, Y_2, \dots, Y_n$  is the joint density function when regarded as a function of the parameter. That is

$$L(\delta: y) = \prod_{i=1}^n f(y_i, \delta)$$

It is more convenient to use the log likelihood.

$$l(\delta: y) = \ln L(\delta, y)$$

The estimate of the parameter can be obtained by taking the derivative of the log likelihood function with respect to the parameter and equating to zero, that is

$$\frac{\partial y}{\partial \delta} \ln L(\delta, y) = 0 \tag{11}$$

### 2.4.2 Maximum Likelihood of Sine-Weibull Distribution

Let  $X_1, X_2, \dots, X_n$  be a random sample of size  $n$  from a Sine-Weibull distribution with a *pdf* given by (1.1), the likelihood function  $L(\lambda: x)$  of this sample is given as

$$L(\lambda: x) = \prod_{i=1}^n f(x_i, \lambda) = \prod_{i=1}^n \frac{\pi}{2} \left(\frac{x_i}{\lambda}\right)^{k-1} e^{-\left(\frac{x_i}{\lambda}\right)^k} \cos \left\{ \frac{\pi}{2} \left[ 1 - e^{-\left(\frac{x_i}{\lambda}\right)^k} \right] \right\}$$

$$L(\lambda: x) = \left(\frac{\pi}{2}\right)^n \sum_{i=1}^n \left(\frac{x_i}{\lambda}\right)^{k-1} \cos \left\{ \frac{\pi}{2} \left[ 1 - e^{-\left(\frac{x_i}{\lambda}\right)^k} \right] \right\} e^{-\sum_{i=1}^n \left(\frac{x_i}{\lambda}\right)^k}$$

Taking the log of the likelihood function gives

$$l(\lambda, x) = \ln \left( \left(\frac{\pi}{2}\right)^n \sum_{i=1}^n \left(\frac{x_i}{\lambda}\right)^{k-1} \cos \left\{ \frac{\pi}{2} \left[ 1 - e^{-\left(\frac{x_i}{\lambda}\right)^k} \right] \right\} e^{-\frac{1}{2} \sum_{i=1}^n \left(\frac{x_i}{\lambda}\right)^k} \right)$$

$$l(\lambda, x) = n \ln \left(\frac{\pi}{2}\right) + (k-1) \ln \sum_{i=1}^n \left(\frac{x_i}{\lambda}\right) + \ln \sum_{i=1}^n \cos \left\{ \frac{\pi}{2} \left[ 1 - e^{-\left(\frac{x_i}{\lambda}\right)^k} \right] \right\} - \sum_{i=1}^n \left(\frac{x_i}{\lambda}\right)^k$$

$$\cos \left\{ \frac{\pi}{2} \left[ 1 - e^{-\left(\frac{x_i}{\lambda}\right)^k} \right] \right\} = 0, \quad \text{because} \quad \cos \left(\frac{\pi}{2}\right) = 0$$

To maximize equation(11), we take the derivative with respect to  $\lambda$  and equate to zero

$$\frac{\partial l}{\partial \lambda} = -(k-1) \sum_{i=1}^n \left(\frac{1}{\lambda}\right) - \sum_{i=1}^n \left(\frac{-kx^k}{\lambda^{k+1}}\right) = 0$$

$$\frac{\partial l}{\partial \lambda} = -(k-1) \sum_{i=1}^n \left(\frac{1}{\lambda}\right) + k \sum_{i=1}^n \left(\frac{x^k}{\lambda^{k+1}}\right) = 0$$

$$k \sum_{i=1}^n \left(\frac{x^k}{\lambda^{k+1}}\right) = (k-1) \sum_{i=1}^n \left(\frac{1}{\lambda}\right)$$

$$k \sum_{i=1}^n x^k = (k-1) \sum_{i=1}^n \lambda^{-1} \lambda^{k+1}$$

$$\sum_{i=1}^n \lambda^k = \frac{k}{(k-1)} \sum_{i=1}^n x^k$$

$$\lambda^k = \frac{1}{n} \frac{k}{(k-1)} \sum_{i=1}^n x^k$$

$$\hat{\lambda} = \sqrt[k]{\frac{1}{n} \frac{k}{(k-1)} \sum_{i=1}^n x^k} \tag{12}$$

Equation (12) gives the maximum likelihood estimator of the parameter  $\lambda$

## 2.5. Some Mathematical Properties

### 2.5.1 Moment

Moments plays a vital role in the field of statistical analysis, particularly when it comes to real applications. Suppose that  $X$  is a random variable and  $r$  is a non-negative integer, the  $r^{\text{th}}$  moment of  $X$  is the quantity  $E(X^k)$  provided its expectation exists. The  $r^{\text{th}}$  is given by:

$$E(x^r) = \int_{x=0}^{\infty} x^r f(x) dx$$

The  $r^{\text{th}}$  moment of proposed Sine-Weibull distribution is derived as follows:

$$E(x^r) = \frac{\pi}{2} \int_{x=0}^{\infty} x^r \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \cos\left\{\frac{\pi}{2}\left[1 - e^{-\left(\frac{x}{\lambda}\right)^k}\right]\right\} dx$$

$$E(x^r) = \frac{\pi}{2\lambda^{k-1}} \int_{x=0}^{\infty} (x)^{k-r+1} e^{-\left(\frac{x}{\lambda}\right)^k} \cos\left\{\frac{\pi}{2}\left[1 - e^{-\left(\frac{x}{\lambda}\right)^k}\right]\right\} dx$$

$$E(x^r) = \frac{\pi}{2\lambda^{k-1}} (\lambda)^{k-1+r} \int_{x=0}^{\infty} \left(\frac{x}{\lambda}\right)^{k-1+r} e^{-\left(\frac{x}{\lambda}\right)^k} \cos\left\{\frac{\pi}{2}\left[1 - e^{-\left(\frac{x}{\lambda}\right)^k}\right]\right\} dx$$

$$E(x^r) = \int_{x=0}^{\infty} \left(\frac{x}{\lambda}\right)^{k-1+r} e^{-\left(\frac{x}{\lambda}\right)^k} \cos\left\{\frac{\pi}{2}\left[1 - e^{-\left(\frac{x}{\lambda}\right)^k}\right]\right\} dx = 1 \text{ (it is a pdf)}$$

$$\Rightarrow E(x^r) = \frac{\pi}{2\lambda^{k-1}} (\lambda)^{k-1+r}$$

$$E(x^r) = \frac{\pi}{2} \lambda^r \tag{13}$$

The first and second moments (when  $r = 1$  and  $r = 2$ ) are therefore given below,

$$E(x) = \frac{\pi}{2} \lambda \tag{14}$$

$$E(x^2) = \frac{\pi}{2} \lambda^2 \tag{15}$$

The variance is given below

$$V(x) = E(x^2) - [E(x)]^2 = \frac{\pi}{2} \lambda^2 - \left(\frac{\pi}{2} \lambda\right)^2$$

$$V(x) = \frac{\pi}{2} \lambda^2 \left(1 - \frac{\pi}{2}\right) \tag{16}$$

$$\text{Standard Deviation (S)} = \sqrt{\frac{\pi}{2} \lambda^2 \left(1 - \frac{\pi}{2}\right)} \tag{17}$$

### 2.5.2 Skewness and Kurtosis of the Sine-Weibull Distributions

The skewness and kurtosis of the sine-Weibull distribution are obtained using the third and fourth moment respectively with the power of the standard deviation of the distribution. These approaches are the measure of kurtosis ( $\alpha_3$ ) and skewness ( $\alpha_4$ ) based on moments

$$(\alpha_3) = \frac{E(x^3)}{S^3}$$

$$(\alpha_3) = \frac{\frac{\pi}{2} \lambda^3}{\left(\sqrt{\frac{\pi}{2} \lambda^2 \left(1 - \frac{\pi}{2}\right)}\right)^3} \tag{18}$$

$$(\alpha_4) = \frac{E(x^4)}{S^4} = \frac{\frac{\pi}{2}\lambda^4}{\left(\sqrt{\frac{\pi}{2}\lambda^2\left(1-\frac{\pi}{2}\right)}\right)^4} = \frac{\frac{\pi}{2}\lambda^4}{\left(\frac{\pi}{2}\lambda^2\left(1-\frac{\pi}{2}\right)\right)^2}$$

$$(\alpha_4) = \frac{1}{\frac{\pi}{2}\left(1-\frac{\pi}{2}\right)^2} \tag{19}$$

### 2.5.3 Entropy

The entropy of is a measure of variation of the uncertainty. There are many entropy measures studied and discussed in literature but the Renyi entropy is perhaps one of the most popular. Renyi entropy of with proposed density function is given by

$$i_{R(\rho)} = \frac{1}{1-\rho} \log \left( \int_0^\infty f(x)^\rho dx \right) \tag{20}$$

where  $\rho > 0$  and  $\rho \neq 0$ . Inserting equation (4) into (20)

$$i_{R(\rho)} = \frac{1}{1-\rho} \log \left( \int_0^\infty \left( \frac{\pi}{2} \left( \frac{x}{\lambda} \right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \cos \left\{ \frac{\pi}{2} \left[ 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \right] \right\} \right)^\rho dx \right) \tag{21}$$

### 2.5.4 Order Statistics

Suppose that  $x_1, x_2, \dots, x_n$  are random samples of size n from probability distribution with pdf  $f(x)$  and cdf  $F(x)$  as defined in (3) and (4) respectively, the  $p^{\text{th}}$  order statistic can be expressed

$$f_n(x) = \frac{n! f(x)}{(p-1)!(n-1)!} F(x)^{p-1} [1-F(x)]^{n-p} \tag{22}$$

The order statistics of the proposed Sine-Weibull distribution is given by:

$$f_n(x) = \frac{n! \left( \frac{\pi}{2} \left( \frac{x}{\lambda} \right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \cos \left\{ \frac{\pi}{2} \left[ 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \right] \right\} \right)^\rho}{(p-1)!(n-1)!} \left\{ \sin \left[ \frac{\pi}{2} \left( 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \right) \right] \right\}^{p-1}$$

$$\times \left\{ 1 - \sin \left[ \frac{\pi}{2} \left( 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \right) \right] \right\}^{n-p} \tag{23}$$

## III. Results

### 3.1 Application

Specifically, AIC is aimed to obtain the best approximating model to the unknown true data generating process. Superficially, BIC differs from AIC only in the first term which depends on sample size n. Models that minimize the BIC are selected. From a Bayesian perspective, BIC is designed to find the most probable model given the data.

### 3.1.1 Dataset

One dataset was considered for illustrative purposes and comparison with the baseline distribution and other competitors. The comparison was done with Weibull distribution and Lomax distribution. We estimated the unknown parameters of the distribution by the maximum-likelihood method. We obtain the values of the Akaike information criterion (AIC), Bayesian information criterion (BIC) and consistent Akaike information criterion (CAIC) for the newly developed distribution as well as the competitors. The dataset consists of thirty successive values of March precipitation (in inches) in Minneapolis/St [16]. The data are as follows:

0.77    1.74    0.81    1.2    1.95    1.2    0.47    1.43    3.37    2.2    3.0    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    2.05

**Table 1:** Summary Statistics of the dataset

Data	Minimum	$Q_1$	Media	Mean	$Q_3$	Maximu
Dataset	0.92	1.302	1.544	1.658	1.814	5.306

Table 1 gives the summary statistics of the data sets such as the mean, the median, the first and third quartile, the minimum and the maximum values.

**Table 2:** MLE, AIC, CAIC, BIC, and HQIC of the data set

Data Set	MLE	AIC	CAIC	BIC	HQIC
Sine-Weibull	55.61173	115.2235	115.3472	120.4388	117.3322
Weibull	150.5514	305.1029	305.2266	310.3132	307.2716
Lomax	150.5514	303.1029	303.1437	310.3132	304.1572

Table 2 presents the results of the analysis of the dataset. The result of the analysis of the Sine-Weibull Distribution was compared with Weibull Distribution and Lomax Distribution to test the efficiency of the model. The proposed Sine-Weibull distribution has proven to be the better model because it has the least AIC, CAIC, BIC and HQIC.

## IV. Discussion

There has been a growing interest among statisticians and applied researchers in developing flexible lifetime models for the betterment of modelling survival data. In this paper, we introduced a two-parameter Sine-Weibull distribution which is obtained by considering a Weibull distribution as the baseline. We study some of its statistical and mathematical properties. Maximum Likelihood Estimation was used in parameter estimation. The usefulness of the new distribution was illustrated via the analysis of real data sets. We hope that the proposed extended model will attract wider applications.

## References

- [1] Alzaatreh, A., Lee, C., & Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, 71(1), 63-79.
- [2] Frechet, M. (1927). Sur la loi de probabilit'e de l'ecart maximum. *Ann. de la Soc. polonaisede Math*, 6, 93–116.
- [3] Rosin, P. and Rammler, E. (1933) The Laws Governing the Fineness of powdered coal. *Journal of the Institute of Fuel*, 7, 29-36
- [4] Muhammad, M., Alshanbari, H.M. Alanzi, A.R.A., Liu, L., Sami, W., Chesneau, C., Jamal, F. A. (2021). New Generator of Probability Models: The Exponentiated Sine-G Family for Lifetime Studies. *Entropy*, 23, 1394.
- [5] Kumar, D., Singh, U., & Singh, S. K. (2015). A new distribution using sine function-its application to bladder cancer patients' data. *Journal of Statistics Applications & Probability*, 4(3), 417.
- [6] Souza, L., Junior, W. R. O., de Brito, C. C. R., Chesneau, C., Ferreira, T. A. E., Soares, L. (2019). General properties for the Cos-G class of distributions with applications. *Eurasian Bulletin of Mathematics*, 2(2), 63–79.
- [7] Souza, L., de Oliveira, W. R., de Brito, C. C. R., Chesneau, C., Fernandes, R., & Ferreira, T. A. (2022). Sec-G class of distributions: Properties and applications. *Symmetry*, 14(2), 299.
- [8] Al-Faris, R. Q., & Khan, S. (2008). Sine square distribution: a new statistical model based on the sine function. *Journal of Applied Probability and Statistics*, 3(1), 163-173.
- [9] Fayomi, A., Algarni, A., & Almarashi, A. M. (2021). Sine Inverse Lomax Generated Family of Distributions with Applications. *Mathematical Problems in Engineering*, 1-11.
- [10] Isa, A. M., Ali, B. A., & Zannah, U. (2022). Sine Burr XII Distribution: Properties and Application to Real Data Sets. *Arid Zone Journal of Basic and Applied Sciences*, 1(3), 48-58.
- [11] Chesneau, C., & Jamal, F. (2020). The sine Kumaraswamy-G family of distributions. *Journal of Mathematical Extension*, 15.
- [12] Al-Babtain, A. A., Elbatal, I., Chesneau, C., & Elgarhy, M. (2020). Sine Topp-Leone-G family of distributions: Theory and applications. *Open Physics*, 18(1), 574-593.
- [13] Isa, A. M., Bashiru, S. O., Ali, B. A., Adepoju, A. A., & Itopa, I. I. (2022). Sine-Exponential Distribution: Its Mathematical Properties and Application to Real Dataset. *UMYU Scientifica*, 1(1), 127-131.
- [14] Nagarjuna, V. B. V., Vardhan, R.V. and Chesneau, C. (2021). On the Accuracy of the Sine Power Lomax Model for Data Fitting. *Modelling*, 2, 78–104.
- [15] Hinkley, D. (1977). On quick choice of power transformation. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 26(1), 67-69.



# RECENT DEVELOPMENTS IN THE COMPUTATION OF THE ROCOF OF MULTI-STATE SYSTEMS AND ITS APPLICATIONS

Guglielmo D'Amico<sup>1</sup> and Fulvio Gismondi<sup>2</sup>

•

<sup>1</sup>Department of Economic Studies, University G. d'Annunzio of Chieti-Pescara, Italy

[g.damico@unich.it](mailto:g.damico@unich.it)

<sup>2</sup>Department of Economic and Business Science, "Guglielmo Marconi" University, Italy

[f.gismondi@unimarconi.it](mailto:f.gismondi@unimarconi.it)

## Abstract

*This paper reviews several theoretical works on the computation of the Rate of Occurrence of Failure (ROCOF) for general multi-state random systems, focusing on recent generalizations. The discussion begins by defining the ROCOF for a Markov process and discussing the main results achieved in the literature, then moves towards the richer framework represented by semi-Markov systems. The paper discusses complications that arise when extending the ROCOF to higher orders so that a measure of the association of failures in time can be obtained. The work then analyzes possible modifications in terms of a conditional version of the ROCOF, which is of special interest in applications. The findings are illustrated by a numerical example from reliability, and the broad applicability is demonstrated by a discussion of different applications in other domains.*

**Keywords:** Markov processes, semi-Markov processes, reliability, applications

## 1. Introduction

Several studies deal with the proposal of new measures of performance for a random system and their computation in applied problems; see e.g. [1]. Among the available indicators, the Rate of Occurrence of Failure (ROCOF) is one of the most frequently used in understanding a system's performance over time. Once a system's failure is defined, the ROCOF is the derivative of the expected number of failures with respect to the time variable. Systems with an increasing path of the ROCOF are expected to deteriorate as time goes on. Contrarily, if the ROCOF shows a decreasing shape, then the system is expected to improve its quality of functioning over time. This seems to be a simple concept, at least in its intuition, but computation and analysis pose relevant questions that have been solved at different moments during the last half century.

In the seventies, some research articles dealing with systems that have already reached the steady state appeared in the literature and showed how to compute the frequency of system failures and their durations [2,3,4].

A few years later, Shi Ding-hua [5] developed a new method for calculating the ROCOF of a system described by a finite-state continuous time-homogeneous Markov chain. The author also discussed the case of a special high-dimensional Markov process equipped with supplementary variables. In the middle of the 1990s, further contributions were given in a couple of research articles by Yeh Lam [6,7]. In the first of his contributions, the author considers a system described by a continuous-time

Markov chain of higher dimension after having introduced additional variables. A formula for evaluating the ROCOF was derived, and an application for a two-component parallel system was presented. In his second contribution, the author enlarged the stage by considering Markovian systems with a denumerable state space.

Markov processes are very frequently adopted for system reliability analysis. Unfortunately, in many circumstances, they are not suitable, either for practical or theoretical reasons. Hence, the need to use more general models, and the semi-Markov ones are a valuable alternative; see e.g. [8,9].

The question of how to compute the ROCOF for a semi-Markov process attracted the attention of Ouhbi, and Limnios [10]. In that work, the authors derived a formula for evaluating the ROCOF for semi-Markov systems and also proposed a statistical estimator of this indicator. A Similar analysis was executed by Georgiadis et al. [11] for the semi-Markov chain, i.e., for semi-Markov models in discrete time.

A new idea was advanced by D'Amico [12] with the concept of ROCOF of order  $n$  (shortly denoted by  $n$ -ROCOF) for Markov processes. This new indicator coincides with the ROCOF when  $n = 1$  and expresses a measure of clustering in the time of failure events. For example, for  $n = 2$  it expresses a measure of association in time of a couple of failures at any couple of times  $(t_1, t_2)$  with  $t_1 < t_2$ . After having defined the  $n$ -ROCOF, the author derived an explicit formula in terms of the matrix generator and initial probability distribution. Next, a nonparametric estimator was advanced, and its asymptotic properties were determined. The  $n$ -ROCOF was applied to the modeling of financial credit ratings, where a conditional version of it was shown to be particularly useful. A few years later, Votsi [13] exploited a conditional version of the ROCOF for semi-Markov chains.

Finally, in [14], the analysis of the  $n$ -ROCOF for semi-Markov processes with finite state space is executed in such a way that the previous quoted articles were generalized. The authors determined an explicit formula for the  $n$ -ROCOF under a general random starting mechanism, considering any possible state and duration of permanence in it. This was done using a mixed continuous-discrete initial probability distribution function. A set of hypotheses on the model parameters was advanced so that the derivation of an explicit formula expressing the  $n$ -ROCOF was obtained. The results were sufficiently general to be of interest not only in the reliability theory field but for every general system and could be applied every time that a partition of the state space can be introduced into working and not-working states.

After giving a thorough analysis of the theoretical findings pertaining to the computation of the ROCOF, we show a numerical example from the dependability area and go on to describe some unusual applications in various fields, spanning from financial mathematics to wind energy generation. Considering that several of these applications have never been mentioned before, they also offer a suggestion for detailed future research.

The paper proceeds as follows: Section 2 describes the basic reliability problem and some of the most important reliability indicators, ROCOF and  $n$ -ROCOF included. Section 3 considers a Markov process for the probabilistic description of the system and shows formulas for the ROCOF and  $n$ -ROCOF. Section 4 provides a summary of the results related to the extension to the semi-Markov framework, showing the latest results in the literature. Section 5 discusses a numerical example for a Markov system and demonstrates the practical usefulness of the considered measures. Moreover, different possible applications from real life problems are detailed. The discussion concludes in Section 6, which reviews the content of the paper and provides general conclusions.

## 2. Basic description of the reliability problem and main indicators

The basic reliability problem can be described assuming that the performance of the system can be identified with one element of a finite set  $E = \{1, 2, \dots, m\}$  called state space. Frequently, an ordering relation on the set  $E$  is considered in such a way that higher ranks  $j \in E$  correspond to a higher system's performance. The state space  $E$  is partitioned into two disjoint subsets  $W$  and  $F$  such that:

$$E = W \cup F, \quad W \cap F = \emptyset, \quad W \neq \emptyset, \quad F \neq \emptyset.$$

The subset  $W$  contains all the elements of  $E$  denoting acceptable working levels of the systems; instead the subset  $F$  contains all the states of  $E$  in which the system is not performing in a satisfactory way or has a fault. Sometimes the state space is divided into three subsets, denoting the working states, the changeable states, and the failure states. The changeable states denote a working system if and only if, before entering the changeable subset, the system was working and will continue to work after leaving it; see [15].

The system evolves in time and changes its state migrating from one state  $i$  to another state  $j$ . The most natural way to study this evolution is by using a stochastic process  $Z = \{Z(t), t \geq 0\}$ . Hence,  $Z(t)$  denotes the state occupied by the system at time  $t$  and if  $Z(t) \in W$  the system is working while if  $Z(t) \in F$  the system is not working.

Specific indicators are used to measure the overall quality of the system; among them, we remember:

- the *availability function*, which is defined by

$$A_i(t) := P[Z(t) \in W | Z(0) = i].$$

It expresses the probability that the system ranked  $i$  at time 0 will be operational at time  $t$  independently of the possible behavior before this time.

- The *reliability function* which is defined by

$$R_i(t) := P[Z(n) \in W, \forall n \in [0, t] | Z(0) = i].$$

This indicator consists of the probability that a system ranked  $i$  at time 0 will not experience a fault (a visit to the subset  $F$ ) from time 0 up to time  $t$ . A generalization of the reliability function considers interval reliability indicators [16], recent results are available in [17,18], and the sequential reliability function [19].

Denote by  $N_f(t)$  the number of failures of the system until time  $t$ , i.e. the number of passages from a state of  $W$  to one of  $F$ . Then,

- the *ROCOF* at time  $t$  for a random system, denoted by  $ro(t)$ , is defined by

$$ro(t) = \lim_{\Delta t \rightarrow 0} \frac{E[N_f(t + \Delta t) - N_f(t)]}{\Delta t}. \quad (1)$$

The ROCOF gives information on whether there are a lot of failures or only a few within a time, and it has a simple probabilistic interpretation that for deteriorating systems shows an increasing behavior and for improving systems, it is decreasing in time.

When studying the reliability of a repairable system, it is of great interest to also measure the relative positioning of tuples of failures. For this reason, the  $n$ -ROCOF was defined D'Amico [12]:

- the  $n$ -ROCOF at times  $\mathbf{t}_1^n = (t_1, t_2, \dots, t_n)$  with  $t_i < t_{i+1}$  for a random system is defined by

$$ro(\mathbf{t}_1^n) = \lim_{(\Delta t_i \rightarrow 0)_{i=1}^n} \frac{E[dN_f(t_1) \cdot dN_f(t_2) \cdot \dots \cdot dN_f(t_n)]}{\Delta t_1 \Delta t_2 \cdot \dots \cdot \Delta t_n}. \quad (2)$$

Clearly, for  $n = 1$  the  $n$ -ROCOF coincides with the ROCOF of the system. For  $n = 2$  we obtain an interesting particular case called the 2-ROCOF which is given by

$$ro(\mathbf{t}_1^2) = \lim_{(\Delta t_i \rightarrow 0)_{i=1}^2} \frac{E[dN_f(t_1)dN_f(t_2)]}{\Delta t_1 \Delta t_2}, \quad (3)$$

and expresses a measure of association of failure events in correspondence of a 2-dimensional vector of times  $\mathbf{t}_1^2 = (t_1, t_2)$ .

### 3. $n$ -ROCOF for Markov processes

A class of models frequently used in the reliability field is that represented by Markov processes; see e.g. [20,21,22,23]. Here we briefly introduce them and show the formula for the  $n$ -ROCOF.

Let consider a continuous time Markov process  $(Z(t), t \in \mathbb{R})$  with a finite state space  $E = \{1, 2, \dots, m\}$  and generator matrix  $\mathbf{Q} = (q_{ij}), i, j \in E$  where  $q_{ij} \geq 0, \forall i \neq j$  and  $q_{ii} = -\sum_{j \neq i} q_{ij}$ . Consider also an initial probability distribution over the states of the process at time zero denoted by the vector  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$  where

$$\alpha_i = P(Z(0) = i).$$

Let  $p_i(t) = P(Z(t) = i), \forall i \in E$  be the state probability at time  $t$ , then it results that

$$p_j(t) = \sum_{i \in E} \alpha_i p_{ij}(t),$$

where  $p_{ij}(t) = P(Z(t) = j | Z(0) = i) = (e^{\mathbf{Q}t})_{ij}$ .

**Theorem** [12] The  $n$ -ROCOF at times  $\mathbf{t}_1^n = (t_1, t_2, \dots, t_n)$  with  $t_i < t_{i+1}$  for a Markov jump process  $(Z(t), t \in \mathbb{R})$  over a finite state space  $E = \{1, 2, \dots, m\}$  and generator matrix  $\mathbf{Q} = (q_{ij}), i, j \in E$  is given by

$$ro(\mathbf{t}_1^n) = \sum_{w, f} \prod_{i=1}^n \alpha_{f_0} \cdot (e^{\mathbf{Q}(t_i - t_{i-1})})_{f_{i-1}w_i} \cdot q_{w_i f_i}, \quad (4)$$

where  $t_0 = 0$  and the symbol  $\sum_{w, f}^n$  is an abbreviate notation for  $\sum_{f_0 \in E} \sum_{w_i \in W, \forall i=1, \dots, n} \sum_{f_i \in F, \forall i=1, \dots, n}$ .

Formula (4) contains interesting particular cases of which we give representation. The ROCOF is simply obtained by setting  $n = 1$  in formula (4) with  $t_0 = 0$ . The result is:

$$\begin{aligned} ro(t_1) &= \sum_{f_0 \in E} \sum_{w_1 \in W} \sum_{f_1 \in F} \alpha_{f_0} \cdot (e^{\mathbf{Q}t_1})_{f_0 w_1} \cdot q_{w_1 f_1} \\ &= \sum_{w_1 \in W} \sum_{f_1 \in F} p_{w_1}(t_1) \cdot q_{w_1 f_1}, \end{aligned} \quad (5)$$

which expresses exactly the formula established by Yeh [7].

Another interesting case is represented by the 2-ROCOF which is simply obtained by setting  $n = 2$  in formula (4) with  $t_0 = 0$ . The result is:

$$\begin{aligned}
 ro(\mathbf{t}_1^2) &= \sum_{w,f} \prod_{i=1}^2 \alpha_{f_0} \cdot (e^{\mathbf{Q}(t_i-t_{i-1})})_{f_{i-1}w_i} \cdot q_{w_i f_i} \\
 &= \sum_{f_0 \in E} \sum_{w_1 \in W} \sum_{f_1 \in F} \sum_{w_2 \in W} \sum_{f_2 \in F} \alpha_{f_0} \cdot (e^{\mathbf{Q}t_1})_{f_0 w_1} \cdot q_{w_1 f_1} \cdot (e^{\mathbf{Q}(t_2-t_1)})_{f_1 w_2} \cdot q_{w_2 f_2} \cdot
 \end{aligned} \tag{6}$$

#### 4. $n$ -ROCOF for semi-Markov processes

Semi-Markov processes are a generalization of Markov processes, allowing for any kind of probability distribution function for the sojourn time in the state of the system. Contrarily, Markov processes require exponentially distributed sojourn times; this assumption can be inadequate in several application fields, reliability theory included; see e.g. [24,25].

The definition of a semi-Markov process needs some preliminary concepts to be introduced. We start by considering a bivariate random sequence  $(J_n, T_n), n \in \mathbb{N}$ . The random variable  $J_n$  denotes the state of the system at its  $n$ -th transition; this variable assumes values in the state space  $E$ . The random variable  $T_n$  denotes the time in which the system enters state  $J_n$ ; this variable assumes any values in the set of positive real numbers. The time the system remains in the state  $J_{n-1}$  before entering state  $J_n$  is called sojourn time. It is denoted by  $X_n = T_n - T_{n-1}$  having set  $X_0 = 0$ . The process  $(J_n, T_n)$  is called Markov Renewal Process (MRP) whenever it satisfies the next assumption:

$$P(J_{n+1} = j, X_n \leq t | J_n, X_{n-1}, J_{n-1}, X_{n-2}, \dots) = P(J_{n+1} = j, X_n \leq t | J_n).$$

The conditional joint probabilities of the MRP are denoted by

$$Q_{i,j}(t) = P(J_{n+1} = j, X_n \leq t | J_n = i)$$

and the matrix  $\mathbf{Q}(t) = (Q_{i,j}(t))$  is called the semi-Markov kernel.

Let  $N(t) = \sup\{n: T_n \leq t\}$  be the counting process of the number of transition up to the time  $t$ . Then the semi-Markov process can be defined by  $Z(t) := J_{N(t)}$ .

Let assume that  $\mathbf{Q}$  is absolutely continuous with respect to the Lebesgue measure on the set of positive real numbers and denote by  $q_{i,j}(t) = \frac{Q_{i,j}(dt)}{dt}$  the corresponding Radon-Nikodym derivatives. Hence, we can consider the hazard rate functions according to the relation

$$\lambda_{ij}(t) = \begin{cases} \frac{q_{i,j}(t)}{1 - H_i(t)} & \text{if } p_{i,j} > 0 \text{ and } H_i(t) < 1 \\ 0 & \text{otherwise} \end{cases}$$

where  $H_i(t) := P(X_n \leq t | J_n = i) = \sum_{j \in E} Q_{i,j}(t)$  and  $p_{i,j} := P(J_{n+1} = j | J_n = i) = \lim_{t \rightarrow +\infty} Q_{i,j}(t)$ .

Let us introduce the backward recurrence time process  $B(t) := t - T_{N(t)}$ . Now we can describe the set of three assumptions used in [14] to derive the formula for the  $n$ -ROCOF of a semi-Markov process.

**Assumption A1:** This first assumption explains a general random starting mechanism for semi-Markov processes. We first consider the vector  $\mathbf{p} = (p_1, p_2, \dots, p_n)$  where

$$p_i = P(J_0 = i)$$

with  $\sum_{i \in E} p_i = 1$ . Moreover we specify a set of cumulative distribution function for the duration in the initial state  $i$ ; i.e.  $F_i(v_0) = P(B(0) \leq v_0 | J_0 = i)$ . Then we assume that

$$F_i(v_0) = \begin{cases} 0 & \text{if } v_0 < 0 \\ a_i & \text{if } v_0 = 0, 0 \leq a_i \leq 1 \\ G_i(v_0) \cdot (1 - a_i) + a_i & \text{if } v_0 > 0 \end{cases}$$

being  $G_i(\cdot)$  an absolutely continuous cumulative distribution function with support in  $(0, \infty)$  and corresponding density function  $g_i(\cdot)$  having finite expectation.

**Assumption A2:** The semi-Markov process has uniformly bounded transition intensities, in formula:

$$\exists c \in \mathbb{R}_+ \text{ such that } \max_{i,j \in E} \sup_{y \in \mathbb{R}_+} \lambda_{ij}(y) \leq c .$$

**Assumption A3:** For each state  $i \in E$  there exist a state  $j \in E$  (depending on  $i$ ) and a non-null subset  $S_{ij}(y)$  of the real numbers such that  $\lambda_{ij}(y) < c$  for all  $y \in S_{ij}$ .

The following main result gives the formula for the  $n$ -ROCOF of a semi-Markov process.

**Theorem [14]** The  $n$ -ROCOF at times  $\mathbf{t}_1^n = (t_1, t_2, \dots, t_n)$  with  $t_i < t_{i+1}$  of a semi-Markov process  $(Z(t), t \geq 0)$  over a finite state space  $E = \{1, 2, \dots, m\}$  and semi-Markov kernel  $\mathbf{Q} = (Q_{ij}(t)), i, j \in E$  is given by

$$\begin{aligned} ro(\mathbf{t}_1^n) = & \sum_{j_0=1}^m \sum_{\mathbf{w} \in W^n, \mathbf{f} \in F^n} a_{j_0} p_{j_0} \prod_{r=1}^n \int_0^{s_r} \psi'_{f_{r-1} w_r}(0; u_r) q_{w_r f_r}(s_r - u_r) du_r \\ & + \sum_{j_0=1}^m \sum_{\mathbf{w} \in W^n, \mathbf{f} \in F^n} (1 - a_{j_0}) p_{j_0} \left( \int_0^\infty \left( \int_{-v_0}^{s_1} \psi'_{j_0 w_1}(v_0; u_1) q_{w_1 f_1}(s_1 - u_1) du_1 \right) g_{j_0}(v_0) dv_0 \right) \\ & \cdot \prod_{r=2}^n \int_0^{s_r} \psi'_{f_{r-1} w_r}(0; u_r) q_{w_r f_r}(s_r - u_r) du_r , \end{aligned} \quad (7)$$

where  $s_r = t_r - t_{r-1}$ ,  $s_1 = t_1$  and  $\psi_{ij}(y; t) := E_{(i,y)}[N_j(t)] = \sum_{n=0}^\infty Q_{ij}^{(n)}(y; t)$ .

The quantity  $Q_{ij}^{(n)}(y; t) = P(J_n = j, T_n \leq t | J_0 = i, B(0) = y)$ .

We observe that if we consider for all  $j_0 \in E$ ,  $a_{j_0} = 1$  and  $n = 1$ , we have a null duration in the initial state and we obtain exactly the formula for the ROCOF as it was established by [10], i.e.

$$ro(t_1) = \sum_{j_0=1}^m \sum_{\mathbf{w} \in W} \sum_{\mathbf{f} \in F} p_{j_0} \int_0^{t_1} \psi'_{j_0 w}(0; u_1) q_{w f}(t_1 - u_1) du_1 . \quad (8)$$

## 5. Applied problems

Applications of the  $n$ -ROCOF measures to the reliability field are clearly of great interest. Here we show a few results related to a numerical example based on Markov processes in subsection 5.1. Anyway, there are many other fields of application in which the same concepts are worth discussing. An example is the financial modeling of credit rating dynamics, which was extensively discussed in [12,26]. A new application to wind power production is discussed later in subsection 5.2 taking inspiration from the problem discussed in [27] and [28].

### 5.1 A numerical example

Let us consider a random system whose state space is given by the set  $E = \{1, 2, 3, 4, 5, 6, 7, 8\}$ . Each number represents different levels of performance of the system, going from perfect functioning

(state 1) to the worst one (state 8). The state space  $E$  is partitioned into two disjoint subsets  $W = \{1,2,3\}$  and  $F = \{4,5,6,7,8\}$ .

We assume that the system evolves dynamically according to the generator matrix

$$Q = \begin{pmatrix} -0.10 & 0.08 & 0.00 & 0.02 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.80 & -1.30 & 0.35 & 0.15 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.12 & -1.33 & 0.90 & 0.31 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.10 & 0.30 & -0.90 & 0.50 & 0.10 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.30 & 0.15 & -0.95 & 0.35 & 0.15 & 0.00 \\ 0.00 & 0.00 & 0.05 & 0.20 & 0.10 & -0.50 & 0.10 & 0.05 \\ 0.00 & 0.05 & 0.08 & 0.09 & 0.15 & 0.33 & -0.90 & 0.20 \\ 0.00 & 0.00 & 0.10 & 0.10 & 0.20 & 0.30 & 0.30 & -1.00 \end{pmatrix}.$$

Using equation (5), we compute the ROCOF of order 1 for the three working states. The results are shown in Figure 1. The continuous line refers to the ROCOF computed starting from state 1, i.e., using the initial probability distribution  $\alpha = (1,0, \dots, 0)$ ; the dashed line refers to the ROCOF computed starting from state 2, i.e., using the initial probability distribution  $\alpha = (0,1,0, \dots, 0)$ ; the dotted line refers to the ROCOF computed starting from state 3, i.e., using the initial probability distribution  $\alpha = (0,0,1,0, \dots, 0)$ . The figure shows that independently of the initial state, the system is going to deteriorate as the ROCOF shows increasing paths. Nonetheless, the differences according to the initial state are remarkable and demonstrate a higher risk for state 3 and a lower risk for state 1.

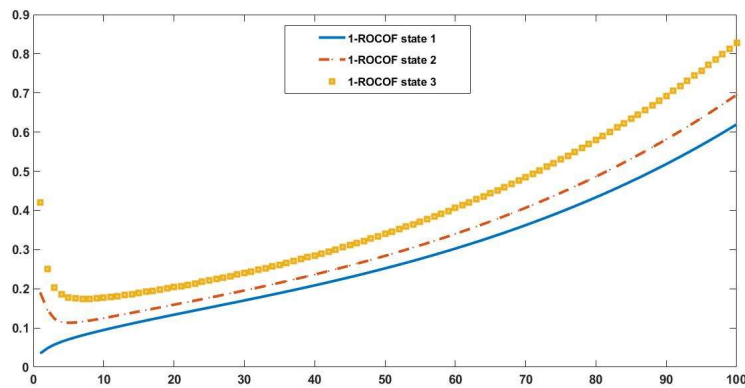


Figure 1. 1-ROCOF for a Markov process

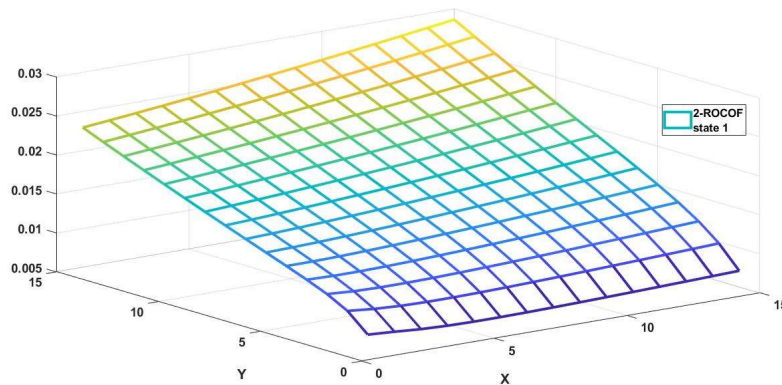


Figure 2. 2-ROCOF for a Markov process starting from state 1

Using equation (6), we compute the 2 –ROCOF starting from state 1. The result is graphically

displayed in figure 2. On the X-axis, we report the time  $t_1$ , while on the Y-axis, we report the time  $t_2 - t_1$ . Hence, the point (10, 5) on the XY-plane corresponds to the choice of  $t_1 = 10$  and  $t_2 = 15$ . Any point on the surface represents the corresponding value of the 2-ROCOF. High values of the surface show evidence for the association of failures at the corresponding times on the X and Y axes. The maximum values (for the times considered in the figure) are concentrated on large values of  $t_1$  and  $t_2 - t_1$ .

We repeat the computations of the 2-ROCOF changing the initial state. The case with state 2 as initial state and that for state 3 are considered in figures 3 and 4, respectively.

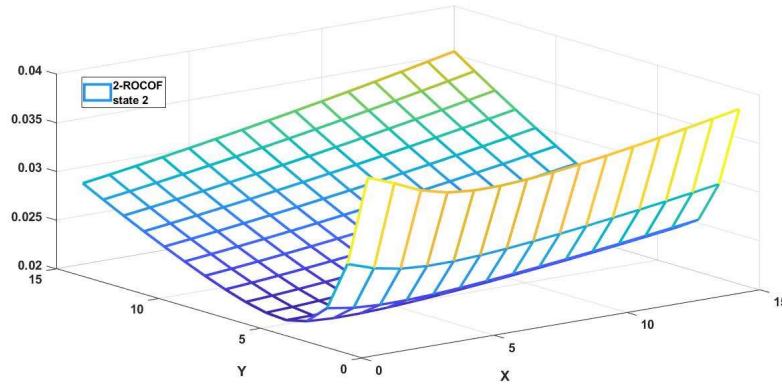


Figure 3. 2-ROCOF for a Markov process starting from state 2

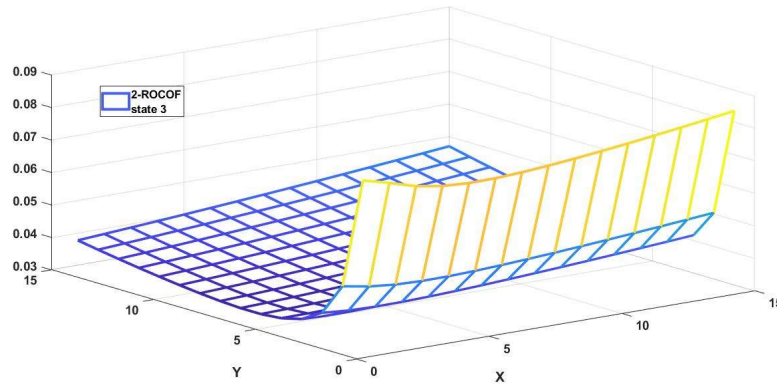


Figure 4. 2-ROCOF for a Markov process starting from state 3

As it is possible to see from these figures, their shapes are completely different from those of figure 2. In figures 3 and 4, the maximum values of association between couples of failures are for combinations of short times  $t_2 - t_1$  independently from time  $t_1$  (which shows a contained variability). This aspect is important because at short values of  $t_2 - t_1$ , the system may show trajectories of alternation between subsets  $W \rightarrow F \rightarrow W \rightarrow F$ , i.e., the presence of two close-in-time failures.

## 5.2 Wind power example

Wind power is one of the most important renewable energy sources. Because wind speed changes very sharply over time, the wind engineer must use mathematical models to predict the power output. Particular care should be given to abrupt interruptions of power production that may be caused by extreme wind speeds. Indeed, on the one hand, low wind is unable to move wind turbine's blades, which determines no energy production. On the other hand, extremely high wind speeds may cause damage to the turbine; hence, the wind engineer must switch it off to avoid



structural breaking. The minimal wind speed necessary to activate the turbine is called the cut-in speed  $v_{ci}$ . The maximal wind speed that the turbine can handle is called the cut-out speed  $v_{co}$ . Finally, it is also important to consider the rated wind speed  $v_r$ , which represents the minimum wind speed value at which the turbine achieves its maximum power production, the so-called rated power. Using the power curve for the wind turbine under consideration, we may calculate the power output as a function of the wind speed. The relationship between wind speed  $v(t)$  and wind power  $Pow(t)$  at any time  $t$  is

$$Pow(t) = \begin{cases} 0 & \text{if } v(t) \leq v_{ci} \\ \frac{P_r \cdot (v^3(t) - v_{ci}^3)}{v_r^3 - v_{ci}^3} & \text{if } v_{ci} < v(t) < v_r \\ P_r & \text{if } v_r < v(t) < v_{co} \\ 0 & \text{if } v(t) > v_{co} \end{cases}$$

where  $P_r$  is the rated power. Essentially, for wind speed lower than the wind cut-in speed, there is no power production. For wind speed between the cut-in speed and the rated speed the output power is a cubic function of the wind speed. For values between the rated speed and the cut-off speed the turbine produces its rated power. Finally, for speed greater than the wind speed cut-off the turbine does not produce power.

The ROCOF and its generalization could be fruitfully used in this applied field as we are going to show. First, we consider the first-order discrete-time Markov chain model proposed in [28] which was applied to a sample of hourly wind speed data collected by Malaysian Meteorological Station located at Mersing. The wind speed data range from 0 to 12m/s; hence the authors adopted a twelve state Markov chain model with  $E = \{1,2, \dots, 12\}$  where the  $i$ -th state collects all wind speed measurements ranging between  $(i - 1) m/s$  and  $i m/s$ . The estimated transition probability matrix is taken from [28]:

$$P = \begin{bmatrix} 0.371 & 0.407 & 0.174 & 0.036 & 0.009 & 0.002 & 0.001 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.166 & 0.446 & 0.312 & 0.059 & 0.012 & 0.004 & 0.000 & 0.001 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.051 & 0.243 & 0.504 & 0.163 & 0.028 & 0.008 & 0.002 & 0.001 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.017 & 0.083 & 0.303 & 0.391 & 0.160 & 0.035 & 0.008 & 0.002 & 0.001 & 0.000 & 0.000 & 0.000 \\ 0.010 & 0.035 & 0.099 & 0.277 & 0.382 & 0.157 & 0.031 & 0.007 & 0.001 & 0.001 & 0.000 & 0.000 \\ 0.006 & 0.021 & 0.043 & 0.108 & 0.295 & 0.343 & 0.146 & 0.031 & 0.004 & 0.003 & 0.000 & 0.000 \\ 0.005 & 0.016 & 0.027 & 0.047 & 0.110 & 0.302 & 0.324 & 0.142 & 0.021 & 0.004 & 0.002 & 0.000 \\ 0.006 & 0.016 & 0.030 & 0.033 & 0.055 & 0.127 & 0.365 & 0.239 & 0.105 & 0.022 & 0.002 & 0.000 \\ 0.009 & 0.019 & 0.014 & 0.018 & 0.042 & 0.065 & 0.140 & 0.326 & 0.269 & 0.079 & 0.014 & 0.005 \\ 0.014 & 0.054 & 0.055 & 0.014 & 0.027 & 0.028 & 0.041 & 0.205 & 0.288 & 0.164 & 0.083 & 0.027 \\ 0.000 & 0.000 & 0.000 & 0.040 & 0.000 & 0.000 & 0.080 & 0.120 & 0.160 & 0.240 & 0.280 & 0.080 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.200 & 0.000 & 0.200 & 0.600 & 0.000 \end{bmatrix}$$

Now consider a commercial wind turbine with a cut-in speed of 4 m/s . This means that when the wind speed process is in one of the first four states of the Markov chain, there is no power production. The state space  $E$  is partitioned into two disjoint subsets  $W$  and  $F$  according to the following:

$$W = \{5,6, \dots, 12\} \quad F = \{1,2,3,4\}.$$

In order to apply the measures discussed in the previous sections, we need to transform the discrete dynamic expressed by the transition matrix estimated by [28] into a continuous-time one by finding a generator matrix that satisfactorily matches, in some sense, the discrete process. This is a well-known and still open problem in the theory of Markov chains that is called the embedding problem. A detailed discussion is provided in [29], and further results and applications are provided in [30]. We consider here a simple strategy to get a generator matrix rendering results "close" to the observed hourly transition probability matrix  $P$ . We observe that the transition probability function for a continuous-time Markov process satisfy the relation

$$P(t) = e^{Qt} = \sum_{n=0}^{\infty} \frac{(Qt)^n}{n!}.$$

Therefore, given a probability matrix  $P$ , we can try to find a generator matrix  $Q$  such that

$$P \approx \frac{(Qt)^0}{0!} + \frac{(Qt)^1}{1!} = I + Qt.$$

From this relation we recover  $Q = \frac{1}{t} \cdot (P - I)$ . and we obtain the initial guess  $Q^* = (P - I)$  by setting  $t = 1$  to denote one hour. Hence, we solve the following optimization problem:

$$\min_{Q \in \Phi} \|P - e^Q\|,$$

which consists of finding, within the set of generator matrices  $\Phi$ , the one that minimizes the previous matrix norm. In our application, we consider the minimization of the Frobenius matrix norm, and we use the software Matlab to solve this optimization problem with an initial guess  $Q^*$ . The result is the following generator matrix:

$$Q = \begin{bmatrix} -3.295 & 0.035 & 0.006 & 0.002 & 0.000 & 0.001 & 0.001 & 0.000 & 0.309 & 1.999 & 0.942 & 0.000 \\ 0.020 & -2.891 & 0.000 & 0.004 & 0.000 & 0.000 & 0.000 & 0.000 & 0.277 & 1.633 & 0.939 & 0.018 \\ 0.033 & 0.028 & -2.362 & 0.001 & 0.000 & 0.000 & 0.000 & 0.000 & 0.085 & 0.776 & 1.239 & 0.200 \\ 0.268 & 0.059 & 0.014 & -1.975 & 0.014 & 0.000 & 0.000 & 0.000 & 0.028 & 0.137 & 0.504 & 0.951 \\ 0.936 & 0.261 & 0.051 & 0.011 & -1.965 & 0.001 & 0.000 & 0.000 & 0.017 & 0.060 & 0.166 & 0.462 \\ 0.492 & 0.572 & 0.243 & 0.052 & 0.007 & -1.663 & 0.000 & 0.000 & 0.010 & 0.035 & 0.072 & 0.180 \\ 0.183 & 0.503 & 0.540 & 0.237 & 0.035 & 0.007 & -1.746 & 0.080 & 0.009 & 0.027 & 0.046 & 0.079 \\ 0.091 & 0.212 & 0.608 & 0.398 & 0.175 & 0.037 & 0.004 & -1.758 & 0.100 & 0.027 & 0.051 & 0.055 \\ 0.070 & 0.108 & 0.535 & 1.091 & 0.728 & 0.002 & 0.003 & 0.003 & -2.542 & 0.001 & 0.000 & 0.001 \\ 0.045 & 0.047 & 0.068 & 0.341 & 0.480 & 0.273 & 0.138 & 0.045 & 0.024 & -1.877 & 0.392 & 0.024 \\ 0.000 & 0.000 & 0.133 & 0.200 & 0.266 & 0.455 & 0.522 & 0.133 & 0.000 & 0.001 & -1.777 & 0.067 \\ 0.001 & 0.000 & 0.000 & 0.532 & 0.000 & 0.740 & 1.550 & 0.000 & 0.001 & 0.006 & 0.002 & -2.832 \end{bmatrix}$$

Now, we can compute the ROCOF using the formula presented in the previous section. To highlight the potentiality of the continuous-time framework, we compute the indicators on a 5-minute time scale. This can be done simply by considering the hourly-based generator matrix  $Q$  and dividing it by a factor of 1/12.

In figure 5, we report the 1-ROCOF corresponding to three choices of the initial distribution over the states of the system. Specifically, the continuous blue line represents the indicator computed starting from state 5, which denotes a wind speed of 5m/s. The dotted red line denotes the 1-ROCOF behavior starting with a wind speed of 9m/s. Finally, the dashed yellow line stands for the 1-ROCOF with an initial wind speed of 1m/s. It is possible to note that for short times, the 1-ROCOF is monotone with respect to the initial speed. Thus, being in a state with strong wind implies a higher chance of moving into one of the failure states where no power production occurs. The overall behavior becomes irrelevant to the initial state around time 20, which corresponds to 20 · 5min = 100min where the system shows the achievement of a stationary value of the 1 - ROCOF equal to 0.0536.

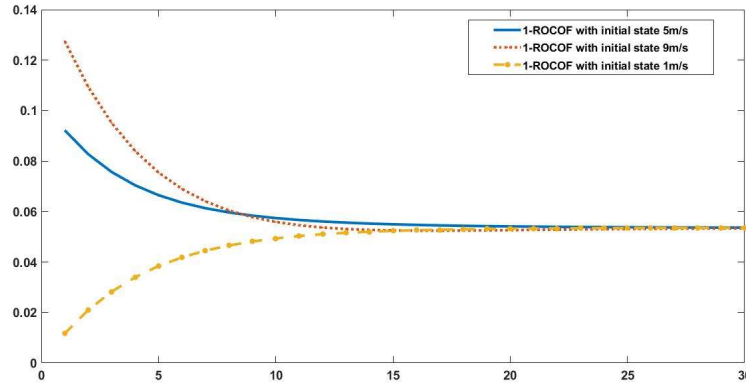


Figure 5. 1-ROCOF for the Markov process of wind speed

For completeness, we proceed by computing the 2-ROCOF for the wind speed Markov process. In figure 6, we display the indicator when the initial state is equal to  $5m/s$  (left panel) and when the initial state is equal to  $9m/s$  (right panel). The panels have similar surfaces. Both indicate a maximum value of the 2-ROCOF for high values of the time  $t_1$  and low values of the time  $t_2 - t_1$ . Hence, the maximum chance for a couple of transitions from working to failure states is for combinations of times as  $(t_1 = 20, t_2 = 21)$ . The indicator is increasing with respect to time  $t_1$  and decreasing with respect to time  $t_2 - t_1$ . The 2-ROCOF assumes higher values for the initial wind speed of  $9m/s$  as compared to the initial wind speed of  $5m/s$  case. In this way, the reliability engineer has a clear idea of when the association between two failure events is high or low. This information can also be used to measure the riskiness of a wind park investment in terms of the intermittency of power production.

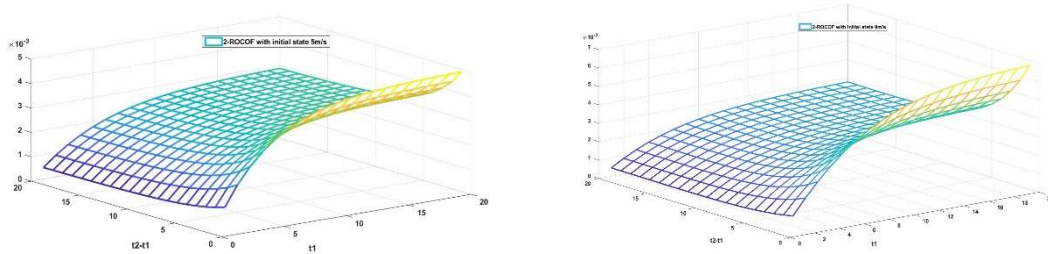


Figure 6. 2-ROCOF for the Markov process of wind speed for two initial wind speed values

#### IV. Discussion

Understanding the rate of occurrence of failures in a random system is of great relevance, both theoretically and practically. This paper offers some general information after commenting on a selection of recent major studies in the field.

- The definition of a recent measure called the  $n$ -ROCOF is reported, and a broad interpretation as a measure of clustering in time of failures is given.
- The computation of the  $n$ -ROCOF under the hypothesis of a continuous-time finite state space Markov chain is explained in detail. The results are also presented in the more general framework represented by semi-Markov processes.

- A numerical example clarifies the results and shows different shapes of the indicators that are flexible enough to represent a great variety of real system behavior. A new application of these concepts is provided in the field of wind engineering and reveals interesting aspects that need an accurate investigation in a specific research article.

## References

- [1] Lisnianski, A., I. Frenkel, and Karagrigoriou, A. eds. Recent advances in multi-state systems reliability: Theory and applications, Springer, 2017.
- [2] Singh, C. and Billinton, R. (1974). A new method to determine the failure frequency of a complex system, *IEEE Transactions on Reliability*, 23(4):231–234.
- [3] Khan, N.M., K. Rajamani, and Banerjee, S.K. (1997). A direct method to calculate the frequency and duration of failures for large networks, *IEEE Transactions on Reliability*, 26(5):318–321.
- [4] Iyer, R.K. and Downs, T. (1978). A moment approach to evaluation and optimization of complex system reliability, *IEEE Transactions on Reliability*, 27(3):226–229.
- [5] Shi D. (1985). A new method for calculating the mean failure numbers of a repairable system during  $(0; t]$ , *Acta Mathematicae Applicatae Sinica*, 8(1):101-110.
- [6] Yeh, L. (1995). Calculating the rate of occurrence of failures for continuous-time Markov chains with application to a two-component parallel system, *Journal of the Operational Research Society*, 46(4):528-536.
- [7] Yeh, L. (1997). The Rate of Occurrence of Failures. *Journal of Applied Probability*, 34(1):234-247.
- [8] Grabski, F. *Semi-Markov processes: applications in system reliability and maintenance*. Elsevier, 2014.
- [9] Vassiliou, P-CG, and Papadopoulou, A.A. (1992). Non-homogeneous semi-Markov systems and maintainability of the state sizes. *Journal of Applied Probability*, 519-534.
- [10] Ouhbi B. and Limnios, N. (2002). The rate of occurrence of failures for semi-Markov processes and estimation. *Statistics and Probability Letters*, 59:245-255.
- [11] Georgiadis, S., N. Limnios, and Votsi, I. (2013). Reliability and Probability of First Occurred Failure for Discrete-Time Semi-Markov Systems, *Applied Reliability Engineering and Risk Analysis: Probabilistic Models and Statistical Inference* 167-179.
- [12] D'Amico G., (2015). Rate of Occurrence of Failures (ROCOF) of Higher-Order for Markov Processes: Analysis, Inference and Application to Financial Credit Ratings, *Methodology and Computing in Applied Probability*, 17:929-949.
- [13] Votsi, I. (2019). Conditional failure occurrence rates for semi-markov chains, *Journal of Applied Statistics*, 46(15):2722-2743.
- [14] D'Amico, G. and Petroni, F. (2023). ROCOF of higher order for semi-Markov processes, *Applied Mathematics and Computation*, 441:127719.
- [15] Yi, H., Cui, L., Shen, J., and Li, Y. (2018). Stochastic properties and reliability measures of discrete-time semi-Markovian systems, *Reliability Engineering & System Safety*, 176:162-173.
- [16] Csenki, A. (1994). On the interval reliability of systems modelled by finite semi-Markov processes, *Microelectron. Reliab.*, 34:1319–1335
- [17] Georgiadis, S. and Limnios, N. (2014). Interval reliability for semi-Markov systems in discrete time, *Journal de la Société Française de Statistique*, 155(3):152-166.
- [18] D'Amico, G., Manca, R., Petroni, F., and Selvamuthu, D. (2021). On the computation of some interval reliability indicators for semi-Markov systems, *Mathematics*, 9(5):575.
- [19] Barbu, V.S., G. D'Amico, and Gkelsinis, T. (2021). Sequential interval reliability for discrete-time homogeneous semi-Markov repairable systems, *Mathematics*, 9(16): 1997.

- [20] Billinton, R., and Bollinger, K.E. (1968). Transmission system reliability evaluation using Markov processes, *IEEE Transactions on power apparatus and systems*, 2:538-547.
- [21] Jain, Vidyottama, Raina Raj, and Dharmaraja, S. (2022). Performability analysis of a M M A P[2]/P H [2]/S model with PH retrial times, *Communications in Statistics-Theory and Methods*, 1-20.
- [22] S., Dharmaraja, Sivama, A. H., Raj, R., and Vishnevskyb, V. (2023). Study of reliability of the on-tether subsystem of a tethered high-altitude unmanned telecommunication platform. *Reliability: Theory & Applications*, 18(1 (72)):172-178.
- [23] Yadav, A.D., N. Nandal, and Malik, S.C. (2023). Markov approach for reliability and availability analysis of a four unit repairable system, *Reliability: Theory & Applications* 18.1(72): 193-205.
- [24] Limnios, N. and G. Oprisan. *Semi-Markov processes and reliability*. Springer Science & Business Media, (2001).
- [25] Dimitrov, Boyan, Vladimir Rykov, and Sahib Esa. Three Approaches in the Study of Recurrent Markovian and Semi-Markovian Processes. *Distributed Computer and Communication Networks: 23rd International Conference, DCCN 2020, Moscow, Russia, September 14–18, 2020, Revised Selected Papers* 23. Springer International Publishing, 2020.
- [26] D'Amico, G., Regnault, P., Scocchera, S., and Storchi, L. (2018). A continuous-time inequality measure applied to financial risk: the case of the European Union, *International Journal of Financial Studies*, 6(3):62.
- [27] Vergine, S., Álvarez-Arroyo, C., D'Amico, G., Escaño, J. M., and Alvarado-Barrios, L. (2022). Optimal management of a hybrid and isolated microgrid in a random setting. *Energy Reports*, 8:9402-9419.
- [28] Shamshad, A., Bawadi, M. A., Hussin, W. W., Majid, T. A., and Sanusi, S. A. M. (2005). First and second order Markov chain models for synthetic generation of wind speed time series. *Energy*, 30(5):693-708.
- [29] Bladt, M., and Sørensen, M. (2005). Statistical inference for discretely observed Markov jump processes. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(3):395-410.
- [30] D'Amico, G., and Regnault, P. (2018). Dynamic measurement of poverty: Modeling and estimation. *Sankhya B*, 80(2):305-340.

# STRIP-PLOT ANALYSIS FOR THE CONSTRUCTION OF COMPLETE TRIPARTITE AND CUBIC GRAPHS

V. Saranya

•

Research Scholar, Department of Statistics, Periyar University, Salem, Tamil Nadu, India

saranya88stat@gmail.com

S. Kavitha

•

Assistant Professor, Department of Statistics, Periyar University, Salem, Tamil Nadu, India

pustatkavitha@gmail.com

## Abstract

*The Strip-Plot Design (SPD) plays an important role in the complete block designs and also using the agricultural, medical and industry fields. SPD is best suited for a two-factor experiment that has more treatments than can be accommodated by a complete block design. In a SPD, one factor is assigned to the horizontal strip plot, one factor is assigned to the vertical – strip plot and one factor is interaction plot. Also, few experimental materials may be rare while other test items may be available in altering doses of other therapeutic factors, which may be expensive or time-consuming. One of the main features of SPD involves three types of experimental errors: row - strip plot error, column – strip plot error and interaction plot error. Experimenting across processing steps is essential for studying the interaction of factors where certain factors come from one step and others arrive from the other. The strip-plot design is a very efficient design for investigating multiple-step processes in terms of both resources and time. Strip-plot designs are economical when the factors are hard to change and the process under research has three discrete stages. When we want to study interactions between factors where some factors are from one step and other factors from another step, it is important to conduct experiments across processing steps. The approach is flexible because it can handle experimental design problems involving factors acting at different levels, unlike the existing method. Graphs are widely used representations of both natural and human-made structures. Graph theory can be used to investigate "things that are connected to other things. "Fits nearly everywhere. Some tough problems become easier to solve when they are represented graphically. We reviewed the agricultural field yield of the strip-plot design and early work on the design of industrial strip-plot design in this paper. We have also described the model of strip-plot design. We, therefore, advise experimenters to ensure that their strip-plot designs contain a sufficient number of rows and columns so that valid statistical inference is possible. A bipartite graph is one in which the edges can be divided into two sets without going into sets. A complete bipartite graph is a bipartite graph that is completed. The complete tripartite graph in which the edges can be divided into three set without going into sets. The cubic graph is a graph in which all vertices have degree three. This paper describes the construction and Statistical Analysis of SPD using some particular types of graphs is discussed through numerical examples.*

**Keywords:** Strip -plot design, complete tripartite graph, cubic graph

## 1. Introduction

India is the third-largest producer of cotton in the world. Cotton grows well in drier parts of the, black soil, red soil and alive soil of the Deccan plateau. It requires high temperature, light rainfall or irrigation, 210 frost-free days and bright sunshine for its growth. It is a Kharif crop and requires 6 to 8 months to mature. The challenge is developing design organizations that meet quality and cost criteria. Every attempt at agricultural science research includes the design of experiments. Suppose to investigate more than one factor simultaneously in a single experiment, which is called the factorial experiment of the design.

Some factors to be tested need bigger plots, and others require smaller plots. Different plots are required in such cases, and the resulting design is known as split plot design (SPD). In 1925, Fisher developed this design for the purpose of agricultural experiments. The cost of the experiment can often be reduced by avoiding complete randomization.

The strip-plot design (SPD) is essential in complete block designs and applications in agriculture, medicine, and industry. One component is assigned to the horizontal strip plot, one to the vertical strip plot, and one to the interaction plot in an SPD.

Graph theory is one of the fastest-growing sciences. Graphs in their applications, are commonly used to represent distinct objects and the relationship between these objects. The visual representation of a graph is the declaration of an object vertex, while the relationship between objects is expressed as an edge. In recent years, graph theory has established itself as an important mathematical tool in various subjects, from available research and chemistry to genetics and linguistics and from electrical engineering and geography to sociology and architecture in its own right. At the same time is mathematical to discipline in its own right. Peter Horak et al. [1] have focused on this result is a special case of a general conjecture made by Erdos and NeSetiil: For each  $d \geq 3$ , the edge set of a graph of maximum degree  $d$  can always be partitioned into  $\lceil 5d^2/4 \rceil$  subsets, each of which induces a matching. Raymond Greenlaw and Rossella Petreschi [2] have developed a new algorithm is presented for cubic graphs.

Arden Miller [3] has focused on using statistical experimental designs Strip-Plot Configurations of Fractional Factorials. George A. et al. [4] have discussed the strip-plot design for two-step processes. Elizabeth J. et al. [5] have reviewed recent developments and provided guidelines for using the Decomposition of complete tripartite graphs into gregarious 4-cycles. Heidi Arnouts et al. [6] have focused on the Strip-plot experiments, and the cost of experimentation can often be reduced by forgoing complete randomization. Antal Ivanyi et al. [7] have developed an exchange algorithm for tripartite graphs with given degree set. Abdollah Khodkar [8] has discussed the signed edge domination numbers of complete tripartite graphs. Sheikh Rashid et al. [9] has discussed the study of cubic graphs with its application and introduced certain concepts, including cubic graphs, internal cubic graphs, and external cubic graphs, and illustrate these concepts by examples. Velimor D. et al. [10] have presented the procedure for complete tripartite graphs with spanning maximal planar subgraphs.

Peter Bradshaw [11] has focused on vertex-disjoint triangles as a "tratching." The problem of finding a tratching that covers all vertices of a tripartite graph can be shown to be NP-complete using a reduction from the three-dimensional matching problem. K Nisa et al. [11] have discussed the Analysis of variance for strip plot design with missing values: bias correction of the mean squares. Hossein Rashmansloua et al. [13] discussed about cubic graphs with novel application and define the direct product. we introduce the notion of complete cubic graphs and present some properties of self-complementary cubic graphs. Peter Goos [14] has reviewed recent developments and provided guidelines for using the fish patty experiment: a strip-plot look. This paper discussed a statistical analysis of SPD using complete tripartite and cubic graphs with a numerical example.

## 2. Preliminaries

### 2.1 Strip – Plot Design

In strip plot design, each block is divided into several vertical and horizontal strips depending on the levels of the individual factors. Therefore, the Analysis of strip plot design is carried out in three parts. The first part is the vertical strip analysis, the second part is the horizontal strip analysis, and the third is the interaction analysis.

### 2.2 Complete Tripartite Graph

A complete tripartite graph is a set of vertices split into three disjoint sets such that no two graph vertices within the same set are adjacent and every vertex in one set is adjacent to every vertex in the other two sets. If the three sets contain  $p$ ,  $q$ , and  $r$  graph vertices, a complete tripartite graph.

### 2.3 Cubic Graph

In the mathematical field of graph theory, a cubic graph is one in which all vertices have degree three. In other words, a cubic graph is a three-regular graph. Cubic graphs are also called trivalent graphs.

## 3. Statistical Analysis of Strip – Plot Design

The linear model for strip-plot design is

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \gamma_k + (\tau\gamma)_{ik} + \varepsilon_{ijk} \quad i=1,2,\dots,r, \quad j=1,2,\dots,v, \quad k=1,2,\dots,n \quad (1)$$

$Y_{ijk}$  is observation corresponds to the  $k^{\text{th}}$  level of factor (A),  $j^{\text{th}}$  level of factor (A) and  $i^{\text{th}}$  replication.  $\mu$  the general mean effect.

$\tau_j$  is  $i^{\text{th}}$  block effect,  $A$  is the  $j^{\text{th}}$  level of factor  $A$ ,  $B$  is the  $k^{\text{th}}$  level of factor  $B$ .

$(\tau\beta)_{ij}$  is the interaction between  $j^{\text{th}}$  level factor  $A$  and  $k^{\text{th}}$  level factor  $B$ , the error components.

$\varepsilon_{ijk}$  and  $\varepsilon_{ijk}$  are independently and normally distributed with means zero and respective variance  $\sigma_a^2$ ,  $\sigma_b^2$  and  $\sigma_e^2$ .

In statistical analysis, separate estimates of error are obtained for the main effects of the factors  $A$  and  $B$  and their interaction  $A.B$ . Thus, three mean error squares will be applicable for testing the significance of the main results of the characteristics and their interaction separately.

The vertical strip plot for the first factor, the horizontal strip plot for the second factor, and the vertical and horizontal bars in the interaction strip plot for the interaction between two factors are always perpendicular to each other. The correlation plot is very small and primarily illustrates the interaction between the two design factors. As a result, we may say that correlation is assessed more precisely in strip plot design.

This is an outline of the variance analysis table:

- Correction factor (C.F.) = —
- Total sum of square (SST) =
- Replication sum of square (SSR) = — - C
- Horizontal factor sum of square (S.S. (H.F.)) = — - C
- Horizontal factor error sum of square (SSE<sub>a</sub>) = — - - S -



- Vertical factor sum of square (S.S. (V.F.)) =  $\frac{1}{r} \sum_{j=1}^b T_{.j}^2 - C$
- Vertical factor error sum of square ( $SSE_b$ ) =  $\frac{1}{r} \sum_{j=1}^b T_{.j}^2 - \frac{1}{r} \sum_{j=1}^b T_{.j}^2 - SSR$
- Interaction effect sum of square =  $\frac{1}{r} \sum_{j=1}^b T_{.j}^2 - \frac{1}{r} \sum_{j=1}^b T_{.j}^2 - SSA$
- Interaction error sum of square ( $SSE_c$ ) =  $SST - (All\ other\ sum\ of\ square)$

**Table 1:** ANOVA table for strip – plot design

Sv	Df	Ss	Mss	F-Ratio
R. (R)	(r-1)	SSR	$\frac{SSR}{r-1}$	$\frac{SSR}{r-1} / \frac{SSE_b}{(r-1)(b-1)}$
H.F. (A)	(a-1)	SSA	$\frac{SSA}{a-1}$	$\frac{SSA}{a-1} / \frac{SSE_a}{(r-1)(a-1)}$
H.F.E. (a)	(r-1)(a-1)	SSEa	$\frac{SSE_a}{(r-1)(a-1)}$	-
V.F. (B)	(b-1)	SSB	$\frac{SSB}{b-1}$	$\frac{SSB}{b-1} / \frac{SSE_b}{(r-1)(b-1)}$
V.F.E.(b)	(r-1)(b-1)	SSEb	$\frac{SSE_b}{(r-1)(b-1)}$	-
I.E. (AB)	(a-1)(b-1)	SSAB	$\frac{SSAB}{(a-1)(b-1)}$	$\frac{SSAB}{(a-1)(b-1)} / \frac{SSE_c}{(r-1)(a-1)(b-1)}$
I.E. (c)	(r-1)(a-1)(b-1)	SSEc	$\frac{SSE_c}{(r-1)(a-1)(b-1)}$	-

#### 4. Construction of Strip – Plot Design using Graphs

##### 4.1. Method for Construction for Tripartite Graph

- Let us consider the horizontal strip, vertical strip, and intersection plots as vertex set Q. This vertex set P can be divided into three subsets: Q1, Q2, and Q3.
- Then the replication is considered as the first subset Q1, variety as the second subset Q2, and Soils as the third subset Q3.
- Now consider the first (replication) vertex ( $R_1$ ) of the first subset, and then  $R_1$  is connected to all the vertices of the second and third subset through edges.
- Next, consider the second replication vertex ( $R_2$ ). It's connected to all the vertices of the second and third subsets through the edges.
- Similarly, all the remaining replication vertices of the first subset are connected to all the vertices of the second and third subsets through the corresponding edges.
- Finally, we get the complete tripartite graph for the vertical strip, horizontal strip, and intersection plots.

##### 4.1.1 Application

In our study, to collect the yields of primary data on cotton cultivation varieties at Salem District of Tamilnadu. Three replicates of various cotton varieties (LRA(P.T.), Supriya, Surabhi) in kilograms and three Soil (Black, Red, and Alive). The four replications of Cotton cultivation in kilograms for

yields per plot, three varieties of crops are tested, the layout being Strip plot design data is given below.

**Table 2:** Replication wise data for yield of cotton (kg/ha)

Replication	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>
Variety	Soil(S <sub>1</sub> )			
V <sub>1</sub>	3328	3258	3400	3128
V <sub>2</sub>	3220	3150	3115	3015
V <sub>3</sub>	2850	2800	2700	2625
	Soil(S <sub>2</sub> )			
V <sub>1</sub>	2814	2750	2915	2963
V <sub>2</sub>	2656	2655	2500	2700
V <sub>3</sub>	2515	2514	2415	2400
	Soil(S <sub>3</sub> )			
V <sub>1</sub>	3050	3118	3250	3150
V <sub>2</sub>	2950	3000	3065	2950
V <sub>3</sub>	2650	2750	2950	2800

**Table 3:** Replication × variety for horizontal factor

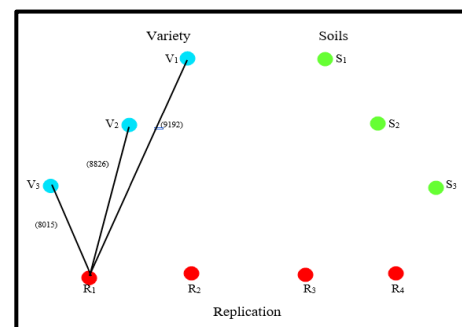
	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	Replication Total
R <sub>1</sub>	9192	8826	8015	26033
R <sub>2</sub>	9126	8805	8064	25995
R <sub>3</sub>	9565	8680	8003	26248
R <sub>4</sub>	9241	8665	7825	25731
Variety Total	37124	34976	31907	104007

The complete tripartite graph construction method for horizontal – strip plot is given below.

- From the above table 3 vertex is fixed as Q, which is divided into three subsets, the figure 1 shows that Q<sub>1</sub> (replication), Q<sub>2</sub> (variety) and Q<sub>3</sub> (soils).
- The figure 2 shows that first replication vertex (R<sub>1</sub>) connected to all the vertices of variety (V<sub>1</sub>, V<sub>2</sub> and V<sub>3</sub>) through the edge values 9192(Y<sub>1</sub>), 8826(Y<sub>2</sub>), and 8015(Y<sub>3</sub>).



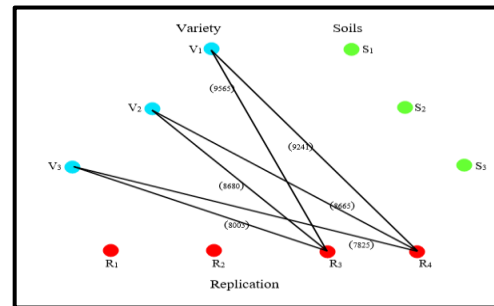
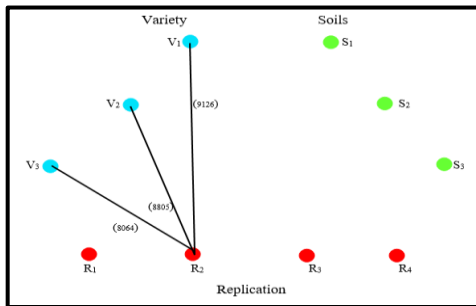
**Figure 1:** Graph of subsets



**Figure 2:** Graph for first replication (R<sub>1</sub>)

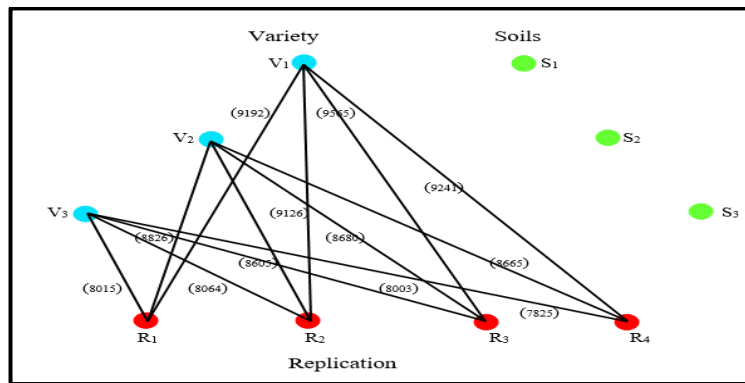
- The figure 3 shows that second replication vertex (R<sub>2</sub>), and it is connected to all the vertices of variety (V<sub>1</sub>, V<sub>2</sub> and V<sub>3</sub>) through the edge values 9126(Y<sub>1</sub>), 8805(Y<sub>2</sub>), and 8064(Y<sub>3</sub>).
- Similarly, the figure 4 shows that third and fourth replication vertices (R<sub>3</sub> and R<sub>4</sub>) are connected to all the vertices of variety (V<sub>1</sub>, V<sub>2</sub> and V<sub>3</sub>) through the

corresponding edge values. ( $Y_1, Y_2$  and  $Y_3$ ) 9565, 8680, and 8003 ( $Y_1, Y_2$  and  $Y_3$ ) 9241, 8665, and 7825.



**Figure 3:** Graph for second replication ( $R_2$ )      **Figure 4:** Graph for third and fourth replication ( $R_3$  and  $R_4$ )

- The figure 5 shows that complete tripartite graph of variety and replication for the horizontal - strip plot.



**Figure 5:** Graph for complete tripartite graph of horizontal – strip plot

**Table 4:** Replication  $\times$  soils for vertical factor

	S1	S2	S3	Replication Total
R1	9398	7885	8650	26033
R2	9208	7919	3868	25995
R3	9215	7830	9203	26248
R4	8768	8063	8900	25731
Soils Total	36589	31797	35621	104007

The construction method of the complete tripartite graph for vertical – strip plot is given below

- From the above table 4 that first replication vertex ( $R_1$ ). The figure 6 shows that first replication vertex is connected to all soils ( $S_1, S_2$  and  $S_3$ ) through the values 9398, 7985, and 8650 ( $Y_1, Y_2$  and  $Y_3$ ).
- The figure 7 shows that second replication vertex ( $R_1$ ). The second replication vertex is connected to all Soils ( $S_1, S_2$  and  $S_3$ ) through the values 9208, 7919, and 8868 ( $Y_1, Y_2$  and  $Y_3$ ).

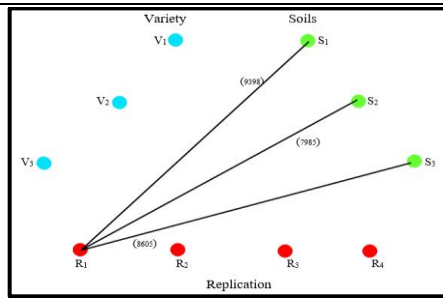


Figure 6: Graph for first replication ( $R_1$ )

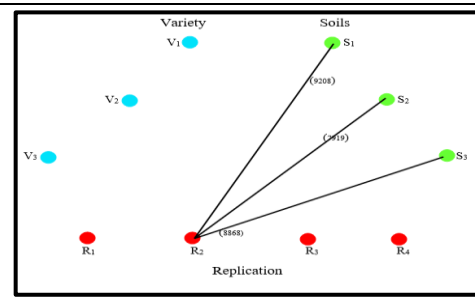


Figure 7: Graph for second replication ( $R_2$ )

- Similarly, the figure 8 shows that third and fourth replication vertices ( $R_3$  and  $R_4$ ) are connected to all the vertices of soils ( $S_1$ ,  $S_2$  and  $S_3$ ) through the corresponding edge values. 9215, 7830 and 9203 ( $Y_1$ ,  $Y_2$  and  $Y_3$ ) 8768, 8063 and 8900 ( $Y_1$ ,  $Y_2$  and  $Y_3$ ).
- The figure 9 shows that complete tripartite graph for replication and soils vertical - strip plot.

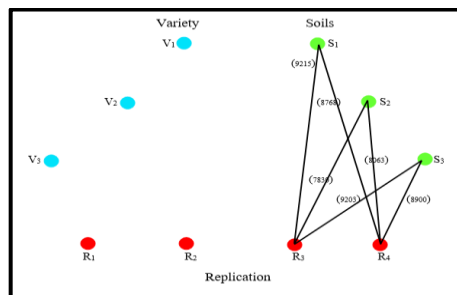


Figure 8: Graph for third and fourth replication ( $R_3$  and  $R_4$ )

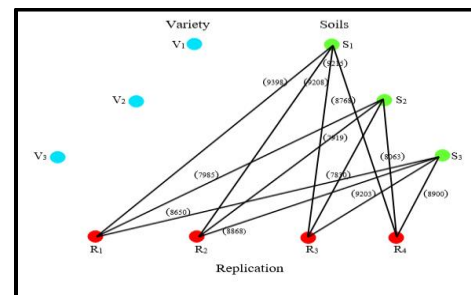


Figure 9: Complete tripartite graph of vertical Strip - plot

Table 5: Variety  $\times$  soils for interaction plot

	$S_1$	$S_2$	$S_3$	Variety Total
$V_1$	13114	11442	12568	37124
$V_2$	12500	10511	11965	34976
$V_3$	10975	9844	11088	31907
Soils Total	36589	31797	35621	104007

The construction method of complete tripartite graph for interaction plot are given below

- The above table 5 that first variety vertex ( $V_1$ ). The first variety vertex is connected to all soils ( $S_1$ ,  $S_2$  and  $S_3$ ) through the values 13114, 11442 and 12568 ( $Y_1$ ,  $Y_2$  and  $Y_3$ ).

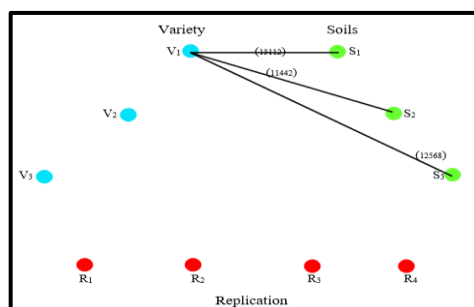


Figure 10: Graph for first variety ( $V_1$ )

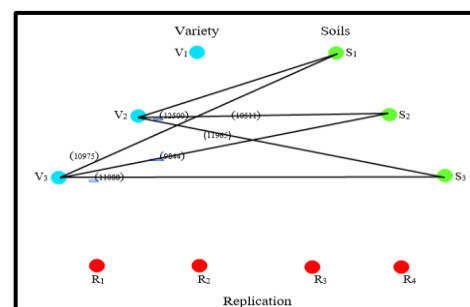
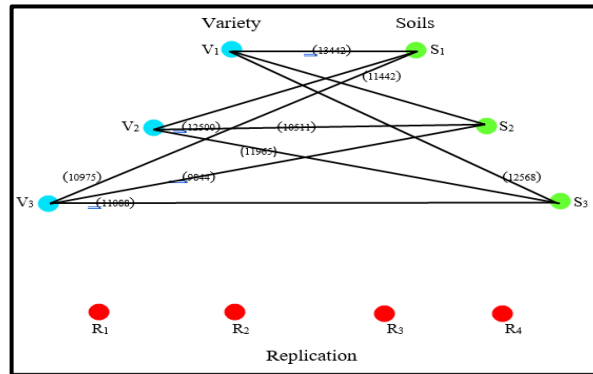


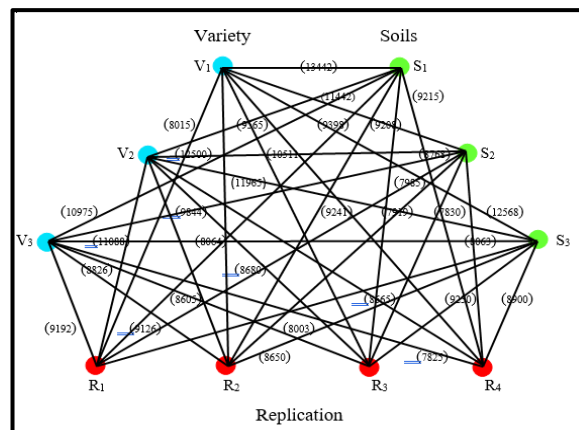
Figure 11: Graph for second and third variety ( $V_2$  and  $V_3$ )

- Similarly, the figure 11 shows that second and third varieties ( $V_2$  and  $V_3$ ) are connected to all the Soils ( $S_1$ ,  $S_2$  and  $S_3$ ), through the corresponding values 12500, 10511 and 11965 ( $Y_1$ ,  $Y_2$  and  $Y_3$ ) 10975, 9844 and 11088 ( $Y_1$ ,  $Y_2$  and  $Y_3$ ).
- The figure 12 shows that complete tripartite graph for the variety and soil interaction plot.



**Figure 12:** Complete tripartite graph for interaction plot

- The figure 13 shows that complete tripartite graph for replication and variety, replication and soils, and variety and soils.



**Figure 13:** Complete tripartite graph for horizontal, vertical and interaction strip plot

Compute the correction factor and sum of squares as

- Correction factor (C.F.) = 300484890.3
- Total sum of square (SST) = 2490006.7

Compute the sum of squares for the horizontal analysis:

- Replication sum of square (SSR) = 14996.256
- Horizontal factor sum of square (S.S. (H.F.)) = 1145826.5
- Horizontal factor error sum of square ( $SSE_a$ ) = 40929.8

Compute the sum of squares for the vertical analysis:

- Vertical factor sum of square (S.S. (V.F.)) = 1070090.6
- Vertical factor error sum of square ( $SSE_b$ ) = 118191.7

Compute the sum of squares for the interaction analysis:

- Interaction effect sum of square = 59701.4
- Interaction error sum of square ( $SSE_c$ ) = 40.271

**Table 6:** ANOVA for strip plot design

Sv	D.f	Ss	Mss	F-Ratio	P-Value
Replication	3	14996.256	4998.752	1.36467	0.26720020
Variety(A)	2	1145826.5	5722913.25	83.9847	-
Error(E <sub>a</sub> )	6	40929.8	6821.633	-	-
Soils(B)	2	1070090.6	535040.33	27.161314	0.00100000
Error(E <sub>b</sub> )	6	118191.7	19698.617	-	-
Interaction(A×B)	4	59701.4	14925.35	4.44747	0.01958176
Error (E <sub>c</sub> )	12	40271	3355.916	-	-
Total	35	-	-	-	-

The table value of replication and variety is greater than the calculated values. So the null hypothesis is accepted. There is no significant difference between the four replications and the three varieties. The table value of soils is greater than the calculated value. So the null hypothesis is accepted. There is no significant difference between the three soil levels. The table value of the interaction effect is also more important than the calculated value. So the null hypothesis is accepted.

There is no significant difference between the interaction effects. The P-value of the above experiment is more significant than the 5% significance level. Therefore the null hypothesis is accepted. There is no significant difference that occurred in the above experiment.

#### 4.2 Method for Construction of Cubic Graph

- Let us consider the horizontal-strip plot, vertical-strip plot, and interaction plot factors as vertex set Q. Then the elements are divided into two subsets, Q<sub>1</sub> and Q<sub>2</sub>.
- Then the replication is considered the first subset Q<sub>1</sub> and variety as the second subset Q<sub>2</sub>.
- Now consider the first (replication) vertex R<sub>1</sub> of the first subset and then R<sub>1</sub> is connected to all the vertices of the second subset through edges.
- Next, consider the second replication vertex R<sub>2</sub> it is connected to all the vertices of the second subset through the edges.
- Similarly, all the remaining replication vertices of the first subset are connected to all the vertices of the second subset through the corresponding edges.
- Finally, we get the cubic graph for horizontal, vertical, and interaction plots.

##### 4.2.1 Application

In our study, to collect the kilometers of primary data on petrol two-wheeler brands at Salem District of Tamilnadu. Three replicates of various two-wheeler brands (Honda, Tvs, Suzuki), in kilometers and three route way of (Hillstration, City, Highways). The four replications of petrol in kilometers per litter, three brands of kilometres are tested, and the layout being Strip plot design data is given below.

**Table 7:** Day wise for kilometres of petrol

Days	D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>
Brand	Route(R <sub>1</sub> )		
B <sub>1</sub>	30	31	31
B <sub>2</sub>	35	34	34
B <sub>3</sub>	33	32	33
	Route(R <sub>2</sub> )		

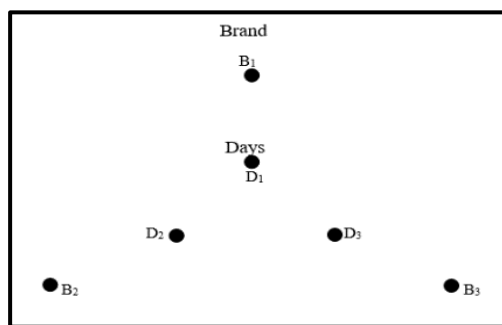
B <sub>1</sub>	35	36	37
B <sub>2</sub>	42	40	41
B <sub>3</sub>	37	38	39
Route(R <sub>3</sub> )			
B <sub>1</sub>	50	51	50
B <sub>2</sub>	57	55	56
B <sub>3</sub>	54	53	54

**Table 8:** Days × brand for horizontal factor

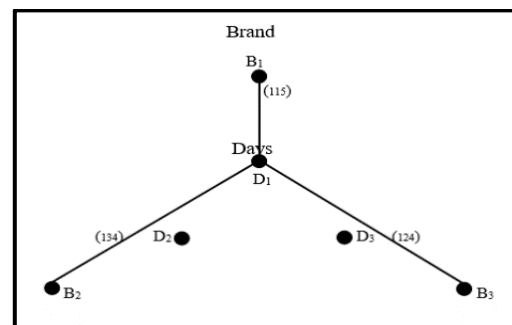
	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>	Days Total
D <sub>1</sub>	115	134	124	373
D <sub>2</sub>	118	129	123	370
D <sub>3</sub>	118	131	126	375
Brand Total	351	394	373	1118

The construction method of cubic graph for horizontal–strip plot is given below.

- From the above table 8 vertex is fixed as Q, which is divided into two subsets, the figure 14 shows that Q<sub>1</sub> (days) and Q<sub>2</sub> (brand).
- The figure 15 shows that first day vertex (D<sub>1</sub>). The first days vertex is connected to all brand (B<sub>1</sub>, B<sub>2</sub> and B<sub>3</sub>) through the values 115(Y<sub>1</sub>), 134(Y<sub>2</sub>), 124(Y<sub>3</sub>).

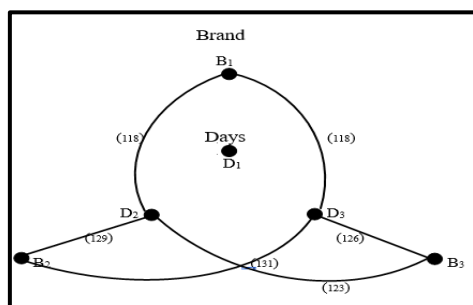


**Figure 14:** Graph of subsets

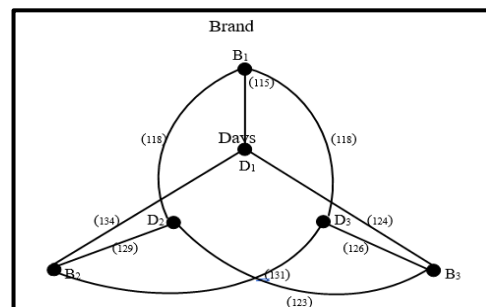


**Figure 15:** Graph of first day (D<sub>1</sub>)

- Similarly, the figure 16 shows that second and third day vertex (D<sub>2</sub> and D<sub>3</sub>). The second and third days vertex is connected to all brand (B<sub>1</sub>, B<sub>2</sub> and B<sub>3</sub>) through the values 118, 129, and 123 (Y<sub>1</sub>, Y<sub>2</sub> and Y<sub>3</sub>), 118, 131 and 126 (Y<sub>1</sub>, Y<sub>2</sub> and Y<sub>3</sub>).
- The figure 17 shows that cubic graph for days and brand.



**Figure 16:** Graph of second and third days (D<sub>2</sub> and D<sub>3</sub>)



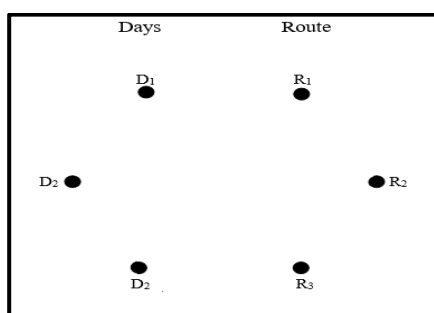
**Figure 17:** Cubic graph for horizontal – strip plot

**Table 9:** Days × route for vertical factor

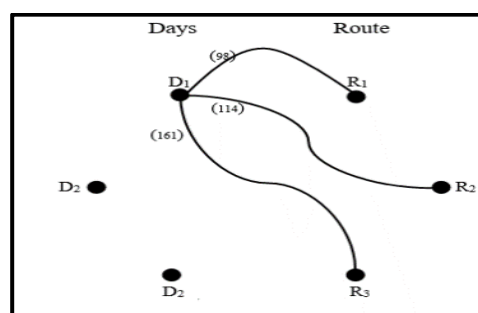
	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	Days Total
D <sub>1</sub>	98	114	161	373
D <sub>2</sub>	97	114	159	370
D <sub>3</sub>	98	117	160	375
Route Total	293	345	480	1118

The construction method of the cubic graph vertical–strip plot is given below.

- From the above table 9 vertex is fixed as Q, which is divided into two subsets, the figure 18 shows that Q<sub>1</sub> (days) and Q<sub>2</sub> (route).
- The figure 19 shows that first day vertex (D<sub>1</sub>). The first days vertex is connected to all Route (R<sub>1</sub>, R<sub>2</sub> and R<sub>3</sub>) through the values 98(Y<sub>1</sub>), 114(Y<sub>2</sub>), 161(Y<sub>3</sub>).

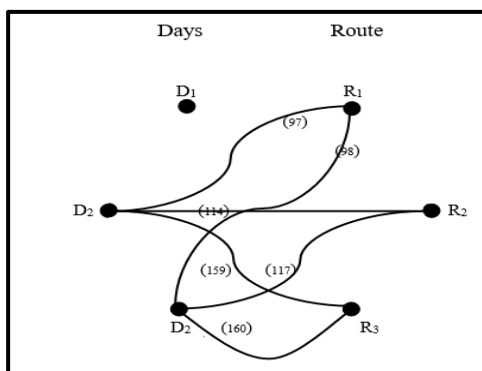


**Figure 18:** Cubic graph for subset

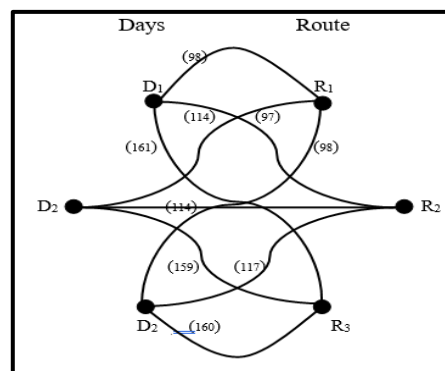


**Figure 19:** Cubic graph for first day (D<sub>1</sub>)

- Similarly, the figure 20 shows that second and third day vertex (D<sub>2</sub> and D<sub>3</sub>). The second and third day vertex is connected to all routes (R<sub>1</sub>, R<sub>2</sub> and R<sub>3</sub>), through the values 97, 114 and 159(Y<sub>1</sub>, Y<sub>2</sub> and Y<sub>3</sub>), 98, 117 and 160 (Y<sub>1</sub>, Y<sub>2</sub>, and Y<sub>3</sub>).
- The figure 21 shows that cubic graph for days and route.



**Figure 20:** Cubic graph for second and third days (D<sub>2</sub> and D<sub>3</sub>)



**Figure 21:** Cubic graph for vertical–strip plot

**Table 10:** Brand × route for Interaction factor

	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	Brand Total
B <sub>1</sub>	92	108	151	351
B <sub>2</sub>	103	123	168	394
B <sub>3</sub>	98	114	161	373
Route Total	293	345	480	1118



The construction method of the cubic graph for the Interaction strip plot is given below.

- From the above table 10 vertex is fixed as Q, which is divided into two subsets, the figure 22 shows that  $Q_1$  (brand) and  $Q_2$  (route).
- The figure 23 shows that first vertex ( $R_1$ ). The first route vertex is connected to all brand ( $B_1, B_2$  and  $B_3$ ) through the values 92( $Y_1$ ), 108( $Y_2$ ), 151( $Y_3$ ).

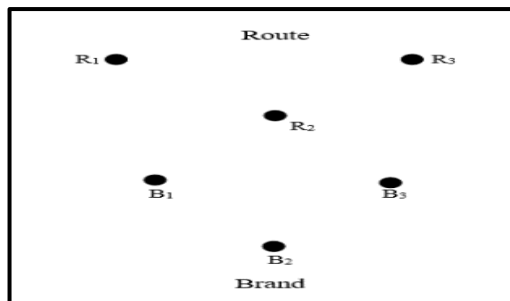


Figure 22: Cubic graph for subset

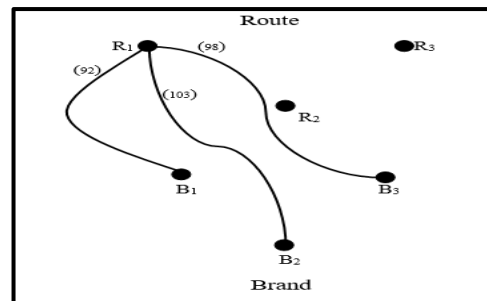


Figure 23: Cubic graph for first brand ( $B_1$ )

- Similarly, the figure 24 shows that second and third route vertex ( $R_2$  and  $R_3$ ). The second and third route vertex is connected to all brand ( $B_1, B_2$  and  $B_3$ ), through the values 103, 123, and 168 ( $Y_1, Y_2$  and  $Y_3$ ), 98, 114, and 161 ( $Y_1, Y_2$  and  $Y_3$ ).
- The figure 25 shows that cubic graph for route and brand.

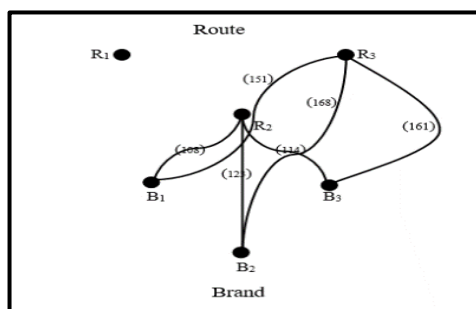


Figure 24: Cubic graph for second and third brand ( $B_1$  and  $B_3$ )

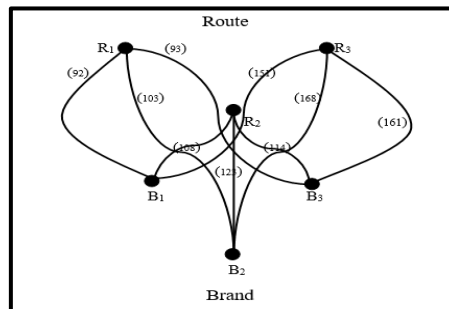


Figure 25: Cubic graph for interaction Strip - plot

Compute the correction factor and sum of squares as

- Correction factor (C.F.) = 46293.48148
- Total sum of square (SST) = 2188.51852

Compute the sum of squares for the horizontal analysis:

- Replication sum of square (SSD) = 1.4074
- Horizontal factor sum of square (S.S. (H.F.)) = 102.7407
- Horizontal factor error sum of square ( $SSE_a$ ) = 6.3704

Compute the sum of squares for the vertical analysis:

- Vertical factor sum of square (S.S. (V.F.)) = 2070.2963
- Vertical factor error sum of square ( $SSE_b$ ) = 1.43096

Compute the sum of squares for the interaction analysis:

- Interaction effect sum of square = 4.1477
- Interaction error sum of square ( $SSE_c$ ) = 2.37456

**Table 11:** ANOVA for strip plot design

Sv	D.f	Ss	Mss	F-Ratio	P-Value
Replication	2	1.4074	0.7037	2.263180	0.12596233
Brand(A)	2	102.7407	51.3704	32.25563	-
Error(E <sub>a</sub> )	4	6.3704	1.5926	-	-
Route(B)	2	2070.2963	1035.14815	2795.884156	0.00000000
Error(E <sub>b</sub> )	4	1.48096	0.37024	-	-
Interaction(A×B)	4	1.48096	1.036925	3.4934472	0.04526749
Error (E <sub>c</sub> )	8	4.1477	0.29682	-	-
Total	26	2.37456	-	-	-

The table values of replication and brand method are more significant than the calculated values. So the null hypothesis is accepted. There is no significant difference between the three replications and the three-route method. The table value of the route method is greater than the calculated value. So the null hypothesis is accepted. There is no significant difference between the three route methods. The table value of the interaction effect is also more effective than the calculated value. So the null hypothesis is accepted. There is no significant difference between the interaction effects.

The P-value of the above experiment is more significant than the 5% significance level. Therefore the null hypothesis is accepted. There is no significant difference that occurred in the above experiment.

## 5. Conclusion

Many real-world experiments deviate from textbook examples and sometimes involve multiple types of structures. Running agricultural and industrial tests in strip plot analysis is an effective method to reduce costs. The strip-plot design is the most efficient design in terms of both the resources required and the time required to study multi-step processes. This paper describes the construction and analysis of strip-plot analysis using some particular type of graphs through numerical examples from different fields, the hypothesis testing is compared by the strip-plot ANOVA method with the software using the method. When comparing the results of these methods, they produce the same results. Here some particular type of graphs is used to construct the SPD. In the future, there is an idea to expand this procedure to other experimental designs, such as Split-Split Plot Designs, Incomplete Block Designs etc.

## Reference

- [1] Horák, P., Qing, H. and Trotter, W. T. (1993). Induced matchings in cubic graphs. *Journal of Graph Theory*, 17(2): 151-160.
- [2] Greenlaw, R. and Petreschi, R. (1995). Cubic graphs. *ACM Computing Surveys (CSUR)*, 27(4): 471-495.
- [3] Robin J. Wilson, Introduction to Graph Theory, Fourth Edition, Addison Wesley Longman Limited, England, 1996.
- [4] Miller, A. (1997). Strip-plot configurations of fractional factorials. *Technometrics*, 39(2):153-161.
- [5] Milliken, G. A., Shi, X., Mendicino, M. and Vasudev, P. K. (1998). Strip-plot design for two-step processes. *Quality and reliability engineering international*, 14(4):197-210.
- [6] Billington, E. J. and Hoffman, D. G. (2003). Decomposition of complete tripartite graphs into regarious 4-cycles. *Discrete Mathematics*, 261(1-3): 87-111.

- [7] Arnouts, H., Goos, P. and Jones, B. (2010). Design and analysis of industrial strip-plot experiments. *Quality and Reliability Engineering International*, 26(2): 127-136.
- [8] J.A. Bondy U.S.R. Murty Graph theory, Springer International Edition, 2013.
- [9] Iványi, A., Pirzada, S. and Dar, F. A. (2015). Tripartite graphs with given degree set. *Acta Universitatis Sapientiae, Informatica*, 7(1): 72-106.
- [10] Douglas C. Montgomery Design and analysis of experiments, Fifth Edition, International Student Version, Arizona State University, 2016.
- [11] Abdollah Khodkar. (2018). Signed Edge Domination Numbers of Complete Tripartite Graphs: Part 2, *Australasian Journal Of Combinatorics*, 71(3):351–368.
- [12] Sheikh Rashid, Naveed Yaqoob, Muhammad Akram and Gulistan, (2018). Cubic graphs with Application, *International Journal of Analysis and Applications*, 16:5.
- [13] Velimor D. Almonte, Severino V. Gervacio and Emmanuel S. Natalio, (2019). Complete Tripartite Graphs with Spanning maximal planar Subgraphs, *Journal of the Mathematical Society of the Philippine*, 42(2):1-10.
- [14] Bradshaw, P. A. (2019). Triangle Packing on Tripartite Graphs Is Hard. *Rose-Hulman Undergraduate Mathematics Journal*, 20(1):7.
- [15] Santoso, K. A., Agustini, I. H., Prihandini, R. M., and Alfarisi, R. (2019). Vertex colouring using the adjacency matrix. In *Journal of Physics: Conference Series*, 1211:1, p. 012019.
- [16] Muhiuddin, G., Takallo, M. M., Jun, Y. B., and Borzooei, R. A. (2020). Cubic graphs and their application to a traffic flow problem. *International Journal of Computational Intelligence Systems*, 13(1):1265-1280.
- [17] Nisa, K., Hamsyiah, N., and Usman, M. (2020). Analysis of variance for strip plot design with missing values: bias correction of the mean squares. In *Journal of Physics: Conference Series*, 1524: 1, p. 012049.
- [18] Rashmanlou, H., Muhiuddin, G., Amanathulla, S. K., Mofidnakhahi, F., and Pal, M. (2021). A study on cubic graphs with novel application. *Journal of Intelligent & Fuzzy Systems*, 40(1):89-101.
- [19] Goos, P. (2022). The fish patty experiment: a strip-plot look. *Journal of Quality Technology*, 54(2): 236-248.

### List of Abbreviation

- Sv - Sources of variance  
D.f - Degrees of freedom  
Ss - Sum of squares  
Mss - Mean sum of squares  
R. (R) - Replication (R)  
H.F. (A) - Horizontal Factor(A)  
H.F.E. (a) - Horizontal Factor Error (a)  
V.F. (B) - Vertical Factor(B)  
V.F.E. (b) - Vertical Factor Error (b)  
I.E. (AB) - Interaction Effect (AB)  
I.E.(c) - Interaction Error(c).

# A HYBRID APPROACH TO SINGLE ACCEPTANCE SAMPLING PLANS FOR LIFETIMES MODELED WITH THE EXPONENTIAL RAYLEIGH DISTRIBUTION

<sup>1</sup>Nandhini M, <sup>2</sup>Radhika A\*, <sup>3</sup>Jeslin J, <sup>4</sup>Manigandan P



<sup>1</sup>Research Scholar, Department of Statistics, Periyar University, Salem-11,  
nandhinichithra9@gmail.com

<sup>2</sup>Assistant Professor, Department of Statistics, Periyar University, Salem-11  
radhisaran2004@gmail.com

<sup>3</sup>Research Scholar, Department of Statistics, Periyar University, Salem-11  
jeslin.statistics@gmail.com

<sup>4</sup>Department of Statistics, Periyar University, Salem-11  
srimanigandan95@gmail.com

## Abstract

*This article explores into the examination of a novel compound distribution termed the "Exponential Rayleigh distribution" in the context of truncated life testing within a sampling plan. It introduces a hybrid single acceptance sampling plan tailored for truncated life testing scenarios where the item's lifespan adheres to the Exponential Rayleigh distribution. One of the primary segments within the domain of product quality control is referred to as "sampling inspection by variables". This category encompasses procedures that involve the selection of multiple individual units based on measurements taken from a sample to assess a specific quality attribute under scrutiny. These plans, used to assess whether to accept or reject a submitted batch of items based on their observed lifetimes, are commonly known as reliability test plans. The article also outlines the development of a test plan to determine when to conclude the experiment given specific parameters like sample size, producer's risk, consumer's risk, and termination criteria. Sampling inspection, or reliability sampling, plays a pivotal role in maintaining product quality. It involves subjecting items to testing, collecting data on their lifespans, and making acceptance or rejection decisions based on the test results. When assessing an item's quality primarily based on its lifespan, which can be suitably described using a continuous probability distribution; such a plan is termed a "life test sampling plan." This article explores the application of the Exponential Rayleigh distribution within the realm of reliability sampling plans, emphasizing the utilization of hybrid censoring for life checks and median lifetime evaluations. This approach is leveraged to formulate reliability single sampling plans applicable to the Exponential Rayleigh distribution. The article utilizes binomial probabilities to compute the parameters of these sampling plans, aiming to strike a balance between protecting the interests of both the producer and the consumer while minimizing producer risks. The study involves calculating the specified median lifetime and determining design parameters like sample size and acceptance thresholds to meet predefined quality standards. The flexibility of the Exponential Rayleigh distribution in analyzing various types of lifetime data is highlighted, owing to its scale and shape parameters. To illustrate the concepts related to sampling strategies, a numerical example is provided in the sampling strategies section of the article.*

**Keywords:** Reliability Sampling, Median life-time, Hybrid Censoring, Exponential-Rayleigh Distribution.

## I. Introduction

Based on the examination of a sample of goods, sampling inspection plans are used to determine the appropriateness of batches that contain finished products. In reliability sampling plans, the lifespan of the tested items is a crucial element when determining the outcome of the batch after the testing of life. As a result, it may be appropriate to conclude a life test by setting a time limit and counting the number of failures that occur before the time limit. This is because the length of inspections may be a significant constraint. In manufacturing industries, the reliability sampling plan serves as a statistical tool for determining the allocation of lots by information gathered through a life check. This method requires a greater amount of sampling cost and inspection time compared to regular sampling plans. To make the inspection cost-effective, censoring schemes, such as Time censoring (Type-I), Product censoring (Type-II), and hybrid censoring, are used frequently throughout the life test. It is appropriate to develop reliability sampling plans with censoring methods when inspection time is constrained and inspection costs are minimal.

The two main distributions in life testing and reliability theory are the Exponential and Rayleigh distributions. They possess important structural properties and mathematical flexibility. One of the fundamental distributions in statistics theory and application is the exponential distribution. It has several important statistical characteristics, but its absence of memory property best describes it. But it shows excellent tractability in mathematics. As a result, the theory and uses of the exponential distribution are extensively covered in the literature [1]. When studying any lifespan data or skewed data, the three-parameter gamma and three-parameter Weibull distributions are frequently used. Both distributions have several favorable characteristics and intriguing physical explanations. Both have quite an amount of versatility for examining various forms of lifetime data because of the scale and shape factors [2]. A two-dimensional random vector of normal variables which has independent, identically distributed coordinates with mean zero is the basis of the Rayleigh distribution, which bears Lord Rayleigh's name. Numerous scenarios where the magnitudes of normal variables are crucial can be addressed using this distribution. A function that appears in the Maximum Likelihood equation is approximated using a hyperbolic approximation rather than a linear approximation in a Modified Maximum Likelihood Estimate of the scale parameter of the Rayleigh distribution [3]. Since they reduce the amount of time and resources needed for testing, these sample programs are very helpful to practitioners, the ability, at various phases of the experiment, to exclude functional test specimens from further testing [4]. Acceptance sampling, also known as sampling inspection is a crucial quality control technique that outlines the policies and steps for deciding whether to accept or reject a batch of goods based on the examination of one or more samples. Consideration is given to the Burr (XII) distribution's application in the reliability sampling plan. Utilizing a set of simulated observations from the Burr (XII) distribution, the evaluation of such plan was discussed [5]. In the industrial sector, reliability sampling plans are used to make disposition decisions for batches based on product life testing. These plans are created while taking into account pertinent probability distributions for the lifespans of the tested products [6]. A new single sampling plan based on ranked data scheme for generalized exponential distribution using median ranked set sampling [7].

The objective of this research is to establish dependable sampling plans based on exponential Rayleigh distribution employing a hybrid censoring scheme that corresponds to producer's and consumer's risk levels. A Lifetime of products follows a specific behavior that is described by a probability distribution. Estimation and inferential part of the developed theory of statistics is the key interest of the researcher and this is fulfilled with the help of these distributions [8]. According to the criteria of the exponential Rayleigh distribution, the study produces the Operating Characteristic (OC) function of the Reliability Single Sampling Plan (RSSP) in part 2. In part 3, it is explained how to create and use the sampling plans. Moreover, part 4 discusses the development of tables that provide optimal sampling plans for certain situations. An example is given to illustrate the selection of a sampling plan. Part 5 summarizes the outcomes of the study.

## II. The Theoretical Perspective on the Rayleigh Distribution

The Rayleigh distribution is a continuous probability distribution widely used in probability theory and statistics, particularly for random variables with non-negative values. It has a connection with the chi distribution, specifically when having two degrees of freedom, albeit involving rescaling. This distribution is named after Lord Rayleigh. It frequently appears when analyzing the overall magnitude of a vector in a plane in relation to its directional components. For instance, in the two-dimensional analysis of wind velocity, the Rayleigh distribution naturally emerges when each component has zero mean, equal variance, and follows a normal distribution. Another scenario where the Rayleigh distribution is relevant is in the context of random complex numbers. When real and imaginary components are independently and identically distributed as Gaussian with equal variance and zero mean, the absolute value of the complex number follows a Rayleigh distribution.

In the field of Magnetic Resonance Imaging (MRI), Rayleigh distribution is applied. MRI images are often interpreted as magnitude images, although they are recorded as complex images. Consequently, the background data in MRI images follows a Rayleigh distribution, allowing for the estimation of noise variance in MRI images using this method. Furthermore, the Rayleigh distribution has found application in the field of nutrition. It has been employed to establish connections between dietary nutrient levels and the physiological responses of both humans and animals. This approach represents a method for computing nutritional response relationships through the utilization of Rayleigh distribution parameters."

## III. Operating Characteristics of RSSPs under Exponential Rayleigh distribution

One technique is known as a single sampling plan for reliability to make decisions about submitted lots by testing randomly selected items from the lot. Mean life is used as a quality metric to calculate the probability of acceptance to determine design parameters like sample size 'n' and acceptance number 'c' [9]. This plan is characterized by four parameters (N, n, c, t), which include the lot size (N), sample size (n), acceptance number (c), and test termination time (t). The implementation of the sampling plan involves using these parameters to make decisions about the lot. Choose a random selection of n products from the submitted lot of size N.

- (1) The supplied lot of size N should be randomly selected to yield a set of n products.
- (2) Execute a life test on the chosen items with t as the test termination time. Count the number of things that failed, X=x.
- (3) If X>C or time t, whichever occurs first, the life test should be terminated.
- (4) Accept the lot, if x≤ c at time t; reject the lot if x>c either at time t or earlier.

Let T be the product's lifespan; it will be distributed using an exponential Rayleigh distribution with a probability density function (PDF)

$$f(x) = \lambda \beta x e^{\frac{\beta}{2}x^2} \cdot e^{-\lambda(e^{\frac{\beta}{2}x^2}-1)} \quad x \in R; \lambda, \beta > 0 \quad (1)$$

The specifications for the scale and shape are indicated here by  $\lambda$  and  $\beta$ . Following are the formulas for the exponential Rayleigh distribution's cumulative distribution function.

$$F(x) = 1 - \lambda e^{-\lambda(e^{\frac{\beta}{2}x^2}-1)} \quad x \in R; \lambda, \beta > 0 \quad (2)$$

With median respectively,

$$m = \sqrt{\frac{2}{\beta} \log \left( 1 + \left( \frac{-\log(1-\mu)}{\lambda} \right) \right)}$$

Estimate to a parameter  $\beta$  respectively

$$\beta = \frac{2}{m^2} \log \left( \frac{1 - \log \left( \frac{1}{2} \right)}{\lambda} \right)$$

From each value of  $1/m$ , the lot fraction non-conforming,  $p$  may be computed.

$$F(X) = F\left(\frac{1}{m}\right) = p$$

Utilizing their OC functions, a sampling plan's effectiveness can be evaluated. A sample plan's OC function is defined by

$$P_a = P(X \leq c) = \sum_{x=0}^c P(X = x)$$

It is reasonable that the probability distribution for  $X$  follows a hyper geometric distribution. The probability distribution of  $X$  can be assumed approximately as hyper geometric distribution. When  $N$  is large, the sampling distribution of  $X$  can be approximated by the Binomial ( $n, p$ ) distribution [10]. Under these circumstances, here, it is proposed that

$$P_a(p) = \sum_{x=0}^c n c_x p^x q^{n-x}$$

#### IV. Plan Parameter determination under the conditions of the Exponential Rayleigh Distribution

Utilizing the OC function stipulated by the Binomial probability distribution, the best reliability single sampling plans are identified under the circumstances of the ER ( $\lambda, \theta$ ) distribution. A modified maximum likelihood estimate for Rayleigh distribution using hyperbolic approximation [11]. A sample strategy is typically developed so that it simultaneously protects the manufacturer and the customer. By designating two points on the OC curve, namely ( $p_1, 1-\alpha$ ) and ( $p_2, \beta$ ), the protection of the producer and the customer is guaranteed. In this case,  $p_1$  stands for the acceptable quality level, for producer risk,  $p_2$  for restricting quality level, and for consumer risk. It is possible to determine an ideal RSSP for points meeting the following criteria.

$$P_a(p_1) \geq 1 - \alpha$$

and

$$P_a(p_2) \leq \beta$$

These conditions may be written as

$$\sum_{x=0}^c n c_x p_1^x q_1^{n-x} \geq 1 - \alpha \quad (3)$$

and

$$\sum_{x=0}^c n c_x p_2^x q_2^{n-x} \leq \beta \quad (4)$$

To find the best values of  $n$  and  $c$  subject to (3) and (4), various techniques may be used. Finding the plan parameters involve using the iterative process outlined below. Therefore, the ideal values of the plan parameters  $n$  and  $c$  for given,  $\lambda, t, m_1, m_2, \alpha, \beta$ , and may be found as follows:

- (1) When  $m_1 > m_2$  with the required values of  $m_1$  and  $m_2$ , calculate  

$$\beta_1 = \frac{2}{m_1^2} \log\left(\frac{1 - \log\left(\frac{t}{2}\right)}{\lambda}\right)$$
 and  $\beta_2 = \frac{2}{m_2^2} \log\left(\frac{1 - \log\left(\frac{t}{2}\right)}{\lambda}\right)$
- (2) Corresponding to  $t$ ,  $\beta_1$  and  $\beta_2$ , determine  $p_1 = F_T(1/m_1)$  and  $p_2 = F_T(1/m_2)$
- (3) Set  $c=0$
- (4) Find the largest  $n$ , say  $n_L$ , such that  $P_a(p_1) \geq 1 - \alpha$
- (5) Find the smallest  $n$ , say  $n_S$ , such that  $P_a(p_2) \leq \beta$
- (6) If  $n_S \leq n_L$ , then the optimum plan is  $(n_S, c)$ ; otherwise increase  $c$  by 1.
- (7) Repeat Steps 4 through 6 until optimum values of  $n$  and  $c$  are obtained.

By the hybrid censoring systems covered in part 2 and after figuring out  $n$  and  $c$ , a submitted lot may undergo sample inspection.

## V. Construction of Tables

Binomial probabilities are used to calculate the values of  $n$  and  $c$  for the best reliability sampling plans for various combinations of  $\lambda$ ,  $t$ ,  $m_1$ ,  $m_2$ ,  $\alpha$ , and  $\beta$ . Plans for acceptance sampling from exponential populations that use the lifetime-performance index both with and without censoring [12]. Both the producer's risk and the consumer's risk are taken into account at two distinct levels, such as  $\alpha=0.05$ ,  $0.05$  and  $\beta=0.05$ ,  $0.10$  respectively. The producer's expectations for the mean lifetime of the products are considered as  $m_1=6000$ ,  $7000$ ,  $8000$ ,  $9000$ , and  $10000$  hours respectively. Assumed values for the shape parameter  $\lambda$  and the test termination times  $t$  are  $300$ ,  $450$ , and  $600$  hours and  $\lambda=1$  correspondingly. The consumer's projected mean product lifespan is taken as  $m_2= 1000$ ,  $1500$ ,  $2000$ ,  $2500$ ,  $3000$ ,  $3500$ , and  $4000$  hours respectively. Tables 1 through Table 3 give the  $n$  and  $c$  values for the best reliability sampling strategies. Each cell entry  $(n, c)$  in every table reflects the ideal value of the pair  $(n, c)$  that corresponds to the given values of  $\lambda$ ,  $t$ ,  $m_1$ ,  $m_2$ ,  $\alpha$ , and  $\beta$ . choosing a plan from these for certain requirements is illustrated in the following example.

### Illustration

Let  $ER(1, \beta)$  is distributing the lifetime of the products that have been submitted for inspection. The average lifespan of products that live up to producer and customer expectations is, respectively,  $m_1=6000$  hours and  $m_2=4000$  hours. Let's say the quality inspector instructs the life test to be censored at  $t=300$  hours. The values of the limiting quality level and the acceptable quality level can therefore be calculated as  $p_1=0.0013$  and  $p_2=0.0029$ , respectively. The plan parameters can be calculated using the binomial probabilities from Table 1 as  $n=8200$  and  $c=16$  if the producer's risk and the consumer's risk are  $\alpha=0.05$  and  $\beta=0.05$ , respectively.

Now, the inspection of the lot-by-lot sampling based on the life test can be done as follows: The submitted lot may have up to  $8200$  products randomly chosen as a sample. All of the sampled goods are eligible for life testing. The life test may be stopped if there have been  $16$  failures or fewer after  $300$  hours. The lot might be taken. However, if the seventeenth failure happens before  $t=300$  hours, the life test should be stopped. The lot could be disregarded.



**Table 1:** Parameters of RSSPs under the conditions of ER ( $\beta, \lambda=1$ ) Distribution with  $\alpha=0.05, \lambda=1$  and  $t=300$  hours.

t=300, $\lambda=1$		$m_1$	6000	7000	8000	9000	10000
		t/ $m_1$	0.05	0.0428	0.0375	0.0333	0.03
$m_2$	t/ $m_2$	P1	0.0013	0.0009	0.0007	0.0005	0.0004
		P2					
1000	0.3	0.0473	(81,1)	(48,0)	(48,0)	(48,0)	(48,0)
			(99,1)	(99,1)	(62,0)	(62,0)	(62,0)
1500	0.2	0.0210	(184,1)	(184,1)	(184,1)	(184,1)	(184,1)
			(224,1)	(224,1)	(224,1)	(224,1)	(224,1)
2000	0.15	0.0118	(448,2)	(327,1)	(327,1)	(327,1)	(327,1)
			(530,2)	(530,2)	(399,1)	(399,1)	(399,1)
2500	0.12	0.0075	(880,3)	(701,2)	(701,2)	(512,1)	(512,1)
			(1021,3)	(829,2)	(829,2)	(829,2)	(624,1)
3000	0.1	0.0052	(1760,5)	(1267,3)	(1010,2)	(1010,2)	(738,1)
			(2246,6)	(1736,4)	(1471,3)	(1194,2)	(1194,2)
3500	0.0857	0.0038	(3357,8)	(2396,5)	(1725,3)	(1375,2)	(1375,2)
			(4057,9)	(3059,6)	(2364,4)	(2002,3)	(1626,2)
4000	0.075	0.0029	(6398,13)	(3972,7)	(3130,5)	(2254,3)	(2254,3)
			(8200,16)	(5299,9)	(3996,6)	(3088,4)	(2616,3)

In each cell, the first pair is the value of (n, c) corresponding to ( $\alpha=0.05, \beta=0.10$ ) and the Second pair corresponding to ( $\alpha=0.05, \beta=0.05$ ).

**Table 2:** Parameters of RSSPs under the conditions of ER ( $\beta, \lambda=1$ ) Distribution with  $\alpha=0.05, \lambda=1$  and  $t=450$  hours.

t=450, $\lambda=1$		$m_1$	6000	7000	8000	9000	10000
		t/ $m_1$	0.075	0.0642	0.0562	0.05	0.045
$m_2$	t/ $m_2$	P1	0.0029	0.0021	0.0016	0.0013	0.0010
		P2					
1000	0.45	0.1064	(36,1)	(21,0)	(21,0)	(21,0)	(21,0)
			(43,1)	(43,1)	(27,0)	(27,0)	(27,0)
1500	0.3	0.0473	(81,1)	(81,1)	(81,1)	(81,1)	(48,0)
			(99,1)	(99,1)	(99,1)	(99,1)	(99,1)
2000	0.225	0.0266	(199,2)	(145,1)	(145,1)	(145,1)	(145,1)
			(235,2)	(235,2)	(177,1)	(177,1)	(177,1)
2500	0.18	0.0170	(390,3)	(311,2)	(311,2)	(227,1)	(227,1)
			(453,3)	(367,2)	(367,2)	(367,2)	(277,1)
3000	0.15	0.0118	(781,5)	(563,3)	(448,2)	(448,2)	(327,1)
			(997,6)	(770,4)	(653,3)	(530,2)	(530,2)
3500	0.1285	0.0087	(1491,8)	(1064,5)	(766,3)	(610,2)	(610,2)
			(1801,9)	(1358,6)	(1049,4)	(889,3)	(722,2)
4000	0.1125	0.0066	(2842,13)	(1764,7)	(1390,5)	(1001,3)	(1001,3)
			(3643,16)	(2354,9)	(1774,6)	(1371,4)	(1162,3)

In each cell, the first pair is the value of (n, c) corresponding to ( $\alpha=0.05, \beta=0.10$ ) and the Second pair corresponding to ( $\alpha=0.05, \beta=0.05$ ).

**Table3:** Parameters of RSSPs under the conditions of ER ( $\beta, \lambda=1$ ) Distribution with  $\alpha=0.05, \lambda=1$  and  $t=600$  hours.

t=600, $\lambda=1$		$m_1$	6000	7000	8000	9000	10000
		t/ $m_1$	0.1	0.0857	0.075	0.0667	0.06
$m_2$	t/ $m_2$	P1	0.0052	0.0386	0.0029	0.0023	0.0018
		P2					
1000	0.6	0.1883	(20,1) (24,1)	(12,0) (24,1)	(12,0) (15,0)	(12,0) (15,0)	(12,0) (15,0)
1500	0.4	0.0841	(45,1) (55,1)	(45,1) (55,1)	(45,1) (55,1)	(45,1) (55,1)	(27,0) (55,1)
2000	0.3	0.0473	(111,2) (131,2)	(81,1) (131,2)	(81,1) (99,1)	(81,1) (99,1)	(81,1) (99,1)
2500	0.24	0.0303	(219,3) (254,3)	(174,2) (206,2)	(174,2) (206,2)	(127,1) (206,2)	(127,1) (155,1)
3000	0.2	0.0210	(439,5) (497,5)	(316,3) (433,4)	(252,2) (366,3)	(252,2) (297,2)	(184,1) (297,2)
3500	0.1714	0.0154	(838,8) (1012,9)	(598,5) (763,6)	(430,3) (589,4)	(343,2) (499,3)	(343,2) (405,2)
4000	0.15	0.0118	(1598,13) (2047,16)	(992,7) (1323,9)	(781,5) (997,6)	(563,3) (770,4)	(563,3) (653,3)

In each cell, the first pair is the value of (n, c) corresponding to ( $\alpha=0.05, \beta=0.10$ ) and the second pair corresponding to ( $\alpha=0.05, \beta=0.05$ ).

## VI. Conclusion

In this article, a new sampling distribution is introduced for testing product quality when conducting acceptance sampling for life tests that follow the Exponential Rayleigh distribution. The paper also outlines reliability sampling plans for conducting life tests through hybrid censoring, specifically for products that follow the Exponential Rayleigh distribution. These plans have been designed to protect the interests of both the producer and consumer and the use of hybrid censoring helps to reduce the amount of time required for implementation. The article also includes tables that provide optimal plans for certain specified strengths.

## References

- [1] Balakrishnan and Basu (1995). The Exponential Distribution: Theory, Methods and Applications. *Taylor and Francis Group an informa business*.
- [2] Gupta D and Kundu (2002). Theory & Methods: Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, 41, 173-188.
- [3] Kaviyarasu V and Sivasankari S (2020). Acceptance sampling Plan for Life Testing under Generalized Exponential-Poisson Distribution. *International Journal of Mathematics Trends and Technology*, 66,148-156.
- [4] Uditha Balasooriya and Sutaip L.C. Saw (1998). Reliability sampling plans for the two-parameter exponential distribution under progressive censoring. *Journal of Applied Statistics*, 707-14.
- [5] Vijayaraghavan R, Saranya C R and Sathya Narayana Sharma K (2021). Reliability Single Sampling Plans under the assumption of burr type XII distribution. *RT and A*, 4(65) -16.
- [6] Loganathan M and Gunasegaran (2017). Determination of Reliability Single Sampling Plans Based on Exponentiated Distribution. *Global and Stochastic Analysis*, 4, 111-118..
- [7] Amjad D. Al-Nasser and Fatima S. Abdullah (2017). On Using the Median Ranked Set Sampling for Developing Reliability Test Plans Under Generalized Exponential Distribution.

*Pakistan Journal of Statistics and Operation Research*, XIII 4, 757-774.

- [8] Abhimanyu Singh Yadav and Mahendra Saha (2021). Reliability Test Plan Based on Logistic-Exponential Distribution and Its Application. *Journal of Reliability and Statistical Studies*, 14, 695-724.
- [9] Aisha Fayomi and Khushnoor Khan (2022). A group acceptance sampling plan for another generalized transmuted exponential distribution based on truncated lifetimes. *Quality and Reliability Engineering International*.
- [10] Weiss (2010). Binomial Approximation to the hyper geometric Distribution. *Statistics and Geospatial Data Analysis (SOGA), Basics of Statistics*.
- [11] Lalitha S and Anand Mishra (2011). Modified maximum likelihood estimation for Rayleigh distribution. *Communication in Statistics-Theory and Methods*, 25.
- [12] Chien-Wei wu, Ming-Hung Shu and Yu-Ning Chang (2018). Variable-sampling plans based on lifetime-performance index under exponential distribution with censoring and its extensions. *Applied Mathematical Modeling*, 55 (81-93).
- [13] Mahesh Kumar and Ramyamol P C (2016). Design of optimal reliability acceptance sampling plans for exponential distribution. *Economic Quality Control*.
- [14] Muhammad Aslam Debasis Kundu and Munir Ahmad (2010). Time truncated acceptance sampling plans for generalized exponential distribution. *Journal of Applied Statistics*.
- [15] Sampath and Lalitha (2016). Economic Reliability Test Plan under Hybrid Exponential Distribution. *International Journal of Computational and Theoretical Statistics*.
- [16] Srinivasa Rao G (2009). A Group Acceptance Sampling Plans for Lifetimes Following a Generalized Exponential Distribution. *Economic Quality Control*, 24, 75-85.
- [17] Srinivasa Rao G and Ghitany M E (2011). An Economic Reliability Test Plan for Marshall-Olkin Extended Exponential Distribution. *Applied Mathematical Sciences*, 5,103-112.
- [18] Srinivasa Rao (2011). A Hybrid Group Acceptance Sampling Plans for Lifetimes Based on Generalized Exponential Distribution. *Nature*, 11(12), 2232-2237.
- [19] Sukhdev Singh and Yogesh Mani Tripathi (2015). Sampling Plans Based on Truncated Life Test for a Generalized Inverted Exponential Distribution. *Industrial Engineering and Management Systems*. 14, 183-195.
- [20] Syed Adil Hussain and Ishfaq Ahmad (2021). Mean ranked acceptance sampling plan under exponential distribution. *Aim Shams Engineering Journal*. 12, 4125-4131.

# CHARACTERIZATION OF SOME CONTINUOUS DISTRIBUTIONS BY CONDITIONAL VARIANCE OF RECORD VALUES

ZAKI ANWAR<sup>1</sup>, MOHD FAIZAN<sup>2</sup>, ZAKIR ALI<sup>3</sup>

•

<sup>1</sup>Department of Statistics and Operations Research (Women's College),  
Aligarh Muslim University, Aligarh, India.

<sup>2,3</sup>Department of Statistics and Operations Research,  
Aligarh Muslim University, Aligarh, India.

<sup>1</sup>zakistats@gmail.com

<sup>2</sup>mdfaizan02@gmail.com

<sup>3</sup>Corresponding Author: zakirali56656@gmail.com

## Abstract

*Characterization of a probability distribution gives a unique property enjoyed by that distribution. Various approaches are available in the literature to characterize distributions through record values. Many researchers have characterized Exponential, Pareto, and Power function distributions using moments, conditional expectation, and some other characteristics of record values. In this paper, we have characterized these three distributions through conditional variance of adjacent record values. The results have been verified using numerical computation.*

**Keywords:** Characterization of continuous distributions, conditional variance, record values.

## 1. INTRODUCTION

Let  $X_1, X_2, \dots$  be a sequence of independent, identically distributed random variables with distribution function  $(df)F(x)$  and probability density function  $(pdf)f(x)$ . Let  $X_{U(r)}$  be the  $r$ th upper record value, then the conditional  $pdf$  of  $X_{U(r+1)}$  given  $X_{U(r)} = x, 1 \leq r < s$  is given by (Ahsanullah, 2004)[1]

$$f\left(X_{U(r+1)} = y \mid X_{U(r)} = x\right) = \frac{f(y)}{\bar{F}(x)} \quad (1.1)$$

where  $\bar{F}(x) = P(X > x) = 1 - F(x)$ .

One can transform the upper record into lower record values by replacing the original sequence of  $(X_j)$  by  $(-X_j, j \geq 1)$  (Ahsanullah, 2004) [1]. Let  $X_{L(r)}$  be the  $r$ -th lower record value, then the conditional  $pdf$  of  $X_{L(r+1)}$  given  $X_{L(r)} = x, 1 \leq r < s$  is given by

$$f\left(X_{L(r+1)} \mid X_{L(r)} = x\right) = \frac{f(y)}{F(x)}. \quad (1.2)$$

The record values have been extensively studied in literature. For an excellent review, one may refer to Ahsanullah (2004) [1]. Arnold et al. (1998) [2] and Nevzorov (2001) [3] amongst others. Characterization of distributions through conditional expectations of record values have been considered, among others, by Nagaraja, H.N. and Nevzorov, V.B. (1997) [4], Franco and Ruiz(1997)

[5], Athar et al. (2003) [6], Khan et al. (2010) [7] and Faizan and Khan (2011) [8].

Beg, M.I. and Kirmani. S.N.U.A. (1978) [9] characterized exponential distribution by a weak homoscedasticity. Khan and Beg (1987) [10] extended the result of Beg and Kirmani (1978) for Weibull distribution. Khan et al. (2008) [11] characterized a general class of distribution by conditional variance of order statistics and Shah et al. (2018) [12] characterized Pareto and power function distributions by conditional variance of order statistics, In this paper we have characterized exponential, Pareto and power function distributions by conditional variance of record values.

## 2. CHARACTERIZATION RESULTS

**Theorem 2.1:** Let  $x$  be a random variable with df  $F(x)$  and  $E(X^2) < \infty$ . Then for  $r < s$

$$V[X_{x(r+1)} | X_{v(r)} = x] = \theta^2 \quad (2.1)$$

for some  $\theta > 0$  if and only if

$$\bar{F}(x) = e^{-re}; \quad x > 0. \quad (2.2)$$

**Proof:** First we will prove (2.2) implies (2.1). It is easy to see that from (1.1) and (2.2)

$$E[X_{u(r+1)} | X_{u(r)} = x] = x + \theta \quad (2.3)$$

and

$$E[X_{u(r+1)}^2 | X_{u(r)} = x] = x^2 + 2x\theta + 2\theta^2 \quad (2.4)$$

Now, using (2.3) and (2.4), we have

$$V[X_{u(r+1)} | X_{u(r)} = x] = \theta^2$$

For sufficiency part, we have from (2.2)

$$\int_x^\infty y^2 \frac{f(y)}{\bar{F}(x)} dy - \left( \int_x^\infty y \frac{f(y)}{\bar{F}(x)} dy \right)^2 = \theta^2$$

$$\bar{F}(x) \int_x^\infty y^2 f(y) dy - \left( \int_x^\infty y f(y) dy \right)^2 = \theta^2 \bar{F}^2(x) \quad (2.5)$$

Differentiating (2.5) twice w.r.t.  $x$  and simplifying, we get

$$\int_x^\infty y f(y) dy = x\bar{F}(x) + \theta^2 f(x) \quad (2.6)$$

Now differentiate (2.6) again w.r.t.  $x$ , we get

$$\bar{F}(x) = -\theta^2 f'(x)$$

and hence the result.

**Theorem 2.2:** Let  $X$  be a random variable with  $dfF(x)$  and  $E(X^2) < \infty$ . Then, for some  $r < s$  and  $0 < p < 1$ , we have

$$V[X_{u(r+1)} | X_{u(r)} = x] = \frac{p}{(p-2)(p-1)^2} x^2 \quad (2.7)$$

if and only if

$$\bar{F}(x) = \left(\frac{\alpha}{x}\right)^p; \quad \alpha \leq x < \infty. \quad (2.8)$$

**Proof:** First we will prove (2.8) implies (2.7). By using (1.1) and (2.8), it is easy to show that

$$E \left[ X_{u(r+1)} \mid X_{u(r)} = x \right] = \frac{p}{p-1}x$$

and

$$E \left[ X_{u(r+1)}^2 \mid X_{u(r)} = x \right] = \frac{p}{p-2}x^2$$

which gives

$$V \left[ X_{u(r+1)} \mid X_{u(r)} = x \right] = \frac{p}{(p-2)(p-1)^2}x^2.$$

Now, to prove (2.7) implies (2.8), we have using (1.1) and (2.7)

$$\bar{F}(x) \int_x^\infty y^2 f(y) dy - \left( \int_x^\infty y f(y) dy \right)^2 = cx^2 \bar{F}^2(x) \quad (2.9)$$

where

$$c = \frac{p}{(p-2)(p-1)^2}.$$

Differentiating (2.9) twice w.r.t.  $x$  and simplifying, we get

$$\int_x^\infty y^2 f(y) dy - 2x \int_x^\infty y f(y) dy = (2c-1)x^2 \bar{F}(x) - 2cx \frac{\bar{F}^2(x)}{f(x)}. \quad (2.10)$$

Now, after differentiating (2.10) w.r.t.  $x$ , we get

$$\int_x^\infty y f(y) dy = cx^2 f(x) - (4c-1)x \bar{F}(x) + c \frac{\bar{F}^2(x)}{f'(x)} - cx \frac{\bar{F}^2(x) f'(x)}{f^2(x)}. \quad (2.11)$$

Again differentiating (2.11), we get

$$\begin{aligned} & 2x \frac{\bar{F}(x) f'^2(x)}{f^3(x)} - 2 \frac{\bar{F}(x) f'(x)}{f^2(x)} - x \frac{\bar{F}(x) f''(x)}{f'^2(x)} + 2x \frac{f'(x)}{f(x)} \\ & + x^2 \frac{f'(x)}{\bar{F}(x)} + 6x \frac{f(x)}{\bar{F}(x)} - 6 + \frac{1}{c} = 0. \end{aligned}$$

Let  $\frac{\bar{F}(x)}{F(x)} = y = y(x)$  bearing in mind that  $f(x) = F'(x)$ ,  $f'(x) = F''(x)$ ,  $f''(x) = F'''(x)$ ,  $\frac{\bar{F}''(x)}{F(x)} = y' + y^2$ ,  $\frac{\bar{F}'''(x)}{F(x)} = y'' + 3yy' + y^3$ , we get

$$x \frac{y'' + 3yy' + y^3}{y^2} - 2x \frac{(y' + y^2)^2}{y^3} - \left( x^2 - \frac{2}{y^2} - \frac{2x}{y} \right) (y' + y^2) - 6xy + p^2 - 4p - \frac{2}{p} - 1 = 0. \quad (2.12)$$

There exists a unique solution of the differential equation (2.12) that satisfies the prescribed initial conditions

that  $y'(a) = -\frac{p}{a^2}$

and

$$y(a) = \frac{p}{a}$$

where  $a$  is any finite point in the support of  $F$ . Thus by the existence and uniqueness theorem (Boyce and Diprima, 2012) [13], we get

$$\frac{\bar{F}'(x)}{\bar{F}(x)} = y = -\frac{p}{x}$$

which implies that

$$\bar{F}(x) = (Ax)^{-p}; \quad \alpha \leq x < \infty.$$

where  $A$  is a constant to be determined and hence the Theorem.

**Theorem 2.3:** Let  $X$  be a random variable with  $dfF(x)$  and  $E(X^2) < \infty$ . Then for  $r < s$

$$V[X_{L(r+1)} | X_{L(r)} = x] = \frac{p}{(p+2)(p+1)^2} x^2$$

if and only if

$$F(x) = \left(\frac{x}{\beta}\right)^p; \quad 0 \leq x < \beta < \infty. \quad (2.13)$$

**Proof:** This can be proved on lines of Theorem 2.2

**Table 1:** Verification of the characterization results in case of Exponential distribution.

$\theta$	$X$	L.H.S.	R.H.S.	$ L.H.S. - R.H.S.  $	$\left  \frac{L.H.S. - R.H.S.}{R.H.S.} \right $
1.5	0.4	2.2499	2.25	0.0001	0.00005
2.5	0.8	6.2497	6.25	0.0003	0.00005
4.5	1.6	20.2504	20.25	0.0004	0.00002
5.5	2.0	30.2454	30.25	0.0046	0.00015
6.5	2.4	42.2430	42.25	0.0070	0.00017
7.5	2.8	56.2436	56.25	0.0064	0.00011
8.5	3.2	72.2603	72.25	0.0103	0.00014
9.5	3.6	90.2302	90.25	0.0198	0.00022
10.5	4.0	110.2485	110.25	0.0015	0.00001
11.5	4.4	132.2380	132.25	0.0120	0.00009
12.5	4.8	156.1910	156.25	0.0590	0.00038
13.5	5.2	182.2197	182.25	0.0303	0.00017
14.5	5.6	210.2026	210.25	0.0474	0.00023
15.5	6.0	240.2801	240.25	0.0301	0.00013

**Table 2:** Verification of the characterization results in case of Pareto distribution.

$\alpha$	$p$	$X$	L.H.S.	R.H.S.	$ L.H.S. - R.H.S.  $	$\left  \frac{L.H.S. - R.H.S.}{R.H.S.} \right $
0.3	3	0.2156	0.0347	0.0348	0.0001	0.0029
0.7	4	0.5797	0.0745	0.0747	0.0002	0.0027
1.1	5	0.8523	0.0756	0.0757	0.0001	0.0013
1.5	6	1.1692	0.0838	0.0820	0.0018	0.0220
1.9	7	1.4536	0.0699	0.0822	0.0123	0.1496
2.3	8	1.7510	0.0836	0.0834	0.0151	0.1530
2.7	9	2.0642	0.0943	0.0856	0.0087	0.1016
3.1	10	2.3171	0.0829	0.0812	0.0017	0.0209
3.5	11	2.6390	0.0850	0.0851	0.0001	0.0011
3.9	12	2.9604	0.0878	0.0869	0.0009	0.0103
4.3	13	3.2612	0.0847	0.0873	0.0026	0.0002
4.7	14	3.5998	0.0899	0.0895	0.0004	0.0045
5.1	15	3.8431	0.0862	0.0869	0.0007	0.0080
5.5	16	4.2212	0.0871	0.0905	0.0034	0.0376
5.9	17	4.5505	0.0917	0.1147	0.0230	0.2005

**Conclusions:**

This paper introduces a study of the Exponential, Pareto, and Power function distributions, showcasing their characterizations based on the conditional variance of adjacent record values. The validity of our findings has been confirmed through some numerical computation.

#### REFERENCES

- [1] M. Ahsanullah, Record values-theory and applications. University Press of America, 2004.
- [2] B. C. Arnold, N. Balakrishnan and H.N. Nagaraja, Records. John Wiley & Sons, 2011.
- [3] V. Nevzorov, "Records: Mathematical theory. transl. math," 2001
- [4] H. Nagaraja and V. Nevzorov, "On characterizations based on record values and order statistics," Journal of statistical planning and inference, vol. 63, no. 2, pp.271-284, 1997.
- [5] M. Franco and J.M. Ruiz, "On characterizations of distributions by expected values of order statistics and record values with gap," Metrica, vol. 45, pp. 107-119, 1997.
- [6] H. Athar, M. Yaqub, and H. Islam, "On characterization of distribution through linear regression of record values and order statistics," Aligarh J. Statist, vol. 23, pp. 97-105, 2003.
- [7] A. H. Khan, M. Faizan, and Z. Haque, "Characterization of continuous distributions through record statistics," Communications of the corean Mathematical society, vol. 25, no. 3, pp. 485-489, 2010.
- [8] M. Faizan and M. Khan, "A characterization of continuous distributions through lower record statistics," in Prob Stat Forum, vol. 4, pp. 39-43, 2011.
- [9] M. Beq and S. Kirmani, "Characterization of the exponential distribution by a weak donoscedasticity," Communications in Statistics- Theory and Methods, vol. 7, no. 3, pp. 307-310, 1978.
- [10] A. Khan and M. Beg, "Characterization of the weibull distribution by conditional variance," Sankhya: The Indian Journal of Statistics, Series A, pp. 268-271, 1987.
- [11] A. Khan, M. Faizan, and M. Khan, " Chacterization of continuous distributions by conditional variance of order statistics," Calcutta Statistical Association Bulletin, vol. 60. no. 3-4, pp. 235-224, 2008.
- [12] I. A. Sah, H. M. Barakat, and A. H. Khan, "Characterization of pareto and power function distributions by conditional variance of order statistics," CR Acad. Bulg. Sci, vol. 71, no. 3, pp. 313-316, 2018.
- [13] W. E. Boyce, R. C. DiPrima, and D. B. Meade, Elementary differential equations and boundary value problems. John Wiley & Sons, 2021.



# SOME INFERENTIAL ASPECTS ON THREE-STATION TANDEM QUEUE

AMBILY JOSE<sup>1</sup>, AGNES JEROME<sup>1</sup>, M. R. IRSHAD<sup>1</sup>



<sup>1</sup>Cochin University of Science and Technology, Kochi-22, Kerala.  
ambilyjose@cusat.ac.in, agnesjerome2000@gmail.com, irshadmr@cusat.ac.in

## Abstract

*Considered is a three-station tandem queue with service times at stations 1, 2, and 3 are exponentially distributed with customers arriving according to the Poisson process at station 1. Given that the stationary distribution is the product of three independent geometric distributions with the intensity parameters, maximum likelihood estimators and Bayes estimators of the intensity parameters based on the number of customers present at different time periods are obtained. Furthermore, the minimal posterior risk and minimum Bayes risk of the estimators are computed. Also, a simulation study is conducted to evaluate the performance of the estimators obtained.*

**Keywords:** Three-station tandem queue, Classical inference, Bayesian inference, MCMC sampling

## 1. INTRODUCTION

Most works on queuing models are restricted to deriving the formulations for transient or stationary (steady state) solutions and do not take into account the related statistical inference issues. Some of the crucial tools to understanding any random phenomenon using stochastic models are classical inference and Bayesian inference. The past has not paid much attention to the analysis of queuing systems in all these directions. Standard parametric models are highly suitable whenever the systems are completely observable in terms of their fundamental random components, such as inter-arrival times and service times.

Estimation of the parameters associated with the queueing models are integral part of queueing theory. Frequently, previous experiments or analyses of the inter-arrival time or service time data have revealed some information about the parameters of the distributions of inter-arrival time or service time. The Bayesian approach offers the framework for formally integrating prior knowledge with the facts currently available.

Here are some of the queueing system research that have been done in the past where the estimate of queueing parameters was done using both classical and Bayesian methods. Inter-arrival and service times were used as the observed data in an empirical Bayesian framework by [9] to estimate the parameters for various queueing systems. Based on the number of customers present at various sampling time points, [5] computed an maximum likelihood estimator (MLE) and Bayes estimator of traffic intensity in an M/M/1 queueing model. Regarding tandem queues with dependent service time structures, [2] studied statistical inferential aspects. Using the classical inference method, they modelled tandem queues and estimated the parameters. The statistical analysis of a tandem queue with blocking was then undertaken by [3] and focused on a two station tandem queue. Again, [1] investigated the Bayesian inference for a two station tandem queue, calculated the traffic intensities for the two stations, and determined the confidence

interval of the estimators. In the M/M/1 queue with bivariate priors, Bayes estimation has been studied by [6]. Then [4] performed a simulation research applying the Markov Chain Monte Carlo (MCMC) approach including the Metropolis-Hastings (M-H) algorithm and explored the Bayesian inference of the Markovian queuing model with two heterogeneous servers.

This paper attempts a detailed study of a three station tandem queue with customers arriving according to the Poisson process, with rate  $\lambda$  for service at station 1 and service times at station 1, station 2, and station 3 being exponentially distributed with service rates  $\mu_1, \mu_2$  and  $\mu_3$  respectively. The maximum likelihood and Bayes estimators of the intensity parameters  $\rho_1, \rho_2$  and  $\rho_3$  are computed using the number of customers present at various sampled time points under the assumption that the stationary distribution is the product of three independent geometric distributions with parameters  $\rho_1, \rho_2$  and  $\rho_3$  accordingly. Additionally, the minimal Bayes risk of the estimators and the minimum posterior risk related to Bayes estimators are derived.

This paper is structured as follows: Section 1 discussed an introduction to tandem queues as well as some early research in this area. Section 2 explored the model, the system description, and the inferential aspects of the model. Section 3 looked at the estimated number of customers in the system and its implications. Section 4 examined the model using simulation. Finally, Section 5 contains the paper's conclusions.

## 2. SYSTEM DESCRIPTION AND STEADY STATE PROBABILITY

Consider a simplified one channel queuing system consisting of three service stations as in the figure 1. A customer that arrives for servicing must pass through station 1, station 2 and station 3



Figure 1: System configuration

before finishing the service. The model's underlying assumptions are as follows:

1. Arrivals occur according to the Poisson distribution with mean rate  $\lambda$  at station 1.
2. Service times at station 1, station 2 and station 3 are exponentially distributed with service rates  $\mu_1, \mu_2$  and  $\mu_3$  respectively.
3. A queue of infinite size is allowed in front of station 1 and station 2 but at most one customer is permitted to wait between station 2 and station 3.
4. Each station is either free or busy.
5. If a customer in station  $i, i = 1, 2$  completes their service before station  $(i + 1), i = 1, 2$  becomes free, then it is said that station  $i, i = 1, 2$  is blocked.

Let  $p_{n_1, n_2, n_3}(t)$  be the probability that there are  $n_1$  customers in station 1,  $n_2$  customers in station 2 and  $n_3$  customers in station 3 at time  $t$  (in queue or in system). In the steady state it can be shown that,

$$p_{n_1, n_2, n_3}(t) = \rho_1^{n_1} (1 - \rho_1) \rho_2^{n_2} (1 - \rho_2) \rho_3^{n_3} (1 - \rho_3), \quad n_1, n_2 = 0, 1, 2, 3, \dots \ \& \ n_3 = 0, 1,$$

where,  $\rho_i = \frac{\lambda}{\mu_i}, i = 1, 2, \& 3$  and steady state results exist provided  $\rho_i < 1$ .

## 2.1. Classical Inference

The likelihood function of the number of customers present at  $n$  different time points  $t_1, t_2, t_3, \dots, t_n$  is given by

$$l((\rho_1, \rho_2, \rho_3) | ((x_1, y_1, z_1), \dots, (x_n, y_n, z_n))) = \rho_1^{\sum_{i=1}^n x_i} (1 - \rho_1)^n \rho_2^{\sum_{i=1}^n y_i} (1 - \rho_2)^n \rho_3^{\sum_{i=1}^n z_i} (1 - \rho_3)^n. \quad (1)$$

Taking logarithms and differentiating the log-likelihood function of (1) with respect to  $\rho_1, \rho_2$  and  $\rho_3$  and equating to zero, we get the MLEs of  $\rho_1, \rho_2, \rho_3$  and are given by

$$\hat{\rho}_1 = \frac{\sum_{i=1}^n x_i}{n + \sum_{i=1}^n x_i}, \hat{\rho}_2 = \frac{\sum_{i=1}^n y_i}{n + \sum_{i=1}^n y_i} \text{ and } \hat{\rho}_3 = \frac{\sum_{i=1}^n z_i}{n + \sum_{i=1}^n z_i}.$$

In other words,

$$\hat{\rho}_1 = \frac{T_1}{n + T_1}, \hat{\rho}_2 = \frac{T_2}{n + T_2} \text{ and } \hat{\rho}_3 = \frac{T_3}{n + T_3},$$

where,

$$T_1 = \sum_{i=1}^n x_i \sim NB(n, 1 - \rho_1), T_2 = \sum_{i=1}^n y_i \sim NB(n, 1 - \rho_2) \text{ and } T_3 = \sum_{i=1}^n z_i \sim NB(n, 1 - \rho_3)$$

and  $T_1, T_2$  and  $T_3$  are independent (see, [8]). Clearly, the probability mass functions (pmfs) of  $T_1, T_2$  and  $T_3$  are given by

$$\begin{aligned} P[T_1 = t_1] &= \binom{t_1 + n - 1}{n - 1} (1 - \rho_1)^n \rho_1^{t_1}, \\ P[T_2 = t_2] &= \binom{t_2 + n - 1}{n - 1} (1 - \rho_2)^n \rho_2^{t_2} \text{ and} \\ P[T_3 = t_3] &= \binom{t_3 + n - 1}{n - 1} (1 - \rho_3)^n \rho_3^{t_3}, \end{aligned}$$

where,  $t_1 = 0, 1, 2, \dots, t_2 = 0, 1, 2, \dots$  and  $t_3 = 0, 1, 2, \dots$ . It can be shown that

$$E(T_1) = \frac{n^2 \rho_1}{1 - \rho_1}, E(T_2) = \frac{n^2 \rho_2}{1 - \rho_2} \text{ and } E(T_3) = \frac{n^2 \rho_3}{1 - \rho_3}.$$

Also

$$Var(T_1) = \frac{n^2 \rho_1}{(1 - \rho_1)^2}, Var(T_2) = \frac{n^2 \rho_2}{(1 - \rho_2)^2} \text{ and } Var(T_3) = \frac{n^2 \rho_3}{(1 - \rho_3)^2}.$$

Since  $\hat{\rho}_1, \hat{\rho}_2$  and  $\hat{\rho}_3$  are one to one functions of  $T_1, T_2$  and  $T_3$  respectively, it is clear that  $\hat{\rho}_1, \hat{\rho}_2$  and  $\hat{\rho}_3$  assume the values  $\frac{t_1}{n+t_1}, \frac{t_2}{n+t_2}$  and  $\frac{t_3}{n+t_3}$  respectively with  $t_1, t_2, t_3 = 0, 1, 2, 3, \dots$ . Further, the joint pmf of  $\hat{\rho}_1, \hat{\rho}_2$  and  $\hat{\rho}_3$  is given by

$$\begin{aligned} P[\hat{\rho}_1 = u, \hat{\rho}_2 = v, \hat{\rho}_3 = w] &= P \left[ \frac{t_1}{n + t_1} = u, \frac{t_2}{n + t_2} = v, \frac{t_3}{n + t_3} = w \right] \\ &= P \left[ t_1 = \frac{nu}{1-u} \right] P \left[ t_2 = \frac{nv}{1-v} \right] P \left[ t_3 = \frac{nw}{1-w} \right] \\ &= \binom{\frac{nu}{1-u} + n - 1}{n - 1} (1 - \rho_1)^n \rho_1^{\frac{nu}{1-u}} \binom{\frac{nv}{1-v} + n - 1}{n - 1} (1 - \rho_2)^n \rho_2^{\frac{nv}{1-v}} \\ &\quad \times \binom{\frac{nw}{1-w} + n - 1}{n - 1} (1 - \rho_3)^n \rho_3^{\frac{nw}{1-w}}. \end{aligned}$$

In the next section, Bayes estimators of  $\rho_1, \rho_2$  and  $\rho_3$  and their Bayes risks are found.

## 2.2. Bayesian Inference

The number of customers present at various sampled time points is used to determine the Bayes estimators of  $\rho_1, \rho_2$  and  $\rho_3$  as well as their Bayes risks. The natural conjugate prior density for  $(\rho_1, \rho_2, \rho_3)$  is taken to be the product of three independent Beta distributions of first kind with the parameters  $(m_1, n_1), (m_2, n_2)$  and  $(m_3, n_3)$ , respectively. As a result, we suppose that  $(\rho_1, \rho_2, \rho_3)$  has a prior distribution that is the product of three separate Beta distributions of the first kind, each with the parameters  $(m_1, n_1), (m_2, n_2)$  and  $(m_3, n_3)$ . That is,

$$\begin{aligned} \tau(\rho|(m_1, n_1), (m_2, n_2), (m_3, n_3)) &= \frac{1}{\beta(m_1, n_1)\beta(m_2, n_2)\beta(m_3, n_3)} \rho_1^{m_1-1} (1 - \rho_1)^{n_1-1} \rho_2^{m_2-1} \\ &\quad \times (1 - \rho_2)^{n_2-1} \rho_3^{m_3-1} (1 - \rho_3)^{n_3-1}, \end{aligned}$$

where  $0 < \rho_1, \rho_2, \rho_3 < 1, \rho = (\rho_1, \rho_2, \rho_3), m' = (m_1, n_1), n' = (m_2, n_2)$  and  $p' = (m_3, n_3)$ . The marginal probability density function (pdf) of  $T = (T_1, T_2, T_3) = (\sum_{i=1}^n x_i, \sum_{i=1}^n y_i, \sum_{i=1}^n z_i)$ , which is called the predictive pdf and is given by

$$\begin{aligned} f^*(t) &= \int_0^1 \int_0^1 \int_0^1 f(t_1, t_2, t_3; \rho_1, \rho_2, \rho_3) \cdot \tau(\rho|(m', n', p')) d\rho_1 \cdot d\rho_2 \cdot d\rho_3 \\ &= \int_0^1 \int_0^1 \int_0^1 P[T_1 = t_1] \cdot P[T_2 = t_2] \cdot P[T_3 = t_3] \tau(\rho|(m', n', p')) d\rho_1 \cdot d\rho_2 \cdot d\rho_3 \\ &= \frac{\beta(t_1 + m_1, n + n_1) \cdot \beta(t_2 + m_2, n + n_2) \cdot \beta(t_3 + m_3, n + n_3)}{\beta(m_1, n_1) \cdot \beta(m_2, n_2) \cdot \beta(m_3, n_3)} \prod_{i=1}^3 \binom{t_i + n - 1}{n - 1}. \end{aligned}$$

Hence the posterior distribution of  $\rho = (\rho_1, \rho_2, \rho_3)$  is given by

$$\begin{aligned} q(\rho|(x, y, z)) &= \frac{f(t_1, t_2, t_3; \rho) \tau(\rho|(m', n', p'))}{\int_0^1 \int_0^1 \int_0^1 f(t_1, t_2, t_3; \rho) \tau(\rho|(m', n', p')) d\rho} \\ &= \frac{1}{\beta(t_1 + m_1, n + n_1)} \rho_1^{(t_1+m_1)-1} (1 - \rho_1)^{(n+n_1)-1} \\ &\quad \times \frac{1}{\beta(t_2 + m_2, n + n_2)} \rho_2^{(t_2+m_2)-1} (1 - \rho_2)^{(n+n_2)-1} \\ &\quad \times \frac{1}{\beta(t_3 + m_3, n + n_3)} \rho_3^{(t_3+m_3)-1} (1 - \rho_3)^{(n+n_3)-1}, \quad 0 < \rho_1, \rho_2, \rho_3 < 1. \end{aligned}$$

It should be pointed out that the posterior distribution of  $\rho = (\rho_1, \rho_2, \rho_3)$  is the result of the pdfs of three independent Beta distributions of first-kind with the parameters  $(t_1 + m_1, n + n_1), (t_2 + m_2, n + n_2)$  and  $(t_3 + m_3, n + n_3)$ , respectively. Therefore, under the squared error loss, the Bayes estimator of  $\rho = (\rho_1, \rho_2, \rho_3)$  is given by

$$\begin{aligned} E[\rho|(x, y, z)] &= \int_0^1 \int_0^1 \int_0^1 \rho_1 \cdot \rho_2 \cdot \rho_3 \cdot q(\rho|(x, y, z)) d\rho \\ &= \frac{t_1 + m_1}{t_1 + m_1 + n + n_1} \frac{t_2 + m_2}{t_2 + m_2 + n + n_2} \frac{t_3 + m_3}{t_3 + m_3 + n + n_3}. \end{aligned}$$

Furthermore, the minimum posterior risk related to this Bayes estimator is provided by

$$V_p[\hat{\rho}^B|(x, y, z)] = \text{diag}(E[\hat{\rho}_1 - \rho_1]^2, E[\hat{\rho}_2 - \rho_2]^2, E[\hat{\rho}_3 - \rho_3]^2),$$

where

$$\begin{aligned} E[\hat{\rho}_1 - \rho_1]^2 &= \int_0^1 \int_0^1 \int_0^1 [\hat{\rho}_1 - \rho_1]^2 q(\rho|(x, y, z)) d\rho_1 d\rho_2 d\rho_3 \\ &= \frac{[n_1(n_1 + 1) + n]t_1^2 + n(n - 2m_1n_1)t_1 + [m_1(m_1 + 1)n^2]}{(n + t_1)^2(t_1 + m_1 + n + n_1)(t_1 + m_1 + n + n_1 + 1)}, \end{aligned}$$

$$E[\hat{\rho}_2 - \rho_2]^2 = \frac{[n_2(n_2 + 1) + n]t_2^2 + n(n - 2m_2n_2)t_2 + [m_2(m_2 + 1)n^2]}{(n + t_2)^2(t_2 + m_2 + n + n_2)(t_2 + m_2 + n + n_2 + 1)} \text{ and}$$

$$E[\hat{\rho}_3 - \rho_3]^2 = \frac{[n_3(n_3 + 1) + n]t_3^2 + n(n - 2m_3n_3)t_3 + [m_3(m_3 + 1)n^2]}{(n + t_3)^2(t_3 + m_3 + n + n_3)(t_3 + m_3 + n + n_3 + 1)}.$$

Therefore,  $E[V_p(\hat{\rho}^B|(x, y, z))]$  gives a minimum Bayes risk of  $\hat{\rho}^B = (\hat{\rho}_1^B, \hat{\rho}_2^B, \hat{\rho}_3^B)$  with respect to the marginal distribution  $h(x, y, z)$  of  $(x, y, z)$ , where  $(x, y, z) = (x_1, y_1, z_1), (x_2, y_2, z_2), \dots, (x_n, y_n, z_n)$  is derived as follows:

The marginal distribution  $h(x, y, z)$  of  $(x, y, z)$  is given by

$$h(x, y, z) = \int_0^1 \int_0^1 \int_0^1 L(\rho|(x, y, z)) \cdot \tau(\rho|(m', n', p')) d\rho_1 \cdot d\rho_2 \cdot d\rho_3$$

$$= \frac{\beta(m_1 + t_1, n + n_1)\beta(m_2 + t_2, n + n_2)\beta(m_3 + t_3, n + n_3)}{\beta(m_1, n_1)\beta(m_2, n_2)\beta(m_3, n_3)}$$

resulting in the minimum Bayes risk factor

$$r_{\tau, \hat{\rho}^B} = E[V_p(\hat{\rho}^B|(x, y, z))] = E[\text{diag}(E[\hat{\rho}_1 - \rho_1]^2, E[\hat{\rho}_2 - \rho_2]^2, E[\hat{\rho}_3 - \rho_3]^2)].$$

### 3. EXPECTED NUMBER OF CUSTOMERS IN THE SYSTEM

The expected number of customers in the system is defined by

$$L_s = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \sum_{n_3=0}^1 (n_1 + n_2 + n_3) p_{n_1, n_2, n_3}(t)$$

$$= \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \sum_{n_3=0}^1 (n_1 + n_2 + n_3) \rho_1^{n_1} (1 - \rho_1) \rho_2^{n_2} (1 - \rho_2) \rho_3^{n_3} (1 - \rho_3)$$

$$= (1 - \rho_1)(1 - \rho_2)(1 - \rho_3) \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \sum_{n_3=0}^1 (n_1 + n_2 + n_3) \rho_1^{n_1} \rho_2^{n_2} \rho_3^{n_3}$$

$$= \frac{\rho_1}{1 - \rho_1} + \frac{\rho_2}{1 - \rho_2} + \frac{\rho_3}{1 - \rho_3}.$$

Therefore,

$$L_s = \lambda \left[ \frac{1}{(\mu_1 - \lambda)} + \frac{1}{(\mu_2 - \lambda)} + \frac{1}{(\mu_3 - \lambda)} \right]. \tag{2}$$

In the next section, we obtain a  $100(1 - \alpha)\%$  asymptotic confidence interval for the expected number of customers in the system.

#### 3.1. Maximum Likelihood Estimator for the expected number of customers in the system

Given an exponential inter-arrival time population with the parameter  $\lambda$ , let  $X_1, X_2, \dots, X_n$  be a random sample of size  $n$ . Let  $Y_{i1}, Y_{i2}, \dots, Y_{in}$  represent a random sample of size  $n$  taken from a population of service times with an exponential distribution and parameter  $\mu_i, i = 1, 2, 3$ . Therefore, it is clear that

$$E[\bar{X}] = \frac{1}{\lambda}, \quad E[\bar{Y}_i] = \frac{1}{\mu_i}, \quad i = 1, 2, 3.$$

Here  $\bar{X}$  and  $\bar{Y}_i, i = 1, 2, 3$ , respectively represents sample means for inter-arrival times and service times. It can be shown that  $\bar{X}$  and  $\bar{Y}_i, i = 1, 2, 3$  are, respectively, the MLEs of  $\frac{1}{\lambda}$  and  $\frac{1}{\mu_i}, i = 1, 2, 3$ .

Let  $\theta_1 = \frac{1}{\mu_1}$ ,  $\theta_2 = \frac{1}{\mu_2}$ ,  $\theta_3 = \frac{1}{\mu_3}$  and  $\theta_4 = \frac{1}{\lambda}$ . Then the the expected number of customers in the system given in (2) reduces to

$$L_s = \frac{\theta_1}{(\theta_4 - \theta_1)} + \frac{\theta_2}{(\theta_4 - \theta_2)} + \frac{\theta_3}{(\theta_4 - \theta_3)}.$$

Therefore, using the invariance property of the MLE, the MLE of  $L_s$  is given by

$$\hat{L}_s = \frac{\bar{Y}_1}{\bar{X} - \bar{Y}_1} + \frac{\bar{Y}_2}{\bar{X} - \bar{Y}_2} + \frac{\bar{Y}_3}{\bar{X} - \bar{Y}_3}.$$

It should be noticed that  $\hat{L}_s$  is a real valued function that is also differentiable in  $\bar{Y}_1, \bar{Y}_2, \bar{Y}_3$  and  $\bar{X}$ .

### 3.2. CAN estimator for expected number of customers

By applying the multivariate central limit theorem, we have

$$\sqrt{n} [(\bar{Y}_1, \bar{Y}_2, \bar{Y}_3, \bar{X}) - (\theta_1, \theta_2, \theta_3, \theta_4)] \xrightarrow{d} N(0, \Sigma) \text{ as } n \rightarrow \infty.$$

The dispersion matrix  $\Sigma = ((\sigma_{ij}))$  is given by  $\Sigma = \text{diag}(\theta_1^2, \theta_2^2, \theta_3^2, \theta_4^2)$ . Again from [7], we have

$$\sqrt{n}(\hat{L}_s - L_s) \xrightarrow{d} N(0, \sigma^2(\theta)) \text{ as } n \rightarrow \infty,$$

where  $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$  and

$$\sigma^2(\theta) = \sum_{i=1}^3 \left( \frac{\partial L_s}{\partial \theta_i} \right)^2 \sigma_{ii} = \theta_4^2 \left( \frac{\theta_1^2}{(\theta_4 - \theta_1)^4} + \frac{\theta_2^2}{(\theta_4 - \theta_2)^4} + \frac{\theta_3^2}{(\theta_4 - \theta_3)^4} \right). \quad (3)$$

Hence it is concluded that,  $\hat{L}_s$  is a CAN estimator of  $L_s$ .

### 3.3. Confidence interval for expected number of customers

Let  $\sigma^2(\hat{\theta})$  be the estimator of  $\sigma^2(\theta)$  obtained by replacing  $\theta$  by a consistent estimator  $\hat{\theta}$ , namely  $\hat{\theta} = (\bar{Y}_1, \bar{Y}_2, \bar{Y}_3, \bar{X})$ . Let  $\hat{\sigma}^2 = \sigma^2(\hat{\theta})$ . Since  $\sigma^2(\theta)$  is a continuous function of  $\theta$ ,  $\hat{\sigma}^2$  is a consistent estimator of  $\sigma^2(\theta)$  (see, [8]), we have

$$\hat{\sigma}^2 \xrightarrow{p} \sigma^2(\theta) \text{ as } n \rightarrow \infty.$$

By Slutsky's theorem (see, [8])  $(X_n \xrightarrow{d} x, Y_n \xrightarrow{p} b \implies \frac{X_n}{Y_n} \xrightarrow{d} \frac{x}{b}, b \neq 0)$ , we have

$$\sqrt{n} \left( \frac{\hat{L}_s - L_s}{\hat{\sigma}} \right) \xrightarrow{d} N(0, 1) \text{ as } n \rightarrow \infty.$$

That is,

$$Pr \left[ -k_{\frac{\alpha}{2}} < \sqrt{n} \left( \frac{\hat{L}_s - L_s}{\hat{\sigma}} \right) < k_{\frac{\alpha}{2}} \right] = (1 - \alpha),$$

where  $k_{\frac{\alpha}{2}}$  is obtained from the standard normal table. Hence,  $100(1 - \alpha)\%$  asymptotic confidence interval for  $L_s$  is given by  $\left( \hat{L}_s \pm k_{\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{n}} \right)$ , where  $\hat{\sigma}$  is obtained from the equation given in equation(3) by replacing  $\theta_1, \theta_2$  and  $\theta_3$  by the corresponding MLEs  $\bar{Y}_1, \bar{Y}_2, \bar{Y}_3$  and  $\bar{X}$  respectively.

#### 4. COMPUTATIONAL CONSIDERATIONS

The Bayes estimator of model parameters of three-station tandem queue with one customer being allowed to wait in the last station using an MCMC (see, [10]) simulation method as is as follows:

1. Defining the likelihood function: The likelihood function is a probability distribution that describes the probability of observing the data given the model parameters. In a queuing model it would be the probability of observing the number of customers in each station, the waiting time, and the service time given the model parameters (such as arrival rate, service rate and observation time).
2. Defining the prior distribution: The prior distribution is a probability distribution that describes the probability distribution of the model parameters before observing the data. In a queuing model it would be the probability of the arrival rate, service rate and observation time.
3. Defining the posterior distribution: The posterior distribution is the probability distribution of the model parameters given the data. It is calculated by multiplying the likelihood function and the prior distribution.
4. Specify the starting values for the MCMC chain: Choose some initial values for the model parameters that we want to estimate.
5. Run the MCMC simulation: Use an MCMC algorithm such as the M-H algorithm to generate a large number of samples from the posterior distribution.
6. Extract the samples from the MCMC chain: Retrieve the samples generated by the MCMC algorithm for each model parameters.
7. Calculate the posterior mean and standard deviation: Compute the mean and standard deviation of the samples for each model parameter. These will be the Bayes estimates of the model parameters.
8. Validate the estimates: Compare the Bayes estimates with the true values of the model parameters (if they are known) or with the estimates obtained using other methods, such as maximum likelihood estimation or method of moments.
9. Assess the convergence of the chain: Check if the chain has converged or not using methods such as trace plots, Gelman-Rubin diagnostic, or effective sample size.

##### 4.1. Simulation

The initial values given for simulation are :

$$\rho_1 = 0.3, \rho_2 = 0.4, \rho_3 = 0.7, m_1 = 5, m_2 = 6, m_3 = 7, n_1 = 10, n_2 = 9, n_3 = 8.$$

**Table 1:** Table 1: Table of MSE and Bias for different sample sizes.

Sample Size	Estimates	MSE	Bias
500	0.2677	0.08643	0.17954
	0.3575	0.00374	0.07256
	0.6794	0.05953	0.09211
1000	0.2730	0.00789	0.00623
	0.3823	0.00043	0.00058
	0.6847	0.00312	0.00085
2000	0.2877	0.000036	0.00032
	0.3956	0.000023	0.000082
	0.7148	0.000016	0.00028
5000	0.3062	0.000006	0.000022
	0.4341	0.000004	0.000039
	0.7232	0.000003	0.000009

From the table 1 it is clear that when sample size increases, the Mean Square Error (MSE) and Bias are decreasing and tending to zero, indicating that the validity of the estimators obtained.

#### 4.2. Histograms for simulation range

The histogram of the simulation range for the traffic intensities  $\rho_1, \rho_2$  and  $\rho_3$  is plotted. The Y axis measures the frequency and the X axis shows the range of values that the corresponding traffic intensity takes with respect to the initial value. From the figure 2, figure 3 and figure 4, it is clear that the simulation results have taken a normal curve shape.

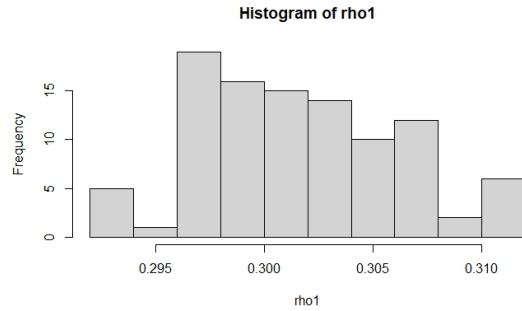


Figure 2: Histogram 1

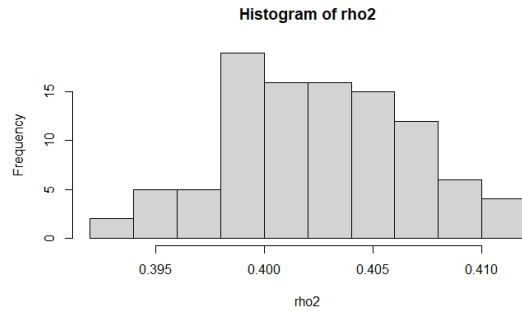


Figure 3: Histogram 2

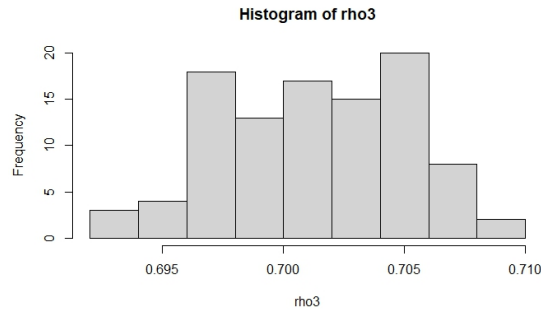


Figure 4: Histogram 3



## 5. CONCLUSIONS

In this study, we used the MLE and Bayesian techniques to estimate the traffic intensity for a three-station tandem queue where only one customer was permitted to wait between the last two stations. The Bayes estimators of  $\rho_1$ ,  $\rho_2$  and  $\rho_3$  were obtained using the beta prior, and the minimal Bayes risk was calculated. We also estimated the expected customers for the system. Then, using Slutsky's theorem, the confidence interval for the expected number of customers was determined. A three-station tandem queue was simulated using MCMC to obtain a Bayes estimators, and the performance of the estimators are verified through a broad simulation study.

## REFERENCES

- [1] Chandrasekhar, P. and Ambily, J. (2009). Bayesian inference for a two station tandem queue. *5th Asian Mathematical Conference, Malaysia*.
- [2] Chandrasekhar, P. and Chandrasekar, B. and Yadavalli, V. S. S. (2006). Statistical inference for a tandem queue with dependent structure for service times. *In Proceedings of the Sixth IASTED International Conference on Modelling, Simulation and Optimization*, 11–13.
- [3] Chandrasekhar, P. and Natarajan, R. and Yadavalli, V.S. S. (2007). Statistical analysis for a tandem queue with blocking. *In Proceedings of the Second National Conference on Management Science and Practice*.
- [4] Joby, K. J. and Deepthi, V. (2020) Bayesian inference of Markovian queueing model with two heterogeneous servers. *Stochastic Modelling and Applications*, 24(1):1–14.
- [5] Mukhvarjee, S. P. and Choudhury, S. (2005). Maximum likelihood and bayes estimators in m/m/1 queue. *Stochastic Modelling and Applications*, 8(2):47–55.
- [6] Mukhvarjee, S. P. and Choudhury, S. (2016). Bayes estimation in m/m/1 queues with bivariate prior. *Journal of Statistics and Manegement systems*, 19:681–699.
- [7] Rao, C. R. Linear Statistical inference and its applications, Wiley Eastern Pvt. Ltd., New Delhi, 1974.
- [8] Rohatgi, V. K. and Saleh, A. K. An introduction to probability and statistics, John Wiley & Sons, 2015.
- [9] Thiruvaiyaru, D. and Ishwar, V. B. (1992). Empirical Bayes estimation for queueing systems and networks. *Queueing Systems*, 11:179–202.
- [10] Tierney, L. (1994). Markov chains for exploring posterior distributions. *The Annals of Statistics*, 22:1701–1728.

# RELIABILITY MODELING OF A BUTTER CHURNER AND CONTINUOUS BUTTER MAKING PRODUCTION SYSTEM

UPASANA SHARMA<sup>1</sup> AND DRISHTI<sup>2\*</sup>

<sup>1,2</sup>Department Statistics, Punjabi University, Patiala- 147002, India  
usharma@pbi.ac.in, drish2796@gmail.com

## Abstract

*In the dairy plant, an investigation into the machine that makes butter was subjected to a reliability study in relation to the seasonal demand. In the process of expanding the butter churner into a machine that can make butter continuously, a more reliable operational model was devised. Both the models and the data acquired with MATLAB have been subjected to availability and reliability testing and analysis. In addition, the graphical analysis was carried out with the help of Code Blocks and Excel. A comparison of the two models was then covered as the final topic. It was discovered that (a) the extended model was superior to the current model, (b) the failure rate of the existing line increased, which implies that a new machine needs to be added to the line to share the load, which results in improved production, and (c) the failure rate of the extended model was lower than the failure rate of the existing model. (c) in order to maximise profits while simultaneously minimising losses The effectiveness of the system ought to be enhanced by performing routine maintenance during both the summer and the winter.*

**Keywords:** Butter churner, continuous butter making, seasons, semi-Markov process, profit.

## 1. INTRODUCTION

As a result of high levels of "lifetime" engineering uncertainty, reliability engineering deals with predicting, preventing, and managing engineering failures. Costs of failures caused by equipment failure, parts costs, repairs, and personnel costs are all taken into account when reliability engineering is conducted. Industry engineers now put their effort on efficiency and high quality production. This can be achieved by improving system performance. When it comes to industrial applications on food production lines, ensuring a high level of reliability is highly important; however, reliability itself can be complex, many interconnected variables must be taken into account when guiding and assessing various levels of reliability.

Using maintenance regimes [9] processed site performance improvement in the dairy industry. [8] presented a case study on optimised performance of butter oil production. Based on real data [5] represented generation of wind power and electric power demand. Reliability analysis where operation is effected by temperature conditions was given by [2] and [1]. RAM analysis for modeling complex engineering systems was used by [6].

Introducing redundancy into a system can enhance its reliability. Redundancy with standby (redundant) units refers to the usage of additional units with the primary unit of the system, with the additional unit(s) becoming operational and performing all the desired functions with equivalent parameters upon the failure of the primary unit. Standby redundancy technique was used by several researchers to enhance system performance namely [3], [4], [7] etc. Work on standby units in a dairy industry was done by [10], [11] and [12].

### Description of the systems

In model 1, the system which we have considered consists of a churner that works in both the seasons i.e., summer and winter. In winters, due to high demand system is always operating

unless a failure occurs that can be due to electricity halt or any fault in the churner. In summers, due to less demand the system sometimes goes to cold standby state when there is no demand. In model 2, the system consists of churner and continuous butter making. Both the units starts to operate to accomodate the demand in winters, on the failure of any one unit the system works on reduced capacity. In summers, the butter churner is operative and CBM is in cold standby state, it operates on the failure of the churner. The system either goes to cold standby or maintenance state when there is no demand.

**Methods**

Both the models have been analyzed using semi-Markov process and regenerative point technique probabilistically.

**2. ANNOTATIONS**

Table 1:

Notations of the model 1	
Notations	Descriptions
$\lambda$	Failure rate of the main unit i.e. Churner.
$\lambda_1$	Rate of electricity failure due to which churner stops operating.
$\gamma$	Rate at which churner goes to down state when demand is less than production.
$\delta$	Rate when churner comes to operative state from a cold standby state.
$\alpha$	Rate of going from winters to summers.
$\beta$	Rate of going from summers to winters.
$ch$	Main unit of the system i.e.ch.
$S$	Summer season.
$W$	Winter season.
$Och$	Main unit of the system is in operating state.
$d > p$	Demand is more than production.
$d < p$	Demand is less than production.
$CSch$	Main unit is in cold standby state.
$Frch$	Main unit is under repair.
$HCSch$	Main unit in cold standby state due to electricity halt.
$G(t), g(t)$	c.d.f. and p.d.f of time to repair of the main unit.
$G_1(t), g_1(t)$	c.d.f. and p.d.f of time to repair the electricity halt.
$G_2(t), g_2(t)$	c.d.f. and p.d.f of time to going back to operating state from down state.

**3. TRANSITION PROBABILITIES AND MEAN SOJOURN TIME**

Various states of the system are shown in figure 3.1 called as state transition diagram. Here, the states  $S_0, S_1, S_2$  are operating states,  $S_5$  is a cold standby state whereas, states  $S_3, S_4, S_6, S_7$  are the failed states.

**Transition Probabilites**

- $dQ_{01}(t) = \beta e^{-(\alpha+\beta)(t)} dt$
- $dQ_{13}(t) = \lambda_1 e^{-(\lambda+\lambda_1)(t)} dt$
- $dQ_{25}(t) = \gamma e^{-(\gamma+\lambda+\lambda_1)(t)} dt$
- $dQ_{02}(t) = \alpha e^{-(\alpha+\beta)(t)} dt$
- $dQ_{14}(t) = \lambda e^{-(\lambda+\lambda_1)(t)} dt$
- $dQ_{26}(t) = \lambda_1 e^{-(\gamma+\lambda+\lambda_1)(t)} dt$

- $dQ_{27}(t) = \lambda e^{-(\gamma+\lambda+\lambda_1)(t)} dt$

The non-zero probabilities  $p_{ij}$  are as follows:

- $p_{ij} = Q_{ij}(\infty) = \int_0^\infty q_{ij} dt$
- $p_{02} = \frac{\alpha}{\alpha+\beta}$
- $p_{14} = \frac{\lambda}{\lambda+\lambda_1}$
- $p_{26} = \frac{\lambda_1}{\gamma+\lambda+\lambda_1}$
- $p_{31} = p_{62} = g * 1(0)$
- $p_{01} = \frac{\beta}{\alpha+\beta}$
- $p_{13} = \frac{\lambda_1}{\lambda+\lambda_1}$
- $p_{25} = \frac{\gamma}{\gamma+\lambda+\lambda_1}$
- $p_{27} = \frac{\lambda}{\gamma+\lambda+\lambda_1}$
- $p_{41} = p_{72} = g * (0)$

From the above transition probabilities it is verified that:

- $p_{01} + p_{02} = 1$
- $p_{13} + p_{14} = 1$
- $p_{25} + p_{26} + p_{27} = 1$

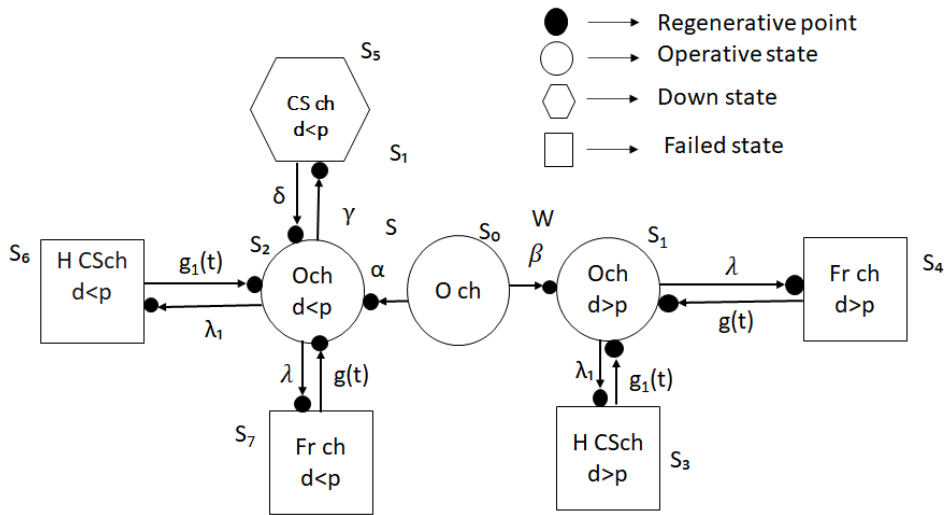


Figure 1: State Transition Diagram

The unconditional mean time taken by the system to transit for any regenerative state  $j$  when time is counted from the epoch of entrance into state  $i$  is mathematically state as:

- $m_{ij} = \int_0^\infty t dQ_{ij}(t) dt = -q_{ij}^*(0)$
- $m_{01} + m_{02} = \mu_0$
- $m_{13} + m_{14} = \mu_1$
- $m_{25} + m_{26} + m_{27} = \mu_2$

The mean sojourn time  $\mu_i$  in the regenerative state  $i$  is defined as time of stay in that state before transition to any other state:

- $\mu_0 = \frac{1}{\alpha+\beta}$
- $\mu_1 = \frac{1}{\lambda+\lambda_1}$
- $\mu_2 = \frac{1}{\gamma+\lambda+\lambda_1}$
- $\mu_3 = \mu_6 = -g_1^*(0)$
- $\mu_4 = \mu_7 = -g^*(0)$
- $\mu_5 = \frac{1}{\delta}$

#### 4. MEAN TIME TO SYSTEM FAILURE

The average duration between successive system failures, i.e. MTSF is defined as the expected time for which the system is in operation before it completely fails. Mean time to system failure

(MTSF) of the system is determined by considering failed state as absorbing state. When the system starts from the state 0, the mean time to system failure is:

$$T_0 = \lim_{s \rightarrow 0} R^*(s) = \lim_{s \rightarrow 0} \frac{1 - \phi_o^{**}(s)}{s} = \frac{N}{D}$$

where,

$$N = (\mu_0 + \mu_1 p_{01})(1 - p_{25}) + (\mu_2 + \mu_5 p_{25})(p_{02})$$

$$D = 1 - p_{25}$$

### 5. AVAILABILITY ANALYSIS OF THE SYSTEM IN SUMMERS

Availability  $A_i(t)$  is a measure that allows for a system to repair when failure occurs. The availability of the system is defined as the probability that the system is successful at time t. The long run availability of the system is given by

$$A_0^s = \lim_{s \rightarrow 0} [sA_0^{*s}(s)] = \frac{N_1}{D_1}$$

where,

$$N_1 = \mu_2 p_{02}$$

$$D_1 = \mu_2 + \mu_5 p_{25} + \mu_0 p_{26} + \mu_7 p_{27}$$

### 6. AVAILABILITY ANALYSIS OF THE SYSTEM IN WINTERS

Availability  $A_i(t)$  is a measure that allows for a system to repair when failure occurs. The availability of a system is defined as the probability that the system is successful at time t. The long run availability of the system is given by

$$A_0^w = \lim_{s \rightarrow 0} [sA_0^{*w}(s)] = \frac{N_2}{D_2}$$

where,

$$N_2 = \mu_1 p_{01}$$

$$D_2 = \mu_1 + \mu_4 p_{14} + \mu_3 p_{13}$$

### 7. BUSY PERIOD ANALYSIS FOR REPAIR IN SUMMERS

Busy period  $B_i(t)$  in summers is defined as the probability that the repairman is busy at time t when the system entered to a regenerative state i. The total time in which the repairman is busy doing repair of the system in steady state is given by:

$$B_0^s = \lim_{s \rightarrow 0} [sB_0^{*s}(s)] = \frac{N_3}{D_1}$$

where,

$$N_3 = p_{02}(p_{26}\mu_6 + p_{27}\mu_7)$$

$D_1$  is already defined above.

### 8. BUSY PERIOD ANALYSIS FOR REPAIR IN WINTERS

Busy period  $B_i(t)$  in winters is defined as the probability that the repairman is busy at time t when the system entered to a regenerative state i. The total time in which the repairman is busy doing repair of the system in steady state is given by:

$$B_0^w = \lim_{s \rightarrow 0} [sB_0^{*w}(s)] = \frac{N_4}{D_2}$$

where,

$$N_4 = p_{01}(W_3 p_{13} + W_4 p_{14})$$

$D_2$  is already defined above.

### 9. EXPECTED NUMBER OF REPAIRS IN SUMMERS

Let  $V_i(t)$  be the expected number of repairs in  $(0, t)$  given that the system entered into regenerative state  $i$  at  $i = 0$ . The expected number of repairs during summers in steady state is given by:

$$V_r = \lim_{s \rightarrow 0} sV_r^{**}(s) = \frac{N_5}{D_1}$$

$$N_5 = p_{02}(1 - p_{25})$$

$D_1$  is already defined above in equation.

### 10. EXPECTED NUMBER OF REPAIRS IN WINTERS

Let  $V_i(t)$  be the expected number of repairs in  $(0, t)$  given that the system entered into regenerative state  $i$  at  $i = 0$ . The expected number of repairs during winters in steady state is given by:

$$V_r = \lim_{s \rightarrow 0} sV_r^{**}(s) = \frac{N_6}{D_2}$$

$$N_6 = p_{01}$$

$D_2$  is already defined above in equation.

### 11. PROFIT ANALYSIS OF THE SYSTEM

Profit incurred to the system model in steady state is given by

$$P = (C_0A_0^s + C_1A_0^w) - (C_2B_0^s + C_3B_0^w + C_4V_0^s + C_5V_0^w)$$

where,

$C_0$ =Revenue per unit up time in summers.

$C_1$ =Revenue per unit up time in winters.

$C_2$ =Cost per unit up time for which the repairman is busy for repair in summers.

$C_3$ =Cost per unit up time for which the repairman is busy for repair in winters.

$C_4$ =Cost per repair in summers.

$C_5$ =Cost per repair in winters.

### 12. GRAPHICAL ANALYSIS AND CONCLUSION

For further numerical and graphical evaluation, let us assume the repair and failure rates to be exponentially distributed

$$g(t) = \theta e^{-\theta(t)}, g_1(t) = \theta_1 e^{-\theta_1(t)}$$

- $p_{01} = \frac{\beta}{\alpha + \beta}$
- $p_{13} = \frac{\lambda_1}{\lambda + \lambda_1}$
- $p_{25} = \frac{\gamma}{\gamma + \lambda + \lambda_1}$
- $p_{27} = \frac{\lambda}{\gamma + \lambda + \lambda_1}$
- $p_{41} = p_{72} = 1$
- $\mu_1 = \frac{1}{\lambda + \lambda_1}$
- $\mu_3 = \mu_6 = \frac{1}{\theta_1}$
- $\mu_5 = \frac{1}{\delta}$
- $p_{02} = \frac{\alpha}{\alpha + \beta}$
- $p_{14} = \frac{\lambda}{\lambda + \lambda_1}$
- $p_{26} = \frac{\lambda_1}{\gamma + \lambda + \lambda_1}$
- $p_{31} = p_{62} = 1$
- $\mu_0 = \frac{1}{\alpha + \beta}$
- $\mu_2 = \frac{1}{\gamma + \lambda + \lambda_1}$
- $\mu_4 = \mu_7 = \frac{1}{\theta}$

The parameters obtained using the original data collected from the Verka Milk Plant, Bathinda, Punjab.

Table 2:

Parameters obtained from data collected	
Parameters for model 1	Values
$\lambda$	.00045892
$\lambda_1$	.0002563
$g_1(t)$	.04213
$g(t)$	.062981
$\alpha$	.0004314
$\beta$	.000526
$\delta$	.000155
$\gamma$	.000955
$C_0$	830000
$C_1$	1030000
$C_2$	10500
$C_3$	12500
$C_4$	12000
$C_5$	15500

System effectiveness measures evaluated are given below:

Table 3:

Parameters obtained from data collected	
Parameters for model 1	Values
Mean time to system failure	9453.77 hrs
Availability in summers	.8975
Availability in winters	.8984
Busy period for repair in summers	.000485
Busy period for repair in winters	.0004204
Expected number of repairs in summers	.000217
Expected number of repairs in winters	.000031

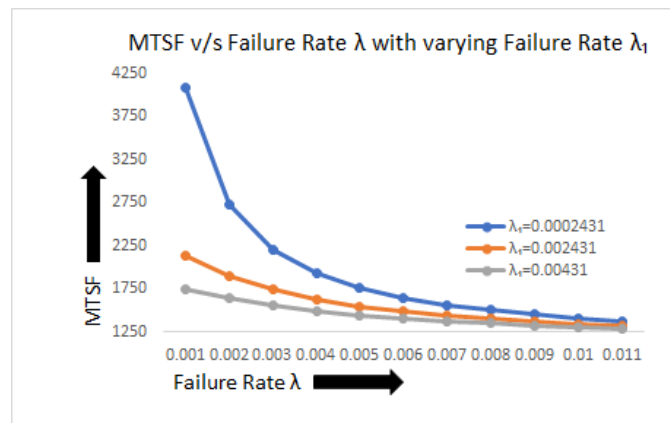


Figure 2: MTSF v/s Failure Rate

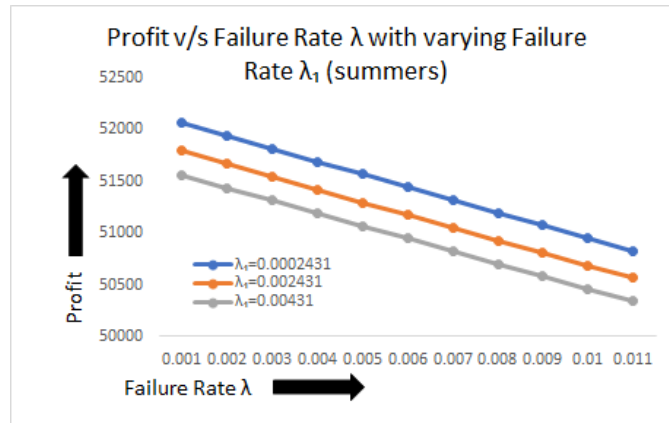


Figure 3: Profit v/s Failure Rate in Summers

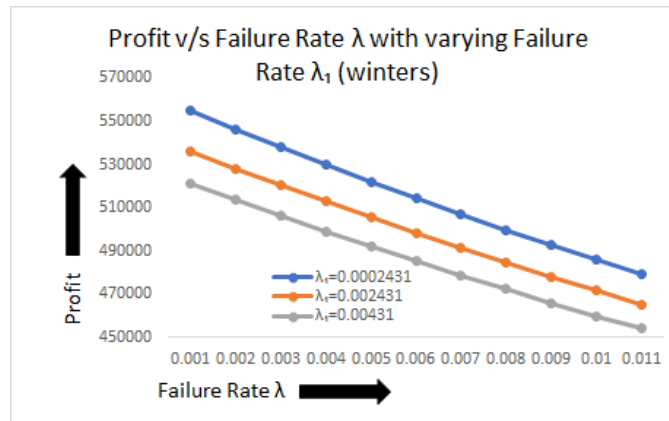


Figure 4: Profit v/s Failure Rate in Winters

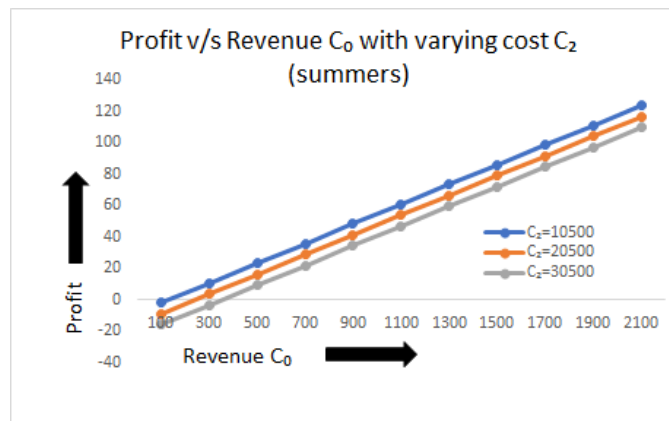


Figure 5: Profit v/s Failure Rate in Winters



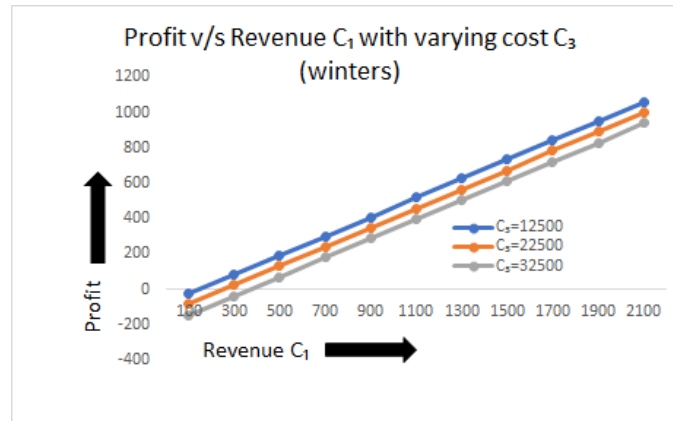


Figure 6: Profit v/s Failure Rate in Winters

Table 4:

Notations of the model	
Figures	Descriptions
5	Profit P1 increases as the revenue C <sub>0</sub> increases. C <sub>2</sub> =10500; Profit >=< according to C <sub>2</sub> , when C <sub>2</sub> is >=<Rs.275.53, similarly for C <sub>2</sub> =20500 where cut off point is Rs.163.577 C <sub>2</sub> =30500; where cut off point is Rs. 452.675
6	Profit P2 increases as the revenue C <sub>1</sub> increases. C <sub>3</sub> =12500; Profit >=< according to C <sub>3</sub> , when C <sub>1</sub> is >=<Rs.251.85, similarly for C <sub>3</sub> =22500 where cut off point is Rs.140.469. C <sub>3</sub> =32500; where cut off point is Rs. 429.089

Figure 3 and figure 4 depicts the trend of mean time to system failure and profit v/s the failure rate. It has been observed that as the failure rate  $\lambda$  of the system increases mean time to system failure and profit decreases. It also decreases on increasing failure rate  $\lambda_1$ . Figure 5,6 states that profit increases as the cost C<sub>1</sub> increases as well it increases with increasing profit C<sub>3</sub>.

**MODEL 2 Assumptions**

Model 2 have the following assumptions:

- The system is operating at the initial stage.
- At the initial stage the churner is operating and continuous butter making is in a cold standby state.
- Both the systems operates during winters due to high demand.
- Only one unit is operating during summers due to less demand.
- In summers it also undergoes maintenance.
- The system sometimes goes to cold standby state in case of no demand in summers.
- The repair is done on the failure of the system.
- Repair rates are assumed to have arbitrary distribution.
- Failure rates are taken to be exponentially distributed.

- After repair the system operates as new.
- The system goes to failed state either on the failure of the churner or due to halt in the electricity.

### 13. ANNOTATIONS FOR MODEL 2

Table 5:

Notations of the model 2	
Notations	Descriptions
$\lambda$	Failure rate of the churner.
$\lambda_1$	Failure rate of the continuous butter making.
$\gamma$	Rate at which churner goes to down state when demand is less than production.
$\delta$	Rate when churner comes to operative state from a cold standby state.
$\alpha$	Rate of going to winters.
$\beta$	Rate of going to summers.
$ch$	Unit churner of the system.
$cbm$	Unit continuous butter making of the system.
$S$	Summer season.
$W$	Winter season.
$Och$	Churner is in operating state.
$Ocbm$	CBM is in operating state.
$d > p$	Demand is more than production.
$d < p$	Demand is less than production.
$CSch$	Main unit is in a cold standby state.
$CScbm$	CBM is in a cold standby state.
$Frch$	Churner is under repair.
$HCSch$	Churner is in cold standby state due to electricity halt.
$G(t), g(t)$	c.d.f. and p.d.f of time to repair of the churner.
$G_1(t), g_1(t)$	c.d.f. and p.d.f of time to repair of CBM.
$G_2(t), g_2(t)$	c.d.f. and p.d.f of time to going back to operating state from maintenance.

### 14. MODEL 2

### 15. ANNOTATIONS FOR MODEL 2

### 16. TRANSITION PROBABILITES AND MEAN SOJOURN TIME

Various states of the system are shown in figure 1.5 called as state transition diagram. Here, the states  $S_0, S_1, S_2, S_3, S_5$  are operating states,  $S_4$  is a cold standby state whereas, states  $S_9, S_{10}$  are the reduced capacity states and rest are failed states.

- $dQ_{01}(t) = \beta e^{-(\alpha+\beta)(t)} dt$
- $dQ_{19}(t) = \lambda_1 e^{-(\lambda+\lambda_1)(t)} dt$
- $dQ_{23}(t) = \lambda_2 e^{(\lambda+\lambda_2+\gamma)t} dt$
- $dQ_{25}(t) = \lambda e^{(\lambda+\lambda_2+\gamma)t} dt$
- $dQ_{3,13}(t) = \lambda e^{-\lambda(t)} G(t) dt$
- $dQ_{02}(t) = \alpha e^{-(\alpha+\beta)(t)} dt$
- $dQ_{1,10}(t) = \lambda e^{-(\alpha+\beta)(t)} dt$
- $dQ_{24}(t) = \gamma e^{(\lambda+\lambda_2+\gamma)t} dt$
- $dQ_{32}(t) = g_2(t) e^{-\lambda(t)} dt$
- $dQ_{37}^{(13)}(t) = (\lambda e^{-\lambda(t)}(c)1)g_2(t) dt$

- $dQ_{42}(t) = \delta e^{-\delta(t)} dt$
- $dQ_{56}(t) = \lambda_1 e^{-\lambda_1(t)} G^-(t) dt$
- $dQ_{67}(t) = g_2(t) dt$
- $dQ_{78}(t) = \lambda e^{-\lambda(t)} G_1^-(t) dt$
- $dQ_{91}(t) = g_1(t) e^{-\lambda(t)} dt$
- $dQ_{9,10}^{(12)}(t) = (\lambda e^{-\lambda(t)}(c)1)g_1(t) dt$
- $dQ_{10,11}(t) = \lambda_1 e^{-\lambda_1(t)} G^-(t) dt$
- $dQ_{13,7}(t) = g_2(t) dt$
- $dQ_{52}(t) = g(t) e^{-\lambda_1(t)} dt$
- $dQ_{57}^{(6)}(t) = (\lambda_1 e^{-\lambda_1(t)})g(t) dt$
- $dQ_{72}(t) = g_1(t) e^{-\lambda(t)} dt$
- $dQ_{75}^{(8)} = (\lambda e^{-\lambda(t)}(c)1)g_1(t) dt$
- $dQ_{9,12}(t) = \lambda e^{-\lambda(t)} G_1^-(t) dt$
- $dQ_{10,1}(t) = g(t) e^{-\lambda_1(t)} dt$
- $dQ_{10,9}^{(11)}(t) = (\lambda_1 e^{-\lambda_1(t)}(c)1)g(t) dt$
- $dQ_{12,10}(t) = g_1(t) dt$

The non-zero probabilities  $p_{ij}$  are as follows:

- $p_{ij} = Q_{ij}(\infty) = \int_0^\infty q_{ij} dt$
- $p_{02} = \frac{\alpha}{\alpha+\beta}$
- $p_{1,10} = \frac{\lambda}{\lambda+\lambda_1}$
- $p_{24} = \frac{\gamma}{\lambda+\lambda_2+\gamma}$
- $p_{32} = g_2^*(\lambda)$
- $p_{52} = g_2^{(*)}(\lambda_1)$
- $p_{72} = g_1^{(*)}(\lambda)$
- $p_{91} = g_1^{(*)}(\lambda)$
- $p_{10,1} = g^{(*)}(\lambda_1)$
- $p_{01} = \frac{\beta}{\alpha+\beta}$
- $p_{19} = \frac{\lambda_1}{\lambda+\lambda_1}$
- $p_{23} = \frac{\lambda_2}{\lambda+\lambda_2+\gamma}$
- $p_{25} = \frac{\lambda}{\lambda+\lambda_2+\gamma}$
- $p_{3,13} = p_{37}^{(13)} = 1 - g_2^*(\lambda)$
- $p_{56} = p_{57}^{(6)} = 1 - g_2^{(*)}(\lambda_1)$
- $p_{78} = p_{75}^{(8)} = 1 - g_1^{(*)}(\lambda)$
- $p_{9,12} = p_{9,10}^{(12)} = 1 - g_1^{(*)}(\lambda)$
- $p_{10,11} = p_{10,9}^{(11)} = 1 - g^{(*)}(\lambda_1)$

From the above transition probabilities it is verified that:

- $p_{01} + p_{02} = 1$
- $p_{23} + p_{24} + p_{25} = 1$
- $p_{32} + p_{37}^{(13)} = 1$
- $p_{52} + p_{57}^{(6)} = 1$
- $p_{72} + p_{75}^{(8)} = 1$
- $p_{91} + p_{9,10}^{(12)} = 1$
- $p_{10,1} + p_{10,9}^{(11)} = 1$
- $p_{19} + p_{1,10} = 1$
- $p_{32} + p_{3,13} = 1$
- $p_{52} + p_{56} = 1$
- $p_{72} + p_{78} = 1$
- $p_{91} + p_{9,12} = 1$
- $p_{10,1} + p_{10,11} = 1$

The unconditional mean time taken by the system to transit for any regenerative state  $j$  when it (time) is counted from the epoch of entrance into state  $i$  is mathematically state as:

- $m_{ij} = \int_0^\infty t dQ_{ij}(t) dt = -q_{ij}^*(0)$
- $m_{19} + m_{1,10} = \mu_1$
- $m_{32} + m_{3,13} = \mu_3$
- $m_{52} + m_{56} = \mu_5$
- $m_{72} + m_{75} = \mu_7$
- $m_{91} + m_{9,12} = \mu_9$
- $m_{10,1} + m_{10,11} = \mu_{10}$
- $m_{01} + m_{02} = \mu_0$
- $m_{23} + m_{24} + m_{25} = \mu_2$
- $m_{32} + m_{37}^{(13)} = K_2$
- $m_{52} + m_{57}^{(6)} = K$
- $m_{72} + m_{75}^{(8)} = K_1$
- $m_{91} + m_{9,10}^{(12)} = K_1$
- $m_{10,1} + m_{10,9}^{(11)} = K$

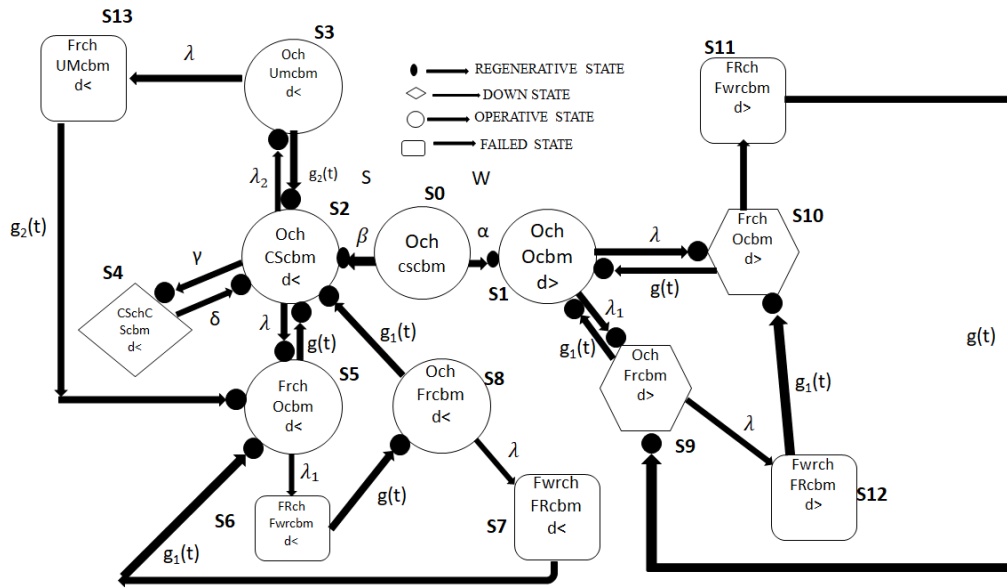


Figure 7: Model 2: State Transition Diagram

The mean sojourn time  $\mu_i$  in the regenerative state  $i$  is defined as time of stay in that state before transition to any other state:

- $\mu_0 = \frac{1}{\alpha + \beta}$
- $\mu_2 = \frac{1}{\gamma + \lambda + \lambda_2}$
- $\mu_4 = \frac{1}{\delta}$
- $\mu_7 = \mu_9 = \frac{1 - g_1^*(\lambda)}{\lambda}$
- $\mu_{11} = \int_0^\infty G(t) dt$
- $\mu_1 = \frac{1}{\lambda + \lambda_1}$
- $\mu_3 = \frac{1 - g_2^*(\lambda_1)}{\lambda_1}$
- $\mu_5 = \frac{1 - g^*(\lambda_1)}{\lambda}$
- $\mu_{10} = \frac{1 - g^*(\lambda_1)}{\lambda_1}$
- $\mu_{12} = \int_0^\infty G_1(t) dt$

### 17. MEAN TIME TO SYSTEM FAILURE FOR MODEL 2

The average duration between successive system failures, i.e. MTSF is defined as the expected time for which the system is in operation before it completely fails. Mean time to system failure (MTSF) of the system is determined by considering failed state as absorbing state. When the system starts from the state 0, the mean time to system failure is:

$$T_0 = \lim_{s \rightarrow 0} R^*(s) = \lim_{s \rightarrow 0} \frac{1 - \phi^{**}(s)}{s} = \frac{N}{D}$$

where,

$$D = p_{19}p_{23}p_{32}p_{91} - p_{24} - p_{25}p_{52} - p_{19}p_{91} - p_{10,1}p_{1,10} - p_{23}p_{32} + p_{19}p_{24}p_{91} + p_{19}p_{25}p_{52}p_{91} + p_{23}p_{32}p_{10,1}p_{1,10} + p_{24}p_{10,1}p_{1,10} + p_{25}p_{52}p_{10,1}p_{1,10} + 1$$

$$N = \mu_0(p_{23}p_{32}p_{91} + p_{25}p_{56} - p_{19}p_{23}p_{32}p_{91} - p_{19}p_{25}p_{56}p_{91} - p_{23}p_{32}p_{10,1}p_{1,10} - p_{25}p_{56}p_{10,1}p_{1,10}) + \mu_1(p_{91} + p_{01}p_{9,12} - p_{23}p_{32}p_{91} - p_{24}p_{42}p_{91} - p_{25}p_{52}p_{91} - p_{02}p_{23}p_{32}p_{91} - p_{02}p_{25}p_{56}p_{91} - p_{01}p_{23}p_{32}p_{9,12} - p_{01}p_{24}p_{42}p_{9,12} - p_{01}p_{25}p_{52}p_{9,12}) + (\mu_2 + \mu_4p_{24})(p_{42} - p_{19}p_{42}p_{91} - p_{42}p_{10,1}p_{1,10} - p_{01}p_{19}p_{42}p_{9,12} - p_{01}p_{42}p_{1,10}p_{10,11}) + \mu_3(p_{02}p_{23} - p_{02}p_{19}p_{23}p_{91} - p_{02}p_{23}p_{10,1}p_{1,10}) + \mu_5(p_{02}p_{25} - p_{02}p_{19}p_{25}p_{91} - p_{02}p_{25}p_{10,1}p_{1,10}) + \mu_9(p_{01}p_{19} - p_{01}p_{19}p_{23}p_{32} - p_{01}p_{19}p_{24}p_{42} - p_{01}p_{19}p_{25}p_{52}) + \mu_{10}(p_{01}p_{1,10} - p_{01}p_{23}p_{32}p_{1,10} - p_{01}p_{24}p_{42}p_{1,10} - p_{01}p_{25}p_{52}p_{1,10})$$

## 18. RELIABILITY MEASURES

### 18.1. Availability Analysis in Summers

Availability  $A_i(t)$  is a measure that allows for a system to repair when failure occurs. The availability of a system is defined as the probability that the system is successful at time  $t$ . The long run availability of the system is given by

$$A_0^s = \lim_{s \rightarrow 0} [sA_0^{*s}(s)] = \frac{N_1}{D_1}$$

where,

$$N_1 = \mu_0 + \mu_2 p_{02} + \mu_3 p_{02} p_{23} + \mu_5 p_{02} p_{25} - \mu_0 p_{23} p_{32} - \mu_0 p_{24} - \mu_0 p_{25} p_{52} - \mu_0 p_{57}^{(6)} p_{75}^{(8)} + \mu_7 p_{02} p_{23} p_{37}^{(13)} + \mu_7 p_{02} p_{25} p_{57}^{(6)} - \mu_0 p_{23} p_{37}^{(13)} p_{72} - \mu_2 p_{02} p_{57}^{(6)} p_{75}^{(8)} - \mu_0 p_{25} p_{57}^{(6)} p_{72} + \mu_5 p_{02} p_{23} p_{37}^{(13)} p_{75}^{(8)} - \mu_3 p_{02} p_{23} p_{57}^{(6)} p_{75}^{(8)} + \mu_0 p_{23} p_{32} p_{57}^{(6)} p_{75}^{(8)} - \mu_0 p_{23} p_{37}^{(13)} p_{52} p_{75}^{(8)} + \mu_0 p_{24} p_{57}^{(6)} p_{75}^{(8)} \\ D_1 = (\mu_2 + \mu_4 p_{24}) (1 - p_{57}^{(6)} p_{75}^{(8)}) + \mu_3 (p_{23} p_{72} + p_{23} p_{52} p_{75}^{(8)}) + \mu_5 (p_{75}^{(8)} + p_{25} p_{72} - p_{23} p_{32} p_{75}^{(8)} - p_{24} p_{75}^{(8)}) + \mu_7 (p_{57}^{(6)} - p_{23} p_{32} p_{57}^{(6)} + p_{23} p_{37}^{(13)} p_{52} - p_{24} p_{57}^{(6)})$$

### 18.2. Availability Analysis in Winters when the System Works at Full Capacity

The availability of a system is defined as the probability that the system is successful at time  $t$ . The long run availability of the system is given by

$$A_0^s = \lim_{s \rightarrow 0} [sA_0^{*s}(s)] = \frac{N_2}{D_2}$$

where,

$$N_2 = \mu_0 + \mu_1 p_{01} - \mu_0 p_{19} p_{91} - \mu_0 p_{10,1} p_{1,10} - \mu_0 p_{10,9}^{(11)} p_{9,10}^{(12)} - \mu_0 p_{91} p_{10,9}^{(11)} p_{1,10} - \mu_1 p_{01} p_{10,9}^{(11)} p_{9,10}^{(12)} - \mu_0 p_{19} p_{10,1} p_{9,10}^{(12)} \\ D_2 = \mu_1 (p_{10,1} + p_{91} p_{10,9}) + \mu_9 (p_{10,9} + p_{19} p_{10,1}) + \mu_{10} (p_{1,10} + p_{19} p_{9,10})$$

### 18.3. Availability Analysis in Winters when the System Operates at Reduced Capacity

Availability of the system when it operates at reduced capacity is given by

$$A_0^w = \lim_{s \rightarrow 0} [sA_0^{*w}(s)] = \frac{N_3}{D_2}$$

where,

$$N_3 = p_{01} (\mu_9 p_{19} + \mu_{10} p_{1,10} + \mu_9 p_{10,9}^{(11)} p_{1,10} + \mu_{10} p_{19} p_{9,10}^{(12)}) \\ D_2 \text{ is already defined above.}$$

### 18.4. Busy Period Analysis for Repair in Summers

Busy period  $B_i(t)$  in summers is defined as the probability that the repairman is busy at time  $t$  when the system entered to a regenerative state  $i$ . The total time in which the repairman is busy doing repair of the system in steady state is given by:

$$B_0^{sr} = \lim_{s \rightarrow 0} [sB_0^{*sr}(s)] = \frac{N_4}{D_1}$$

where,

$$N_4 = p_{02} (\mu_5 p_{25} + \mu_7 p_{23} p_{37}^{(13)} + \mu_7 p_{25} p_{57}^{(6)} + \mu_5 p_{23} p_{37}^{(13)} p_{75}^{(8)}) \\ D_2 \text{ is already defined above.}$$

### 18.5. Busy Period for Maintenance in Summers

Busy period  $B_i(t)$  in summers for maintenance is obtained. The total time in which the repairman is busy doing repair of the system in steady state is given by:

$$B_0^{sm} = \lim_{s \rightarrow 0} [sB_0^{*sm}(s)] = \frac{N_5}{D_2}$$

where,

$$N_5 = -\mu_3 p_{02} p_{23} (p_{57}^{(6)} p_{75}^{(8)} - 1) \\ D_2 \text{ is already defined above.}$$

### 18.6. Busy Period Analysis for Repair in Winters

Busy period for repair in winters is obtained as given below:

The total time in which the repairman is busy doing repair of the system in steady state is given by:

$$B_0^{wr} = \lim_{s \rightarrow 0} [sB_0^{*wr}(s)] = \frac{N_6}{D_2}$$

where,

$$N_6 = p_{01}(\mu_9 p_{19} + \mu_{10} p_{1,10} + \mu_9 p_{10,9}^{(11)} p_{1,10} + \mu_{10} p_{19} p_{9,10}^{(12)})$$

$D_2$  is already defined above.

### 18.7. Expected Number of Repairs in Summers

Let  $V_i(t)$  be the expected number of repairs in  $(0, t)$  given that the system entered into regenerative state  $i$  at  $i = 0$ .

The expected number of repairs during summers in steady state is given by:

$$V^{sr} = \lim_{s \rightarrow 0} sV^{sr}(s) = \frac{N_7}{D_1}$$

$$N_7 = p_{02}(p_{25} p_{52} + p_{25} p_{57}^{(6)} + p_{23} p_{37}^{(13)} p_{72} + p_{23} p_{37}^{(13)} p_{75}^{(8)} + p_{25} p_{57}^{(6)} p_{72} + p_{25} p_{57}^{(6)} p_{75}^{(8)} + p_{23} p_{37}^{(13)} p_{52} p_{75}^{(8)} + p_{23} p_{37}^{(13)} p_{57}^{(6)} p_{75}^{(8)})$$

$D_1$  is already defined above.

### 18.8. Expected Number of Maintenances in Summers

Let  $V_i(t)$  be the expected number of maintenances. The expected number of repairs during summers in steady state is given by:

$$V^{sm} = \lim_{s \rightarrow 0} sV^{sm}(s) = \frac{N_8}{D_1}$$

$$N_8 = -(p_{32} + p_{37}^{(13)}) p_{02} p_{23} (p_{57}^{(6)} p_{75}^{(8)} - 1)$$

$D_1$  is already defined above.

### 18.9. Expected Number of Repairs in Winters

Let  $V_i(t)$  be the expected number of repairs in winters. The expected number of repairs during summers in steady state is given by:

$$V^{wr} = \lim_{s \rightarrow 0} sV^{wr}(s) = \frac{N_9}{D_2}$$

$$N_9 = p_{01}(p_{19} p_{91} + p_{10,1} p_{1,10} + p_{10,9} p_{1,10} + p_{19} p_{9,10}^{(12)} + p_{91} p_{10,9} p_{1,10} + p_{19} p_{10,1} p_{9,10}^{(12)} + p_{19} p_{10,9} p_{9,10}^{(12)} + p_{10,9} p_{1,10} p_{9,10})$$

$D_2$  is already defined above.

## 19. PROFIT ANALYSIS OF THE SYSTEM

Profit incurred to the system model in steady state is given by

$$P = (C_0 A_0^s + C_1 A_0^{wf} + C_2 A_0^{wr}) - (C_3 B_0^s + C_4 B_0^w + C_5 B_0^{sm} + C_6 V_0^{sr} + C_7 V_0^w + C_8 V_0^{sm})$$

where,

$C_0$ =Revenue per unit up time in summers.

$C_1$ =Revenue per unit up time in winters when the system operates at full capacity.

$C_2$ =Revenue per unit up time in winters when the system operates at reduced capacity.

$C_3$ =Cost per unit up time for which the repairman is busy for repair in summers.

$C_4$ =Cost per unit up time for which the repairman is busy for repair in winters.

$C_5$ =Cost per unit up time for which the repairman is busy for maintenance in summers.

$C_6$ =Cost per repair in summers.

$C_7$ =Cost per repair in winters.

$C_8$ =Cost per maintenance in summers.

## 20. GRAPHICAL ANALYSIS AND CONCLUSION

For further numerical and graphical evaluation, let us assume the repair and failure rates to be exponentially distributed

$$g(t) = \theta e^{-\theta(t)}, g_1(t) = \theta_1 e^{-\theta_1(t)}, g_2(t) = \theta_2 e^{-\theta_2(t)}$$

- $p_{01} = \frac{\beta}{\alpha + \beta}$
- $p_{19} = \frac{\lambda_1}{\lambda + \lambda_1}$
- $p_{23} = \frac{\lambda_2}{\lambda + \lambda_2 + \gamma}$
- $p_{25} = \frac{\lambda}{\lambda + \lambda_2 + \gamma}$
- $p_{37}^{(13)} = p_{3,13} = \frac{\theta_2}{\lambda_1 + \theta_2}$
- $p_{57}^{(6)} = p_{56} = \frac{\theta}{\lambda_1 + \theta}$
- $p_{75}^{(8)} = p_{78} = \frac{\theta_1}{\lambda + \theta_1}$
- $p_{9,10}^{(12)} = p_{9,12} = \frac{\theta_1}{\lambda + \theta_1}$
- $p_{10,11}^{(12)} = p_{10,12} = \frac{\theta}{\lambda_1 + \theta}$
- $mu_1 = \frac{1}{\lambda + \lambda_1}$
- $\mu_3 = \frac{\theta_2}{\lambda_1(\lambda_1 + \theta_2)}$
- $\mu_5 = \mu_{10} = \frac{\theta}{\lambda_1(\lambda_1 + \theta)}$
- $\mu_{13} = \frac{1}{\theta_2}$
- $\mu_6 = \frac{1}{\theta}$
- $p_{02} = \frac{\alpha}{\alpha + \beta}$
- $p_{1,10} = \frac{\lambda}{\lambda + \lambda_1}$
- $p_{24} = \frac{\gamma}{\lambda + \lambda_2 + \gamma}$
- $p_{32} = \frac{\lambda_1}{\lambda_1 + \theta_2}$
- $p_{52} = \frac{\lambda_1}{\lambda_1 + \theta}$
- $p_{72} = \frac{\lambda}{\lambda + \theta_1}$
- $p_{91} = \frac{\lambda}{\lambda + \theta_1}$
- $p_{10,1} = \frac{\lambda_1}{\lambda_1 + \theta}$
- $\mu_0 = \frac{1}{\alpha + \beta}$
- $\mu_2 = \frac{1}{\gamma + \lambda + \lambda_2}$
- $\mu_4 = \frac{1}{\delta}$
- $\mu_7 = \mu_9 = \frac{\theta_1}{\lambda(\lambda + \theta_1)}$
- $\mu_8 = \mu_{12} = \frac{1}{\theta_1}$

The parameters obtained using the original data collected from the Verka Milk Plant, Bathinda, Punjab.

Table 6:

Parameters obtained from data collected	
Parameters for model 1	Values
$\lambda$	.00045892
$\lambda_1$	.0004567
$\lambda_2$	0.000246572
$g_1(t)$	.06312
$g(t)$	.062981
$g_2(t)$	0.002628867
$\alpha$	.000562
$\beta$	.0004314
$\delta$	.000955
$\gamma$	.000155
$C_0$	830000
$C_1$	1030000
$C_2$	61660
$C_3$	10500
$C_4$	12500
$C_5$	15500

$C_6$	19500
$C_7$	6400
$C_8$	7000

System effectiveness measures evaluated are given below:

Table 7:

Parameters obtained from data collected	
Parameters for model 2	Values
Mean time to system failure	99682.28 hrs
Availability in summers	0.985
Availability in winters when system operates at full capacity	.989
Availability in winters when system operates at reduced capacity	.001435
Busy period for repair in summers	.003814
Busy period for maintenance in summers	.038744
Busy period for repair in winters	.007864
Expected number of repairs in summers	.000242
Expected number of maintenances in summers	.000120
Expected number of repairs in winters	.000499

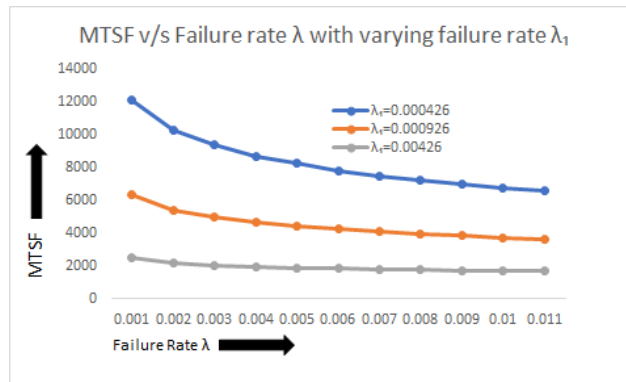


Figure 8: MTSF v/s Failure Rate

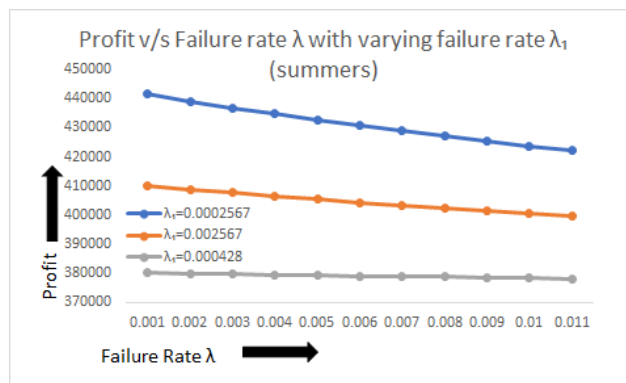


Figure 9: Profit v/s Failure Rate in Summers



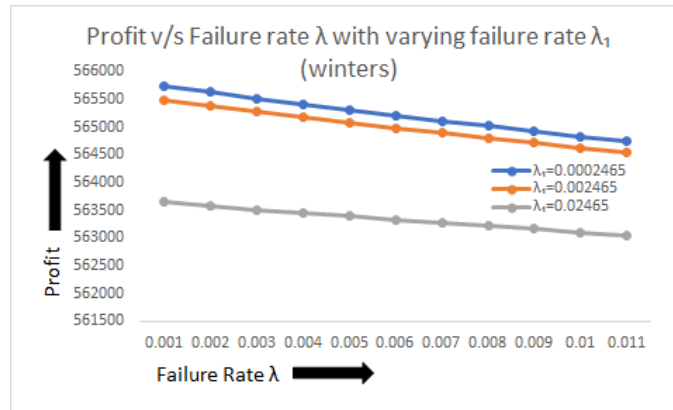


Figure 10: Profit v/s Failure Rate in Winters

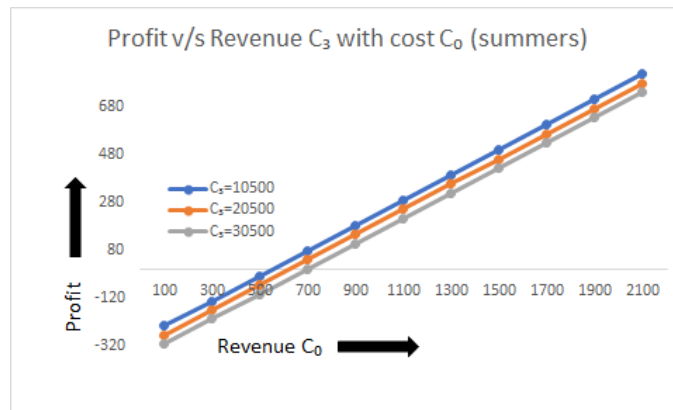


Figure 11: Profit v/s Failure Rate in Winters

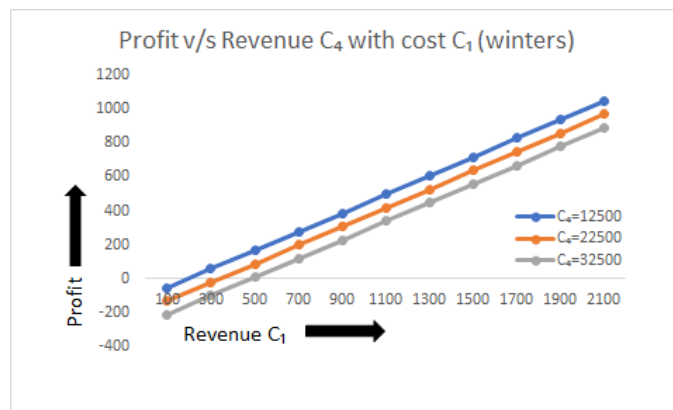


Figure 12: Profit v/s Failure Rate in Winters

Table 8:

Notations of the model	
Figures	Descriptions
11	Profit P1 increases as the revenue $C_0$ increases. $C_3=10500$ ; Profit $\geq$ according to $C_3$ , when $C_3$ is $\geq$ Rs.645.34, similarly for $C_3=20500$ where cut off point is Rs.573.039 $C_3=30500$ ; where cut off point is Rs. 500.7389
12	Profit P2 increases as the revenue $C_1$ increases. $C_4=12500$ ; Profit $\geq$ according to $C_4$ , when $C_1$ is $\geq$ Rs.203.345, similarly for $C_4=22500$ where cut off point is Rs.460.203. $C_4=32500$ ; where cut off point is Rs. 317.061

The MTSF, profit in the summers (P1), and profit in the winters (P2) graphs 8,9,10 exhibit a similar trend with failure rate  $\lambda$  and  $\lambda_1$ , which means that as the failure rate rises, the MTSF and profit fall.

## 21. CONCLUSION

The significance of implementing dependability in verka milk plant is analysed and concluded upon in this study. Using the parameters laid out in tables above, it has been shown that the second model generates more money after CBM is put into effect. Results from mathematical measurements and graphs showing that MTSF and Profit drop with increasing values of failure rates must be used to gain a more in-depth understanding of the essential real influencing elements and, in turn, enhance the reliability model. But the equations derived for MTSF, assessments of the system's functionality, and profit can be used to find alternative cut-off points related to the required rates, costs, and probabilities involved. The formulas for the proposed system can then be generated by plugging in the actual numbers for the relevant rates and costs. Important decisions about the system's dependability and profitability can be made with the help of graphs showing cut-off points for key rates, costs, and revenue.

## REFERENCES

- [1] Sheetal, Taneja G and Singh, D. Reliability and Profit Analysis of a System with Effect of Temperature on Operation, 2018.
- [2] Sheetal, Dalip Singh and Taneja, Gulshan. Reliability Analysis of a System Working in High Temperature Zones with Fault-Dependent Repair during Night hours, 2018.
- [3] Behboudi, Zohreh and Borzadaran, GR Mohtashami and Asadi, Majid. Reliability modeling of two-unit cold standby systems: a periodic switching approach, 2021.
- [4] Ram, Mangey and Singh, Suraj Bhan and Singh, Vijay Vir. Stochastic analysis of a standby system with waiting repair strategy, 2013.
- [5] Verdejo, Humberto and Awerkin, Almendra and Saavedra, Eugenio and Kliemann, Wolfgang and Vargas, Luis. Stochastic modeling to represent wind power generation and demand in electric power system based on real data, 2016.
- [6] Sharma, Rajiv Kumar and Kumar, Sunand. Performance modeling in critical engineering systems using RAM analysis, 2008.
- [7] Kumar, Ashok and Agarwal, Manju. A review of standby redundant systems, 1980.
- [8] Aggarwal, Anil Kr and Singh, Vikram and Kumar, Sanjeev. Availability analysis and performance optimization of a butter oil production system: a case study, 2017.
- [9] Arthur, Neil. Dairy processing site performance improvement using reliability centered maintenance, 2004.

- [10] Sharma, Upasana and Kaur, Jaswinder. Study of two units standby system with one essential unit to increase its functioning,2016.
- [11] Sharma, Upasana and Kaur, Jaswinder. Cost benefit analysis of a compressor standby system with preference of service, repair and replacement is given to recently failed unit, 2016.
- [12] Sharma, Upasana and Kaur, Jaswinder. Evaluation of Various Reliability Measures of Three Unit Standby System Consisting of One Standby Unit and One Generator, 2016.

# STREAMLINING PRODUCT DEPLOYMENT: ENHANCING EFFICIENCY THROUGH KITTING PROCESSES

G. AYYAPPAN, S. SANKEETHA



Department of Mathematics, Puducherry Technological University, India.

Department of Mathematics, Saradha Gangadharan College, India.

ayyappanpec@hotmail.com      sangeetha.sivarajp@gmail.com

## Abstract

*Considering a single server with two queues that is prone to unreliability. The server offers a kitting process and performs necessary checks and rectifications when required. The arrival of items follows a Markovian arrival process, while the service is distributed based on a phase type distribution. The incoming products may exhibit issues such as poor quality or defects. If either of the queues is empty, the server is unable to provide the requested service and remains inactive. Furthermore, if all queues are empty, the server goes into a vacation mode. Breakdowns, repairs, instances of customers leaving without service (reneging), and vacation periods are all modeled using an exponential distribution. To gain insights into the performance of the queueing model, various performance metrics are analyzed and represented through 2D and 3D graphs.*

**Keywords:** Markovian Arrival Process, PH distribution, Vacation, Optional service, Breakdown and Repair.

## 1. INTRODUCTION

The Markov arrival process (MAP) is a widely employed modeling approach that captures the dynamic Markov structure underlying point processes. It offers adaptability and versatility, making it suitable for probabilistic models that employ matrix analysis techniques. Neuts [15] made significant contributions by proposing and extensively investigating the flexible nature of Markov point processes. MAP shares similarities with other point processes, including Markov-modulated Poisson processes, phase-like updating processes, and semi-Markov point processes. It enables the simulation of both updating and non-updating models, making it a valuable tool for studying arrival patterns. Chakravathy [7] has provided in-depth insights and extensive discussions on MAP, specifically focusing on its m-dimensional parameter matrix  $(D_0, D_1, D_2)$ , where  $D_0$  governs transitions associated with no arrivals and  $D_1$  and  $D_2$  controls alternations related to arrival events. This parameterization allows for effective control and analysis of arrival dynamics in various systems.

Wang et al. [28] presented a framework for optimizing the kitting process in manufacturing. It addresses the challenges of efficiently organizing and sequencing materials required for assembly operations. The authors propose a mathematical model to minimize the overall kitting time, reduce material handling, and improve productivity in manufacturing settings. Yadav et al. [26] focused on optimizing the kitting process in an automotive assembly line. It investigates the challenges associated with kitting and proposes a mathematical model for optimizing the allocation of parts to kits. A hybrid optimization approach is applied that combines genetic

algorithms and simulated annealing to minimize the total distance traveled by workers during the kitting process. The study provides insights into improving the efficiency of the kitting process in automotive manufacturing. Ayyappan and Nithya [6] studied a retrieval feature that allows customers who experience service unavailability to reattempt service after a certain period. The model considers priority services, where one type of customer is given priority over the other in terms of service. Breakdowns and repairs are differentiated, meaning that the server may require different amounts of time to recover from different types of failures. Synchronized renegeing is taken into account, which means that customers may abandon the queue simultaneously if their waiting time exceeds a specific threshold. Additionally, the model incorporates an optional vacation, allowing the server to take breaks during certain periods.

Zhang and Fang [29] introduce a novel optimization algorithm designed to enhance the efficiency of bulk service systems. These systems are frequently encountered in various industries, including manufacturing and transportation, where multiple units of work or customers are processed simultaneously. The primary objective of the proposed algorithm is to minimize service time and decrease waiting times for customers within bulk service systems. To achieve this, the algorithm combines two powerful optimization techniques: stochastic optimization and reinforcement learning. The algorithm works in iterations, continuously refining its policies based on feedback from the system. It collects data on customer arrival patterns, service times, and queue lengths, which are then used to update the stochastic optimization models and reinforce the learned policies. This iterative process allows the algorithm to adapt to dynamic changes in the system and continuously optimize its performance. Li and Li [13] focused on optimizing bulk service systems that involve parallel servers. It addresses the challenges associated with efficiently allocating and coordinating multiple servers to improve system performance. The authors propose novel optimization algorithms and strategies to minimize service time and reduce waiting times for customers.

Smith and Johnson [22] investigated the influence of bulk service providers on the overall performance of supply chains. Also examines how the involvement of bulk service providers affects various aspects of supply chain operations, including efficiency, cost, and customer satisfaction. The impact of bulk service providers on key performance indicators are analyzed such as order fulfillment, inventory management, and lead times. Additionally, it highlights the importance of establishing effective collaboration and coordination mechanisms between bulk service providers and other supply chain stakeholders. Also emphasize on the significance of information sharing, communication, and performance monitoring to ensure optimal supply chain performance. Wang et al. [27] presents a hybrid optimization approach specifically tailored for bulk service systems in e-commerce warehouses. The authors combine mathematical modeling, simulation, and metaheuristic algorithms to enhance the efficiency of warehouse operations, such as order picking, packing, and shipping. The proposed approach aims to reduce order fulfillment time and improve customer satisfaction in e-commerce fulfillment centers.

Arun et al.[2] analyzed a bulk service queue with server breakdowns, balking, and renegeing. It provides a detailed analysis of the system's performance measures, such as the expected waiting time and the expected queue length, under different scenarios. Sun and Zhang [23] focused on the development of a bulk service system specifically designed for autonomous mobile robots, the growing demand for efficient and flexible service systems in industries where autonomous mobile robots are utilized. These systems involve the simultaneous processing of multiple tasks or requests, and efficient management is crucial to optimize performance and resource utilization. A comprehensive design framework for a bulk service system is proposed that integrates autonomous mobile robots. They outline the key components of the system, including task allocation, robot navigation, and coordination mechanisms. The findings of the study demonstrate the advantages of incorporating autonomous mobile robots into bulk service systems. The proposed design framework provides a blueprint for developing efficient and

scalable systems that can adapt to changing demands and optimize resource allocation.

Arivudainambi and Arivudainambi [1] studied a mathematical model for analyzing a bulk service queue with multiple vacations, server breakdowns, and general service times. It provides a detailed analysis of the system's performance measures, such as the expected waiting time and the expected queue length. Li and Zhang [14] proposed an optimal control policies for a bulk service queue with impatient customers and time-varying arrival rates. The proposed policies are designed to minimize the total expected cost, including waiting costs and service costs, under different operating conditions. Saroja and Saravanarajan [20] studied bulk service queueing models with server vacations and feedback controls. It provides a detailed analysis of the system's performance measures, such as the expected waiting time and the expected queue length, under different scenarios. Ayyappan and Meena [5] examined the service rate that gradually declines until degradation is fixed. After completing a certain number of services ( $K$ ), the degradation is addressed. During the service period, the server may experience a breakdown at any moment, triggering an immediate repair process. Once the service is complete, the server transitions to the close-down process. If there are no customers in the system when the server returns from vacation, the server will wait until a customer arrives. If a customer arrives without a starting failure, the server provides service. However, if there is a starting failure, the server immediately goes into the repair process.

Thottan and DeVeciana [24] presented a vacation model that incorporates autonomous server vacations and customer impatience. The research focuses on analyzing the performance of queueing systems under such conditions and investigates the impact of autonomous server vacations and customer impatience on system efficiency. Huang and Li [8] investigated on optimization of vacation queues that involve multiple vacation periods and general service times. The authors investigate the problem of determining optimal control policies for allocating vacation time and managing service rates in order to optimize various performance measures. They consider system characteristics such as queue length, waiting time, and system utilization. By analyzing the impact of different control policies on the system's performance, the authors provide insights into the efficient management of vacation queues. Their research contributes to the development of strategies for optimizing service allocation and improving the overall efficiency of queueing systems with multiple vacation periods and general service times. Anis et al. [4] explored the analysis of a finite-buffer queue that incorporates server vacations and customer impatience. It investigates the performance measures of the queueing system, including queue length, waiting time, and server utilization. The study provides understanding the enhancement of buffer size, vacation policies, and customer impatience management.

Srinivasan and Sriram [21] analyzed on studying vacation queues where the server is subject to breakdowns and repair. The authors analyze the impact of server breakdowns on the performance of the queueing system. They investigate various performance measures such as queue lengths, waiting times, and server utilization during both normal operation and breakdown periods. The study provides insights into the optimization of repair policies to minimize system downtime and improve overall system performance. By considering the combined effect of vacations and server breakdowns, the authors contribute to the understanding of real-world queueing systems where service interruptions due to breakdowns are common. Kim et al.[10] researched on vacation models that consider customer abandonments. It investigates the impact of customer abandonment behavior on queueing systems during vacation periods. The study provides perception on optimization of vacation policies and customer abandonment management.

Rakesh Kumar et al. [19] examined a single-server Markovian queueing model that incorporated customer impatience, including balking and renegeing, alongside a threshold mechanism and customer retention. They employed probability generating functions to analyze the model's transient behavior. Kalyanaraman and Janani [9] addresses a finite population Poisson queue em-

ploying a fixed batch service rule. Following each service, the server goes on vacation, regardless of queue size, providing service at a reduced rate during this period. The research calculates system size probabilities, derives performance metrics, and also explores an infinite population model with limited waiting room capacity as a secondary model. Krishnamurthy et al. [11] centers on the examination of a queuing system characterized by its multi-stage bulk service approach and the availability of service in batches. Within this system, incoming customers are initially grouped into batches before undergoing bulk servicing. The research extensively presents mathematical derivations pertaining to performance metrics, including system size, mean waiting time, and mean service time.

Raina Raj and Selvamuthu Dharmaraja [17] introduces an architectural framework that prioritizes energy efficiency within the SAT network, with a particular focus on HAPs. Furthermore, a stochastic model is proposed to account for three distinct states of energy conservation for HAPs, including modes of power conservation, standby, and rest, where energy consumption is minimal or negligible. Upon the arrival of a data packet, HAPs promptly transition to active service mode, ensuring the entire system operates in an active state. Anilkumar and Jose [3] examines a discrete-time inventory model  $(s, S)$  is investigated, featuring Bernoulli process customer arrivals and geometrically distributed service and replenishment times. When inventory drops to zero due to customer service or lack of replenishment, the system can accommodate a maximum of  $k$  customers, with any excess customers considered lost until replenishment occurs. Rakesh Kumar et al. [18] conducted a comprehensive study examining the utilization of queuing theory in the analysis of cloud computing systems. Their research specifically delved into the phenomenon of task reneging, where requests are dropped from the queue due to user impatience, deadlines, security protocols, or active queue management strategies.

## 2. MOTIVATION

In a software development company, a team is working on creating a new application that consists of multiple modules and features. Rather than developing and delivering each module individually, they adopt a kitting process to streamline the deployment process and improve efficiency. In this kitting process, each module or feature is treated as a separate item and is placed in a dedicated queue. The server, which represents the deployment team, retrieves the modules from the queues and starts assembling the software kit. They integrate the modules, perform necessary configurations, and ensure compatibility between different components.

Once the kit is assembled, the server performs thorough testing and quality assurance checks to verify the functionality and stability of the software. If any issues are identified, such as bugs or compatibility conflicts, the server rectifies them before proceeding. Once the kit passes the testing phase, it is packaged for release to the end-users or clients. The server ensures that all required documentation, user guides, and support materials are included in the kit before delivering it. By employing the kitting process in software development, the company streamlines the deployment process, reduces errors, and ensures that the end-users receive a comprehensive and well-tested software package.

## 3. MATHEMATICAL FORMULATION

This model considers two types of arrivals within a system. The first type follows a Markovian arrival process and has infinite capacity, while the second type has a finite capacity of  $K$ . The server is responsible for the packing service, which follows a phase type distribution denoted as  $(\alpha_1, T_1)$ . The equation  $T_1^0 + T_1 e = 0$  holds true, where  $T_1^0$  represents a column vector. Once the packing is completed using the kitting process, the server proceeds to verify the checklist for the packed product.

If the checklist is satisfied, the product is deemed ready for the outlet. Otherwise, the server initiates the rechecking and rectification process. This rechecking process follows a phase type distribution denoted as  $(\alpha_2, T_2)$ . The equation  $T_2^0 + T_2 e = 0$  holds true, where  $T_2^0$  represents a column vector. If either of the queues becomes empty, the server remains idle. However, when both queues are empty, the server goes on vacation, with the vacation parameter  $\eta$  following an exponential distribution.

Additionally, the server is subject to breakdown during both the packing service and rechecking, with a breakdown parameter  $\zeta$  following an exponential distribution. When the server experiences a breakdown while serving, it completes the ongoing service and then enters a repair process with a parameter  $\gamma$  following an exponential distribution. Moreover, the products in both queues are susceptible to renegeing, indicated by parameters  $\delta_1$  and  $\delta_2$ , respectively, following an exponential distribution. Reneging can occur due to factors such as lack of quality or defects.

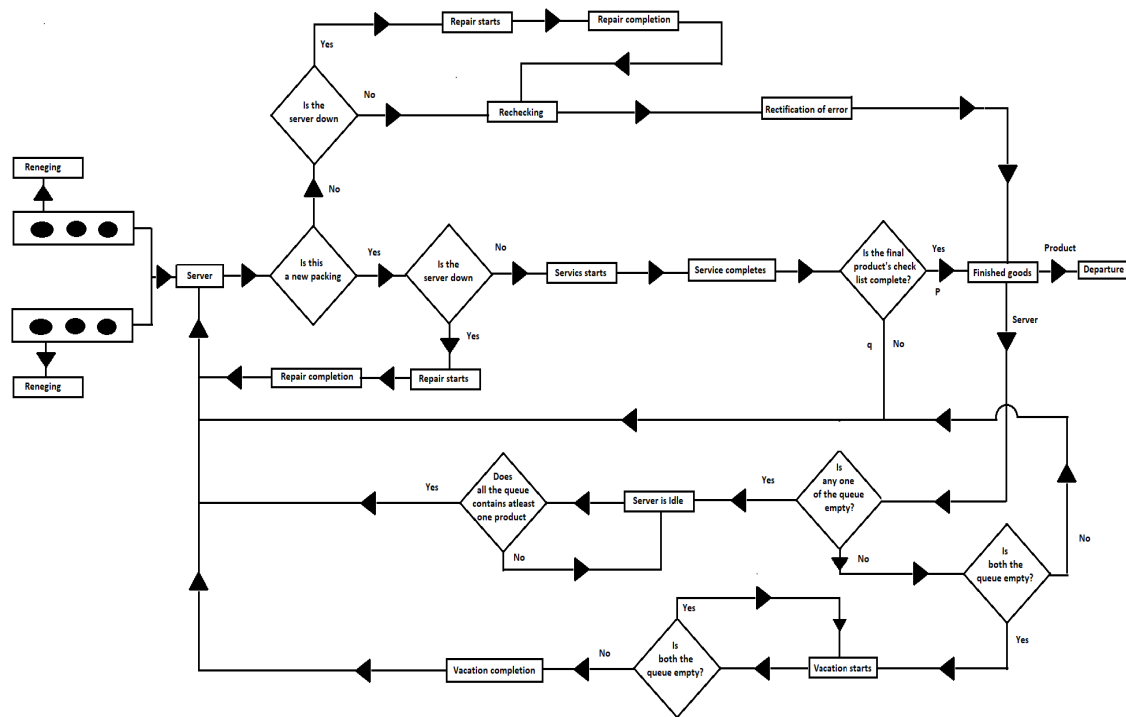


Figure 1: Schematic Representation of Our Model

In pursuit of a matrix-geometric solution, the model is explored within the framework of a QBD (Quasi-Birth-Death) process. For a comprehensive exploration of Matrix Analytic Methods, refer to the works of Neuts [16] and Latouche and Ramaswami [12]. The QBD model's state space is formally defined, and an examination of the infinitesimal generator's structure is carried out, leveraging the subsequent notational conventions.

Let

- $I_j$  is the identity matrix of dimension  $j$ .
- $e_1$  is the column vector of dimension  $m_1 m_2 [(n_1 + n_2)(1 + K) + (4 + 3K)]$  with its entries 1.
- $N_1(t)$  indicates the total number of items in the type I queue.
- $N_2(t)$  indicates the total number of items in the type II queue.
- $S(t)$  indicates the position of the server.



where

$$S(t) = \begin{cases} 0, & \text{if server is idle} \\ 1, & \text{if the server is engaged with packing} \\ 2, & \text{if the server is engaged with rework} \\ 3, & \text{if the server faces breakdown while packing} \\ 4, & \text{if the server faces breakdown while rework} \\ 5, & \text{if server is on vacation} \end{cases}$$

- $J_1(t)$  indicates the service phase when the server is engaged with packing.
- $J_2(t)$  indicates it the service phase when the server is engaged with packing.
- $M_1(t)$  indicates the phase of the Markovian Arrival Process for type I queue.
- $M_2(t)$  indicates the phase of the Markovian Arrival Process for type II queue.

Let  $\{(N_1(t), N_2(t), S(t), J_1(t), J_2(t), M_1(t), M_2(t)); t \geq 0\}$  represent the continuous time Markov chain for the QBD process with the state space.

$$\Omega = l(0) \cup l(i)$$

where,

$$l(0) = \{(0, j, 0, s_1, s_2) : 0 \leq j \leq K, 1 \leq s_1 \leq m_1, 1 \leq s_2 \leq m_2\}$$

For  $i \geq 0$ ,

$$l(i) = \cup \{(0, j, 1, r_1, s_1, s_2) : 1 \leq j \leq K, 1 \leq r_1 \leq n_1, 1 \leq s_1 \leq m_1, 1 \leq s_2 \leq m_2\} \\ \cup \{(0, j, 2, r_2, s_1, s_2) : 1 \leq j \leq K, 1 \leq r_2 \leq n_2, 1 \leq s_1 \leq m_1, 1 \leq s_2 \leq m_2\} \\ \cup \{(0, j, l, s_1, s_2) : 0 \leq j \leq K, 3 \leq l \leq 5, 1 \leq s_1 \leq m_1, 1 \leq s_2 \leq m_2\}$$

$$\text{For } i \geq 1, \quad l(i) = \{(i, 0, 0, s_1, s_2) : 1 \leq s_1 \leq m_1, 1 \leq s_2 \leq m_2\}$$

The infinitesimal matrix generation of the QBD process is given by

$$Q = \begin{bmatrix} B_{00} & B_{01} & 0 & 0 & 0 & 0 & \dots \\ B_{10} & A_1 & A_0 & 0 & 0 & 0 & \dots \\ 0 & A_2 & A_1 & A_0 & 0 & 0 & \dots \\ 0 & 0 & A_2 & A_1 & A_0 & 0 & \dots \\ \dots & \dots & \dots & \ddots & \ddots & \ddots & \dots \end{bmatrix}$$

where each of its block matrix are as follows,

$$B_{00} = \begin{bmatrix} b_{00}^{11} & b_{00}^{12} & 0 & 0 & \dots & 0 & 0 \\ b_{00}^{21} & b_{00}^{22} & b_{00}^{23} & 0 & \dots & 0 & 0 \\ 0 & b_{00}^{32} & b_{00}^{22} & b_{00}^{23} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & b_{00}^{32} & b_{00}^{M+1M+1} \end{bmatrix}$$

$$b_{00}^{11} = \begin{bmatrix} b_{11}^{011} & \alpha_1 q T_1^0 \otimes I_{m_1 m_2} & \zeta I_{m_1 m_2} & 0 & b_{14}^{011} \\ 0 & b_{22}^{011} & 0 & \zeta I_{m_1 m_2} & b_{24}^{011} \\ \gamma I_{m_1 m_2} & 0 & b_{33}^{011} & 0 & 0 \\ 0 & \gamma I_{m_1 m_2} & 0 & b_{33}^{011} & 0 \\ 0 & 0 & 0 & 0 & I_{m_2} \otimes D_0 \otimes I_{m_1} \end{bmatrix}$$

$$\begin{aligned}
 b_{11}^{011} &= I_{m_1} \otimes (T_1 \oplus (D_0 - \zeta I_{m_2})) \\
 b_{14}^{011} &= e_{m_2} \otimes \alpha_1 p T_1^0 \otimes I_{m_1} \\
 b_{22}^{011} &= I_{m_2} \otimes (T_2 \oplus (D_0 - \zeta I_{m_1})) \\
 b_{24}^{011} &= e_{m_2} \otimes \alpha_2 T_2^0 \otimes I_{m_1} \\
 b_{33}^{011} &= I_{m_2} \otimes (D_0 - \gamma I_{m_1})
 \end{aligned}$$

$$b_{00}^{12} = \begin{bmatrix} 0 & I_{n_1 m_1} \otimes D_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{n_2 m_1} \otimes D_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{m_1} \otimes D_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{m_1} \otimes D_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{m_1} \otimes D_2 \end{bmatrix}$$

$$b_{00}^{21} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \delta_2 I_{m_1 m_2} \\ \delta_2 I_{n_1 m_1 m_2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \delta_2 I_{n_2 m_1 m_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \delta_2 I_{m_1 m_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \delta_2 I_{m_1 m_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \delta_2 I_{m_1 m_2} \end{bmatrix}$$

$$b_{00}^{22} = \begin{bmatrix} b_{11}^{022} & 0 & 0 & 0 & 0 & 0 \\ \alpha_1 p T_1^0 \otimes I_m & b_{22}^{022} & \alpha_1 q T_1^0 \otimes I_m & \zeta I_{mn} & 0 & 0 \\ \alpha_2 T_2^0 \otimes I_m & 0 & b_{33}^{022} & 0 & \zeta I_{mn} & 0 \\ 0 & \gamma I_{mn} & 0 & b_{44}^{022} & 0 & 0 \\ 0 & 0 & \gamma I_{mn} & 0 & b_{44}^{022} & 0 \\ 0 & 0 & 0 & 0 & 0 & b_{11}^{022} \end{bmatrix}$$

$$\begin{aligned}
 b_{11}^{022} &= I_{m_2} \otimes (D_0 - \delta_2 I_{m_1}) \\
 b_{22}^{022} &= I_{m_1} \otimes (T_1 \oplus (D_0 - (\zeta + \delta_2) I_{m_2})) \\
 b_{33}^{022} &= I_{m_2} \otimes (T_2 \oplus (D_0 - (\zeta + \delta_2) I_{m_1})) \\
 b_{44}^{022} &= I_{m_2} \otimes (D_0 - (\gamma + \delta_2) I_{m_1})
 \end{aligned}$$

$$b_{00}^{23} = \begin{bmatrix} I_{m_1} \otimes D_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{n_1 m_1} \otimes D_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{n_2 m_1} \otimes D_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{m_1} \otimes D_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{m_1} \otimes D_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{m_1} \otimes D_2 \end{bmatrix}$$

$$b_{00}^{32} = \begin{bmatrix} \delta_2 I_{m_1 m_2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \delta_2 I_{n_1 m_1 m_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \delta_2 I_{n_2 m_1 m_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \delta_2 I_{m_1 m_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta_2 I_{m_1 m_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \delta_2 I_{m_1 m_2} \end{bmatrix}$$

$$b_{00}^{M+1M+1} = \begin{bmatrix} b_{11}^{0M+1M+1} & 0 & 0 & 0 & 0 & 0 \\ b_{14}^{011} & b_{22}^{0M+1M+1} & b_{23}^{0M+1M+1} & \zeta I_{n_1 m_1 m_2} & 0 & 0 \\ b_{24}^{011} & 0 & b_{33}^{0M+1M+1} & 0 & \zeta I_{m_1 m_2} & 0 \\ 0 & \gamma I_{m_1 m_2} & 0 & b_{44}^{0M+1M+1} & 0 & 0 \\ 0 & 0 & \gamma I_{m_1 m_2} & 0 & b_{44}^{0M+1M+1} & 0 \\ 0 & 0 & 0 & 0 & 0 & b_{55}^{0M+1M+1} \end{bmatrix}$$

$$\begin{aligned}
 b_{11}^{0M+1M+1} &= I_{m_2} \otimes (D_0 + D_2 - \delta_2 I_{m_1}) \\
 b_{22}^{0M+1M+1} &= I_{m_1} \otimes (T_1 \oplus (D_0 + D_2 - (\xi + \delta_2) I_{m_2})) \\
 b_{23}^{0M+1M+1} &= e_{m_2} \otimes \alpha_1 q T_1^0 \otimes I_{m_1} \\
 b_{33}^{0M+1M+1} &= I_{m_2} \otimes (T_2 \oplus (D_0 + D_2 - (\xi + \delta_2) I_{m_1})) \\
 b_{44}^{0M+1M+1} &= I_{m_2} \otimes (D_0 + D_2 - (\gamma + \delta_2) I_{m_1}) \\
 b_{55}^{0M+1M+1} &= I_{m_2} \otimes D_0 + D_2 - \delta_2 I_{m_1}
 \end{aligned}$$

$$B_{01} = \begin{bmatrix} b_{01}^{11} & 0 & 0 & \cdots & 0 \\ 0 & b_{01}^{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & b_{01}^{22} \end{bmatrix}$$

$$b_{01}^{11} = \begin{bmatrix} 0 & I_{n_1 m_2} \otimes D_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{n_2 m_2} \otimes D_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{m_2} \otimes D_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{m_2} \otimes D_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{m_2} \otimes D_1 \end{bmatrix}$$

$$b_{01}^{22} = \begin{bmatrix} I_{m_2} \otimes D_1 & 0 & 0 & 0 & 0 \\ I_{n_1 m_2} \otimes D_1 & 0 & 0 & 0 & 0 \\ 0 & I_{n_2 m_2} \otimes D_1 & 0 & 0 & 0 \\ 0 & 0 & I_{m_2} \otimes D_1 & 0 & 0 \\ 0 & 0 & 0 & I_{m_2} \otimes D_1 & 0 \\ 0 & 0 & 0 & 0 & I_{m_2} \otimes D_1 \end{bmatrix}$$

$$B_{10} = \begin{bmatrix} b_{10}^{11} & 0 & 0 & \cdots & 0 & 0 \\ b_{10}^{21} & b_{10}^{22} & 0 & \cdots & 0 & 0 \\ 0 & b_{10}^{32} & b_{10}^{22} & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & b_{10}^{32} & b_{10}^{22} \end{bmatrix}$$

$$b_{10}^{11} = \begin{bmatrix} 0 & 0 & 0 & 0 & \delta_1 I_{m_1 m_2} \\ \delta_1 I_{n_1 m_1 m_2} & 0 & 0 & 0 & 0 \\ 0 & \delta_1 I_{n_2 m_1 m_2} & 0 & 0 & 0 \\ 0 & 0 & \delta_1 I_{m_1 m_2} & 0 & 0 \\ 0 & 0 & 0 & \delta_1 I_{m_1 m_2} & 0 \\ 0 & 0 & 0 & 0 & \delta_1 I_{m_1 m_2} \end{bmatrix}$$

$$b_{10}^{21} = \begin{bmatrix} e_{m_2} \otimes \alpha_1 p T_1^0 \otimes I_{m_1} & 0 & 0 & 0 & 0 \\ e_{m_2} \otimes \alpha_2 T_2^0 \otimes I_{m_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$b_{10}^{22} = \begin{bmatrix} 0 & \delta_1 I_{n_1 m_1 m_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \delta_1 I_{n_2 m_1 m_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \delta_1 I_{m_1 m_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta_1 I_{m_1 m_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \delta_1 I_{m_1 m_2} \end{bmatrix}$$

$$b_{10}^{32} = \begin{bmatrix} 0 & e_{m_2} \otimes \alpha_1 p T_1^0 \otimes I_{m_1} & 0 & 0 & 0 & 0 \\ 0 & e_{m_2} \otimes \alpha_2 T_2^0 \otimes I_{m_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} a_1^{11} & a_1^{12} & 0 & 0 & \cdots & 0 & 0 \\ a_1^{21} & a_1^{22} & a_1^{23} & 0 & \cdots & 0 & 0 \\ 0 & b_1^{32} & a_1^{22} & a_1^{23} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & a_1^{32} & a_1^{M+1M+1} \end{bmatrix}$$

$$a_1^{11} = \begin{bmatrix} a_{11}^{11} & 0 & 0 & 0 & 0 & 0 \\ b_{14}^{011} & a_{22}^{11} & b_{23}^{0M+1M+1} & \zeta I_{m_1 m_2} & 0 & 0 \\ b_{24}^{011} & 0 & a_{33}^{11} & 0 & \zeta I_{m_1 m_2} & 0 \\ 0 & \gamma I_{m_1 m_2} & 0 & a_{44}^{11} & 0 & 0 \\ 0 & 0 & \gamma I_{m_1 m_2} & 0 & a_{44}^{11} & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{66}^{11} \end{bmatrix}$$

$$\begin{aligned} a_{11}^{11} &= I_{m_2} \otimes (D_0 - \delta_1 I_{m_1}) \\ a_{22}^{11} &= I_{m_1} \otimes (T_1 \oplus D_0 - (\zeta + \delta_1) I_{m_2}) \\ a_{33}^{11} &= I_{m_2} \otimes (T_2 \oplus D_0 - (\zeta + \delta_1) I_{m_1}) \\ a_{44}^{11} &= I_{m_2} \otimes D_0 - (\gamma + \delta_1) I_{m_1} \\ a_{66}^{11} &= I_{m_2} \otimes D_0 - \delta_1 I_{m_1} \end{aligned}$$

$$a_1^{12} = \begin{bmatrix} I_{m_1} \otimes D_2 & 0 & 0 & 0 & 0 \\ I_{n_1 m_1} \otimes D_2 & 0 & 0 & 0 & 0 \\ 0 & I_{n_2 m_1} \otimes D_2 & 0 & 0 & 0 \\ 0 & 0 & I_{m_1} \otimes D_2 & 0 & 0 \\ 0 & 0 & 0 & I_{m_1} \otimes D_2 & 0 \\ 0 & 0 & 0 & 0 & I_{m_1} \otimes D_2 \end{bmatrix}$$

$$a_1^{21} = \begin{bmatrix} 0 & \delta_2 I_{n_1 m_1 m_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \delta_2 I_{n_2 m_1 m_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \delta_2 I_{m_1 m_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta_2 I_{m_1 m_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \delta_2 I_{m_1 m_2} \end{bmatrix}$$

$$a_1^{22} = \begin{bmatrix} a_{11}^{122} & e_{m_2} \otimes \alpha_1 p T_1^0 \otimes I_{m_1} & \zeta I_{n_1 m_1 m_2} & 0 & 0 \\ 0 & a_{22}^{122} & 0 & \zeta I_{n_1 m_1 m_2} & 0 \\ \gamma I_{m_1 m_2} & 0 & a_{33}^{122} & 0 & 0 \\ 0 & \gamma I_{m_1 m_2} & 0 & a_{44}^{122} & 0 \\ \eta I_{m_1 m_2} & 0 & 0 & 0 & a_{44}^{122} \end{bmatrix}$$

$$\begin{aligned} a_{11}^{122} &= I_{m_1} \otimes (T_1 \oplus D_0 - (\zeta + \delta_1 + \delta_2) I_{m_2}) \\ a_{22}^{122} &= I_{m_2} \otimes (T_2 \oplus D_0 - (\zeta + \delta_1 + \delta_2) I_{m_1}) \\ a_{33}^{122} &= I_{m_1} \otimes D_0 - (\gamma + \delta_1 + \delta_2) I_{m_2} \\ a_{44}^{122} &= I_{m_1} \otimes D_0 - (\eta + \delta_1 + \delta_2) I_{m_2} \end{aligned}$$

$$a_1^{23} = \begin{bmatrix} I_{n_1 m_1} \otimes D_2 & 0 & 0 & 0 & 0 \\ 0 & I_{n_2 m_1} \otimes D_2 & 0 & 0 & 0 \\ 0 & 0 & I_{m_1} \otimes D_2 & 0 & 0 \\ 0 & 0 & 0 & I_{m_1} \otimes D_2 & 0 \\ 0 & 0 & 0 & 0 & I_{m_1} \otimes D_2 \end{bmatrix}$$

$$a_1^{M+1M+1} = \begin{bmatrix} a_{11}^{1M+1M+1} & e_{m_2} \otimes \alpha_1 p T_1^0 \otimes I_{m_1} & \zeta I_{n_1 m_1 m_2} & 0 & 0 \\ 0 & a_{22}^{1M+1M+1} & 0 & \zeta I_{n_2 m_1 m_2} & 0 \\ \gamma I_{m_1 m_2} & 0 & a_{33}^{1M+1M+1} & 0 & 0 \\ 0 & \gamma I_{m_1 m_2} & 0 & a_{33}^{1M+1M+1} & 0 \\ \eta I_{m_1 m_2} & 0 & 0 & 0 & a_{44}^{1M+1M+1} \end{bmatrix}$$

$$a_{11}^{1M+1M+1} = I_{m_1} \otimes (T_1 \oplus D_0 + D_2 - (\zeta + \delta_1 + \delta_2)) I_{m_2}$$

$$a_{22}^{1M+1M+1} = I_{m_2} \otimes (T_2 \oplus D_0 + D_2 - (\zeta + \delta_1 + \delta_2)) I_{m_1}$$

$$a_{33}^{1M+1M+1} = I_{m_1} \otimes D_0 + D_2 - (\gamma + \delta_1 + \delta_2) I_{m_2}$$

$$a_{44}^{1M+1M+1} = I_{m_1} \otimes D_0 + D_2 - (\eta + \delta_1 + \delta_2) I_{m_2}$$

$$A_0 = \begin{bmatrix} a_0^{11} & 0 & 0 & \cdots & 0 \\ 0 & a_0^{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & a_0^{22} \end{bmatrix}$$

$$a_0^{11} = \begin{bmatrix} I_{m_2} \otimes D_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{n_1 m_2} \otimes D_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{n_2 m_2} \otimes D_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{m_2} \otimes D_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{m_2} \otimes D_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{m_2} \otimes D_1 \end{bmatrix}$$

$$a_0^{22} = \begin{bmatrix} I_{n_1 m_2} \otimes D_1 & 0 & 0 & 0 & 0 \\ 0 & I_{n_2 m_2} \otimes D_1 & 0 & 0 & 0 \\ 0 & 0 & I_{m_2} \otimes D_1 & 0 & 0 \\ 0 & 0 & 0 & I_{m_2} \otimes D_1 & 0 \\ 0 & 0 & 0 & 0 & I_{m_2} \otimes D_1 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} a_2^{11} & 0 & 0 & \cdots & 0 & 0 \\ a_2^{21} & a_2^{22} & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_2^{21} & a_2^{22} \end{bmatrix}$$

$$a_2^{11} = \begin{bmatrix} \delta_1 I_{m_1 m_2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \delta_1 I_{n_1 m_1 m_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \delta_1 I_{n_2 m_1 m_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \delta_1 I_{m_1 m_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta_1 I_{m_1 m_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \delta_1 I_{m_1 m_2} \end{bmatrix}$$

$$a_2^{21} = \begin{bmatrix} e_{m_2} \otimes \alpha_1 p T_1^0 \otimes I_{m_1} & 0 & 0 & 0 & 0 \\ e_{m_2} \otimes \alpha_1 T_2^0 \otimes I_{m_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$a_2^{22} = \begin{bmatrix} \delta_1 I_{n_1 m_1 m_2} & 0 & 0 & 0 & 0 \\ 0 & \delta_1 I_{n_2 m_1 m_2} & 0 & 0 & 0 \\ 0 & 0 & \delta_1 I_{m_1 m_2} & 0 & 0 \\ 0 & 0 & 0 & \delta_1 I_{m_1 m_2} & 0 \\ 0 & 0 & 0 & 0 & \delta_1 I_{m_1 m_2} \end{bmatrix}$$

#### 4. ANALYSIS OF THE STABILITY CONDITION

Determining the stability of a system is crucial to ensure its smooth operation and efficient handling of incoming arrivals. The concept of traffic intensity serves as a key metric in assessing system stability. By comparing the average arrival rate with the average service rate over the long run, we can gauge whether the system is capable of managing the workload effectively. For stability, it is desirable that the traffic intensity remains below 1, indicating that the system can handle the incoming arrivals without becoming overwhelmed.

Analyzing the stability of a Markovian arrival process (MAP) presents unique challenges compared to simpler arrival processes like the Poisson process. This is due to the diverse inter arrival time distributions that MAPs can exhibit. To explore stability conditions in MAPs, researchers employ matrix-analytic methods and simulation-based methods. These approaches involve analyzing matrices and eigenvalues to ascertain the system's stability. Simulation-based methods, in particular, prove valuable when dealing with complex systems that lack analytical solutions, enabling researchers to simulate and study system behavior under varying conditions.

Let  $A$  be an irreducible infinitesimal generator matrix of order  $m_1 m_2 [(n_1 + n_2)(1 + K) + (4 + 3K)]$ . We can decompose  $A$  as  $A = A_0 + A_1 + A_2$ . The vector  $\wp = (\wp_0, \wp_1, \wp_2, \dots, \wp_{K+1})$  represents an invariant probability vector. It satisfies the conditions  $\wp A = 0$  and  $\wp e = 1$ , where  $\wp e$  denotes the dot product between  $\wp$  and the vector  $e$ .

$$\begin{aligned} \wp_0 [a_0^{11} + a_1^{11} + a_2^{11}] + \wp_1 [a_1^{21} + a_2^{21}] &= 0. \\ \wp_0 [a_1^{21}] + \wp_1 [a_0^{22} + a_1^{22} + a_2^{22}] + \wp_2 [b_1^{32} + a_2^{21}] &= 0. \\ \wp_{i-1} [a_1^{23}] + \wp_i [a_0^{22} + a_1^{32} + a_2^{32}] + \wp_{i+1} [b_1^{32} + a_2^{21}] &= 0, \text{ for } i = 1 \text{ to } K - 1. \\ \wp_K [a_1^{23}] + \wp_{K+1} [a_0^{22} + a_1^{K+1, K+1} + a_2^{22}] &= 0. \end{aligned}$$

Given the normalizing condition  $\wp e = 1$ , in a stable system, it is necessary that

$$\begin{aligned} \wp A_0 e_{mn[l(K+1)+1]} &< \wp A_2 e_{mn[l(K+1)+1]}. \\ \wp_0 a_0^{11} + (\wp_1 + \wp_2 + \dots + \wp_{K+1}) a_0^{22} &< \wp_0 a_2^{11} + (\wp_1 + \wp_2 + \dots + \wp_K) a_2^{21} + (\wp_1 + \wp_2 + \dots + \wp_{K+1}) a_2^{22}. \end{aligned}$$

#### 5. THE VECTOR OF INVARIANT PROBABILITIES

The crucial role of capturing the system's steady-state behavior is played by the invariant probability vector, which is symbolically represented as  $X$ . In order to obtain the vector  $X$ , it is necessary to solve the system of equations represented as  $XQ = 0$ , while simultaneously ensuring the normalization condition  $Xe = 1$ . Once the stability requirements are fulfilled, the remaining components of  $X$  can be computed using an iterative approach. It is important to emphasize that  $X$  can be partitioned into sub-vectors, including  $X_0$  and  $X_i$  for  $i \geq 1$ , which have specific dimensions based on the system's characteristics. The dimension of  $X_0$  is  $m_1 m_2 [(n_1 + n_2)(1 + K) + (3 + 4K)]$ , while  $X_i$  for  $i \geq 1$  has a dimension of  $m_1 m_2 [(n_1 + n_2)(1 + K) + (4 + 3K)]$ . Precisely calculating the values of  $X_0$  and  $X_i$  involves considering the unique properties and parameters of the system at hand. The expression for  $X_i$  can be represented as:

$$X_i = X_1 R^{i-1}, \quad i = 2, 3, 4, \dots,$$

Here,  $R$  refers to the rate matrix, which serves as the minimal non-negative solution to the matrix quadratic equation.

$$R^2A_2 + RA_1 + A_0 = 0$$

The boundary states, represented as  $X_0$  and  $X_1$ , are determined by solving the following equations:

$$\begin{aligned} X_0B_{00} + X_1B_{10} &= 0 \\ X_0B_{01} + X_1(A_1 + RA_2) &= 0 \end{aligned}$$

These equations are subject to the normalizing condition:

$$X_0e + X_1(I - R)^{-1}e = 1$$

It's worth noting that Latouche and Ramaswamy [12] have improved the computation of the rate matrix  $R$  by introducing the Logarithmic Reduction Algorithm. This algorithm simplifies the process of obtaining  $R$ , making it more efficient and straightforward.

$$\text{Step 1 : } H \leftarrow (-A_1)^{-1}A_0, L \leftarrow (-A_1)^{-1}A_2, G = L \text{ and } T = H.$$

$$\begin{aligned} \text{Step 2 : } U &= HL + LH; \\ M &= H^2; \\ H &= (I - U)^{-1}M; \\ M &= L^2; \\ L &= (I - U)^{-1}M; \\ G &= G + TL; \\ T &= TH; \end{aligned}$$

continue Step 1 until  $\|e - Ge\|_\infty < \epsilon$ .

$$\text{Step 3 : } R = -A_0(A_1 + A_0G)^{-1}.$$

## 6. EXAMINATION OF BUSY PERIOD

In the context of queueing theory, an essential aspect is the analysis of the busy period. This term refers to the duration that starts when a customer enters an empty queue and concludes when the queue once again becomes vacant. However, when dealing with Quasi-Birth-Death (QBD) processes, a different concept known as the "fundamental period" emerges. The fundamental period characterizes the duration needed for the system to shift from level  $i$  to level  $i - 1$ , where  $i$  assumes a value of 2 or greater. It's worth noting that special considerations are needed for boundary states, particularly when  $i$  takes on values of 0 or 1. Furthermore, when examining all levels  $i$  greater than or equal to 2, it becomes evident that there is a total of  $mn[l(1 + K) + 1]$  states. This expression quantifies the number of states associated with each level within the queueing model.

Notations:

- $G_{vv'}(k, x)$  corresponds to the likelihood that the QBD process enters level  $u - 1$  at time  $t = 0$  after undergoing precisely  $k$  leftward transitions and arriving at state  $(u, v')$ , under the condition that it initially commenced in state  $(u, v)$  at time  $t = 0$ .
- The transition matrix  $\bar{G}_{vv'}(z, s)$  is defined as  $\sum_0^\infty z^k \int_0^\infty e^{-sx} dG_{vv'}(k, x)$ , where the conditions are  $|z| \leq 1$  and  $Re(s) \geq 0$ . This matrix incorporates a combination of infinite series and integrals to capture the intricate transitions inherent in the QBD process.
- $\bar{G}(z, s)$  takes the form of a matrix  $(G_{vv'}(z, s))$  and adheres to the equation  $\bar{G}(z, s) = z[sI - A_1]^{-1}A_2 + [sI - A_1]^{-1}A_0\bar{G}^2(z, s)$ , representing the interplay among various elements of the QBD process.

- In the context of the first passage time analysis,  $G = G_{vv'} = \bar{G}(0, 1)$  captures the behavior of the process in the absence of boundary states, providing insights into its performance without considering boundary effects.
- $\bar{G}_{(vv')}^{(1,0)}(k, x)$  is the conditional probability that enters the level 0 from 1 at time  $t = 0$ .
- $\bar{G}_{(vv')}^{(0,0)}(k, x)$  is the first conditional probability returning to level 0.
- $\mathfrak{R}_{1v}$  denotes the anticipated duration for the first passage between levels  $u$  and  $u - 1$  when the process is in state  $(u, v)$  at time  $t = 0$ .
- $\bar{\mathfrak{R}}_1$  is a column vector composed of the entries  $\mathfrak{R}_{1v}$ , representing the expected first passage times for different states.
- $\mathfrak{R}_{2v}$  stands for the average number of customers who receive service during the initial passage between levels  $u$  and  $u - 1$  when the process begins in state  $(u, v)$  at time  $t = 0$ .
- $\bar{\mathfrak{R}}_2$  is a column vector composed of the entries  $\mathfrak{R}_{2v}$ , signifying the average number of service completions during the first passage time for different states.
- $\bar{\mathfrak{R}}_1^{(1,0)}$  represents the average duration for the first passage from level 1 to 0 within the QBD process.
- $\bar{\mathfrak{R}}_2^{(1,0)}$  signifies the average number of completed services during the initial passage from level 1 to 0.
- $\bar{\mathfrak{R}}_1^{(0,0)}$  denotes the average time taken for the first return to level 0 within the QBD process.
- $\bar{\mathfrak{R}}_2^{(0,0)}$  represents the average number of completed services during the initial return to level 0.

The  $G$  matrix can be computed using the following expression, utilizing the previously determined rate matrix  $R$  obtained through the Logarithmic Reduction Algorithmic technique:

$$G = -[A_1 + RA_2]^{-1}A_2$$

For the boundary states, specifically 1 and 0, we can establish equations satisfied by  $\bar{G}^{(1,0)}(z, s)$  and  $\bar{G}^{(0,0)}(z, s)$ , respectively:

$$\begin{aligned}\bar{G}^{(1,0)}(z, s) &= z[sI - A_1]^{-1}B_{10} + [sI - A_1]^{-1}A_0\bar{G}(z, s)\bar{G}^{(1,0)}(z, s). \\ \bar{G}^{(0,0)}(z, s) &= z[sI - B_{00}]^{-1}B_{01}\bar{G}^{(1,0)}(z, s).\end{aligned}$$

Since  $G$ ,  $\bar{G}^{(1,0)}(z, s)$ , and  $\bar{G}^{(0,0)}(z, s)$  are stochastic in nature, we can readily compute moments as follows.

$$\begin{aligned}\mathfrak{R}_1 &= -\frac{\partial}{\partial s}\bar{G}(z, s)|_{s=0, z=1} = -[A_0(G + 1) + A_1]^{-1}e \\ \mathfrak{R}_2 &= \frac{\partial}{\partial z}\bar{G}(z, s)|_{s=0, z=1} = -[A_0(G + 1) + A_1]^{-1}A_2e \\ \mathfrak{R}_1^{(1,0)} &= -\frac{\partial}{\partial s}\bar{G}^{(1,0)}(z, s)|_{s=0, z=1} = -[A_1 + A_0G]^{-1}[A_0\mathfrak{R}_1 + e] \\ \mathfrak{R}_2^{(1,0)} &= \frac{\partial}{\partial z}\bar{G}^{(1,0)}(z, s)|_{s=0, z=1} = -[A_1 + A_0G]^{-1}[B_{10}e + A_0\mathfrak{R}_2] \\ \mathfrak{R}_1^{(0,0)} &= -\frac{\partial}{\partial s}\bar{G}^{(0,0)}(z, s)|_{s=0, z=1} = -B_{00}^{-1}[e + B_{01}\mathfrak{R}_1^{(1,0)}] \\ \mathfrak{R}_2^{(0,0)} &= \frac{\partial}{\partial z}\bar{G}^{(0,0)}(z, s)|_{s=0, z=1} = -B_{00}^{-1}B_{01}\mathfrak{R}_2^{(1,0)}.\end{aligned}$$



## 7. PERFORMANCE MEASURES

When a system reaches a steady-state, it signifies that the system has achieved stability and performance measures can be derived and examined. These performance measures play a vital role in evaluating the various aspects of system performance and determining its efficiency and effectiveness. By analyzing these measures, we can gain valuable insights into the system's behavior and identify areas that require improvement to enhance overall performance. Performance measures serve as quantitative indicators that shed light on important system characteristics such as throughput, response time, resource utilization, and reliability. They provide a comprehensive view of how well the system is functioning and can help in assessing its overall effectiveness in meeting desired objectives. By closely monitoring and analyzing performance measures, decision-makers can identify potential bottlenecks, inefficiencies, or areas of improvement within the system. This enables them to make informed decisions and take appropriate actions to optimize system performance, increase productivity, and enhance customer satisfaction.

- Probability the server is idle .  
 $P_I = \sum_{j=1}^K x_{0j0} + \sum_{i=1}^{\infty} x_{i00}$ .
- Probability the server is busy with packing.  
 $P_{BP} = \sum_{i=0}^{\infty} \sum_{j=0}^K x_{ij1}$ .
- Probability the server is busy with rework.  
 $P_{BR} = \sum_{i=0}^{\infty} \sum_{j=0}^K x_{ij2}$ .
- Probability the server is in breakdown while busy with packing.  
 $P_{BDP} = \sum_{i=0}^{\infty} \sum_{j=0}^K x_{ij3}$ .
- Probability the server is in breakdown while busy with rework.  
 $P_{BDR} = \sum_{i=0}^{\infty} \sum_{j=0}^K x_{ij4}$ .
- Probability the server is on vacation.  
 $P_V = \sum_{i=0}^{\infty} \sum_{j=0}^K x_{ij5}$ .
- Expected system size  
 $E_{System} = x_1[(I - R)^{-2}]e_1$ .

## 8. COST ANALYSIS

Let us introduce a cost associated with different system management metrics for our model of interest. We can then formulate a cost function, TC, which takes these metrics into account.

$$TC = CH * E_{system} + P_V * CV + P_I * CI + P_{BP} * CBP + P_{BR} * CBR + P_{BDP} * CBDP + P_{BDR} * CBDR + \mu_1 * C1 + \mu_2 * C2 + \gamma * C3$$

where

- TC-Total cost of the system per unit time.
- CH-Customer holding cost in the system per unit time.
- CV - Cost when the server is on vacation per unit time.
- CI - Cost when the server is idle per unit time.
- CBP - Cost when the server is busy with packing per unit time.
- CBR - Cost when the server is busy with rework per unit time.
- CBDP- Cost when the server faces breakdown while packing per unit time.

- CBDR- Cost when the server faces breakdown while rework per unit time.
- C1 -Cost afforded for packing service by the server per unit time per unit time.
- C2 - Cost afforded for rework service by the server per unit time per unit time.
- C3 - Cost afforded for carrying out the repair process per unit time.

## 9. NUMERICAL ANALYSIS

In this section, we will delve into the qualitative behavior of the model through a series of illustrations that include both numerical and graphical representations. By manipulating various model parameters, such as the arrival process and service time distribution, we aim to gain a deeper understanding of how these parameters affect the model's behavior. Input data for these parameters will be drawn from three sets of values available in the literature, allowing us to examine a wide range of scenarios and explore the model's response to different parameter settings. Through these illustrations, we will shed light on the dynamics and trends exhibited by the model as we vary the model parameters, helping us gain insights into its behavior in different scenarios.

### Erlang of order 2 (ERL-A)

$$D_0 = \begin{bmatrix} -5 & 5 & 0 & 0 & 0 \\ 0 & -5 & 5 & 0 & 0 \\ 0 & 0 & -5 & 5 & 0 \\ 0 & 0 & 0 & -5 & 5 \\ 0 & 0 & 0 & 0 & -5 \end{bmatrix}; D_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \end{bmatrix}; D_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \end{bmatrix}$$

### Exponential (Exp-A)

$$D_0 = [-1]; D_1 = [0.6] D_2 = [0.4]$$

### Hyperexponential (HYP-EXP-A)

$$D_0 = \begin{bmatrix} -1.90 & 0 \\ 0 & -0.19 \end{bmatrix}; D_1 = \begin{bmatrix} 1.026 & 0.114 \\ 0.1026 & 0.0114 \end{bmatrix} D_2 = \begin{bmatrix} 0.684 & 0.076 \\ 0.0684 & 0.0076 \end{bmatrix}$$

Given that Varghese et al. [25] has suggested three phase type distributions for the service process, we will consider these distributions in our analysis. These phase type distributions, which have been proposed by Chakravarthy [7] and documented in the literature, will serve as the basis for our examination of the model's behavior. By incorporating these distributions into our analysis, we aim to gain a deeper understanding of how the model performs under different service time distribution settings and how it responds to varying parameters associated with these distributions. This will enable us to assess the qualitative behavior of the model and uncover any patterns or trends that emerge as we explore these three phase type distributions.

### Erlang of order 2 (ERL-S)

$$\alpha_1 = \alpha_2 = (1, 0); T_1 = T_2 = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}$$

### Exponential (Exp-A)

$$\alpha_1 = (1); T_1 = [-1]$$

$$\alpha_2 = (1); T_2 = [-1]$$

### Hyperexponential (HYP-EXP-A)

$$\alpha_1 = (0.3, 0.7); T_1 = \begin{bmatrix} -9 & 3 \\ 2 & -8 \end{bmatrix}$$

$$\alpha_2 = (0.4, 0.6); T_2 = \begin{bmatrix} -12 & 6 \\ 5 & -10 \end{bmatrix}$$

**Illustration 1:**

In this analysis, we examine the implications of the renege rate ( $\delta_1$ ) of the customers on the expected system size ( $E_{system}$ ) for various combinations of service and arrival times. We consider specific parameter values, including  $\lambda = 2, \mu_1 = 5, \mu_2 = 6, \zeta = 1, \gamma = 3, \delta_2 = 1, \eta = 4, p = 0.3,$  and  $q = 0.7$ . The observations derived from Table 1 to 3 are outlined below.

- As the renege rate increases, more customers choose to leave the system without completing their service requests. This results in a lower number of customers in the system at any given time, leading to an decrease in the expected system size.
- When customers renege at a higher rate, the system experiences a shorter average waiting time and lower congestion due to customers leaving before being served. This decrease congestion leads to only few customers remaining in the system, resulting in a lower expected system size.

**Illustration 2:**

In this analysis, we examine the effects of the vacation rate ( $\zeta$ ) and service rate ( $\mu_1$ ) of the server on the expected system size ( $E_{system}$ ). We consider various combinations of service and arrival times and use specific parameter values, including  $\lambda = 2, \mu_1 = 6, \gamma = 3, \delta_1 = 1, \delta_2 = 1, \eta = 4, p = 0.3,$  and  $q = 0.7$ . The observations derived from Figure 29-37 are outlined below.

- When both the vacation rate ( $\zeta$ ) and service rate ( $\mu_1$ ) increase, it generally leads to a decrease in the expected system size. This means that, on average, there will be fewer customers present in the system at any given time.
- An increase in the vacation rate ( $\zeta$ ) implies that the availability of the server increases. Similarly, an increase in the service rate ( $\mu_1$ ) means that the server can process customer requests at a faster pace. When both the vacation rate and service rate increase, the server has a reduced overall availability for serving customers due to more frequent breaks.
- These observations highlight the varying impacts of vacation rate and service rate on the projected system size across different arrival and service times. Erlang arrivals show the most significant reduction in system size, followed by exponential arrivals, while hyper exponential arrivals display a slower rate of decrease.

**Table 1:** Renege rate ( $\delta_1$ ) vs Expected System Size - ERL-A

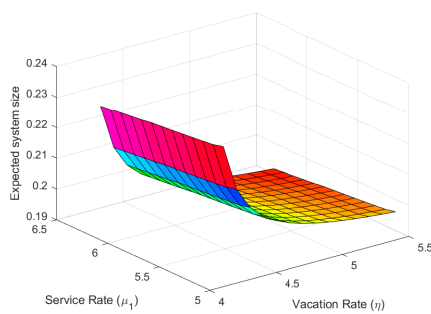
$\delta_1$	service		
	Erlang	Exponential	Hyperexponential
1.0	2.689867713	2.764059228	2.776027059
1.1	2.680737416	2.74942792	2.765549157
1.2	2.672576038	2.734797612	2.755075026
1.3	2.665225627	2.720166304	2.744593354
1.4	2.658647922	2.705534995	2.724115453
1.5	2.647557358	2.690903687	2.703637552
1.6	2.640837634	2.676272379	2.683159655
1.7	2.625331925	2.661641071	2.672681749
1.8	2.611886634	2.647009763	2.662203847
1.9	2.591670138	2.632378454	2.651725946

**Table 2: Renege rate ( $\delta_1$ ) vs Expected System Size - EXP-A**

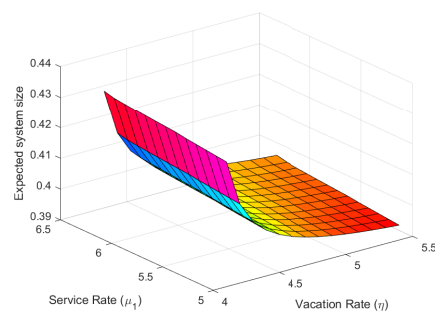
$\delta_1$	service		
	Erlang	Exponential	Hyperexponential
1.0	2.732538773	2.776538896	2.815784456
1.1	2.721217834	2.766061989	2.794520162
1.2	2.710069838	2.755584087	2.777635494
1.3	2.709039454	2.745106186	2.763847123
1.4	2.698116376	2.734628284	2.752344555
1.5	2.689888555	2.724150383	2.742586259
1.6	2.681785353	2.713672482	2.734194049
1.7	2.673681963	2.703194585	2.726894449
1.8	2.665578647	2.699716679	2.720483688
1.9	2.657475344	2.682238777	2.714806803

**Table 3: Renege rate ( $\delta_1$ ) vs Expected System Size - HYP-EXP-A**

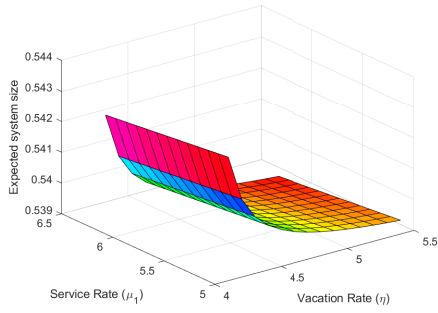
$\delta_1$	service		
	Erlang	Exponential	Hyperexponential
1.0	2.857181106	2.902141938	3.011784445
1.1	2.832316046	2.882431771	2.969464924
1.2	2.813250935	2.866914417	2.945390736
1.3	2.798124579	2.854332628	2.926737224
1.4	2.785797988	2.843905244	2.911866499
1.5	2.775536658	2.835114586	2.899730666
1.6	2.766845678	2.827598651	2.889632399
1.7	2.759379139	2.821103707	2.881091624
1.8	2.752887573	2.815428859	2.873767852
1.9	2.747186259	2.797494693	2.867413281



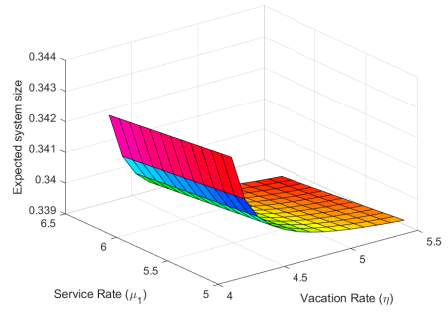
**Figure 2: Vacation rate ( $\eta$ ), Service rate ( $\mu_1$ ) vs Expected system size - Ek/Ek/1**



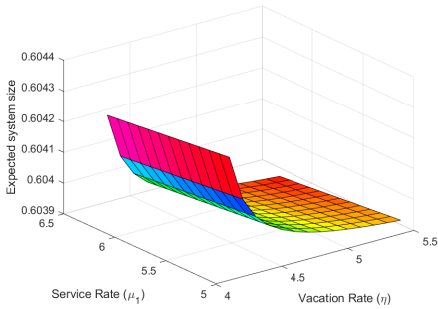
**Figure 3: Vacation rate ( $\eta$ ), Service rate ( $\mu_1$ ) vs Expected system size - Ek/M/1**



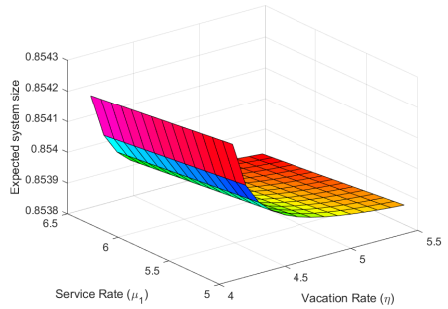
**Figure 4:** Vacation rate ( $\eta$ ), Service rate ( $\mu_1$ ) vs Expected system size -  $Ek/Hk/1$



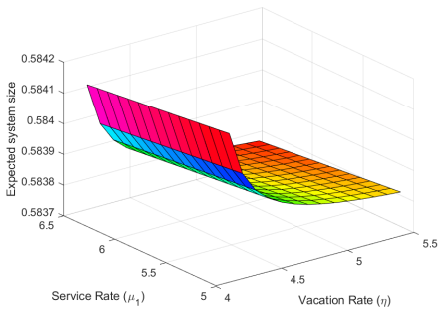
**Figure 5:** Vacation rate ( $\eta$ ), Service rate ( $\mu_1$ ) vs Expected system size -  $M/Ek/1$



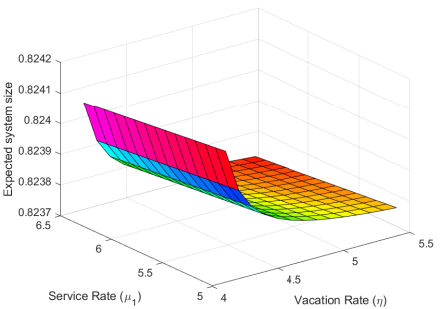
**Figure 6:** Vacation rate ( $\eta$ ), Service rate ( $\mu_1$ ) vs Expected system size -  $M/M/1$



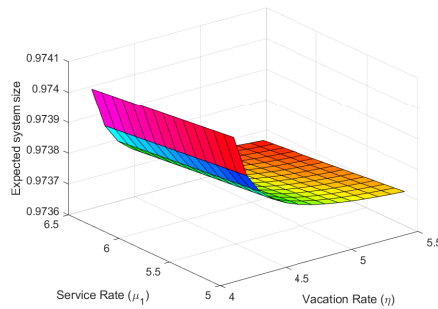
**Figure 7:** Vacation rate ( $\eta$ ), Service rate ( $\mu_1$ ) vs Expected system size -  $M/Hk/1$



**Figure 8:** Vacation rate ( $\eta$ ), Service rate ( $\mu_1$ ) vs Expected system size -  $Hk/Ek/1$



**Figure 9:** Vacation rate ( $\eta$ ), Service rate ( $\mu_1$ ) vs Expected system size -  $Hk/M/1$



**Figure 10:** Vacation rate ( $\eta$ ), Service rate ( $\mu_1$ ) vs Expected system size -  $Hk/Hk/1$

## 10. CONCLUSION

In conclusion, our model encompasses a complex system involving a single server managing two queues, each susceptible to various uncertainties. We have meticulously detailed the arrival and service processes, highlighting the critical phases and distribution patterns that govern them. The model accounts for the inherent unpredictabilities, such as server breakdowns, repairs, customer renegeing, and vacation periods. By examining both infinite and finite capacity arrivals, we have provided a comprehensive framework for analyzing the performance and reliability of this intricate system. This model can serve as a valuable tool for optimizing operations, enhancing service quality, and minimizing disruptions in scenarios where such intricate dynamics are at play.

Broadening the system's scope to accommodate intricate service time patterns mirroring real-world complexities holds the potential for a more profound comprehension of service dynamics. Upcoming research endeavors will center on refining scheduling strategies and computational methods for handling batch arrivals, server disruptions, repair processes, bulk services, and the involvement of multiple service providers. These initiatives seek to minimize customer waiting intervals, optimize resource distribution, and elevate overall system effectiveness, with the ultimate goal of enhancing the applicability of such systems across diverse domains.

## REFERENCES

- [1] Arivudainambi, K. and Arivudainambi, V. (2018). Performance analysis of a bulk service queue with multiple vacations, server breakdowns, and general service times, *Journal of Industrial Engineering and Management Science*, 1(1):49–60.
- [2] Arun, P., Arivudainambi, K., and Arivudainambi, V. (2017). Analysis of a bulk service queue with server breakdowns, balking, and renegeing, *Journal of Industrial Engineering International*, 13(4):523–532.
- [3] Anilkumar, M. P. and Jose, K. P. (2022). Analysis is a discrete time queueing-inventory model with back-order of items, *3C Empresa. Investigación y pensamiento crítico.*, 11(2):50–62.
- [4] Anis, B., Arbia, Y. and Hamdi, M. (2021). Analysis of a finite-buffer queue with server vacations and impatience, *International Journal of Systems Science: Operations and Logistics*, 8(1):111–126.
- [5] Ayyappan, G. and Meena, S. (2023). Phase type queueing model of server vacation, repair, and degrading service with breakdown, starting failure and closedown, *Reliability: Theory and Applications*, 18(2):464–483
- [6] Ayyappan, G. and Nithya, S. (2023). Analysis of  $M^{[X_1]}, M^{[X_1]}/G_1, G_2/1$  retail queue with priority service, differentiate breakdown, repair, synchronized renegeing and optional vacation, *Reliability: Theory and Applications*, 18(2):376–391
- [7] Chakravarthy, SR. (2010). Markovian Arrival Processes, *Wiley Encyclopedia of Operations Research and Management Science*, DOI.org/10.1002/9780470400531.eorms0499
- [8] Huang, H. and Li, W. (2020). Optimal control of vacation queues with multiple vacations and general service times, *Applied Mathematics and Computation*, DOI: 10.1016/j.amc.2020.125220
- [9] Kalyanaraman, R. and Janani, G. (2023). Finite Population and Finite Capacity Single Server Batch Service Queues With Compulsory Working Vacation, *International Journal of Mathematics in Operational Research*, DOI: 10.1504/IJMOR.2023.10057466 .
- [10] Kim, D., Kim, J. and Whitt, W. (2021). Vacation models with abandonments, *Management Science*, 67(4):2361–2385.
- [11] Krishnamoorthy, A., Joshua, A.N. and Vishnevsky, V. (2021). Analysis of a k-Stage Bulk Service Queueing System with Accessible Batches for Service, *Mathematics* , DOI.org/10.3390/math9050559
- [12] Latouche, G. and Ramaswami, V. (1999). Introduction to Matrix Analytic Methods in Stochastic Modeling, *American Statistical Association, Virginia and the Society for Industrial and Applied Mathematics, Pennsylvania*, DOI.org/10.1137/1.9780898719734.

- [13] Li, H. and Li, X. (2021). Optimization of Bulk Service Systems with Parallel Servers, *Computers and Industrial Engineering*, DOI.org/10.1016/j.cie.2020.107207
- [14] Li, Y., and Zhang, G. (2017). Optimal control of a bulk service queue with impatient customers and time-varying arrival rates, *Journal of Systems Science and Complexity*, 30(5):1165–1180.
- [15] Neuts, M. F. (1979). A versatile Markovian point process, *Journal of Applied Probability*, 16(4):764–779
- [16] Neuts, M.F. (1981). Matrix-Geometric Solutions in Stochastic Models An Algorithmic Approach, *Johns Hopkins University Press, Baltimore and London*.
- [17] Raina Raj and Selvamuthu Dharmaraja (2023). Stochastic Modelling of Multi-Layer Hap-Leo Systems in 6g for Energy Saving: An Analytical Approach, *Computer Communications*, DOI:10.1016/j.comcom.2023.07.037
- [18] Rakesh Kumar et al. (2021). Performance Analysis of a Cloud Computing System using Queuing Model with Correlated Task Reneging, *Journal of Physics Conference Series*, 2091(1):1–6
- [19] Rakesh Kumar, Sapana Sharma and Sherif I. Ammar (2019). Transient and steady-state analysis of a queuing system having customers' impatience with threshold, *RAIRO - Operations Research*, 53(5):1861–1876
- [20] Saroja, V. and Saravananarajan, S. (2018). Bulk service queueing models with server vacations and feedback controls, *Journal of Industrial Engineering and Management Science*, 1(2):89–100.
- [21] Srinivasan, A. and Sriram, S. S. (2020). Analysis of vacation queues with server subject to breakdowns and repair, *International Journal of Applied Mathematics and Computer Science*, 30(4):637–652.
- [22] Smith, E. and Johnson, J. (2022). An Analysis of the Impact of Bulk Service Providers on Supply Chain Performance, *Journal of Business Logistics*, 43(1):1–16.
- [23] Sun, Y. and Zhang, J. (2022). Designing a Bulk Service System for Autonomous Mobile Robots, *IEEE Transactions on Robotics*, 38(1):247–258.
- [24] Thottan, M. and DeVeciana, G. (2021). A vacation model with autonomous server vacations and customer impatience, *INFORMS Journal on Computing*, 33(3):982–999.
- [25] Varghese, J., Chakravarthy, SR. and Krishnamoorthy, A. (2012). On a customer induced interruption in a service system, *J Stoch Anal Appl*, 30(6):949–962.
- [26] Yadav, S. et al. (2020). Optimization of Kitting Process in an Automotive Assembly Line, *Production and Manufacturing Research*, 8(1):551–570.
- [27] Wang, Y. et al. (2021). A Hybrid Approach for Optimizing Bulk Service Systems in E-commerce Warehouses, *Computers and Industrial Engineering*, DOI.org/10.1016/j.cie.2021.107272
- [28] Wang, L. et al. (2021). A Framework for Kitting Process Optimization in Manufacturing, *Journal of Manufacturing Systems*, 60:487–501.
- [29] Zhang, X. and Fang, Y. (2022). Efficient Optimization Techniques for Bulk Service Systems, *IEEE Transactions on Automation Science and Engineering*, 19(1):265–277.

# ON SOME PROPERTIES AND APPLICATIONS OF THE TYPE II HALF -LOGISTIC EXPONENTIATED FRECHET DISTRIBUTION

Olalekan Akanji Bello <sup>1\*</sup>

Sani Ibrahim Doguwa<sup>1</sup>

Abukakar Yahaya<sup>1</sup>

Haruna Mohammed Jibril<sup>2</sup>

Department of Statistics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria, Nigeria<sup>1</sup>.

Department of Mathematics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria,  
Nigeria<sup>2</sup>.

[olalekan4sure@gmail.com](mailto:olalekan4sure@gmail.com)<sup>1\*</sup>

[sidoguwa@gmail.com](mailto:sidoguwa@gmail.com)<sup>1</sup>

[ensiliyu2@yahoo.co.uk](mailto:ensiliyu2@yahoo.co.uk)<sup>1</sup>

[alharun2004@yahoo.com](mailto:alharun2004@yahoo.com)<sup>2</sup>

## Abstract

*As the dimensions of available data for analysis continues to grow rapidly, it becomes imperative to develop new probability distributions that can more accurately represent various phenomena. In this research paper, we introduce a novel continuous probability distribution known as the Type II Half-Logistic Exponentiated Frechet Distribution, characterized by four positive parameters. This distribution expands upon the traditional Frechet distribution by introducing two additional parameters. We derive a significant density representation for this distribution. Furthermore, we delve into several statistical and mathematical properties associated with the Type II Half-Logistic Exponentiated Frechet distribution. This includes explicit expressions for key metrics such as the quantile function, probability weighted moments, moments, moments generating function, reliability function, hazard function, and order statistics. To estimate the model parameters effectively, we employ a maximum likelihood estimation technique and present the results of a simulation study. Our research underscores the superiority of this new distribution by applying it to two real-world datasets. Notably, the findings demonstrate that the Type II Half-Logistic Exponentiated Frechet distribution outperforms other considered distributions in fitting the two real datasets.*

**Keywords:** Type II Half-Logistic Exponentiated-G, Frechet distribution, Moments function, Reliability function, Maximum likelihood, Order Statistics.

## 1. Introduction

Many types of univariate continuous distributions exist, but research in various fields, including engineering, environmental science, finance, and medicine, has shown that real-world data often does not follow the classical distributions. To address this issue, extended forms of these distributions have been developed to provide more flexibility in data modeling. The Frechet



distribution, also called the type II extreme value distribution, plays a vital role in extreme value theory and has numerous applications. There have been several modifications and enhancements to the Frechet distribution proposed in the statistical literature to further improve its usefulness. In recent times, various extensions of the Frechet distribution have been introduced by several researchers in the academic literature. Nadarajah and Kotz [16] were the pioneers of the exponentiated Frechet distribution, while Nadarajah and Gupta [17] introduced the beta Frechet distribution. Mahmoud and Mandouh [12] put forth the transmuted Frechet distribution, Da Silva *et al.*, [7] defined the gamma extended Frechet distribution, Krishna *et al.*, [10] introduced the Marshall-Olkin Frechet distribution, and Mead and Abd-Eltawab [13] introduced the Kumaraswamy Frechet distribution. Elbatal *et al.*, [8] conducted a study on the transmuted exponentiated Frechet distribution, Afify *et al.*, [1] investigated the transmuted Marshall-Olkin Frechet distribution, Afify *et al.*, [3] proposed the Kumaraswamy Marshall-Olkin Frechet distribution, Afify *et al.*, [2] explored the Weibull Frechet distribution, Tablada and Cordeiro [20] defined the modified Frechet distribution, and Mead *et al.*, [15] introduced the beta exponential Frechet distribution.

In a recent study, Bello *et al.*, [4] proposed a new distribution family called the Type II Half-Logistic Exponentiated-G (TIIHLEt-G). This distribution family is defined by two positive shape parameters, denoted by  $\lambda$  and  $\alpha$ , and can be applied to any arbitrary cumulative distribution function (cdf)  $H(x, \mathcal{G})$ . The cumulative distribution function (cdf) and the probability density function for TIIHLEt-G are detailed as follows:

$$F_{TIIHLEt-G}(x; \lambda, \alpha, \boldsymbol{\beta}) = \frac{2H^{\alpha\lambda}(x; \boldsymbol{\beta})}{[1 + H^{\alpha\lambda}(x; \boldsymbol{\beta})]}, \quad x > 0, \lambda, \alpha > 0 \quad (1)$$

and

$$f_{TIIHLEt-G}(x; \lambda, \alpha, \boldsymbol{\beta}) = \frac{2\lambda\alpha h(x; \boldsymbol{\beta})H^{\alpha-1}(x; \boldsymbol{\beta})[H^{\alpha(\lambda-1)}(x; \boldsymbol{\beta})]}{[1 + H^{\alpha\lambda}(x; \boldsymbol{\beta})]^2}, \quad x > 0, \lambda, \alpha > 0 \quad (2)$$

The cdf and pdf of the Frechet distribution are given as

$$H(x; \theta, \delta) = e^{-\left(\frac{\theta}{x}\right)^\delta}, \quad x > 0, \theta, \delta > 0, \quad (3)$$

$$h(x; \theta, \delta) = \delta\theta^\delta x^{-\delta-1} e^{-\left(\frac{\theta}{x}\right)^\delta}, \quad x > 0, \theta, \delta > 0 \quad (4)$$

The most important goal of this paper is to enhance the flexibility of a statistical model by extending the conventional two-parameter Frechet distribution. This novel model is referred to as the Type II Half Logistic Exponentiated Frechet (TIIHLEtF) distribution. The structure of this paper is organised as follows: In Section 2, we introduce and define the TIIHLEtF distribution. Section 3 presents valuable representations for the TIIHLEtF distribution. Section 4 focuses on deriving statistical properties such as probability-weighted moments, ordinary moments, moments-generating function, quartile function, reliability function, hazard function, and order statistics. In Section 5, we estimate the parameters of the new model using the maximum likelihood estimation (MLE) approach. To demonstrate the efficiency and consistency of MLE, we conducted a simulation study in Section 6. In Section 7, we apply the new model to two real datasets to illustrate its practical utility. Finally, Section 8 provides a conclusion for the paper.

## 2. Type II Half-Logistic Exponentiated Frechet (TIIHLEtF) Distribution

In this section, we introduce a novel model referred to as the TIIHLEtF distribution. A random variable  $X$  is considered to follow the TIIHLEtF distribution if its cumulative distribution function

(cdf) is derived by substituting equation (3) into equation (1) in the following approach:

$$F_{TIIHLEtF}(x; \lambda, \alpha, \theta, \delta) = \frac{2e^{-\alpha\lambda\left(\frac{\theta}{x}\right)^\delta}}{1 + e^{-\alpha\lambda\left(\frac{\theta}{x}\right)^\delta}}, x > 0, \lambda, \alpha, \theta, \delta > 0 \quad (5)$$

and its corresponding pdf is

$$f_{TIIHLEtF}(x; \lambda, \alpha, \theta, \delta) = \frac{2\lambda\alpha\delta\theta^\delta x^{-\delta-1} e^{-\left(\frac{\theta}{x}\right)^\delta} e^{-(\alpha-1)\left(\frac{\theta}{x}\right)^\delta} e^{-\alpha(\lambda-1)\left(\frac{\theta}{x}\right)^\delta}}{\left[1 + e^{-\alpha\lambda\left(\frac{\theta}{x}\right)^\delta}\right]^2}, x > 0, \lambda, \alpha, \theta, \delta > 0 \quad (6)$$

where  $\theta$  is a scale parameter and  $\lambda, \alpha, \delta$  are shape parameters.

### 3. Expansion of Density

In this section, we have obtained a valuable expression for the probability density function (pdf) and cumulative distribution function (cdf) of the TIIHLEtF distribution. This achievement is attributed to our utilization of the generalized binomial series given as:

$$(1+Z)^{-\beta} = \sum_{i=0}^{\infty} (-1)^i \binom{\beta+i-1}{i} z^i \quad (7)$$

For  $|z| < 1$  and  $\beta$  is a positive real non integer. The density function of the TIIHLEtF distribution is derived by applying the binomial theorem from equation (7) to equation (6).

$$f_{TIIHLEtF}(x; \lambda, \alpha, \theta, \delta) = 2\lambda\alpha\delta\theta^\delta x^{-\delta-1} e^{-\left(\frac{\theta}{x}\right)^\delta} e^{-(\alpha-1)\left(\frac{\theta}{x}\right)^\delta} e^{-\alpha(\lambda-1)\left(\frac{\theta}{x}\right)^\delta} \left[1 + e^{-\alpha\lambda\left(\frac{\theta}{x}\right)^\delta}\right]^{-2}$$

Now, using the generalized binomial theorem, we can write

$$\left[1 + e^{-\alpha\lambda\left(\frac{\theta}{x}\right)^\delta}\right]^{-2} = \sum_{i=0}^{\infty} (-1)^i \binom{1+i}{i} \left[e^{-\left(\frac{\theta}{x}\right)^\delta}\right]^{\alpha\lambda i}$$

Then, the pdf can be written as:

$$f_{TIIHLEtF}(x; \lambda, \alpha, \theta, \delta) = \sum_{i=0}^{\infty} \eta_p \left[e^{-\left(\frac{\theta}{x}\right)^\delta}\right]^{\alpha\lambda(i+1)} \quad (8)$$

$$\text{where } \eta_p = 2\lambda\alpha\delta\theta^\delta x^{-\delta-1} (-1)^i \binom{1+i}{i}$$

In addition, an expansion for the  $[F(x; \lambda, \alpha, \theta, \delta)]^h$  is produced, with  $h$  being an integer, and the binomial expansion is worked out once more.

$$\begin{aligned} [F(x; \lambda, \alpha, \theta, \delta)]^h &= 2^h \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^h \left[ 1 + \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha\lambda} \right]^{-h} \\ &= \sum_{j=0}^h (-1)^j \binom{h+j-1}{j} \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha\lambda j} \end{aligned}$$

The cdf can be written as:

$$[F(x; \lambda, \alpha, \theta, \delta)]^h = \sum_{j=0}^h \varphi_j \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha\lambda(j+h)} \quad (9)$$

where  $\varphi_j = 2^h (-1)^j \binom{h+j-1}{j}$

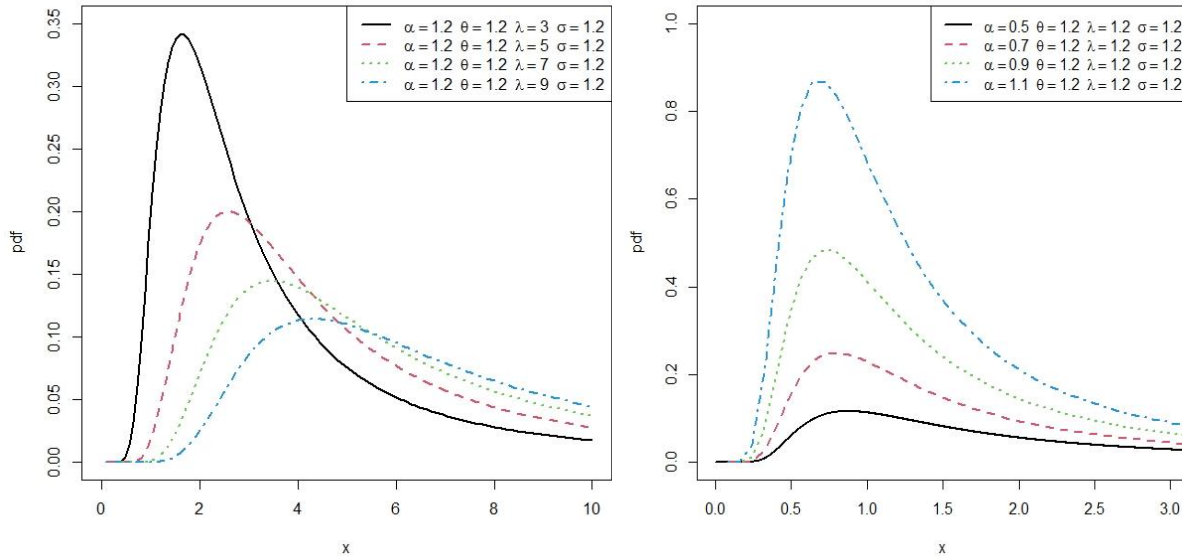


Figure 1: Plots of Pdf of TIIHLEtF distribution for different values of parameters.

#### 4. Statistical Properties

In this section, we derived some statistical properties of the new of distribution.

##### 4.1. Probability weighted moments

Greenwood et al. [10] introduced a concept known as probability weighted moments (PWMs). This technique is employed to create estimators in the inverse form for both distribution parameters and quantiles. The notations used for probability weighted moments is  $\tau_{r,s}$ , and these moments can be computed for a random variable X by utilizing the relationship outlined below.

$$\tau_{r,s} = E[X^r F(X)^s] = \int_{-\infty}^{\infty} x^r f(x) (F(x))^s dx \quad (10)$$

The PWMs for the TIIHLEtF distribution are obtained by inserting equations (8) and (9) into (10), and then replacing h with s in the following manner.

$$\tau_{r,s} = \sum_{i=0}^{\infty} \sum_{j=0}^h \eta_p \varphi_j \int_0^{\infty} x^r \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha\lambda(i+1+j+s)} dx \quad (11)$$

Consider the integral

$$\int_0^{\infty} x^r \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha\lambda(i+1+j+s)} dx$$

Let  $y = \alpha\lambda(i+1+j+s)\left(\frac{\theta}{x}\right)^\delta \Rightarrow x = \left[ \frac{\alpha\lambda(i+1+j+s)\theta^\delta}{y} \right]^{\frac{1}{\delta}}; dx = \frac{dyx^{\delta-1}}{\delta\theta^\delta\alpha\lambda(i+1+j+s)}$

Then

$$\int_0^{\infty} \left[ \frac{\alpha\lambda(i+1+j+s)\theta^\delta}{y} \right]^{\frac{r}{\delta}} e^{-y} \frac{dyx^{\delta-1}}{\delta\theta^\delta\alpha\lambda(i+1+j+s)}$$

$$\int_0^{\infty} y^{-\frac{r}{\delta}} e^{-y} dy = \Gamma\left(1 - \frac{r}{\delta}\right)$$

Hence, the PWMs of TIIHLEtF can be expressed in the following manner.

$$\tau_{r,s} = \sum_{i=0}^{\infty} \sum_{j=0}^s (\alpha\lambda)^\frac{r}{\delta} \theta^r (i+1+j+s)^\frac{r}{\delta-1} \eta_p \varphi_i \Gamma\left(1 - \frac{r}{\delta}\right) \tag{12}$$

Now,

$$\varphi_i = 2^s (-1)^j \binom{s+j-1}{j}$$

and

$$\eta_p = 2(-1)^i \binom{1+i}{i}$$

#### 4.2. Moments

As moments play a crucial role in statistical analysis, particularly in practical applications, we proceed to derived the  $r^{\text{th}}$  moment for the newly introduced distribution.

$$\mu'_r = E(x^r) = \int_0^{\infty} x^r f(x) dx \tag{13}$$

By using the expansion of the pdf in equation (8), we have

$$E(X^r) = \sum_{i=0}^{\infty} \eta_p \int_0^{\infty} x^r \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha\lambda(i+1)} dx \tag{14}$$

Consider the integral

$$\int_0^{\infty} x^r \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha\lambda(i+1)} dx$$

Let  $w = \alpha\lambda(i+1)\left(\frac{\theta}{x}\right)^\delta \Rightarrow x = \left[\frac{\alpha\lambda(i+1)\theta^\delta}{w}\right]^{\frac{1}{\delta}}; dx = \frac{dwx^{\delta-1}}{\alpha\lambda(i+1)\theta^\delta \delta}$

Then

$$\int_0^\infty \left[\frac{\alpha\lambda(i+1)\theta^\delta}{w}\right]^{\frac{r}{\delta}} e^{-w} \frac{dwx^{\delta-1}}{\alpha\lambda(i+1)\theta^\delta}$$

$$\int_0^\infty w^{-\frac{r}{\delta}} e^{-w} dw = \Gamma\left(1 - \frac{r}{\delta}\right)$$

The  $r^{\text{th}}$  moment for TIIHLEtF distribution can be written as follows

$$E(X^r) = \sum_{i=0}^\infty \eta_p \theta^r (\alpha\lambda)^\frac{r}{\delta} (i+1)^\frac{r}{\delta-1} \Gamma\left(1 - \frac{r}{\delta}\right) \tag{15}$$

Now

$$\eta_p = 2(-1)^i \binom{1+i}{i}$$

The mean and variance of TIIHLEtF distribution are as follows

$$E(X) = \sum_{i=0}^\infty \eta_p \theta (\alpha\lambda)^\frac{1}{\delta} (i+1)^\frac{1}{\delta-1} \Gamma\left(1 - \frac{1}{\delta}\right) \tag{16}$$

and

$$\text{var}(X) = \sum_{i=0}^\infty \eta_p \theta (\alpha\lambda)^\frac{1}{\delta} (i+1)^\frac{1}{\delta-1} \Gamma\left(1 - \frac{1}{\delta}\right) - \left[ \sum_{i=0}^\infty \eta_p \theta (\alpha\lambda)^\frac{1}{\delta} (i+1)^\frac{1}{\delta-1} \Gamma\left(1 - \frac{1}{\delta}\right) \right]^2 \tag{17}$$

### 4.3. Moment generating function (mgf)

The Moment Generating Function of  $x$  is given as:

$$M_x(t) = E(e^{tx}) = \int_0^\infty e^{tx} f(x) dx \tag{18}$$

where the expansion of  $e^{tx} = \sum_{m=0}^\infty \frac{t^m x^m}{m!}$

The moment generating function of TIIHLEtF distribution is given by

$$M_x(t) = \sum_{i=0}^\infty \sum_{m=0}^\infty \frac{t^m \eta_p \theta^m (\alpha\lambda)^\frac{m}{\delta} (i+1)^\frac{m}{\delta-1} \Gamma\left(1 - \frac{m}{\delta}\right)}{m!} \tag{19}$$

### 4.4. Reliability function

The reliability function, also referred to as the survivor function, provides the probability that an individual or patient will endure beyond certain specified duration of time. In other words, it gives the likelihood of survival beyond a particular time point. It's defined as

$$R(x; \lambda, \alpha, \theta, \delta) = \frac{1 - \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha\lambda}}{1 + \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha\lambda}} \quad (20)$$

#### 4.5. Hazard function

The hazard function represents the likelihood of an event of interest happening within a relatively brief time interval is defined as follow:

$$T(x; \lambda, \alpha, \theta, \beta) = \frac{2\lambda\alpha\delta\theta^\delta x^{-\delta-1} e^{-\left(\frac{\theta}{x}\right)^\delta} \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha-1} \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha(\lambda-1)}}{1 - \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{2\alpha\lambda}} \quad (21)$$

#### 4.6. Quantile Function

The quantile function plays a crucial role in generating random variables from continuous probability distributions, making it a key element in probability theory. Specifically, for a given value 'x,' the quantile function is denoted as  $F(x) = u$ , where 'u' follows a uniform distribution between 0 and 1 ( $U(0,1)$ ). To simulate the TIIHLEtF distribution, one can readily achieve this by reversing equation (5), resulting in the definition of the quantile function  $Q(u)$ .

$$x = Q(u) = \frac{\theta}{\left[ -\log \left[ \frac{U}{2-U} \right]^{\frac{1}{\alpha\lambda}} \right]^\delta} \quad (22)$$

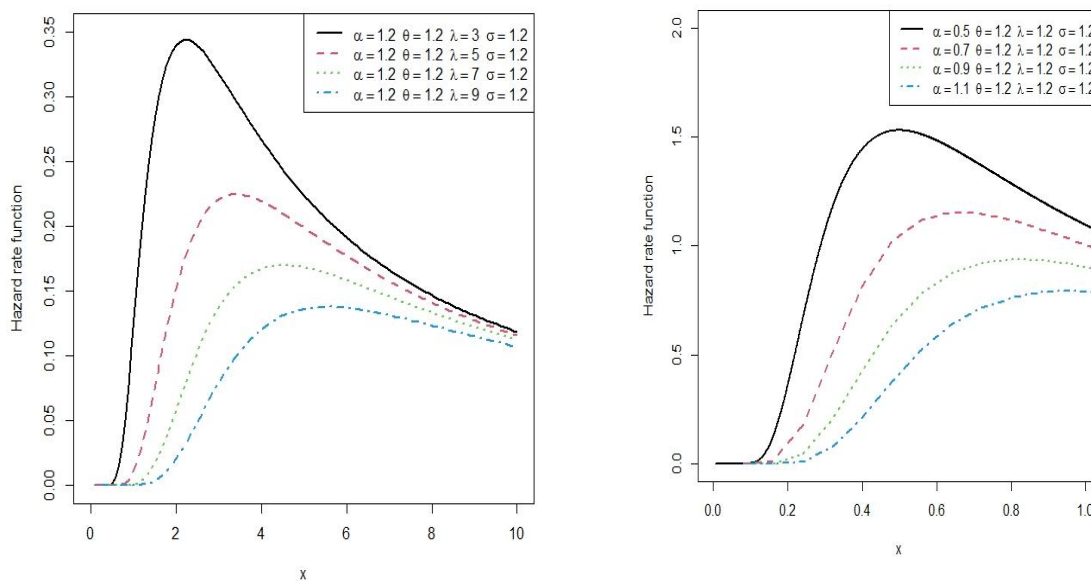


Figure 2: Plots of hazard of the TIIHLEtF distribution for different valves of parameters.

#### 4.7. Order Statistics

Order statistics are widely applied in various statistical fields, including reliability and life testing. Consider a set of  $n$  independent and identically distributed random variables represented as  $X_1, X_2, \dots, X_n$ , each following a continuous distribution function  $F(x)$ . If these random variables are drawn from the TIIHLEtF distribution, we can denote the cumulative distribution function (cdf) as  $F_{r:n}(x)$  and the probability density function (pdf) as  $f_{r:n}(x)$  for the  $r^{\text{th}}$  order statistic, where  $r$  ranges from 1 to  $n$ . In a study by David [1979], the probability density function of  $X_{r:n}$  was provided.

$$f_{r:n}(x) = \frac{f(x)}{B(r, n-r+1)} \sum_{v=0}^{n-r} (-1)^v \binom{n-r}{v} F(x)^{v+r-1} \quad (23)$$

By substituting equation (8) and equation (9) into equation (23), also replacing  $h$  with  $v+r-1$  in equation (9). We have

$$f_{r:n}(x; \lambda, \alpha, \theta, \delta) = 2^{(r+v)} \lambda \alpha \delta \theta^\delta x^{-\delta-1} \frac{1}{B(r, n-r+1)} \sum_{v=0}^{n-r} \sum_{i=0}^{\infty} \sum_{j=0}^{r+v-1} (-1)^{i+j+v} \binom{n-r}{v} \binom{1+i}{i} \binom{r+v+j-2}{j} \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha \lambda (i+j+r+v)} \quad (24)$$

The equation above is called the  $r^{\text{th}}$  order statistics for the TIIHLEtF distribution.

Let  $r = n$ , then the probability density function of the maximum order statistics of TIIHLEtF distribution is

$$f_{n:n}(x; \lambda, \alpha, \theta, \delta) = 2^{(n+v)} n \lambda \alpha \delta \theta^\delta x^{-\delta-1} \sum_{i=0}^{\infty} \sum_{j=0}^{n+v-1} (-1)^{i+j+v} \binom{1+i}{i} \binom{n+v+j-2}{j} \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha \lambda (i+j+n+v)} \quad (25)$$

Also, let  $r = 1$ , then the probability density function of the minimum order statistics of TIIHLEtF distribution is

$$f_{1:n}(x; \lambda, \alpha, \theta, \delta) = 2^{(v+1)} n \lambda \alpha \delta \theta^\delta x^{-\delta-1} \sum_{v=0}^{n-1} \sum_{i=0}^{\infty} \sum_{j=0}^v (-1)^{i+j+v} \binom{n-1}{v} \binom{1+i}{i} \binom{v+j-1}{j} \left[ e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha \lambda (i+j+1+v)} \quad (26)$$

### 5. Parameter Estimation

In this research paper, we investigate the application of the maximum likelihood technique to estimate the unknown parameters of the TIIHLEtF distribution when dealing with complete data. Maximum likelihood estimates (MLEs) possess advantageous characteristics that can be utilized to establish confidence intervals and provide straightforward approximations that perform well with finite data samples. In the realm of distribution theory, these approximations for MLEs can be conveniently managed, either through analytical or numerical methods. Consider a random sample of size  $n$ , denoted as  $x_1, x_2, x_3, \dots, x_n$ , drawn from the TIIHLEtF distribution. Then, the likelihood function, based on the observed sample, for the parameter vector  $(\lambda, \alpha, \theta, \delta)^T$  is defined as follows.

$$\log L = n \log(2) + n \log(\lambda) + n \log(\alpha) + n \log(\delta) + n\delta \log(\theta) - (\delta + 1) \sum_{i=1}^n \log(x_i) - \alpha \sum_{i=1}^n \left(\frac{\theta}{x_i}\right)^\delta - \alpha(\lambda - 1) \sum_{i=1}^n \left(\frac{\theta}{x_i}\right)^\delta - 2 \sum_{i=1}^n \log \left[ 1 + \left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha\lambda} \right] \quad (27)$$

The components of score vector  $\Delta L(\phi) = \left( \frac{\partial L(\phi)}{\partial \lambda}, \frac{\partial L(\phi)}{\partial \alpha}, \frac{\partial L(\phi)}{\partial \theta}, \frac{\partial L(\phi)}{\partial \delta} \right)^T$  are given as

$$\frac{\partial \log L}{\partial \lambda} = \frac{n}{\lambda} - \alpha \sum_{i=1}^n \left(\frac{\theta}{x_i}\right)^\delta - 2 \sum_{i=1}^n \frac{\left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha\lambda} \log \left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha}{1 + \left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha\lambda}} = 0 \quad (28)$$

$$\frac{\partial \log L}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^n \left(\frac{\theta}{x_i}\right)^\delta - (\lambda - 1) \sum_{i=1}^n \left(\frac{\theta}{x_i}\right)^\delta - 2 \sum_{i=1}^n \frac{\lambda \left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha(\lambda-1)} \left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \log \left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]}{1 + \left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha\lambda}} = 0 \quad (29)$$

$$\frac{\partial \log L}{\partial \delta} = \frac{n}{\delta} + n \log(\theta) - \sum_{i=1}^n x_i - \alpha \sum_{i=1}^n \left(\frac{\theta}{x_i}\right)^\delta \log\left(\frac{\theta}{x_i}\right) - \alpha(\lambda - 1) \sum_{i=1}^n \left(\frac{\theta}{x_i}\right)^\delta \log\left(\frac{\theta}{x_i}\right) - 2 \sum_{i=1}^n \frac{\lambda \left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha(\lambda-1)} \alpha \left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha-1} e^{-\left(\frac{\theta}{x_i}\right)^\delta} \left(\frac{\theta}{x_i}\right)^\delta \log\left(\frac{\theta}{x_i}\right)}{1 + \left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha\lambda}} = 0 \quad (30)$$

$$\frac{\partial \log L}{\partial \theta} = \frac{n\delta}{\theta} + \alpha\delta \sum_{i=1}^n \left(\frac{\theta}{x_i}\right)^{\delta-1} \frac{1}{x_i} - \delta\alpha(\lambda - 1) \sum_{i=1}^n \left(\frac{\theta}{x_i}\right)^{\delta-1} \frac{1}{x_i} - 2 \sum_{i=1}^n \frac{\lambda \left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha(\lambda-1)} \alpha \left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha-1} e^{-\left(\frac{\theta}{x_i}\right)^\delta} \left(\frac{\theta}{x_i}\right)^{\delta-1} \frac{1}{x_i}}{1 + \left[ e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha\lambda}} = 0 \quad (31)$$

The MLEs are obtained by setting  $\frac{\partial L(\phi)}{\partial \lambda}, \frac{\partial L(\phi)}{\partial \alpha}, \frac{\partial L(\phi)}{\partial \theta}$  and  $\frac{\partial L(\phi)}{\partial \delta}$  to zero and solving these equations simultaneously. These equations cannot be solved analytically, so we have to appeal to numerical method.



## 6. Simulation Study

In this section, a numerical analysis will be conducted to evaluate the performance of MLE for TIIHLEtF Distribution.

**Table 1:** MLEs, biases and RMSE for some values of parameters

n	Parameters	(1,1,2.5,2.5)			(1,0.5,2,1)		
		Estimated Values	Bais	RMSE	Estimated Values	Bais	RMSE
20	$\lambda$	1.9103	0.9103	0.1823	1.0140	0.0140	0.0450
	$\alpha$	1.0472	0.0472	0.1458	0.5247	0.0247	0.0667
	$\theta$	2.5845	0.0845	0.1504	2.0140	0.0140	0.0434
	$\delta$	2.5845	0.0845	1.5894	2.0140	1.0140	1.0149
50	$\lambda$	1.7149	0.7149	0.1330	1.0076	0.0076	0.0320
	$\alpha$	1.0384	0.0384	0.0930	0.5058	0.0058	0.0303
	$\theta$	2.5401	0.0401	0.1281	2.0113	0.0113	0.0392
	$\delta$	2.5401	0.0401	1.3058	2.0113	1.0113	1.0120
100	$\lambda$	1.5193	0.5193	0.1157	1.0031	0.0031	0.0208
	$\alpha$	1.0382	0.0382	0.0696	0.5008	0.0008	0.0096
	$\theta$	2.5389	0.0389	0.1213	2.0068	0.0068	0.0294
	$\delta$	2.5389	0.0389	0.6298	2.0068	1.0068	1.0072
250	$\lambda$	1.4246	0.4246	0.0980	1.0001	0.0001	0.0018
	$\alpha$	1.0376	0.0376	0.0587	0.5000	0.0000	0.0000
	$\theta$	2.5120	0.0120	0.1202	2.0008	0.0008	0.0102
	$\delta$	2.5120	0.0120	0.6126	2.0008	1.0008	1.0009
500	$\lambda$	1.3032	0.3032	0.0849	1.0000	0.0000	0.0000
	$\alpha$	1.0303	0.0303	0.0470	0.5000	0.0000	0.0000
	$\theta$	2.5110	0.0110	0.1170	2.0000	0.0000	0.0000
	$\delta$	2.5110	0.0110	0.5114	2.0000	1.0000	1.0000
1000	$\lambda$	1.1332	0.1332	0.0762	1.0000	0.0000	0.0000
	$\alpha$	1.0282	0.0282	0.0396	0.5000	0.0000	0.0000
	$\theta$	2.5032	0.0032	0.1092	2.0000	0.0000	0.0000
	$\delta$	2.5032	0.0032	0.5036	2.0000	1.0000	1.0000

The table above shows the values of biases and RMSEs approach zero and the estimates tend to the initial (true) values as the sample increases, which indicates that the estimates are efficient and consistent.

## 7. Applications to Real Data

In this section, we apply the TIIHLEtF distribution to two real datasets and perform a comparative analysis by contrasting it with fits to other distribution models. Specifically, we compare it with the Exponentiated Half-Logistic Frechet (EHLF) distribution proposed by Cordeiro *et al.*, [6], Kumaraswamy Frechet (KExF) distribution by Mead and Abd-Eltawab [14], the Gompertz Frechet (GoFr) distribution by Oguntunde *et al.*, [18], the Exponentiated Frechet (ExFr) distribution by Nadaraja and Kotz [16], and the Frechet distribution introduced by Frechet [9]. This comparison is carried out for illustrative purposes.

The EHLF distribution developed by Cordeiro *et al.* [6] has pdf defined as:

$$f(x; \alpha, \lambda, \theta, \beta) = 2\alpha\lambda\theta\beta^\theta x^{-(\theta+1)} e^{-\left(\frac{\beta}{x}\right)^\theta} \left[ 1 - e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^{\lambda-1} \left[ 1 - \left[ 1 - e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^\lambda \right]^{\alpha-1} \left[ 1 + \left[ 1 - e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^\lambda \right]^{-(\alpha+1)} \quad (32)$$

The KExF distribution developed by Mead and Abd-Eltawab [14] has pdf defined as:

$$f(x; \alpha, \lambda, \theta, \beta) = \alpha\lambda\beta\theta^\beta x^{-\beta-1} e^{-\alpha\left(\frac{\theta}{x}\right)^\beta} \left[ 1 - e^{-\alpha\left(\frac{\theta}{x}\right)^\beta} \right]^{\lambda-1} \quad (33)$$

The GoFr distribution proposed by Oguntunde *et al.*, [18] has pdf given as:

$$f(x; \alpha, \lambda, \theta, \beta) = \theta\beta\alpha^\beta x^{-\beta-1} e^{-\left(\frac{\alpha}{x}\right)^\beta} \left[ e^{-\left(\frac{\alpha}{x}\right)^\beta} \right]^{\lambda-1} e^{-\left[ \frac{\theta}{\lambda} \left( 1 - \left[ 1 - e^{-\left(\frac{\alpha}{x}\right)^\beta} \right] \right)^{-\lambda} \right]} \quad (34)$$

The ExFr Distribution proposed by Nadaraja and Kotz [16] has pdf given as:

$$f(x; \alpha, \lambda, \sigma) = \alpha\lambda\sigma^\lambda \left[ 1 - e^{-\left(\frac{\sigma}{x}\right)^\lambda} \right]^{\alpha-1} x^{-(1+\lambda)} e^{-\left(\frac{\sigma}{x}\right)^\lambda} \quad (35)$$

The Frechet distribution developed by Frechet [9] has pdf defined as:

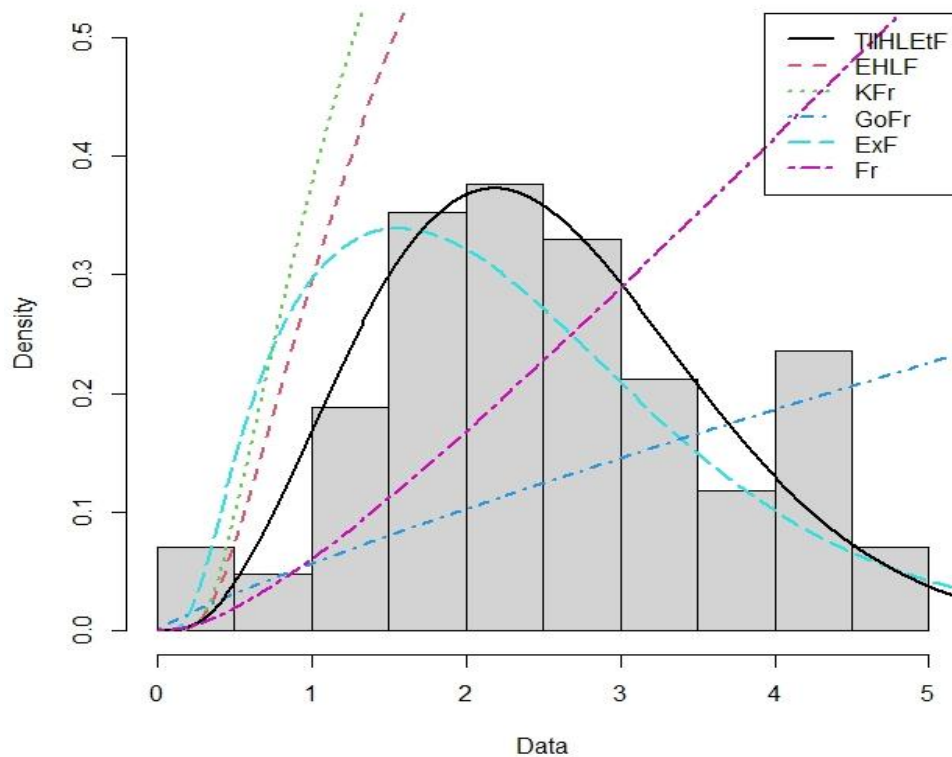
$$f(x; \theta, \sigma) = \delta\theta^\sigma x^{-\sigma-1} e^{-\left(\frac{\theta}{x}\right)^\sigma} \quad (36)$$

The two datasets utilized as illustrative examples in this application showcase the enhanced distribution flexibility and suitability of the newly proposed distribution. It also demonstrates its ability to provide the "best fit" when empirically modeling these datasets, surpassing the previously mentioned comparator distributions. All calculations were carried out using the R programming language.

### Data set 1

The first dataset provided below contains information about the times at which 84 aircraft windshields experienced failures. This dataset was previously utilized in a study by Tahir *et al.*, [21].

0.040, 1.866, 2.385, 3.443, 0.301, 1.876, 2.481, 3.467, 0.309, 1.899, 2.610, 3.478, 0.557, 1.911, 2.625, 3.578, 0.943, 1.912, 2.632, 3.595, 1.070, 1.914, 2.646, 3.699, 1.124, 1.981, 2.661, 3.779, 1.248, 2.010, 2.688, 3.924, 1.281, 2.038, 2.82,3, 4.035, 1.281, 2.085, 2.890, 4.121, 1.303, 2.089, 2.902, 4.167, 1.432, 2.097, 2.934, 4.240, 1.480, 2.135, 2.962, 4.255, 1.505, 2.154, 2.964, 4.278, 1.506, 2.190, 3.000, 4.305, 1.568, 2.194, 3.103, 4.376, 1.615, 2.223, 3.114, 4.449, 1.619, 2.224, 3.117, 4.485, 1.652, 2.229, 3.166, 4.570, 1.652, 2.300, 3.344, 4.602, 1.757, 2.324, 3.376, 4.663.



**Figure 3:** Fitted pdfs for the TIIHLEtF, EHLF, KFr, GoFr, ExF, and Fr distributions to the data set 1

**Table 2: MLEs, Log-likelihoods and Goodness of Fits Statistics for the Data Set 1**

Distributions	$\alpha$	$\lambda$	$\theta$	$\delta$	$\beta$	LL	AIC
TIIHLEtF	1.4258	1.4258	1.0593	0.7272	-	-60.4319	128.8638
EHLF	0.6829	22.7797	0.5953	-	16.8659	-152.1688	312.3376
KFr	13.1105	1.9176	0.1052	-	0.8131	-63.4185	134.837
GoFr	1.3750	1.5499	5.3750	-	1.3750	-186.4972	380.9943
ExF	5.7603	0.6018		7.1979	-	-167.5459	341.0917
Fr	-	-	19.5745	0.3347	-	-146.065	296.1299

Table 2 displays the outcomes of maximum likelihood estimation for estimating the parameters of both the newly proposed distribution and five comparator distributions. Evaluating goodness of fit, the new proposed distribution exhibited the lowest AIC value, with the KFr distribution coming in a close second. A visual assessment of the fit, as shown in Figure 3, further reinforces the superiority of the proposed distribution when compared to the comparator distributions. Consequently, the newly proposed distribution is deemed the most suitable choice for modeling an aircraft windshields failure dataset from the assortment of distributions under consideration.

Data set 2

The second dataset presented below records both the instances of failure and the periods of service for a windshield. This dataset was previously employed in a study conducted by Kundu and Raqab [12].

0.046, 1.436, 2.592, 0.140, 1.492, 2.600, 0.150, 1.580, 2.670, 0.248, 1.719, 2.717, 0.280, 1.794, 2.819, 0.313, 1.915, 2.820, 0.389, 1.920, 2.878, 0.487, 1.963, 2.950, 0.622, 1.978, 3.003, 0.900, 2.053, 3.102, 0.952, 2.065, 3.304, 0.996, 2.117, 3.483, 1.003, 2.137, 3.500, 1.010, 2.141, 3.622, 1.085, 2.163, 3.665, 1.092, 2.183, 3.695, 1.152, 2.240, 4.015, 1.183, 2.341, 4.628, 1.244, 2.435, 4.806, 1.249, 2.464, 4.881, 1.262, 2.543, 5.140.

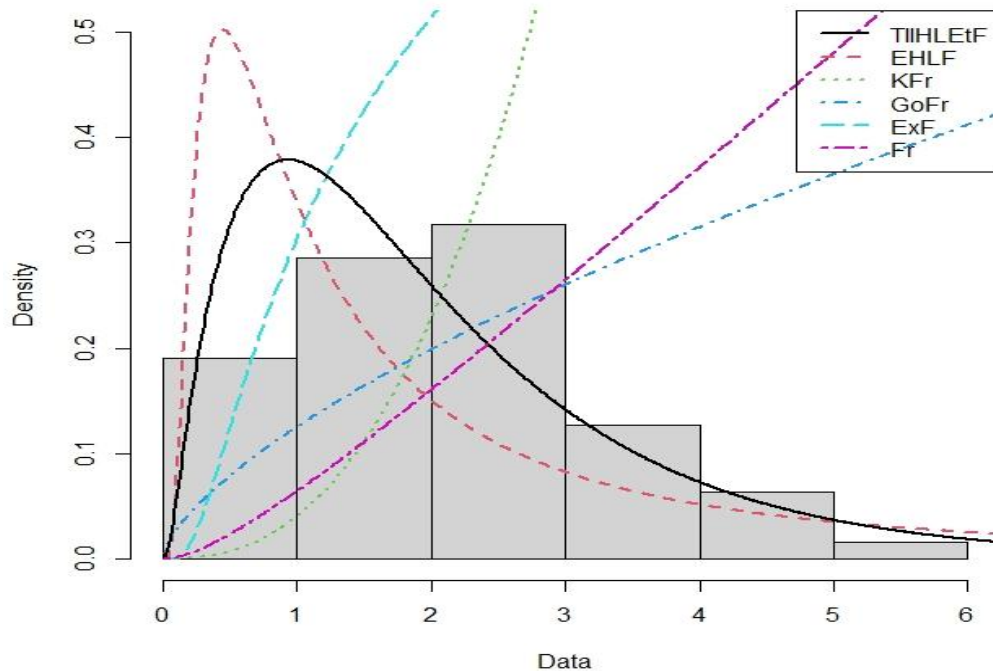


Figure 4: Fitted pdfs for the TIIHLEtF, EHLF, KFr, GoFr, ExF, and Fr distributions to the data set 2

Table 3: MLEs, Log-likelihoods and Goodness of Fits Statistics for the Data Set 2

Distributions	$\alpha$	$\lambda$	$\theta$	$\delta$	$\beta$	LL	AIC
TIIHLEtF	0.8728	2.1344	0.7247	0.6787	-	- 75.4283	158.8566
EHLF	1.3995	5.6071	0.5085	-	3.1002	- 115.1316	238.2632
KFr	0.0046	0.0248	0.0084	-	2.5480	- 130.8708	269.7416
GoFr	3.3750	2.0249	5.3750	-	3.3750	-111.8307	231.6613
ExF	6.5403	0.3166	-	9.4231	-	- 108.8879	223.7758
Fr	-	-	19.4876	0.2848	-	-139.9228	283.8457

In Table 3, you can find the outcomes of Maximum Likelihood Estimation for estimating the parameters of the TIIHLEtF distribution and five other comparator distributions. When considering the goodness of fit statistic AIC, it's worth noting that the new distribution displayed the lowest AIC value, indicating that it is the most appropriate fit for the hypertension patients' dataset. Furthermore, a visual examination of the fit, as depicted in Figure 4, reinforces the superiority of the new distribution over its comparator counterparts. Hence, the new distribution is confirmed as the optimal choice for modeling the data of instances of failure and the periods of service for a windshield.

## 8. CONCLUSION

In this article, we introduced and explored a novel distribution known as the Type II Half-Logistic Exponentiated Frechet Distribution, building upon the distribution family originally proposed by Bello *et al.*, [4]. We conducted a thorough examination of various statistical components associated with this new distribution, including the explicit quantile function, probability-weighted moments, moments, generating function, reliability function, hazard function, and order statistics. The estimation of its parameters was carried out using the maximum likelihood technique. We presented simulation results to assess the performance of this new distribution, and we also compared it to well-established models. Furthermore, we applied it to analyze two real datasets to underscore the significance and versatility of the new distribution. The findings suggest that the new distribution outperforms the existing models considered, indicating its potential applicability in a wide range of practical applications for modeling data.

## References

- [1] Afify, A. Z., Hamedani, G. G., Ghosh, I. and Mead, M. E. (2015). The transmuted Marshall-Olkin Frechet distribution: properties and applications. *International Journal of Statistics and Probability*, 4: 132-184.
- [2] Afify, A. Z., Yousof, H. M., Cordeiro, G. M., M. Ortega, E. M., and Nofal, Z. M. (2016). The Weibull Frechet distribution and its applications. *Journal of Applied Statistics*, 43(14): 2608-2626.
- [3] Afify, A. Z., Yousof, H. M., Cordeiro, G. M., Nofal, Z. M., and Ahmad, M. (2016). The Kumaraswamy Marshall-Olkin Frechet Distribution with Applications. *Journal of ISOSS*, 2(2): 151-168.
- [4] Bello, O. A., Doguwa, S. I., Yahaya, A. and Jibril, H. M. (2021). A Type II Half-Logistic Exponentiated-G Family of Distributions with Applications to Survival Analysis, *FUDMA Journal of Sciences*, 5(3): 177-190.
- [6] Cordeiro, G. M., Alizadeh, M. and Ortega, E. M. M. (2014). The exponentiated half logistic family of distributions: Properties and applications. *Journal of Probability and Statistics*, 1-21.
- [7] Da Silva, R. V., de Andrade, T. A., Maciel, D. B., Campos, R. P., and Cordeiro, G. M. (2013). A New Lifetime Model: The Gamma Extended Frechet Distribution. *Journal of Statistical Theory and Applications*, 12(1): 39-54.
- [8] Elbatal, I., Asha, G., and Raja, A. V. (2014). Transmuted exponentiated Frechet distribution: properties and applications. *Journal of Statistics Applications & Probability*, 3(3): 379.
- [9] Frechet, M. (1924). Sur la Loi des Erreurs d'Observation. *Bulletin de la Soci et e Math ematique de Moscou*, 33(1): 5-8.
- [10] Greenwood, J. A., Landwehr, J. M., Matalas, N. C., and Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water resources research*, 15(5): 1049-1054.
- [11] Krishna, E., Jose, K. K., Alice, T., and Ristic, M. M. (2013). The Marshall-Olkin Fréchet distribution. *Communications in Statistics-Theory and Methods*, 42(22): 4091-4107.

- [12] Kundu, D., and Raqab, M. Z. (2009). Estimation of  $R = P(Y < X)$  for three-parameter Weibull distribution, *Stat. Prob. Lett.*, 79(6): 1839-1846.
- [13] Mahmoud, M. R., and Mandouh, R. M. (2013). On the transmuted Fréchet distribution. *Journal of Applied Sciences Research*, 9(10): 5553-5561.
- [14] Mead, M. E. and Abd-Eltawab A. R. (2014). A note on Kumaraswamy Frechet distribution. *Australian Journal of Basic and Applied Sciences*, 8: 294-300.
- [15] Mead, M. E., Afify, A. Z., Hamedani, G. G., and Ghosh, I. (2017). The beta exponential Frechet distribution with applications. *Austrian Journal of Statistics*, 46(1): 41-63.
- [16] Nadaraja, S., and Kotz, S. (2003). The Exponentiated Frechet distribution, *Interstat Electronic Journal*, 1-7.
- [17] Nadarajah, S., and Gupta, A. K. (2004). The beta Frechet distribution. *Fareast journal of theoretical statistics*, 14(1): 15-24.
- [18] Oguntunde, P. E., Khaleel, M. A., Ahmed, M. T., and Okagbue, H. I. (2019). The Gompertz frechet distribution: properties and applications. *Cogent Mathematics & Statistics*, 6(1), 1568662.
- [19] Ramos, M.W. A., Marinho, P. R. D., Da Silva, R. V., and Cordeiro, G. M. (2013). The exponentiated Lomax Poisson distribution with an application to lifetime data. *Advances and Applications in Statistics*, 34(2), 107-135.
- [20] Tablada, C. J., and Cordeiro, G. M. (2017). The modified Frechet distribution and its properties. *Communications in Statistics-Theory and Methods*, 46(21), 10617-10639.
- [21] Tahir, M. H., Cordeiro, G. M., Mansoor, M., and Zubair, M. (2015). The Weibull-Lomax distribution: Properties and applications. *Hacettepe Journal of Mathematics and Statistics*, 44, 461-480.

# AVAILABILITY OPTIMIZATION OF A PAINT MANUFACTURING PLANT USING GREY WOLF OPTIMIZATION: A METAHEURISTIC APPROACH

Ashish Kumar<sup>1</sup>, Vijay Singh Maan<sup>2</sup>, Monika Saini<sup>3</sup>

<sup>1,2,3</sup>Department of Mathematics and Statistics, Manipal University Jaipur, Jaipur

<sup>1</sup> ashishbarak2020@gmail.com,

<sup>2</sup> vsmaan06@gmail.com,

<sup>3</sup> drmnksaini4@gmail.com

## Abstract

*The primary objective of present research work is to evaluate and improve the performance and availability of the paint manufacturing plant. Paint manufacturing plant consists of five subsystem naming mixer, grinder, thinner, labelling, and filling unit. Among them labelling and filling unit have two machines in parallel configuration and both are working simultaneously. All failure and repair rates are distributed exponentially. Markov birth-death process is utilized to model the dynamic behavior of the system and its sub-components, enabling a quantitative analysis of system availability. Grey wolf optimization (GWO), a swarm-based optimization technique is used to optimize the availability of the system. Moreover, the research conducts a thorough comparison between the outcomes derived from the Markov birth-death process and the GWO technique. By harnessing the power of GWO, the study aims to further enhance the plant's overall performance.*

**Keywords:** Paint Manufacturing Plant, Markov Birth-death Process, Availability, Grey Wolf Optimization

## I. Introduction

In the contemporary industrial landscape, the pursuit of enhanced operational efficiency and availability remains a paramount concern for manufacturing facilities across various sectors. The paint manufacturing industry plays a pivotal role in sectors such as automotive, construction, and consumer goods. However, the intricacies of operating a paint manufacturing plant entail multifaceted challenges that impact both production efficiency and overall plant availability. The convergence of factors including equipment breakdowns, maintenance scheduling, and process bottlenecks can lead to undesirable downtime and reduced performance. Thus, a systematic investigation into optimizing plant performance is not only a scientific pursuit but a practical necessity.

Historically, the paint manufacturing industry has undergone significant transformations, mirroring advancements in technology, materials, and process optimization. As a result, the industry's journey has been marked by shifts in production methodologies, ingredient formulations, and quality assurance practices. Over the years, the industry's evolution has been propelled by the growing demand for superior quality coatings, environmental sustainability, and cost-effective

production. The past era of paint manufacturing was characterized by conventional batch processes and manual labor-intensive operations. These approaches often introduced variability in product quality and production efficiency. However, with the advent of automation, computer-aided design, and advanced process control systems, the industry witnessed substantial improvements in reliability and productivity. Automation minimized human errors, enhanced process repeatability, and facilitated real-time monitoring and control of critical process parameters.

The increasing complexity of paint manufacturing processes, coupled with the demand for higher product quality, has driven the need for sophisticated analytical and optimization tools. In response to this demand, researchers and practitioners have explored various methodologies to enhance the operational reliability and productivity of manufacturing plants. One prominent avenue of exploration has been the integration of metaheuristic techniques, which offer innovative approaches to tackle complex optimization problems. Soltanali et al. [12] aimed to enhance automotive manufacturing productivity and reliability using RAM methodologies. It identified bottlenecks in the vehicle body conveying process and optimized maintenance intervals to improve operational performance. Dahiya and Kumar [4] introduced a novel method for assessing a paint manufacturing plant's performance and availability analysis by employing fuzzy reliability and coverage factors. Ostadi [6] employed a general preventive maintenance model to optimize maintenance costs while ensuring reliability and availability in a flexible manufacturing system (FMS). An optimal preventive maintenance framework was applied to a robot paint sprayer, providing maintenance plans and reliability parameters. Omoregbe and Eniola [7] investigated maintenance practices' impact on competitive advantage in the paint manufacturing industry, revealing a positive relationship between preventive maintenance and competitive advantage. Chanda and Naskar [8] focused on assessing reliability of paint manufacturing plant by collecting breakdown and maintenance data, identifying worker inefficiency and component degradation as primary failure factors. Schultmann et al. [11] addressed challenges faced by small and medium sized companies in supply chains, focusing on reliable throughput times amid uncertainties. It proposed a fuzzy scheduling approach for hybrid flow shops and validated it through a case study in paint manufacturing.

Metaheuristic approaches are widely used in availability optimization problems to find near-optimal solutions for complex problems. Saini et al. [10] assessed cloud infrastructure's availability, crucial for its operation in healthcare and business. Utilizing both, dragonfly algorithm (DA) and grey wolf algorithms (GWO), a stochastic model was optimized, emphasizing the superior performance of the GWO. Saini et al. [9] aimed to create an innovative, efficient irrigation system (EIS) using a series-configured setup with internal cold standby redundancy for sensor units and optimization was performed with GWO and DA to enhance system efficiency and performance. Kumar et al. [2] employed metaheuristic algorithms genetic algorithm (GA) and particle swarm optimization (PSO), to optimize performance of cooling tower. A novel stochastic model for a six-subsystem cooling tower was developed using Markovian processes, considering factors like random variables, repair, and failure rates. Saini et al. [8] aimed to develop a novel stochastic model for optimizing the availability of embedded life-critical systems by using DA and GWO algorithms. Yadav et al. [13] analyzed the reliability and availability of a repairable system using the Markov approach. The impact of failure rate, repair rate, and operating time on reliability, MTSF, and availability was also discussed. Saini et al. [7] aimed to assess the availability and performance of a sewage treatment plant's primary unit using redundancy. Mirjalili et al. [3] introduced the Grey Wolf Optimizer (GWO), a metaheuristic inspired by grey wolves' social structure and hunting behavior. It outperformed other metaheuristics on various test functions and successfully tackled engineering design problems.

The whole manuscript is divided into five sections. Section 1 includes the introduction of proposed system and previous work done in related area. section 2 provides the insights into used materials and methods for investigation. In section 3, mathematical modelling, steady state diagram



and availability analysis of the system is mentioned. Numerical and graphical representation of results is appended in section 4. Section 5 cover the conclusion part of the research.



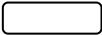
## II. Material and Methods

This section contains the notations and methodology used for the availability investigation of paint manufacturing plant.

### I. Notations

The following nomenclature is used to develop the state transition diagram and mathematical modelling of system.

**Table 1:** Notations for paint manufacturing plant's sub-system

Sr. no.	Sub-systems and notations	Notations for different states function			Failure rates ( $\alpha_i$ )	Repair rates ( $\beta_j$ )
		Operative state	Degraded states	Complete failed state		
1	Mixer (U)	U	-	u	$\alpha_1$	$\beta_1$
2	Grinder (V)	V	-	v	$\alpha_2$	$\beta_2$
3	Dilution/Thinner (W)	W	-	w	$\alpha_3$	$\beta_3$
4	Labelling unit ( $X^2$ ) (Two parallel machine)	$X^2$	$X^1$	x	$\alpha_4, \alpha_6$	$\beta_4, \beta_6$
5	Filling unit ( $Y^2$ ) (Two parallel machine)	$Y^2$	$Y^1$	y	$\alpha_5, \alpha_7$	$\beta_5, \beta_7$
6	$P_i(t)$	Probability that the system is in $i^{\text{th}}$ state at time t				
7		Operative states				
8		Degraded states				
9		Completely failed states				

### II. System Description

The proposed paint manufacturing system comprises five sub-systems like mixer, grinder, thinner, labelling unit, and filling unit. The failure and repair rates of all the subsystems follow exponential distribution. All the subsystem arranged in a series configuration and work-flow diagram of system is append in figure 1.

#### i) Subsystem U (Mixer)

In paint manufacturing, a mixer unit plays a crucial role in blending and homogenizing various raw materials to create consistent and high-quality paint products. The unit's primary purpose is to create a homogeneous mixture by effectively dispersing and combining the ingredients. The failure of this unit can result in the entire system's breakdown.

#### ii) Subsystem V (Grinder)

A grinder unit serves the essential purpose of reducing solid particles, such as pigments and

fillers, into finer particles to achieve the desired texture and consistency in the final paint product. The grinder unit plays a crucial role in breaking down aggregates and achieving uniform particle size distribution, which directly influences the paint's color, opacity, gloss, and overall quality. The failure of this subsystem can impact the overall functionality of the system.

### iii) Subsystem W (Thinner/Diluter)

Thinner or diluter plays a pivotal role in paint manufacturing as a vital solvent used to modify the viscosity and consistency of paint formulations. Thinner is employed to reduce the thickness of paint, making it easier to apply and ensuring a smooth, even coat. Failure of subsystem can disrupt and compromise the entire operation of the system. Subsystem failures have the potential to disrupt and compromise the entire system's operation.

### iv) Subsystem X (Labelling unit)

A labelling unit plays a pivotal role in ensuring that each container bears essential information, including product details, batch numbers, safety warnings, and regulatory compliance. This system comprises two labelling machines working together in parallel configuration with different failure and repair rates.

### v) Subsystem Y (Filling unit)

A filling unit in a paint manufacturing plant is responsible for accurately filling paint into containers, such as cans or buckets. Its importance lies in ensuring precise and consistent product quantities, which are essential for quality control and cost efficiency. The system consists of two filling machines operating in parallel, each with its own distinct rates of failure and repair.

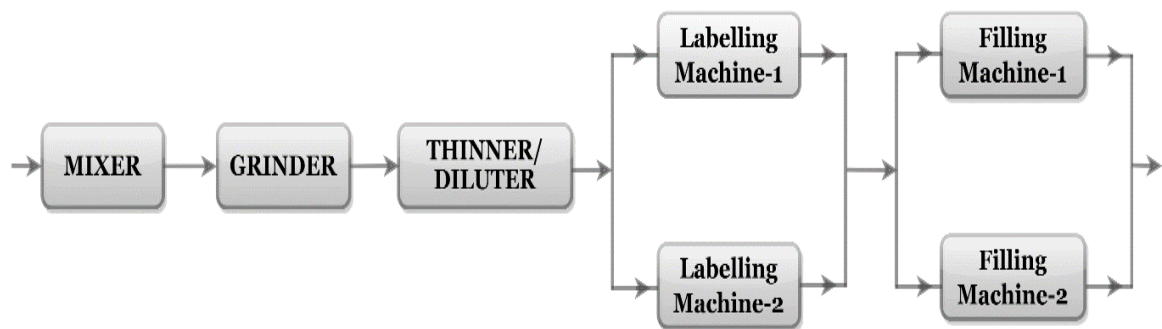


Figure 1: Work-flow diagram of system

## III. Assumptions

- At time  $t=0$ , all subsystems are in good working condition without any failure.
- The rates of failure and repair are exponentially distributed and are equally and independently distributed.
- All subsystems of the paint manufacturing plant are configured in a series format while labelling unit and filling unit have two unit working together in parallel configuration.
- Subsystems works as flawlessly as new after repair.
- An adequate repair facility is always available at operational time.

### III. Mathematical Modelling and Analysis

In this section, a mathematical model for paint manufacturing plant is developed using Markov birth-death process. The Chapman-Kolmogorov differential difference equations derived based on figure 2.

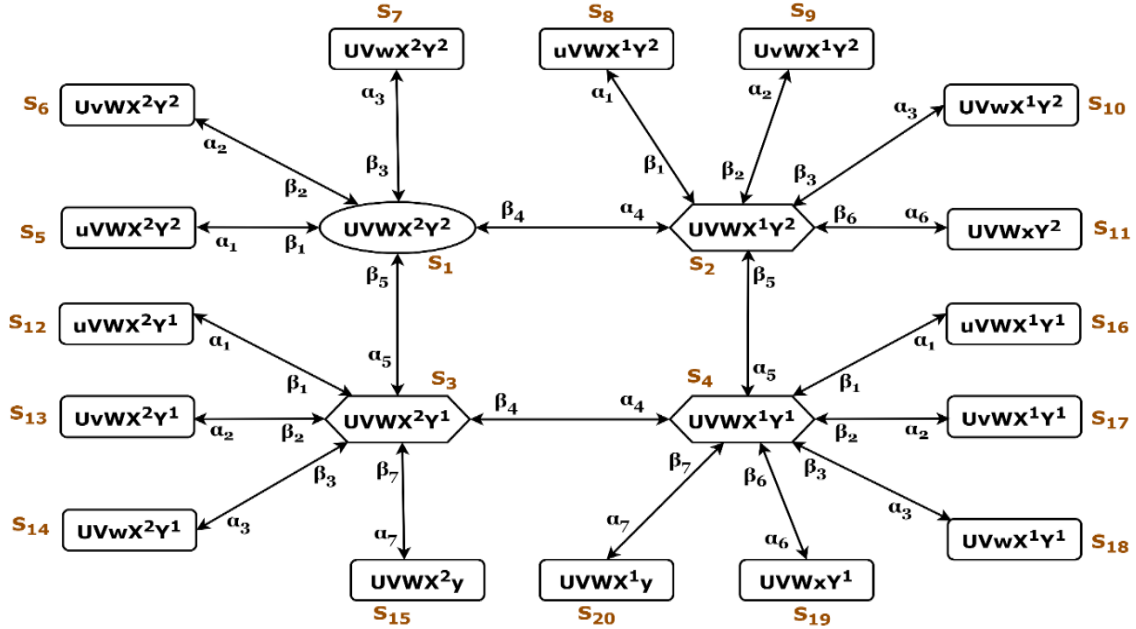


Figure 2. State transition diagram of paint manufacture plants

#### I. Transition Probabilities

$$P_1(t + \Delta t) = (1 - \alpha_1 \Delta t - \alpha_2 \Delta t - \alpha_3 \Delta t - \alpha_4 \Delta t - \alpha_5 \Delta t)P_1(t) + \beta_1 P_5(t) \Delta t + \beta_2 P_6(t) \Delta t + \beta_3 P_7(t) \Delta t + \beta_4 P_2(t) \Delta t + \beta_5 P_3(t) \Delta t$$

$$P_1(t + \Delta t) = P_1(t) - (\alpha_1 \Delta t + \alpha_2 \Delta t + \alpha_3 \Delta t + \alpha_4 \Delta t + \alpha_5 \Delta t)P_1(t) + \beta_1 P_5(t) \Delta t + \beta_2 P_6(t) \Delta t + \beta_3 P_7(t) \Delta t + \beta_4 P_2(t) \Delta t + \beta_5 P_3(t) \Delta t$$

$$\lim_{\Delta t \rightarrow 0} \frac{P_1(t+\Delta t) - P_1(t)}{\Delta t} = -(\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5)P_1(t) + \beta_1 P_5(t) + \beta_2 P_6(t) + \beta_3 P_7(t) + \beta_4 P_2(t) + \beta_5 P_3(t)$$

$$P_1'(t) = -(\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5)P_1(t) + \beta_1 P_5(t) + \beta_2 P_6(t) + \beta_3 P_7(t) + \beta_4 P_2(t) + \beta_5 P_3(t)$$

Taking limit  $\lim_{t \rightarrow \infty}$ , we get

$$\lim_{t \rightarrow \infty} P_1'(t) = -(\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5)P_1(t) + \beta_1 P_5(t) + \beta_2 P_6(t) + \beta_3 P_7(t) + \beta_4 P_2(t) + \beta_5 P_3(t)$$

$$(\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5)P_1 = \beta_1 P_5 + \beta_2 P_6 + \beta_3 P_7 + \beta_4 P_2 + \beta_5 P_3 \quad (1)$$

Similarly for others states,

$$(\alpha_1 + \alpha_2 + \alpha_3 + \beta_4 + \alpha_5 + \alpha_6)P_2 = \beta_1 P_8 + \beta_2 P_9 + \beta_3 P_{10} + \alpha_4 P_1 + \beta_5 P_4 + \beta_6 P_{11} \quad (2)$$

$$(\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \beta_5 + \alpha_7)P_3 = \beta_1 P_{12} + \beta_2 P_{13} + \beta_3 P_{14} + \beta_4 P_4 + \alpha_5 P_1 + \beta_7 P_{15} \quad (3)$$

$$(\alpha_1 + \alpha_2 + \alpha_3 + \beta_4 + \beta_5 + \alpha_6 + \alpha_7)P_4 = \beta_1 P_{16} + \beta_2 P_{17} + \beta_3 P_{18} + \alpha_4 P_3 + \alpha_5 P_2 + \beta_6 P_{19} + \beta_7 P_{20} \quad (4)$$

$$\sum_{i=1}^3 \alpha_i P_1 = \sum_{j=1}^3 \beta_j P_{j+4} \quad (5)$$

$$\sum_{k=1}^3 \alpha_k P_2 = \sum_{l=1}^3 \beta_l P_{l+7} \quad (6)$$

$$\alpha_6 P_2 = \beta_6 P_{11} \quad (7)$$

$$\sum_{m=1}^3 \alpha_m P_3 = \sum_{n=1}^3 \beta_n P_{n+11} \quad (8)$$

$$\alpha_7 P_3 = \beta_7 P_{15} \quad (9)$$

$$\sum_{q=1}^3 \alpha_q P_4 = \sum_{r=1}^3 \beta_r P_{r+15} \quad (10)$$

$$\sum_{s=6}^7 \alpha_s P_4 = \sum_{t=6}^7 \beta_t P_{t+13} \quad (11)$$

Initial conditions,

$$P_{\xi}(0) = \begin{cases} 1 & \text{if } \xi = 0 \\ 0 & \text{if } \xi \neq 0 \end{cases} \quad (12)$$

Solving the linear system of equations (1-11) by using initial conditions mentioned in equation (12), the following probabilities derived at various states and solve them in terms of  $P_1$ , we get

$$\begin{aligned} P_2 &= GP_1, P_3 = HP_1, P_4 = IP_1, P_5 = \frac{\alpha_1}{\beta_1} P_1, P_6 = \frac{\alpha_2}{\beta_2} P_1, P_7 = \frac{\alpha_3}{\beta_3} P_1, P_8 = \frac{\alpha_1}{\beta_1} P_2, P_9 = \frac{\alpha_2}{\beta_2} P_2, \\ P_{10} &= \frac{\alpha_3}{\beta_3} P_2, P_{11} = \frac{\alpha_6}{\beta_6} P_2, P_{12} = \frac{\alpha_1}{\beta_1} P_3, P_{13} = \frac{\alpha_2}{\beta_2} P_3, P_{14} = \frac{\alpha_3}{\beta_3} P_3, P_{15} = \frac{\alpha_7}{\beta_7} P_3, P_{16} = \frac{\alpha_1}{\beta_1} P_4, \\ P_{17} &= \frac{\alpha_2}{\beta_2} P_4, P_{18} = \frac{\alpha_3}{\beta_3} P_4, P_{19} = \frac{\alpha_6}{\beta_6} P_4, P_{20} = \frac{\alpha_7}{\beta_7} P_4 \end{aligned} \quad (13)$$

Here,

$$G = \left(\frac{\alpha_4}{B} + \frac{I^* \beta_5}{B}\right) P_1, H = \left(\frac{\alpha_5}{C} + \frac{I^* \beta_4}{C}\right) P_1, I = \left[\frac{\alpha_4^* \alpha_5 \left(\frac{1}{C} + \frac{1}{B}\right)}{\left(D - \frac{\alpha_4^* \beta_4 + \alpha_5^* \beta_5}{B}\right)}\right], A = (\alpha_4 + \alpha_5), B = (\beta_4 + \alpha_5), \\ C = (\alpha_4 + \beta_5), D = (\beta_4 + \beta_5) \text{ and } '*' \text{ represent the multiplication.}$$

By using normalization condition,

$$\sum_{z=1}^{20} P_z = 1 \quad (14)$$

The expression of  $P_1$  derived by using equations (13-14) and shown in equation (15) as follows:

$$P_1 + P_2 + P_3 + \dots + P_{20} = 1 \\ P_1 = \frac{1}{[1+G+H+I] * \left[1 + \left(\frac{\alpha_1}{\beta_1}\right) + \left(\frac{\alpha_2}{\beta_2}\right) + \left(\frac{\alpha_3}{\beta_3}\right) + \left(\frac{\alpha_6}{\beta_6}\right) * I + \left(\frac{\alpha_7}{\beta_7}\right) * (H+I)\right]} \quad (15)$$

The depiction of system availability involves the addition of probabilities in the upstate. Mathematical expression for system availability is formulated as follows:

$$A_{\theta} = P_1 + P_2 + P_3 + P_4 \quad (16)$$

By putting the values and determine the final availability expression, is as below:

$$A_{\theta} = \frac{[1+G+H+I]}{[1+G+H+I] * \left[1 + \left(\frac{\alpha_1}{\beta_1}\right) + \left(\frac{\alpha_2}{\beta_2}\right) + \left(\frac{\alpha_3}{\beta_3}\right) + \left(\frac{\alpha_6}{\beta_6}\right) * I + \left(\frac{\alpha_7}{\beta_7}\right) * (H+I)\right]} \quad (17)$$

#### IV. Numerical Results and Discussion

In this section, the availability of paint manufacturing system is derived by using the expression given in equation (17) and is found 0.950145478. The arbitrary values of failure and repair rates are taken on the behalf of the previous studies and are append in table 2. For enhancement of availability of the system swarm-intelligence based algorithm named GWO is used. For execution of optimization the possible search space for failure and repair rates are append in table 3 and the optimum availability of the system for different iterations and populations are presented in table 4.

**Table 2:** Failure and repair rates for subsystems of paint manufacturing plant

Sr. No.	Name of subsystem	Failure rates ( $\alpha_i$ )	Repair rates ( $\beta_j$ )
1	Mixer	$\alpha_1=0.005$	$\beta_1=0.889$
2	Grinder	$\alpha_2=0.051$	$\beta_2=1.397$
3	Dilution/ Thinner	$\alpha_3=0.0052$	$\beta_3=0.998$
4	Labelling	$\alpha_4=0.0727$	$\beta_4=1.232$
5	Filling	$\alpha_5=0.0954$	$\beta_5=1.244$
6	Standby labelling machine	$\alpha_6=0.0778$	$\beta_6=1.374$
7	Standby filling machine	$\alpha_7=0.0955$	$\beta_7=1.387$

In figure 3 and 4, the effect of change in failure rate is shown on the other sub-systems availability with increase an 50% in the failure rates and repair rates. It is shown that while varying the failure rate of  $\alpha_1$  from 0.001 to 0.007, the availability of subsystems decreases. Subsystem grinder is fluctuated very much by increasing 50% in failure rates and repair rates both. While floating the value of  $\beta_1$  from 0.001 to 0.007 and 50% increase in other subsystems repair rates, then the availability is also increase.

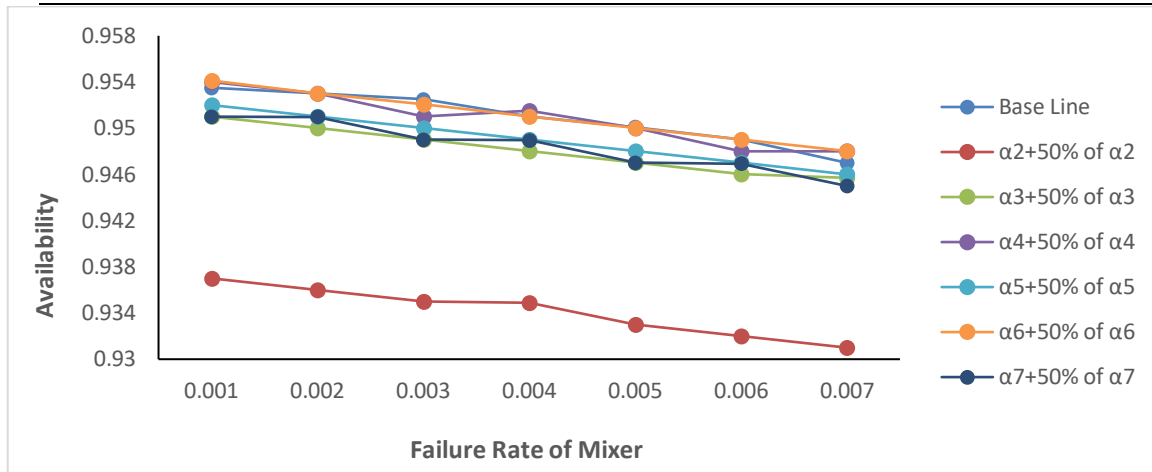


Figure 3: System availability with variation in  $\alpha_1$  and subsequent changes in failure rates of subsystems

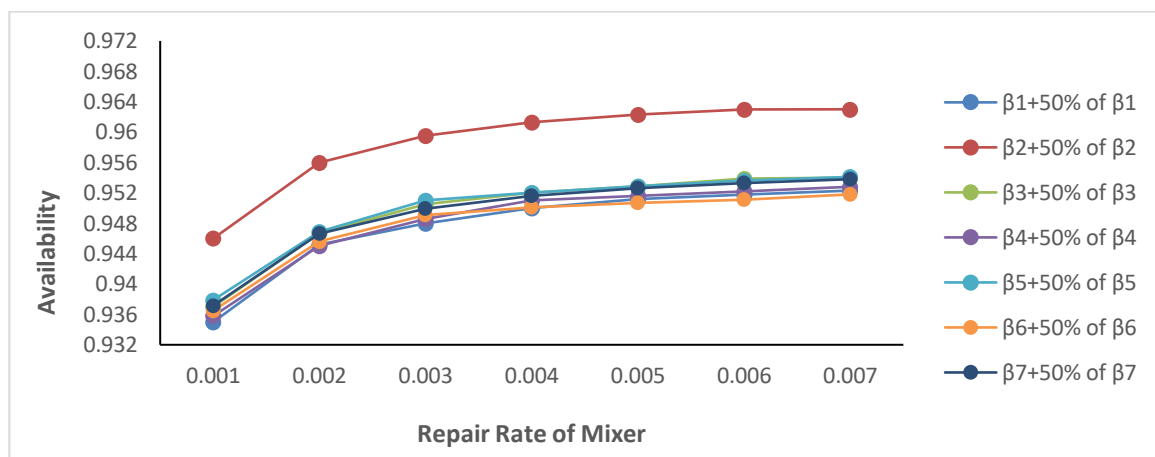


Figure 4: System availability with variation in  $\beta_1$  and subsequent changes in repair rates of subsystems

Table 3: Range of search space for grey wolf optimization

Sr. No.	Subsystem	Range of failure rates ( $\alpha_i$ )	Range of repair rates ( $\beta_j$ )
1	Mixer	[0.0025, 0.0075]	[0.45, 1.34]
2	Grinder	[0.0260, 0.0770]	[0.70, 2.10]
3	Dilution/ Thinner	[0.0028, 0.0082]	[0.50, 1.50]
4	Labelling	[0.0360, 0.1090]	[0.62, 1.85]
5	Standby labelling machine	[0.0480, 0.1440]	[0.63, 1.87]
6	Filling	[0.0390, 0.1170]	[0.69, 2.06]
7	Standby filling machine	[0.0480, 0.1440]	[0.71, 2.09]

Table 4: Optimum availability of system at different iterations with varying population sizes

Population \ Iteration	100	150	200	250	300
10	0.983572	0.983576	0.983575	0.983574	0.983570
30	0.983573	0.983571	0.983571	0.983575	0.983571
50	0.983576	0.983575	0.983549	0.983575	0.983572
70	0.983572	0.983575	0.983569	0.983563	0.983568
90	0.983577	0.983555	0.983570	0.983570	0.983573

## V. Conclusion

In this study, a comparative analysis is performed and it provides insights into the strengths and limitations of each methodology. It is shown that the metaheuristic optimization techniques perform better than the traditional techniques. The overall availability of paint manufacturing plant is improved by 0.9501454 to 0.983577 using GWO. Ultimately, the paper offers valuable insights into both the theoretical and practical dimensions of improving paint manufacturing plant performance and availability. The combined usage of Markov analysis and GWO presents a robust approach for achieving the desired goals, contributing to the advancement of industrial reliability and efficiency.

**Acknowledgement:** This research is funded by Manipal University Jaipur, under the scheme of Dr. Ramdas Pai scholarship.

**Conflict of interest:** There is no conflict of interest between authors.

## References

- [1] Chanda, A. and Naskar, S. K. (2103). Reliability Analysis of a Paint-Manufacturing Plant. In E-proceedings (p. 296).
- [2] Kumar, A., Saini, M., Gupta, N., Sinwar, D., Singh, D., Kaur, M. and Lee, H. N. (2022). Efficient stochastic model for operational availability optimization of cooling tower using metaheuristic algorithms. *IEEE Access*, 10:24659-24677.
- [3] Mirjalili, S., Mirjalili, S. M. and Lewis, A. (2014). Grey wolf optimizer. *Advances in Engineering Software*, 69:46-61.
- [4] Ombirdahiya, M. S. and Kumar, A. (2019). The mathematical modelling and performance evaluation of paint manufacturing system using fuzzy reliability approach. *International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)*, 9:869-886.
- [5] Omoregbe, O. and Eniola, Y. T. (2017). Production facilities maintenance practices and sustainable competitive advantage in the paint manufacturing industry, Benin City, Nigeria. *Annals of the University of Petrosani, Economics*, 17(1):209-222.
- [6] Ostadi, B. (2018). An optimal preventive maintenance model to enhance availability and reliability of flexible manufacturing systems. *Journal of Industrial and Systems Engineering*, 11(2):47-61.
- [7] Saini, M., Goyal, D. and Kumar, A. (2023). Availability and Performance Analysis of Primary Treatment Unit of Sewage Plant. *Thailand Statistician*, 21(2):383-396.
- [8] Saini, M., Kumar, A. and Maan, V. S. (2022). Mathematical modeling and availability optimization of embedded life critical systems. *Advanced Mathematical Models & Applications*, 7(3).
- [9] Saini, M., Kumar, A., Maan, V. S. and Sinwar, D. (2022). Efficient and Intelligent Decision Support System for Smart Irrigation. *Journal of the Nigerian Society of Physical Sciences*, 945-945.
- [10] Saini, M., Maan, V. S., Kumar, A. and Saini, D. K. (2023). Cloud infrastructure availability optimization using Dragonfly and Grey Wolf optimization algorithms for health systems. *Journal of Intelligent & Fuzzy Systems*, (Preprint),1-19.
- [11] Schultmann, F., Fröhling, M. and Rentz, O. (2006). Fuzzy approach for production planning and detailed scheduling in paints manufacturing. *International Journal of Production Research*, 44(8):1589-1612.
- [12] Soltanali, H., Garmabaki, A. H. S., Thaduri, A., Parida, A., Kumar, U., & Rohani, A. (2019). Sustainable production process: An application of reliability, availability, and maintainability methodologies in automotive manufacturing. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 233(4):682-697.
- [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 & Applications*, 18(1 (72)):193-205.

# DESIGNING AND EVALUATION OF SKIP-LOT SAMPLING PLAN OF TYPE SkSP-T WITH SINGLE SAMPLING PLAN AS REFERENCE PLAN UNDER THE CONDITION OF INTERVENED POISSON DISTRIBUTION

**S. Suganya**

•

Assistant Professor,  
Department of Statistics,  
PSG College of Arts & Science,  
Coimbatore, Tamil Nadu, India  
suganyas@psgcas.ac.in

**K. Pradeepa Veerakumari**

•

Assistant Professor,  
Department of Statistics,  
Bharathiar University,  
Coimbatore, Tamil Nadu, India  
pradeepaveerakumari@buc.edu.in

## Abstract

*This paper describes the scheming technique of new system of skip lot sampling plan of type SkSP-T with Single Sampling Plan as Reference plan under the condition of Intervened Poisson Distribution. The designing methodology includes the evaluation of Acceptable Quality Level, Limiting Quality Level, Operating Ratio, and Operating Characteristic curves. Tables are simulated by changing various parametric values of SkSP-T, SSP and IPD and operating characteristic curves are drawn by using R language.*

**Keywords:** Skip-lot sampling plan of type SkSP-T, Intervened Poisson Distribution, Single Sampling Plan.

## I. Introduction

Maintenance of quality is decided to improve the production. Good qualities of products facilitate to reduce both producer and consumer risks. Additionally, it manages the production cost and consumer satisfaction. The determination of designing every sampling method is to find out a succession of the process to be tested in a sequence of lots is defined the quality. Statistical Quality control (SQC) is one of the processing techniques throughout that the production quality is sustained and too reduced the production errors. Every quality control technique defines the defective items and the defective items are replaced by good once. Acceptance sampling (AS) is one of the imperative method used in SQC through judgment a lot concerning its quality of 100% inspection and no inspection. The major purpose of acceptance sampling is towards constructing a sampling plan that is mainly characterized by sample size (n) and acceptance number(c); also it is able to minimize the inspection cost and sampling error.

The most important areas of Acceptance Sampling plan is classified into four broad categories.

It includes lot-by-lot sampling plan by attribute and by variables, Continuous sampling plan, and special purpose plans. The special purpose plan includes Skip-lot sampling plan. Skip-lot sampling plans are inspected only the fraction of submitted lots. Such process of sampling is in reducing the cost in provisions of minimizing the time and exertion. On the other hand skip-lot sampling supposed to only be used instantaneously, it has been established that the excellent quality of the submitted lots are very good.

Dodge [5] introduced the skip-lot sampling plan of type SkSP-1 based on the concept of continuous sampling plan of type CSP-1. Perry [13] developed some specified level of Operating Ratio (OR) for corresponding producers and consumer's risk, OC and ASN function (operating characteristic) of the SkSP-2 plan using Markov chain techniques. SkSP-3 is based on the concept of Continuous Sampling Plan of type CSP-2 of Dodge and Terry [8]. SkSP-3 is developed by Vijayaraghavan [16] using Markov chain technique. The multilevel continuous sampling plans are derived Lieberman and Solomon[7]. The CSP-T plans are tightened multilevel plans that include three levels developed by Fordice [6]. Kandasamy and Govindaraju [11] developed the performance measures of CSP-T plan. Balamurali [1] developed Modified Tightened Three level Continuous sampling plan. Balamurali and Chi-Hyuck Jun [2] developed a modified CSP-T sampling procedure.

Pradeepa Veerakumari and Suganya [15] introduced Skip-lot sampling plan of type SkSP-T (T-tightened) based on the concept of continuous sampling plan of type CSP-T, CSP-M, MMLP-T-2, and SkSP-2. Sampling levels are fixed by using CSP-M procedure; sampling fractions are taken from the CSP-T procedure and other concepts are taken by modified CSP-T and SkSP-2 procedures. The main advantage of skip-lot sampling plan of type SkSP-T plan if there is a defect found in skipping the level, and then there is a normal inspection in that fraction level. The stopping rule parameter S is introduced for the tightening inspection which makes the plan convenient. In the proposed plan sampling frequency (f) is minimized by every skipping inspection level. The Operating Characteristic functions for this SkSP-T plan are also derived with single sampling plan as the reference plan under the condition of Intervened Poisson Distribution. SkSP-T plan vary among normal inspection and skipping inspection with three levels. Pradeepa Veerakumari and Suganya [23] developed skip-lot sampling plan of SkSP-T based on fuzzy logic techniques. Suganya and Pradeepa Veerakumari [22] developed SkSP-T plan based on Burr type XII distribution.

Shanmugam [18] introduced the Intervened Poisson Distribution which is designated as IPD. It is a moderated adaptation of the zero-truncated Poisson distribution (ZTPD). IPD is compared with ZTPD it concludes that IPD produces good quality of products and provide an additional report about the capability of the intervention made in the manufacturing or production process. It is supportive of accepting the result of the corresponding process. The area of IPD has been applied by various product control, process control, manufacturing industries, biologists, and etc. Much real-time (cholera disease, health improvement for before and after treatment) examples can establish in Shanmugam [18,19]. Huang and Fung [9] developed intervened truncated poisson distribution. In Scollnik [20] introduced the new concept it is called intervened generalized Poisson distribution (IGPD) and it is an extension of IPD. Scollnik [21] developed the Bayesian analysis for IPD using Gibbs sampling approach. Dhanavanthan [3,4] introduced the Compound Intervened Poisson distribution (CIPD) also estimated its characteristic parameters for using the concept of statistical inference and probability. Satheesh and Shibu [17] introduced the modified intervened Poisson distribution (MIPD). MIPD parameters can be estimated by using the method of factorial moment, mixed moment, likelihood estimators and uniformly best estimators. Pradeepa Veerakumari and Azarudheen [14] developed various attribute acceptances sampling plan using SSP under the condition of IPD as a reference plan. Jayakumar and Rehana [10] develop and Characterizations , Different Methods of Estimation and Applications of Exponential Intervened Poisson Distribution. Muhammed Rasheed Irshad et.al [12] developed Intervened Poisson Distribution by Lagrangian Approach.



## II. Design of SkSP-T plan and its Operating Characteristic function

Operating procedure of the SkSP-T plan is stated as follows:

Step1: Initiate SkSP-T procedure with normal inspection using the single sampling plan as a reference plan under the condition of Intervened Poisson Distribution.

Step2: When  $i$  successive lots are received on normal inspection, terminate the normal inspection and change to skipping inspection.

Step3: On skipping inspection, inspect only a fraction  $f$  of the lots selected at random, level 1.

Step4: After  $i$  consecutive lots in succession have been founded without a non-conforming at level 1, the system then switches to skipping inspection with a fraction of  $f/2$ , level 2.

Step5: After  $i$  consecutive lots in succession have been founded without a non-conforming at level 2, the system then switches to skipping inspection with a fraction of  $f/4$ , level 3.

Step6: If a non-conforming lot is found on either skipping level, the system reverts to normal inspection.

The Operating Characteristics Function of SkSP-T plan is given by

$$P_a(p) = \frac{P^i(f_2 f_3 (1-P^i) + f_1 f_3 P^i (1-P^i) + f_1 f_2 P^{2i})}{f_1 f_2 f_3 (1-P^i) + P^i(f_2 f_3 (1-P^i) + f_1 f_3 P^i (1-P^i) + f_1 f_2 P^{2i})} \quad (1)$$

## III. Origin of Intervened Poisson Distribution (IPD)

Let us consider the number of defectives in a lot as  $Y_1$ . Implementation of, the number of defectives in a lot produced as of a progression in no way to be perfect due to random inconsistency, which implies the event  $Y_1 > 0$ .  $Y_1$  is a random variable and it is an infrequent event. Then the zero-truncated Poisson distribution with pdf is

$$P(Y_1 = y_1) = \frac{\theta^{y_1}}{(e^\theta - 1)^{y_1}}; y_1 = 1, 2, \dots \quad (2)$$

Where  $\theta > 0$  is called incidence parameter. The above equation can be used this example if the manufactures make any modification in a manufacturing system in order to produce the better quality of products. In this position,  $\rho\theta$  can be changed to  $\theta$ . Where  $\rho$  ( $\rho \geq 0$ ) is called an intervention parameter (IP) and  $Y_2$  be the no. of defectives that occurred after modify in the production process. And  $Y_2$  is denoted that a Poisson random variable with mean  $\rho\theta$ .

A random variable  $Y = Y_1 + Y_2$  i.e., the total no. of defectives occurred. The random variable  $Y$  is formed by Intervened Poisson Distribution with probability function,

$$P(Y = y) = \frac{[(1+\rho)^y - \rho^y] \theta^y}{e^{\rho\theta} (e^\theta - 1)^y}; x = 1, 2, \dots \quad (3)$$

Mean and Variance of the Intervened Poisson Distribution (IPD) is

$$\mu = E(X) = \theta \left[ (1 + \rho) + \frac{1}{(e^\theta - 1)} \right] \quad (4)$$

And

$$\sigma^2 = Var(X) = \mu - e^\theta \left( \frac{\theta}{e^\theta - 1} \right)^2 \quad (5)$$

From the Intervened Poisson Distribution (IPD) than the mean ( $\mu$ ) is greater than its variance ( $\sigma^2$ ). In equation (2) substitute  $\rho = 0$  then the Intervened Poisson Distribution is reduced to zero-truncated Poisson distribution (ZTPD).

The Operating Characteristic function for Single Sampling Plan under the conditions of Intervened Poisson D ( $\theta, \rho$ ) can be defined as,

$$P_a(p) = \sum_{x=1}^c P(Y = y|\theta, \rho) = \sum_{y=1}^c \frac{[(1 + \rho)^y - \rho^y] \theta^y}{e^{\rho\theta} (e^\theta - 1)^y}$$

Where  $\theta = np$  and  $\rho$  is measured in percentage.

#### IV. Designing of Single Sampling Plan under the conditions of Intervened Poisson Distribution (IPD)

The attribute Single Sampling Plan for the necessary parameters are sample size (n), acceptance number (c), the lot size (N), the number of defective in the sample (d) and proportion defective (p). In general, then the defective item  $d \leq c$  (acceptance number) then the lot will be accepted, otherwise, the lot will be rejected.

The skip-lot sampling plan of type SkSP-T, Single Sampling plan and intervened Poisson distribution parameters are determined with the primary objective to carry out both the consumer's and the producer's risk. Both the risk's can be satisfying the subsequent conditions for the particular strength ( $p_1, \alpha, p_2, \beta$ ),

$$\begin{aligned} P_a(p_1) &= 1 - \alpha \\ P_a(p_2) &= \beta \end{aligned}$$

Where,  $p_1$  = is the proportion defective for that the risk of rejection is to be  $\alpha$

And  $p_2$  = is the proportion defective for that the risk of acceptance to be  $\beta$

1. Specify  $p_1$  = Acceptable Quality Level at  $\alpha = 0.05$  or  $0.01$ .
2. Specify  $p_2$  = Limiting Quality Level at  $\beta = 0.10$  or  $0.05$ .
3. Obtain the corresponding ratio  $OR = p_2 / p_1$  at a different combination of  $\alpha$  and  $\beta$ .
- 4.

#### V. Numerical Illustration

The following examples to obtain the new system of skip lot sampling plan of type SkSP-T with Single Sampling Plan as reference plan under the conditions of Intervened Poisson Distribution (IPD) for calculating the Probability of Acceptance, Operating Characteristic (OC) Curve, Average Sample Number (ASN), Average Outgoing Quality (AOQ) and Average Total Inspection (ATI). The SkSP-T, SSP, and IPD parameters  $n$  - sample size,  $c$  - acceptance number,  $i$  - clearance interval,  $f$  - sampling frequency,  $N$  - lot size,  $p$  - proportion or fraction defective,  $P_a(p)$  - Probability of Acceptance  $\theta$  - incidence parameter and  $q$  - intervention parameter.

In the production process assume there is a 1% of intervention occurred in the experimental session and it is preferred to establish a proposed sampling plan for a certain set of values say,  $\alpha=0.05, \beta=0.10$  and  $p_1=0.028741, p_2=0.1428$ . Then the Operating Ratio  $OR=p_2/p_1= 0.142879/0.028741 = 4.97$ . And  $np_1$  value is determined from table 1 as 1.437025 and the corresponding sample size  $n$  is computed as  $n = np_1/p_1 = 1.437025/0.028741= 49.99 \approx 50$ . Hence the parameters of SkSP-T with Single Sampling Plan as reference plan under the Condition of Intervened Poisson Distribution indexed through Acceptable and Limiting Quality Levels.

Table 1 considers the new system of skip-lot sampling plan of type SkSP-T with Single Sampling plan as reference plan under the condition of Intervened Poisson Distribution (IPD) parameter values are estimated and derived its Acceptable Quality Level (AQL) and Limiting Quality Level (LQL). And also calculated the Operating Ratio (OR) it explains the ratio of limiting quality level to acceptable quality level. The new proposed plan system is designing several combinations of OC curves for various parameter values of SkSP-T, SSP and IPD.

In Acceptance Sampling plan five important basic measures are defined in ISO standard (2006). The measures of Acceptance Sampling plan is 1. Probability of Acceptance ( $P_a(p)$ ) 2.Average Sample Number (ASN) 3.Average Outgoing Quality (AOQ) 4.Average Outgoing Quality Limit (AOQL) and 5.Average Total Inspection (ATI).

Figure 1 represents the OC curves for SkSP-T with SSP as reference plan under the conditions of IPD for  $c$  and  $n$  are fixed and by changing the  $q$  values. From this figure conclude that smaller intervention occurs in the production process then the producer's risk will be minimized. Also the acceptance number  $c = 1$  it concludes that proportion defective  $p$  is decreases.

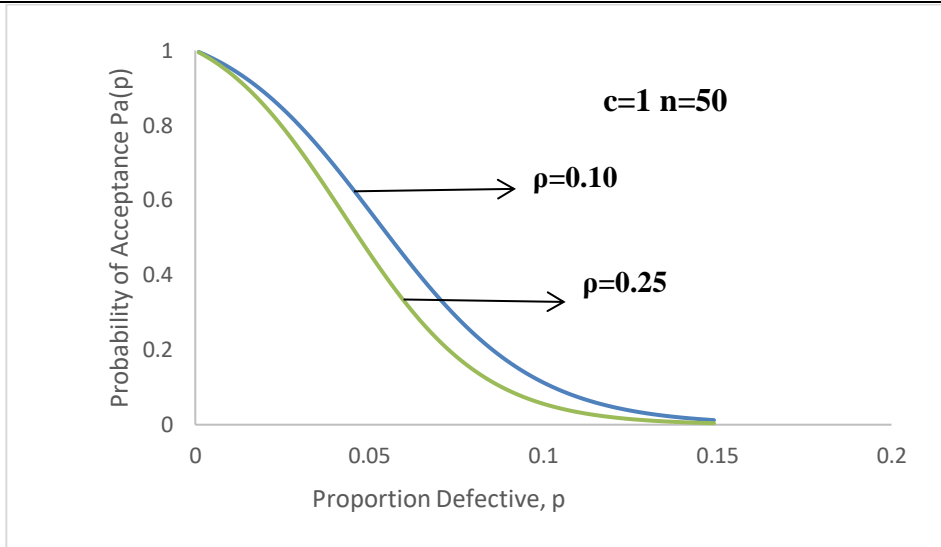
From figure 2 for  $n$  and  $q$  values are fixed and by changing the  $c$  values. It concludes that then the probability of acceptance ( $P_a(p)$ ) is increased although the acceptance number  $c$  also

increases. Figure 3 design the OC curve for fixed  $c$  and  $q$  and by changing then  $n$  (sample size) values. From this figure denotes then the sample size ( $n$ ) increases consumers are safeguarded. However, the sample size is smaller the producers are safeguarded.

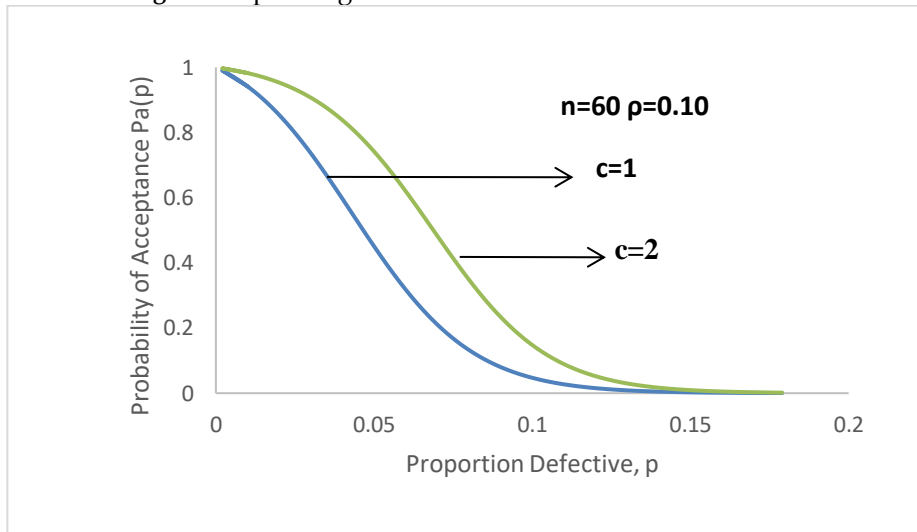
Figure 4 specifies the OC curve for proposed sampling plan. It is observed that  $p$  (proportion defective) increases  $P_a(p)$  (Probability of Acceptance) decreases. OC curves specify the manufacturing to the producer's and consumer's quality level. The Characteristic curve is used for the purpose of Probability of Acceptance depends upon the fraction defective and also conclude that the producer risk ( $\alpha$ ) and consumer risk ( $\beta$ ). It concludes that sample size increase, producer risk is maximized and consumer risk minimized. It is the inequity of Sampling Plan between Good lots and Bad Lots.

**Table 1:** Optimal parameters of SkSP-T plan with single sampling plan as Reference plan under the condition of Intervened Poisson Distribution (IPD)

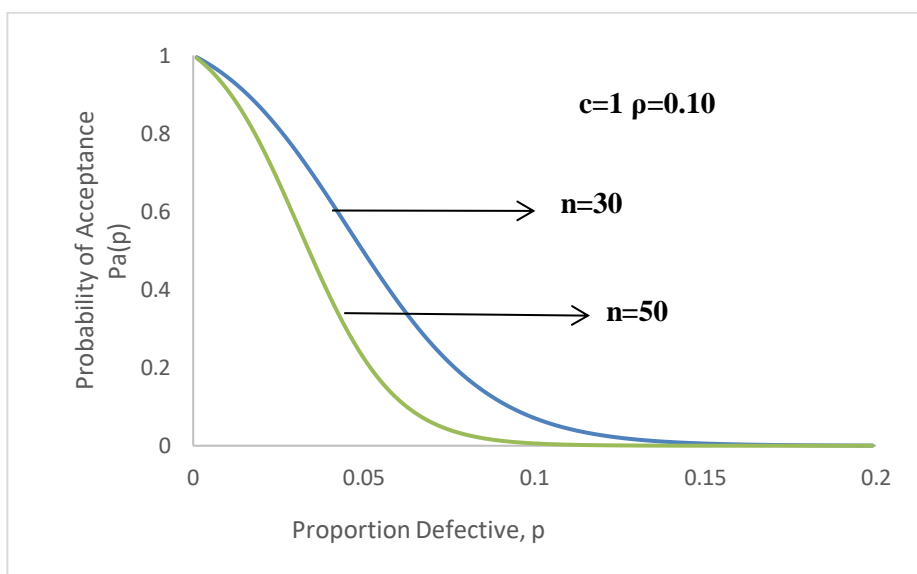
q	i	f <sub>1</sub>	f <sub>2</sub>	f <sub>3</sub>	c	Probability of Acceptance P <sub>a</sub> (p)							Operating Ratio (OR)		
						0.99	0.95	0.90	0.50	0.10	0.05	0.01	$\alpha=0.01$ $\beta=0.01$	$\alpha=0.05$ $\beta=0.05$	$\alpha=0.05$ $\beta=0.10$
0.01	1	1/2	1/4	1/8	1	0.148	0.624	1.053	3.140	5.711	6.577	8.386	56.662	10.540	9.152
		1/3	1/6	1/12		0.215	0.833	1.356	3.605	6.164	7.022	8.810	40.977	8.4298	7.400
		1/5	1/10	1/20		0.340	1.181	1.804	4.203	6.759	7.593	9.427	27.726	6.4300	5.723
		1/7	1/14	1/28		0.445	1.450	2.127	4.594	7.140	7.985	9.778	21.973	5.5069	4.924
		1/9	1/18	1/36		0.552	1.659	2.387	4.891	7.428	8.251	10.07	18.243	4.9735	4.477
0.05	2	1/2	1/4	1/8	2	0.070	0.297	0.510	1.610	3.014	3.485	4.432	63.314	11.734	10.15
		1/3	1/6	1/12		0.102	0.401	0.668	1.863	3.265	3.734	4.737	46.441	9.3117	8.142
		1/5	1/10	1/20		0.161	0.579	0.898	2.190	3.582	4.034	5.011	31.124	6.9672	6.187
		1/7	1/14	1/28		0.211	0.711	1.066	2.404	3.788	4.241	5.192	24.607	5.9648	5.328
		1/9	1/18	1/36		0.262	0.821	1.205	2.567	3.943	4.383	5.346	20.405	5.3386	4.803
0.10	3	1/2	1/4	1/8	3	0.042	0.182	0.317	1.030	1.972	2.289	2.985	71.071	12.577	10.84
		1/3	1/6	1/12		0.061	0.248	0.415	1.196	2.140	2.456	3.119	51.131	9.9032	8.629
		1/5	1/10	1/20		0.098	0.357	0.563	1.415	2.360	2.666	3.334	34.02	7.4678	6.611
		1/7	1/14	1/28		0.128	0.444	0.672	1.559	2.500	2.810	3.459	27.023	6.3288	5.631
		1/9	1/18	1/36		0.161	0.514	0.762	1.669	2.606	2.907	3.564	22.137	5.6556	5.070
0.15	4	1/2	1/4	1/8	4	0.029	0.127	0.221	0.729	1.418	1.653	2.159	74.448	13.016	11.17
		1/3	1/6	1/12		0.042	0.172	0.289	0.850	1.544	1.779	2.270	54.048	10.343	8.977
		1/5	1/10	1/20		0.068	0.249	0.395	1.009	1.707	1.936	2.436	35.824	7.7751	6.855
		1/7	1/14	1/28		0.089	0.311	0.473	1.116	1.812	2.043	2.531	28.438	6.5691	5.826
		1/9	1/18	1/36		0.112	0.360	0.537	1.195	1.891	2.116	2.610	23.304	5.8778	5.253
0.20	5	1/2	1/4	1/8	5	0.022	0.093	0.164	0.549	1.077	2.261	1.688	76.727	24.312	11.58
		1/3	1/6	1/12		0.028	0.128	0.216	0.641	1.176	1.360	1.753	62.607	10.625	9.188
		1/5	1/10	1/20		0.049	0.186	0.296	0.763	1.304	1.483	1.876	38.286	7.9731	7.011
		1/7	1/14	1/28		0.064	0.232	0.354	0.845	1.386	1.567	1.950	30.469	6.7543	5.974
		1/9	1/18	1/36		0.084	0.269	0.404	0.907	1.447	1.625	2.002	23.833	6.0409	5.379



**Figure 1:** Operating Characteristic Curve for Fixed  $c$  and  $n$



**Figure 2:** Operating Characteristic Curve for Fixed  $n$  and  $\rho$



**Figure 3:** Operating Characteristic Curve for Fixed  $c$  and  $\rho$

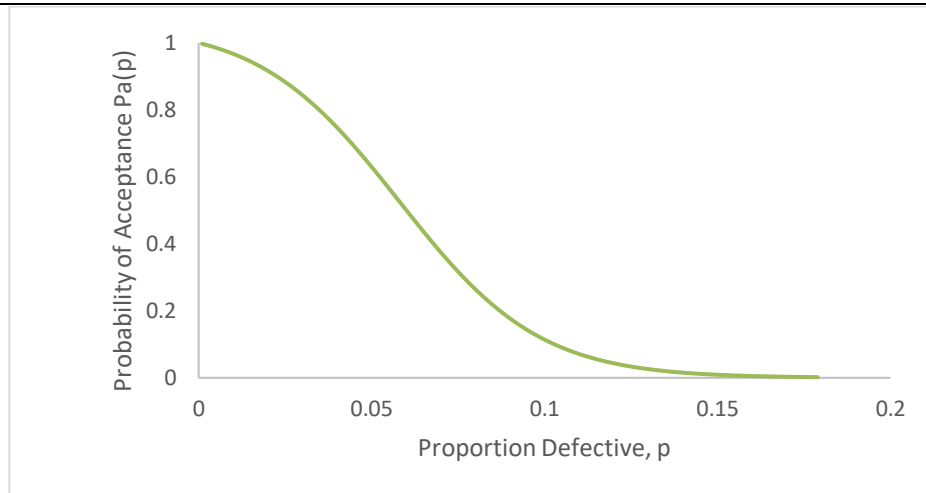


Figure 4: Operating Characteristic Curve for SkSP-T With SSP as reference plan under the condition of IPD

## VI. Conclusion

The new proposed skip-lot sampling plan of type SkSP-T with single sampling plan as reference plan under the condition of intervened poisson distribution is apply during the production process to improve the quality of products to produce. The comparison results have specified that the SkSP-T with IPD is more efficient than the conventional sampling plans. The necessary tables and examples are contributed and applied for the formulation of the new proposed sampling plan.

## References

- [1] Balamurali, S., (2002). "Modified Tightened Three level Continuous sampling plan", *Economic Quality Control*, Vol. 17, p. 221-234.
- [2] Balamurali, S., and Chi-Hyuck Jun, (2004). "Modified CSP-T sampling procedures for continuous production process", *Quality Technology and Quantitative Management* Vol. 1, No. 2, pp. 175-188.
- [3] Dhanavanthan, P. (1998). Compound intervened Poisson distribution, *Biometrical Journal* 40: pp. 941-946.
- [4] Dhanavanthan, P. (2000). Estimation of the parameters of compound intervened Poisson distribution, *Biometrical Journal* 42: pp. 315-320.
- [5] Dodge H.F, (1955). "Skip-Lot sampling plan", *Industrial Quality Control*, 11(5), pp. 3-5.
- [6] Fordice, J.J. (1972). "A Tightened Multi-Level Continuous Sampling Plan CSP-T", Report No.QEM 21230-10, Ammunition Procurement and Supply Agency, Joliet, Illinois.
- [7] G. J. Lieberman and H. Solomon (1954). "Multi-level continuous sampling plans," Technical Report No. 17, Applied Mathematics and Statistics Laboratory, Stanford University.
- [8] H. F. Dodge and M. N. Torrey (1951). "Additional continuous sampling inspection plans," *Industrial Quality Control*, Vol. 7 (1951), pp. 7-12.
- [9] Huang, M., Fung, K.Y. (1989). Intervened truncated Poisson distribution, *Sankhya* 51:pp.302-310.
- [10] Jayakumar. K and Rehana. C.J. (2022). "Exponential Intervened Poisson Distribution: Characterizations, Different Methods of Estimation and Applications", *Journal of the Indian Society for Probability and Statistics*, Vol. 24, pp 75-92.
- [11] Kandasamy, C. and Govindaraju, K., (1993). "Selection of CSP-T plans", *Communication in Statistics - Simulation and Computation*, Vol. 22, No.1, pp. 265-283.
- [12] Muhammed Rasheed Irshad et.al (2023). "A Novel Flexible Class of Intervened Poisson Distribution by Lagrangian Approach", *Stats*, Vol.6, pp.150-168.

- [13] Perry. R. L, (1973). Skip lot Sampling Plans, *Journal of Quality Technology*, 5 (3), pp. 123-130.
- [14] Pradeepa Veerakumari. K and Azarudheen. S (2016). "Evaluation of Single Sampling plan under the conditions of Intervened Poisson Distribution", *Communications in Statistics-Simulation and Computation* , pp.1-9
- [15] Pradeepa Veerakumari. K and Suganya. S (2016). "A New System of SkSP-T with Single Sampling Plan as Reference Plan", *Research Journal of Mathematics and Statistics*, Vol.4 (4), 1-6.
- [16] R. Vijayaraghavan (2000). "Design and evaluation of skip-lot sampling plans of type SkSP-3", *Journal of Applied Statistics*, Vol. 27, No. 7, p. 901-908.
- [17] Satheesh Kumar, C., Shibu, D.S. (2011). Modified Intervened Poisson Distribution, *Statistica* 4: pp.489-498.
- [18] Shanmugam, R. (1985). An intervened Poisson distribution and its medical application, *Biometrics* 41:pp. 1025-1029.
- [19] Shanmugam, R. (1992). An inferential procedure for the Poisson intervention parameter, *Biometrics* 48:pp. 559-565.
- [20] Scollnik, D.P.M. (2006). On the intervened generalized Poisson distribution, *Communication in Statistics-Theory & Methods* 35: pp. 953-963.
- [21] Scollnik, D.P.M. (1995). Bayesian analysis of an intervened Poisson distribution, *Communications in Statistics-Theory & Methods* 24: pp. 735-754.
- [22] Suganya. S and Pradeepa Veerakumari. K (2022). "Skip-lot Sampling Plan of Type SkSP-T with Group Acceptance Sampling Plan as Reference Plan Under Burr-Type XII Distribution", *Reliability: Theory & Applications*, Vol. 17, Issue 1 (67), pp.240-251.
- [23] Suganya. S and Pradeepa Veerakumari. K (2022). "Selection of Skip-lot Sampling plan of Type SkSP-T Using Special Type Double Sampling Plan as Reference Plan Based On Fuzzy Logic Techniques Using R programming Language", *Reliability: Theory & Applications*, Vol.17, Issue 3 (69), pp. 97-108.

# EJAZ DISTRIBUTION A NEW TWO PARAMETRIC DISTRIBUTION FOR MODELLING DATA

AIJAZ AHMAD

•

Department of Mathematics, Bhagwant University, Ajmer, India  
aijazahmad4488@gmail.com

M. A. LONE\*

•

Department of Statistics, University of Kashmir, Srinagar, India  
murtazastat@gmail.com

AAFAQ. A. RATHER

•

Symbiosis Statistical Institute Symbiosis International (Deemed University), Pune, India  
aafaq7741@gmail.com

## Abstract

*This paper introduces a novel probability distribution known as the Ejaz distribution (ED), which is characterized by two parameters. The study offers a comprehensive analysis of this distribution, including an examination of key properties such as moments, moment-generating functions, order statistics, and reliability functions. Additionally, the paper explores the graphical representation of essential functions like the probability density function, cumulative distribution function, and hazard rate function, enhancing our visual understanding of their behavior. The distribution's parameters are estimated using the widely accepted method of maximum likelihood estimation. Through real-world examples, the paper highlights the practical applicability of the Ejaz distribution, demonstrating its performance and relevance in diverse scenarios.*

**Keywords:** Moments, Reliability analysis, order statistics, maximum likelihood estimation, Data analysis.

## 1. INTRODUCTION

In numerous fields such as economics, engineering, finance, insurance, demography, biology, and environmental and medical sciences, various statistical distributions have been widely utilized to describe and predict observed phenomena. However, the data encountered in these disciplines often exhibit complex behaviors and diverse shapes, characterized by varying degrees of skewness and kurtosis. Consequently, many of the conventional standard distributions have limitations when it comes to accurately representing these data. As a result, the application of these classical distributions may not yield satisfactory fits. Hence, numerous researchers have endeavored to enhance these established classical distributions to achieve greater adaptability in modeling data from a wide array of academic domains. In recent times, researchers have been actively engaged in the development of new families of continuous probability distributions known for their remarkable flexibility. This innovation involves the incorporation of extra parameters into the

foundational distributions. These novel families of lifetime distributions have gained prominence, particularly in fields like economics, engineering, finance, insurance, demography, biology, and environmental and medical sciences, where data frequently exhibit intricate behaviors, diverse shapes, skewness, and kurtosis variations. The integration of additional parameters empowers these distributions to offer a more adaptable and versatile framework for modeling complex data. By doing so, they overcome the limitations of traditional standard distributions, enabling researchers to better capture and predict real-world phenomena with precision. Thus, these newly proposed lifetime distribution families have become invaluable tools for data analysis and modeling in a wide range of disciplines. In recent years, researchers have introduced modifications to enhance the adaptability of conventional distributions when interpreting diverse datasets. These changes aim to improve the accuracy of data analysis across different fields by tailoring distribution characteristics to specific dataset requirements. For reference Aijaz et al. [1-3], Terna Godfrey Ieren [18], Albert Luguterah [4], Topp-Leone Rayleigh distribution by Fatoki olayode [9], Amal S. Hassan et al. [5], Frank Gomes-silva et al. [10], Brito et al.[7], Morad Alizadeh et al. [15], Shanker et al. [17], Lindley [14], Flaih, A et al. [11], Akhter, Z et al. [6], G.M. Corderio et al. [13]. The formulated distribution is versatile and suitable for modeling various data types, including left-skewed, right-skewed, and symmetric datasets. This versatility is evident when examining probability density function (PDF) plots, as they demonstrate that this distribution can offer the most optimal fit for complex datasets. Whether the data exhibits a pronounced tail on the left, a tail on the right, or a balanced symmetry, this distribution's flexibility allows it to adapt and provide a robust representation. Its ability to accommodate a wide range of data patterns makes it a valuable tool for statistical modeling and analysis, ensuring accurate and meaningful insights across diverse data scenarios.

Let us suppose  $F(x; \alpha, \beta)$  be cdf of a random variable  $x$  with  $\alpha, \beta$  parameters, then the cumulative distribution function of Ejaz distribution is described as.

$$F(x; \alpha, \beta) = 1 - e^{-\alpha(e^{\beta x} - 1)} \left( 2 - e^{-\alpha(e^{\beta x} - 1)} \right) \quad ; \quad x > 0, \alpha, \beta > 0 \quad (1)$$

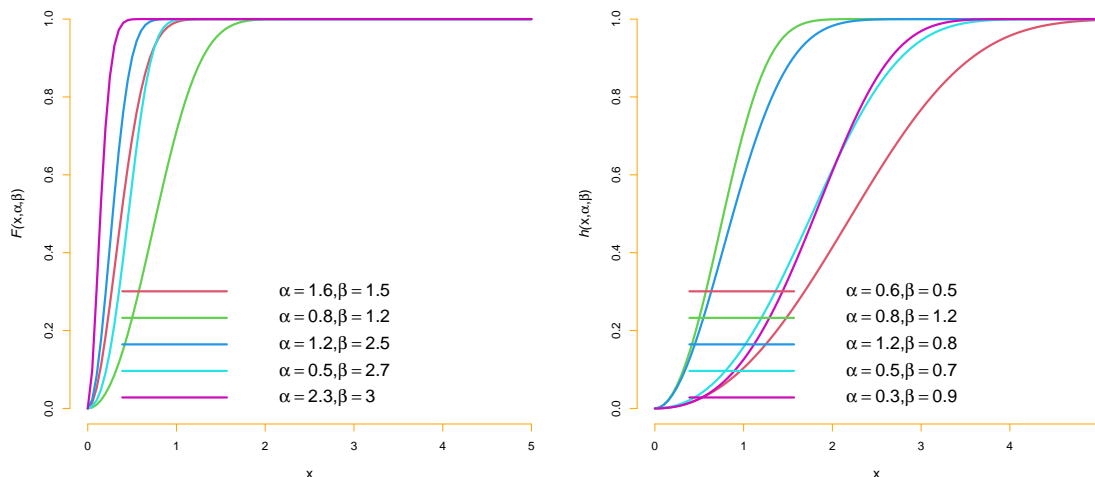


Figure 1: The cdf plots of Ejaz distribution for distinct parameter values.

The corresponding probability density function is described as

$$f(x; \alpha, \beta) = 2\alpha\beta e^{-\alpha(e^{\beta x} - 1) + \beta x} \left( 1 - e^{-\alpha(e^{\beta x} - 1)} \right) \quad ; \quad x > 0, \alpha, \beta > 0 \quad (2)$$



Here we examine the validity of pdf

$$\begin{aligned} \int_0^{\infty} f(x; \alpha, \beta) dx &= 1 \\ &= \int_0^{\infty} 2\alpha\beta e^{-\alpha(e^{\beta x}-1)+\beta x} (1 - e^{-\alpha(e^{\beta x}-1)}) dx \end{aligned}$$

On substituting  $e^{\beta x} - 1 = z$ , so that  $0 < z < \infty$  we have

$$\begin{aligned} &= 2\alpha \int_0^{\infty} e^{-\alpha z} (1 - e^{-\alpha z}) dz \\ &= 2\alpha \left( \frac{1}{2\alpha} \right) = 1 \end{aligned}$$

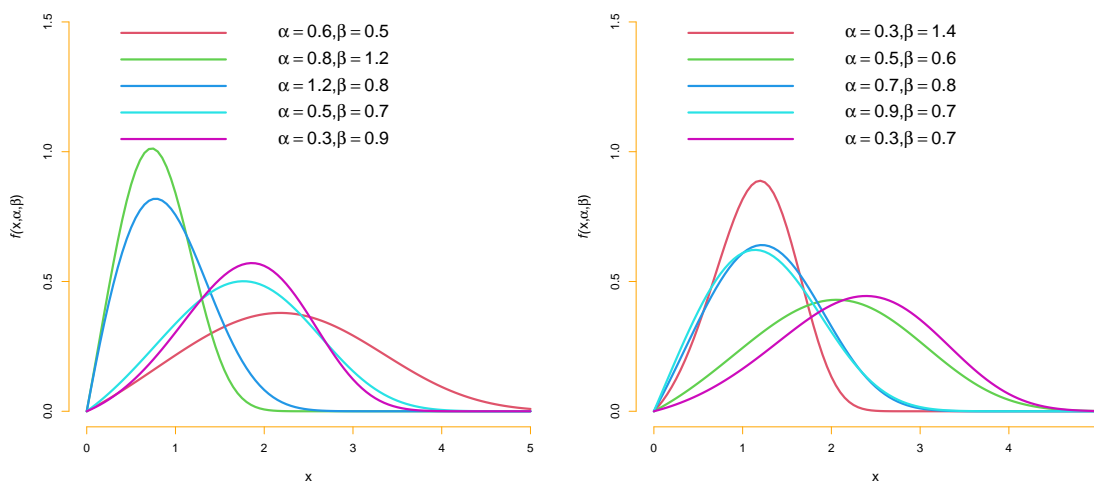


Figure 2: The pdf plots of Ejaz distribution for distinct parameter values.

## 2. MOMENTS

To understand and characterize the properties of the formulated distribution, we perform a moment analysis about the origin. This analysis allows us to derive essential statistical measures such as skewness, kurtosis, and other relevant properties. By examining these moments, we gain valuable insights into the distribution's shape, central tendency, and the presence of any outliers or heavy tails, aiding in its comprehensive statistical characterization and interpretation.

Suppose  $x$  denotes a random variable follows Ejaz distribution. Then  $k^{th}$  moment about origin denoted as  $\mu'_k$  can be obtained as

$$\begin{aligned} \mu'_k &= E(x^k) = \int_0^{\infty} x^k f(x; \alpha, \beta) dx \\ &= 2\alpha\beta \int_0^{\infty} x^k e^{-\alpha(e^{\beta x}-1)+\beta x} (1 - e^{-\alpha(e^{\beta x}-1)}) dx \end{aligned}$$

Making substitution  $e^{\beta x} = z$  so that  $1 < z < \infty$ , we have

$$\mu'_k = \frac{2\alpha}{\beta^k} \left\{ e^{\alpha} \int_1^{\infty} (\log(z))^k e^{-\alpha z} dz - e^{2\alpha} \int_1^{\infty} (\log(z))^k e^{-2\alpha z} dz \right\}$$

Applying integro-Exponential function by Milgram [16].

$$E_s^j(\lambda) = \frac{1}{j+1} \int_1^{\infty} (\log(t))^j t^{-s} e^{-\lambda t} dt$$

$$\mu'_k = \frac{2\alpha e^\alpha \Gamma(k+1)}{\beta^k} \left( E_0^k(\alpha) - e^\alpha E_0^k(2\alpha) \right)$$

Substituting  $k = 1, 2, 3, 4$  we obtain first four moments of the distribution about origin. The variance  $\sigma^2$ , skewness  $\sqrt{\beta_1}$ , kurtosis  $\beta_2$ , coefficient of variation (C.V) and index of dispersion  $\gamma$ .

Let  $x$  be a random variable follows Ejaz distribution. Then the moment generating function of the distribution denoted by  $M_X(t)$  is given by

$$\begin{aligned} M_X(t) &= E(e^{tx}) = \int_0^\infty e^{tx} f(x; \alpha, \beta) dx \\ &= \sum_{k=0}^{\infty} \frac{t^k}{k!} x^k f(x; \alpha, \beta) = \sum_{k=0}^{\infty} \frac{t^k}{k!} E(x^k) \\ &= \sum_{k=0}^{\infty} \frac{t^k}{k!} \frac{2\alpha e^\alpha \Gamma(k+1)}{\beta^k} \left( E_0^k(\alpha) - e^\alpha E_0^k(2\alpha) \right) \end{aligned}$$

### 3. RELIABILITY INDICATORS

This section is focused on researching and developing distinct ageing indicators for the formulated distribution.

#### 3.1. Survival function

Let us suppose  $x$  be a continuous random variable with cdf  $F(x)$ . Then its Survival function which is also known as reliability function is stated as

$$S(x) = p_r(X > x) = \int_x^\infty f(x) dx = 1 - F(x)$$

Therefore, the survival function for Ejaz distribution is given by

$$\begin{aligned} S(x; \alpha, \beta) &= 1 - F(x; \alpha, \beta) \\ &= e^{-\alpha(e^{\beta x} - 1)} \left( 2 - e^{-\alpha(e^{\beta x} - 1)} \right) \end{aligned} \tag{3}$$

#### 3.2. Hazard rate function

The hazard rate function of a random variable  $x$  is denoted as

$$h(x; \alpha, \beta) = \frac{f(x; \alpha, \beta)}{S(x; \alpha, \beta)} \tag{4}$$

using equation (1) and (3) in equation (4), then the hazard rate function of Ejaz distribution is given as

$$h(x; \alpha, \beta) = \frac{2\alpha\beta e^{\beta x} \left( 1 - e^{-\alpha(e^{\beta x} - 1)} \right)}{\left( 2 - e^{-\alpha(e^{\beta x} - 1)} \right)}$$

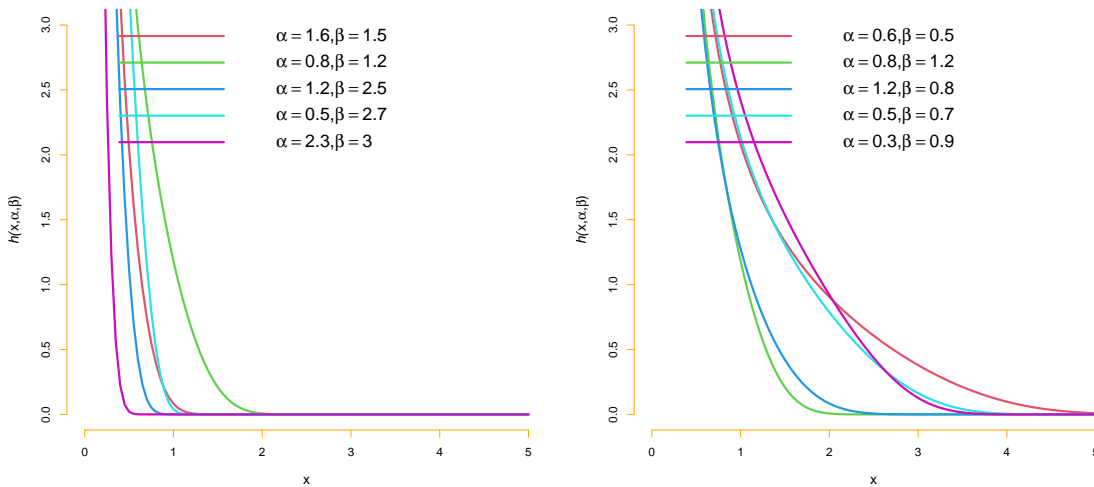


Figure 3: The hrf plots of Ejaz distribution for distinct parameter values.

### 3.3. Cumulative hazard rate function

The cumulative hazard rate function of a random variable  $x$  is given as

$$H(x, \alpha, \beta) = -\ln[\bar{F}(x; \alpha, \beta)] \quad (5)$$

using equation (1) in equation (5), then we obtain cumulative hazard rate function of Ejaz distribution as

$$H(x; \alpha, \beta) = \alpha \left( e^{\beta x} - 1 \right) - \log \left( 2 - e^{-\alpha(e^{\beta x} - 1)} \right)$$

### 3.4. Reverse Hazard rate function

The reverse hazard rate function of random variable  $x$  is described as

$$r(x; \alpha, \beta) = \frac{f(x; \alpha, \beta)}{F(x; \alpha, \beta)} \quad (6)$$

using equation (1) and (2) in equation (6), then the reverse hazard rate function of Ejaz distribution is given as

$$r(x; \alpha, \beta) = \frac{2\alpha\beta e^{-\alpha(e^{\beta x} - 1) + \beta x} \left( 1 - e^{-\alpha(e^{\beta x} - 1)} \right)}{1 - e^{-\alpha(e^{\beta x} - 1)} \left( 2 - e^{-\alpha(e^{\beta x} - 1)} \right)}$$

## 4. ORDER STATISTICS

Let us suppose  $x_1, x_2, \dots, x_n$  be random samples of size  $n$  from Ejaz distribution with pdf  $f(x)$  and cdf  $F(x)$ . Then the probability density function of the  $k^{th}$  order statistics is given as

$$f_x(k) = \frac{n!}{(k-1)!(n-1)!} f(x) [F(x)]^{k-1} [1 - F(x)]^{n-k} \quad (7)$$

Using equation (1) and (2) in equation (7), we have

$$f_x(k) = \frac{n!}{(k-1)!(n-1)!} 2\alpha\beta e^{-\alpha(e^{\beta x} - 1) + \beta x} \left( 1 - e^{-\alpha(e^{\beta x} - 1)} \right) \left[ 1 - e^{-\alpha(e^{\beta x} - 1)} \left( 2 - e^{-\alpha(e^{\beta x} - 1)} \right) \right]^{k-1} \\ \times \left[ e^{-\alpha(e^{\beta x} - 1)} \left( 2 - e^{-\alpha(e^{\beta x} - 1)} \right) \right]^{n-k}$$

The pdf of the first order statistics  $X_1$  of Ejaz distribution is given by

$$f_x(1) = 2n\alpha\beta e^{-\alpha(e^{\beta x}-1)+\beta x} \left(1 - e^{-\alpha(e^{\beta x}-1)}\right) \left[e^{-\alpha(e^{\beta x}-1)} \left(2 - e^{-\alpha(e^{\beta x}-1)}\right)\right]^{n-1}$$

The pdf of the first order statistics  $X_n$  of Ejaz distribution is given by

$$f_x(n) = 2n\alpha\beta e^{-\alpha(e^{\beta x}-1)+\beta x} \left(1 - e^{-\alpha(e^{\beta x}-1)}\right) \left[1 - e^{-\alpha(e^{\beta x}-1)} \left(2 - e^{-\alpha(e^{\beta x}-1)}\right)\right]^{n-1}$$

## 5. MAXIMUM LIKELIHOOD ESTIMATION

Let the random samples  $x_1, x_2, x_3, \dots, x_n$  are drawn from Ejaz distribution. The likelihood function of  $n$  observations is given as

$$L = \prod_{i=1}^n \left(2\alpha\beta e^{-\alpha(e^{\beta x_i}-1)+\beta x_i} \left(1 - e^{-\alpha(e^{\beta x_i}-1)}\right)\right)$$

The log-likelihood function is given as

$$l = n\log(2) + n\log(\alpha) + n\log(\beta) - \alpha \left(e^{\beta x} - 1\right) + \beta x + \sum_{i=1}^n \log \left(1 - e^{-\alpha(e^{\beta x_i}-1)}\right) \quad (8)$$

The partial derivatives of the log-likelihood function with respect to  $\alpha$  and  $\beta$  are given as

$$\frac{\partial l}{\partial \alpha} = \frac{1}{n} - e^{\beta x_i} + 1 + \sum_{i=1}^n \frac{(e^{\beta x_i} - 1) e^{-\alpha(e^{\beta x_i}-1)}}{1 - e^{-\alpha(e^{\beta x_i}-1)}} \quad (9)$$

$$\frac{\partial l}{\partial \beta} = \frac{n}{\beta} - \alpha x_i e^{\beta x_i} + x_i - \alpha \sum_{i=1}^n \frac{x_i e^{\beta x_i} e^{-\alpha(e^{\beta x_i}-1)}}{1 - e^{-\alpha(e^{\beta x_i}-1)}} \quad (10)$$

For interval estimation and hypothesis tests on the model parameters, an information matrix is required. The 2 by 2 observed matrix is

$$I(\psi) = \frac{-1}{n} \begin{bmatrix} E \left( \frac{\partial^2 \log l}{\partial \alpha^2} \right) & E \left( \frac{\partial^2 \log l}{\partial \alpha \partial \beta} \right) \\ E \left( \frac{\partial^2 \log l}{\partial \beta \partial \alpha} \right) & E \left( \frac{\partial^2 \log l}{\partial \beta^2} \right) \end{bmatrix}$$

The elements of above information matrix can be obtain by differentiating equations (9) and (10) again partially. Under standard regularity conditions when  $n \rightarrow \infty$  the distribution of  $\hat{\psi}$  can be approximated by a multivariate normal  $N(0, I(\hat{\psi})^{-1})$  distribution to construct approximate confidence interval for the parameters. Hence the approximate  $100(1 - \zeta)\%$  confidence interval for  $\alpha$  and  $\beta$  are respectively given by

$$\hat{\alpha} \pm Z_{\frac{\zeta}{2}} \sqrt{I_{\alpha\alpha}^{-1}(\hat{\psi})} \text{ and } \hat{\beta} \pm Z_{\frac{\zeta}{2}} \sqrt{I_{\beta\beta}^{-1}(\hat{\psi})}$$

## 6. SIMULATION ANALYSIS

The bias, variance and MSE were all addressed to simulation analysis. From Ejaz distribution taking  $N=500$  with samples of size  $n=25, 50, 150, 200, 250$  and  $400$ . For various parameter combinations, simulation results have been achieved. The bias, variance and MSE values are calculated and presented in table 1 and 2. As the sample size increases, this becomes apparent that these estimates are relatively consistent and approximate the actual values of parameters. Interestingly, with all parameter combinations, the bias and MSE reduce as the sample size increases.

**Table 1:** Bias, variance and their corresponding MSE's for different parameter values  $\alpha = 1.2, \beta = 0.8$

Sample size	Parameters	Bias	Variance	MSE
25	$\alpha$	0.01130	0.00432	0.01473
	$\beta$	0.01251	0.00154	0.00165
50	$\alpha$	0.00314	0.00413	0.00514
	$\beta$	0.00103	0.00071	0.00061
150	$\alpha$	-0.00021	0.00301	0.00201
	$\beta$	0.00406	0.00049	0.00051
200	$\alpha$	-0.00201	0.00156	0.00205
	$\beta$	0.00237	0.00027	0.00028
250	$\alpha$	0.00120	0.00206	0.00203
	$\beta$	0.00255	0.00025	0.00022
300	$\alpha$	0.00177	0.00203	0.00201
	$\beta$	0.00066	0.00021	0.00020

**Table 2:** Bias, variance and their corresponding MSE's for different parameter values  $\alpha = 2.2, \beta = 1.5$

Sample size	Parameters	Bias	Variance	MSE
25	$\alpha$	0.01230	0.03553	0.03573
	$\beta$	0.02003	0.01832	0.01031
50	$\alpha$	0.01214	0.01105	0.01132
	$\beta$	0.01121	0.00506	0.00420
150	$\alpha$	0.00672	0.00668	0.00607
	$\beta$	0.00146	0.00224	0.00216
200	$\alpha$	0.00265	0.00416	0.00506
	$\beta$	0.01076	0.00232	0.00214
250	$\alpha$	0.0027	0.00360	0.00361
	$\beta$	0.00208	0.00145	0.00145
300	$\alpha$	0.00150	0.00301	0.00211
	$\beta$	0.00063	0.00130	0.00130

## 7. DATA ANALYSIS

This subsection evaluates a real-world data sets to demonstrate the Ejaz distribution's applicability and effectiveness. The Ejaz distribution (ED) adaptability is determined by comparing its efficacy to the following conventional distributions.

1:- Weibull distribution having pdf

$$f(x; \alpha, \beta) = \alpha \beta x^{\beta-1} e^{-\alpha x^\beta}; \quad x > 0, \alpha, \beta > 0$$

2:- Fréchet distribution having pdf

$$f(x; \alpha, \beta) = \alpha \beta x^{-\beta-1} e^{-\alpha x^{-\beta}}; \quad x > 0, \alpha, \beta > 0$$

3:- Inverse Burr distribution having pdf

$$f(x; \alpha, \beta) = \alpha \beta (1 - x^{-\alpha})^{-\beta-1}; \quad x > 0, \alpha, \beta > 0$$

4:- Lomax distribution having pdf

$$f(x; \alpha, \beta) = \alpha \beta (1 + \alpha x)^{-\beta-1}; \quad x > 0, \alpha, \beta > 0$$

5:- Exponentiated Rayleigh distribution having pdf

$$f(x; \alpha, \beta) = 2\alpha\beta x e^{-\alpha x^2} (1 - e^{-\alpha x^2})^{\beta-1}; \quad x > 0, \alpha, \beta > 0$$

6:- Lindley distribution having pdf

$$f(x; \alpha) = \frac{\alpha^2}{(1 + \alpha)} (1 + x) e^{-\alpha x}; \quad x > 0, \alpha > 0$$

7:- Inverse Rayleigh distribution having pdf

$$f(x; \alpha, \beta) = \frac{2\alpha}{x^3} e^{-\alpha x^{-2}}; \quad x > 0, \alpha > 0$$

To compare the versatility of the explored distribution, we consider the criteria like AIC (Akaike information criterion), CAIC (Consistent Akaike information criterion), BIC (Bayesian information criterion) and HQIC (Hannan-Quinn information criterion). Distribution having lesser AIC, CAIC, BIC and HQIC values is considered better.

$$AIC = -2l + 2p, \quad AICC = -2l + 2pm / (m - p - 1), \quad BIC = -2l + p(\log(m))$$

$$HQIC = -2l + 2p \log(\log(m)), \quad K.S = \max_{1 \leq j \leq m} \left( F(x_j) - \frac{j-1}{m}, \frac{j}{m} - F(x_j) \right)$$

Where 'l' denotes the log-likelihood function, 'p' is the number of parameters and 'm' is the sample size.

**Data set 1:** The following observations are due to Caramanis et al and Mazmumdar and Gaver [12], where they compare the two distinct algorithms called SC16 and P3 for estimating unit capacity factors. The values resulted from the algorithm SC16 are 2.01, 6.32, 3.52, 2.15, 5.42, 2.04, 2.77, 2.26, 1.95, 1.00, 2.45, 0.74, 0.98, 1.27, 2.77, 3.68, 1.18, 1.09, 1.60, 0.57, 3.33, 0.91, 7.14, 2.08, 3.85, 1.99, 7.76, 2.52, 1.57, 4.67, 4.22, 1.92, 1.59, 4.08, 2.02, 0.84, 6.85, 2.18, 2.04, 1.05, 2.91, 1.37, 2.43, 2.28, 3.74, 1.30, 1.59, 1.83, 3.85, 6.30, 4.83, 0.50, 3.40, 2.33, 4.25, 3.49, 2.12, 0.83, 0.54, 3.23, 4.50, 0.71, 0.48, 2.30, 7.73.

**Data set 2:** The following observations are due to Caramanis et al and Mazmumdar and Gaver [12], where they compare the two distinct algorithms called SC16 and P3 for estimating unit capacity factors. The values resulted from the algorithm SC16 are 0.1, 0.33, 0.44, 0.56, 0.59, 0.59, 0.72, 0.74, 0.92, 0.93, 0.96, 1, 1, 1.02, 1.05, 1.07, 1.07, 1.08, 1.08, 1.08, 1.09, 1.12, 1.13, 1.15, 1.16, 1.2, 1.21, 1.22, 1.22, 1.24, 1.3, 1.34, 1.36, 1.39, 1.44, 1.46, 1.53, 1.59, 1.6, 1.63, 1.68, 1.71, 1.72, 1.76, 1.83, 1.95, 1.96, 1.97, 2.02, 2.13, 2.15, 2.16, 2.22, 2.3, 2.31, 2.4, 2.45, 2.51, 2.53, 2.54, 2.78, 2.93, 3.27, 3.42, 3.47, 3.61, 4.02, 4.32, 4.58, 5.55, 2.54, 0.77.

The ML estimates with corresponding standard errors in parenthesis of the unknown parameters are presented in Table 3 and Table 5. Also the comparison statistics, AIC, BIC, CAIC, HQIC and the goodness-of-fit statistic for the data sets are displayed in Table 4 and Table 6.

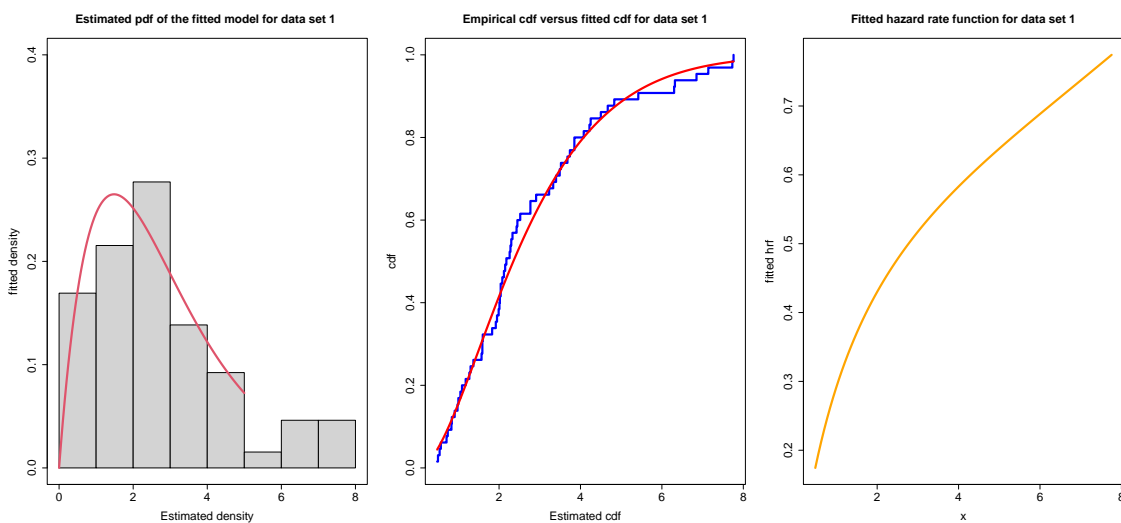
It is observed from the findings that ED provides best fit than other competitive models based on the measures of statistics, AIC, BIC, AICC, HQIC and K-S statistic. Along with p-values of each model.

**Table 3:** The ML Estimates (standard error in parenthesis) for data set 1

Model	$\hat{\alpha}$	$\hat{\beta}$
ED	3.45342 (1.92236)	0.19311 (0.08456)
WD	0.16787 (0.04105)	1.59666 (0.15017)
FD	1.82550 (0.22717)	1.42975 (0.12938)
IBD	1.79634 (0.15843)	2.85966 (0.36236)
LXD	0.00769 (0.00464)	48.2182 (29.232)
ERD	0.07499 (0.01347)	0.73015 (0.11406)
LD	0.59651 (0.05424)	...
IRD	1.78914 (0.22191)	...

**Table 4:** Comparison criterion and goodness-of-fit statistics for data set 1

Model	-2l	AIC	AICC	BIC	HQIC	K.S statistic	p-value
<b>ED</b>	<b>239.11</b>	<b>243.11</b>	<b>243.30</b>	<b>247.45</b>	<b>244.82</b>	<b>0.07732</b>	<b>0.8319</b>
WD	240.85	244.85	245.04	249.20	246.56	0.0955	0.5927
FD	250.87	254.87	255.07	259.22	256.59	0.1491	0.1111
IBD	245.36	249.36	249.56	253.71	251.08	0.12373	0.2726
LXD	130.57	265.14	265.33	269.49	266.86	1.00	2.2e-16
ERD	243.000	247.00	247.19	251.34	248.71	0.12333	0.2762
LD	249.59	251.59	251.65	253.77	252.45	0.11653	0.3406
IRD	267.49	269.49	269.56	271.67	270.35	0.27703	9.293e-05



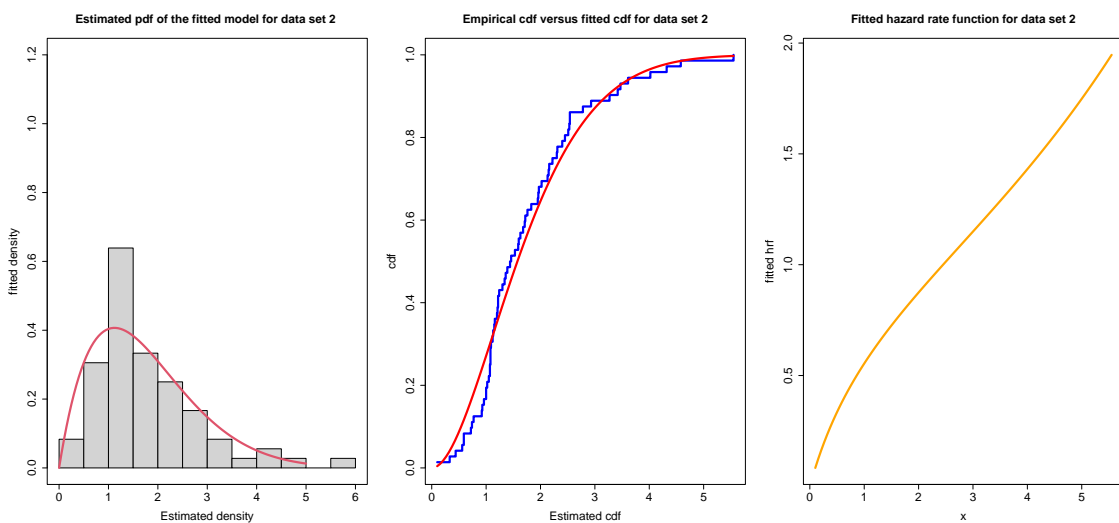
**Figure 4:** Fitted pdf, cdf and hrf for data set 1.

**Table 5:** The ML Estimates (standard error in parenthesis) for data set 2

Model	$\hat{\alpha}$	$\hat{\beta}$
ED	3.45342 (1.92236)	0.19311 (0.08456)
WD	0.29347 (0.05540)	1.796236 (0.15662)
FD	1.04750 (0.13022)	1.17538 (0.08496)
IBD	2.31897 (0.21444)	1.85769 (0.21925)
LXD	0.00841 (0.00579)	68.4375 (47.3451)
ERD	0.22565 (0.03649)	0.90717 (0.14049)
LD	0.87441 (0.07718)	.... ...
IRD	0.45560 (0.05369)	... ...

**Table 6:** Comparison criterion and goodness-of-fit statistics for data set 2

Model	-2l	AIC	AICC	BIC	HQIC	K.S statistic	p-value
<b>ED</b>	<b>191.80</b>	<b>195.80</b>	<b>195.97</b>	<b>200.35</b>	<b>197.61</b>	<b>0.06414</b>	<b>0.6053</b>
WD	192.05	196.05	196.22	200.60	197.86	0.098266	0.4902
FD	234.65	238.65	238.82	243.20	240.46	0.18994	0.01109
IBD	195.21	199.21	199.38	201.02	203.76	0.10925	0.3565
LXD	225.58	229.58	229.75	234.13	231.39	0.28959	1.139e-05
ERD	193.26	197.26	197.44	201.82	199.08	0.10202	0.4419
LD	213.05	215.05	215.10	217.32	215.95	0.2356	0.000671
IRD	323.71	325.71	325.77	327.99	326.61	0.4674	4.352e-14



**Figure 5:** Fitted pdf, cdf and hrf for data set 2.



## 8. CONCLUSION

In this research paper, we introduce a novel two-parameter lifetime distribution, named the "Ejaz distribution." We delve into various mathematical properties associated with this distribution, including its shape, moments, hazard rate, and order statistics. Furthermore, we discuss the utilization of the maximum likelihood estimation method for estimating the distribution's parameters. To illustrate the practical effectiveness and superiority of the Ejaz distribution in comparison to existing alternatives such as the Weibull, Fréchet, Inverse Burr, Lomax, Exponentiated Rayleigh, Lindley, and inverse Rayleigh distributions, we conduct goodness-of-fit tests employing criteria such as the Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC), and Bayesian Information Criterion (BIC) on real-life lifetime datasets. Additionally, we perform a simulation analysis, which reveals an intriguing trend: as the sample size increases, there is a reduction in bias and mean squared error (MSE) across all parameter combinations.

## REFERENCES

- [1] Aijaz.A, Muzamil. J, Syed Quratul Ain. and Rajnee.T. The Hamza distribution with statistical properties and applications. *Asian Journal of Probability and Statistics*, 8(1),(2020), 28-42.
- [2] Aijaz.A, Muzamil. J, Syed Quratul Ain. and Rajnee.T. The Burhan distribution with statistical properties and applications in distinct areas of science. *Earthline journal of mathematical science*, 7(2),(2021), 429-445.
- [3] Aijaz .A, Afaq Ahmad and R. Tripathi, Sauleh distribution with statistical properties and applications. *Annal Biostat. and Biomed Appli.*, 4(1) 2020, 1-5
- [4] Albert., L. Odd generalizd exponential Rayleigh distribution. *Advances and applications in statistics*, vol 48(1)(2016), 33-48.
- [5] Amal H and Said G. N . The inverse Weibull-G familiy. *Journal of data science*, (2018), 723-742.
- [6] Akhter, Z.; Almetwally, E.M.; Chesneau, C. On the Generalized Bilal Distribution: Some Properties and Estimation under Ranked Set Sampling. *Axioms.*, 8 (2022), 11-173.
- [7] Brito E, Cordeiro G. M, Yousuf H. M, Alizadeh M, Silva G.O . The Topp-Leone odd log-logistic family of distributions. *J stat comput simul*,87(15),(2017), 3040-3058.
- [8] Fatoki. O. The Topp-Leone Rayleigh distribution with application. *American journal of mathematics and statistics*, 9(6),(2019), 215-220.
- [9] Frank G. S, Ana P, Edleide D.B. The odd Lindley-G family of distributions. *Austrian journal of statistics*,10,(2016), 1-20.
- [10] Flaih, A., Elsalloukh, H., Mendi, E., Milanova, M. The exponentiated inverted Weibull distribution. *Appl. Math. Inf. Sci*, 6 (2),2012, 167-171.
- [11] M. Caramanis, J. Stremel, W. Fleck and S. Daneil. Probabilistic production costing: an investigation of alternative algorithms. *Internation journal of electrical power and energy system*,5(2),(1983), 75-86.
- [12] G.M. Corderio and M. de Castro. A new family of generalized distribution. *Journal of statistical computation and simulation*,81(7) (2011),75-86.
- [13] Lindley, D, V. Fiducial distributions and Bayes™ theorem. *Journal of Royal Statistical Society, Series B*, 20,1958, 102-107.
- [14] Morad A, Gauss M. C. The Gompertz-G family of distributions. *Journal of statistical theory and practice*, 11(1),(2017), 179-207.
- [15] Milgram, M. S. "The generalized integro-exponential function". *Mathematics of Computation*,44 (170),1985: 443-458.
- [16] Shanker, R. Akash distribution and Its Application. *International Journal of Probability and Statistics*, 4 (3),2015, 65-75.
- [17] Terna G .I, Sauta A, Issa A.A. Odd Lindley- Rayleigh distribution its properties and applications to simulated and real life datasets. *Journal of advances in mathematics and computer science*, 35(1),2020,63-88.

# RELIABILITY INVESTIGATION OF THE SPIRULINA PRODUCTION PLANT USING GUMBEL-HOUGAARD FAMILY COPULA

Priya Chaudhary<sup>1</sup>, Shikha Bansal<sup>2\*</sup>

•

<sup>1</sup>Research Scholar, Department of Mathematics, SRMIST Delhi NCR campus Modi Nagar,  
Ghaziabad, 201204, India

<sup>1</sup>techpriya.20@gmail.com

<sup>2\*</sup>Assistant Professor, Department of Mathematics, SRMIST Delhi NCR campus Modi Nagar,  
Ghaziabad, 201204, India

<sup>2\*</sup>srbansal2008@gmail.com

\*Corresponding Author

## Abstract

*This study examines the consistency metrics used to evaluate the durability of a spirulina production plant, which consists of seven subsystems: cultivation pond, paddlewheel, filter unit, washing unit, spray dryer, ribbon blender, and packaging. By studying the spirulina firm, we can repair it by discovering future failures. We can increase spirulina production so that untimely failure can be prevented and production can be increased. There are two types of system failures: partial and total. While a full failure renders the system incapable of operating, a partial failure is thought to degrade the system. In contrast, repair rates follow two different types of distributions: an ordinary and an exponential distribution. The system in a partially failed or degraded condition is thought to be repaired using general time distribution. In contrast, fully failed systems are thought to be fixed using the Gumbel-Hougaard family copula distribution. Using the supplementary variable approach, the system is examined. A Chapman-Kolmogorov differential equation is created and solved by applying the Gumbel-Hougaard family Copula approach, employing the schematic representation of the system's state. supplementary variable approaches are applied to develop and resolve the differential equations related to transition diagrams, which are significant to this research. Reliability, availability, profitability, and MTTF are the critical performance metrics for the spirulina production plant. Moreover, sensitivity analysis is carried out for MTTF.*

**Keywords:** Laplace transformation, MATLAB tool, Sensitivity, Spirulina production plant

## I. Introduction

The fundamental idea behind reliability is failure-free operation, which refers to an item's capacity to operate as intended without a fault for a predetermined amount of time under predetermined circumstances. Every technology system in the present scientific era depends on dependability to

some extent. A high level of dependability is required for defenses, businesses, and space research projects. The designers, engineers, and manufacturers in both the public and private sectors emphasize the dependable operation of their systems or equipment. Maximizing profit frequently arises in many reliability models of practical utility. The price a repairman must pay to fix the system's failure stage determines the profit that may be made from an operational system. As a result, the primary focus of research on repairable complex systems is anticipating and calculating the costs associated with maintaining a system. In comparison to what is typically found as availability/reliability of the system, the concept of determining the cost necessary to run a procedure involves a thorough understanding of the system's behavior.

Much work has been done to increase reliability while connecting the components in parallel and series. Agarwal and Bansal [1] carried out top-of-the-line repair disciplines with an environmental impact to determine the system's dependability. Xie et al. [2] As reliability and performance analysis of networked computers with opaque bridges have received little to no attention in prior research on networked computers; this study examines the reliability and efficiency analysis of complicated series-parallel networked computers with visible bridges. Agarwal et al. [3] The efficiency of a redundant cold-standby device. Yusuf and Hussaini [4] Evaluate a system consisting of three redundant units, three different forms of failure, and general repair.[5] Using generic stochastic Wiener processes as the foundation, a unique regression estimation technique for deterioration analysis. Agarwal and Bansal [6] Evaluated the solar thermal electric generation facilities' cost study. Bansal and Tyagi [9] Production of leaf springs is modeled mathematically, and availability is examined. Arora and Kumar [7] A thermal power plant's ash management system's stochastic behavior analysis and maintenance planning was provided again using the Markov technique. The probabilistic method must be revised to address the ambiguous and uncertain failure/repair data. Thus, FM has been utilized by several academics in other fields to address such variability in the failure/repair data. Bansal [8] Preemptive-Resume Repair Discipline Availability Analysis of a Repairable Redundant System. Chaudhary and Bansal [11] Assessment of Hydroelectric Power Station Reliability Performance. Bansal et al. [10] Manufacturing Plant for Screws Performance Modeling and Availability Analysis. Chauhan and Malik [12] studied the series-parallel circuits' dependability for the given variable. Fouladirad et al. [13] By reducing the traditional premise that the extent of depreciation may expand forever, which is frequently impractical for specialized units, we build a novel, limited, modified gamma process model to describe and anticipate degrading occurrences. A set of wear measurements of the cylinder liners used in a diesel engine for maritime propulsion are features related to the suggested model's application. Godara and Bansal [14] Boolean function technique and neural network approach are used to analyze the performance of reliability factors in steam turbine generator power plants. Kabiru et al. [15] have concentrated on the sophisticated system's combined distribution, including two reliability evaluation components. Uswarman and Rushdi [16] used multimodal criterion systems for the reliability assessment of rooftop solar photovoltaic panels. Lai and Zwetsloot [17] provide an ensemble rating system for the quality of products that is data-driven and is verified by recognizing high-risk situations firms in a research study of the solar sector. The last two articles focus on repairable equipment' dependability and maintenance. Tyagi and Bansal [18] Wastewater Treatment Process Optimization Model. The apparatus fails if at least  $k$  continuous units fail. A continuous  $k$ -out-of- $n$ : F system comprises  $n$ -ordered units arranged in a line or circle. Several experts have delved deeply into the  $k$ -out-of- $n$  scheme. Maihulla et al. [19] The Gumbel-Hauggaard Family Copula examines a modest solar photovoltaic system's function and cost. Meynaoui et al. [20] Using universal examination of the distribution of the input parameters' sensitivity employed

in quantitative simulators to mimic physical activities and cope with an unintentional situation during a sodium cooling fast nuclear reactor.

Vitamins E, C, and B6 are just a few vitamins and minerals abundant in spirulina that support a robust immune system. According to research, spirulina increases the body's ability to produce white blood cells and antibodies that help your body fight against infections and infections. There are several possibilities for medicinal and therapeutic uses in addition to its significance as a food additive for supplemental human nutrition. The giant spirulina plant in the world today is Earthrise farm, which was founded in 1976 and was the first spirulina farm in North America. Earthrise has produced high-quality and secure spirulina for customers worldwide with over 40 years of expertise and a 108-acre facility. The author's goal is that the model made by the author should be able to produce maximum production without any failure. The author has prepared a model keeping in mind the benefits of Spirulina so that we can get maximum production without failure. Seven subsystems have been chosen. Subsystem two has taken three units, one on hot standby and two on cold standby, while subsystem four has taken two units, one on hot standby and the other on cold standby, and other subsystems have been single units. The copula distribution has been used to correct these states whenever the system partially failed, i.e., operating less efficiently than it should. Because a completely failed state is required for a quick repair, a general repair cannot be used in these situations. The different interests and necessary system dependability measures have been discussed. The findings were obtained using various failure and repair rate numbers. The following are the sections of the paper: an introduction, a spirulina production process, a mathematical modeling, a solution of the model, and an analytical section in which various reliability measures, such as availability, reliability, MTTF, sensitivity to MTTF, and cost analysis, have been calculated using different parameter settings. And the last interpretation of results with the help of tables and graphs.

## II. Methods

### I. Spirulina Production Process

The Spirulina production plant consists of seven subsystems, i.e., cultivation pond, paddlewheel, filter chamber, washing chamber, spray dryer, ribbon blender, and packaging.

#### (a) Cultivation Pond

Cultivation may begin by feeding water to the chamber at the necessary height. The water must have the proper pH and be alkaline by adding the necessary salts at the correct rate. After the water has a typical nutritional makeup, the chamber is ready for spirulina planting. For optimal development and harvesting, 30 grams of dry spirulina should be applied for every 10 liters of water. It is made up of one unit connected in sequence. Thus, further, this unit fails, and the system fails.

#### (b) Paddlewheel

This fan has a paddle wheel or propeller installed on a spinning shaft inside a ring, panel, or cage. The most common applications for propeller fans are light- to medium-duty ones, including ventilation systems where air may be propelled in any direction. These wheels produce oxygen so that the algae can get proper nutrition, and the sun's light can reach the bottom layer so that more and more spirulina accumulate above. It consists of three parallel units. This system's capacity would be reduced with a partial failure. Only when three units fail does a severe failure occur.

(c) Filtration Unit

The spirulina is separated in this chamber by filtering using powerful vacuums. Spirulina with specific contaminants is produced when a filter drains water by sucking it out with a vacuum. It is made up of one unit connected in series. Thus, further, this unit fails, and the system fails.

(d) Washing unit

In this chamber, a high stream of pure water is used for flushing out contaminants. moreover, spirulina cream is available. It consists of two parallel components. This system's capacity would be reduced with a partial failure. Only when two units fail may a severe failure occur.

(e) Spray Dryer

In spray drying, a solution, fluid, or emulsion comprising one or more components of the desired product is atomized into droplets by spraying. Then, the droplets are quickly evaporated into the compound by superheated steam at a specific temperature and pressure. It is made up of one unit connected in sequence. Thus, further, this unit fails, and the system fails.

(f) Ribbon Blender

Spirulina we receive in solid form is processed via a crusher into dry powder. It is made up of one unit connected in sequence. Thus, further, this unit fails, and the system fails.

(g) Packaging

Spirulina is ground into a fine powder and then utilized to manufacture tablets and capsules. Items are measured, sealed, and packed with care. It is prepared to be sold on the market for various uses. This part has yet to be considered for analysis because it hardly ever fails.

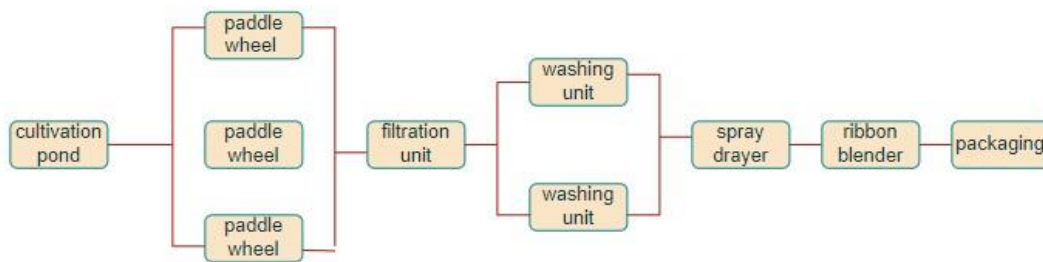


Figure 1: Flow Diagram of Spirulina Production Plant

II. State Description

$S_0$ : All subsystems are in good operating order in state  $S_0$ . The system is fully functional and in excellent condition.

$S_1$ : Due to the breakdown of subsystem one,  $S_1$  is a catastrophic failure. The system is being repaired, and the failing status is being addressed with copula repair.

$S_2$ : The initial unit of subsystem-two failed; the state  $S_2$  reflects a degraded condition with a small partial failure in subsystem-two. The system operates, the state is undergoing general repair, and total repair time is  $(x, t)$ .

$S_3$ : The first and second units of subsystem-two failed; the state  $S_3$  reflects a degraded condition with minor partial failure in subsystem-two. The system is operating, and the state is undergoing general repair. And the elapsed repair time is  $(x, t)$ .

$S_4$ : After failing every unit of subsystem two, the state  $S_4$  reflects an entire state of failure. The system

is being repaired, and the failing status is being addressed with copula repair.

$S_5$ : Due to the breakdown of subsystem three, the state  $S_5$  is a fully failed state. The system is being repaired, and the failing status is being addressed with copula repair.

$S_6$ : The initial unit of subsystem-four failed; the state  $S_6$  reflects a degraded condition with a small partial failure in subsystem-four. The system is operating, and the state is undergoing general repair. And the elapsed repair time is  $(x, t)$ .

$S_7$ : After failing both units of subsystem two, the state  $S_7$  reflects an entire failed state. The system is being repaired, and the failing status is being addressed with copula repair.

$S_8$ : The initial units of subsystems two and four have failed. when the second units of subsystems 2 and 4 are in use. When the third unit of subsystem two is on standby. The system is operating, and the state is undergoing general repair. And the elapsed repair time is  $(x, t)$ .

$S_9$ : The second units of subsystem two and the first unit of subsystem four have failed. when the third unit of subsystem two and the second unit of subsystem four are in use. When the second unit of subsystem four is on standby. The system is operating, and the state is undergoing general repair. And the elapsed repair time is  $(x, t)$ .

$S_{10}$ : Due to the breakdown of subsystem five, the state  $S_{10}$  is a fully failed state. The system is being repaired, and the failing status is being addressed with copula repair.

$S_{11}$ : Due to the breakdown of subsystem six, the state  $S_{11}$  is a fully failed state. The system is being repaired, and the failing status is being addressed with copula repair.

$S_{12}$ : Due to the breakdown of subsystem seven, the state  $S_{12}$  is a fully failed state. The system is being repaired, and the failing status is being addressed with copula repair.

### III. Assumptions

- At first, every system component is in a good functioning state.
- For operational mode, one unit from subsystem one, subsystem two, subsystem three, subsystem four, subsystem five, subsystem six, and subsystem seven is required.
- Moreover, subsystems 1, 3, 5, 6, and 7 will all be inoperative if one of their corresponding units fails.
- If three units from subsystem 2 fail, the system will not function.
- The subsystem will not function if any of its two parts fail.
- When a system component is inoperable or failed condition, it can still be repaired.
- Once a unit in a subsystem completely fails, copula (Gumbel-Haugard Family) repair is necessary.
- The failed unit can execute the function as soon as it has been repaired.
- A system healed via copula operates precisely like an entire system, and no harm is thought to occur during restoration.

### IV. Notations

t: Variable time on a time scale.

s: Laplace transforms variables for all expressions.

$\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7$ : sub system failure rates 1,2,3,4,5,6 and 7 respectively.

$\eta_1(x), \eta_2(x), \eta_3(x), \eta_4(x), \eta_5(x), \eta_6(x), \eta_7(x)$ : Subsystem repair rates 1,2,3,4,5,6 and 7 respectively.

$\Psi_1(x), \Psi_2(x), \Psi_3(x), \Psi_4(x), \Psi_5(x), \Psi_6(x), \Psi_7(x)$ : Unit in a subsystem 1,2,3,4,5,6,7 that completely failed was repaired by a copula.

$P_k(x, t)$ : The possibility that the system is  $S_k$  state for  $k=0$  to 12. The system is being repaired, and the time since the last repair is  $x, t$ .

$\bar{P}(s)$ : Laplace transform of state probability  $P(t)$ .

$E_p(t)$ : expected profit for the period  $[0, t)$ .

$Z_1, Z_2$ : respectively, revenue and operating cost per unit of time.

$S_\alpha(x)$ :  $S_\alpha(x) = \alpha(x) e^{\int_0^x -\alpha(x) dx}$  with repair distribution function  $\alpha(x)$ .

$L[S_\alpha(x)]$ :  $\int_0^\infty e^{-sx} \alpha(x) e^{\int -\alpha(x) dx} dx = \bar{S}_\alpha(s)$ , is the Laplace transform of  $S_\alpha(x)$

$L\left[\frac{1-S_\alpha(x)}{s}\right]$ :  $\int_0^\infty e^{-sx} e^{\int -\alpha(x) dx} dx = \frac{1-\bar{S}_\alpha(s)}{s}$  is the Laplace transform of  $\frac{1-S_\alpha(x)}{s}$

$\mu_0(x) = C_\theta(u_1(x), u_2(x))$ , The Gumbel-Hougaard family copula's expression for joint probability is provided as  $C_\theta(u_1(x), u_2(x)) = e^{[x^\theta + \{\log \eta(x)\}^\theta]^{1/\theta}}$ , where  $u_1 = \eta(x)$ , and  $u_2 = e^x$  where  $\theta$  as a parameter,  $1 \leq \theta \leq \infty$ .

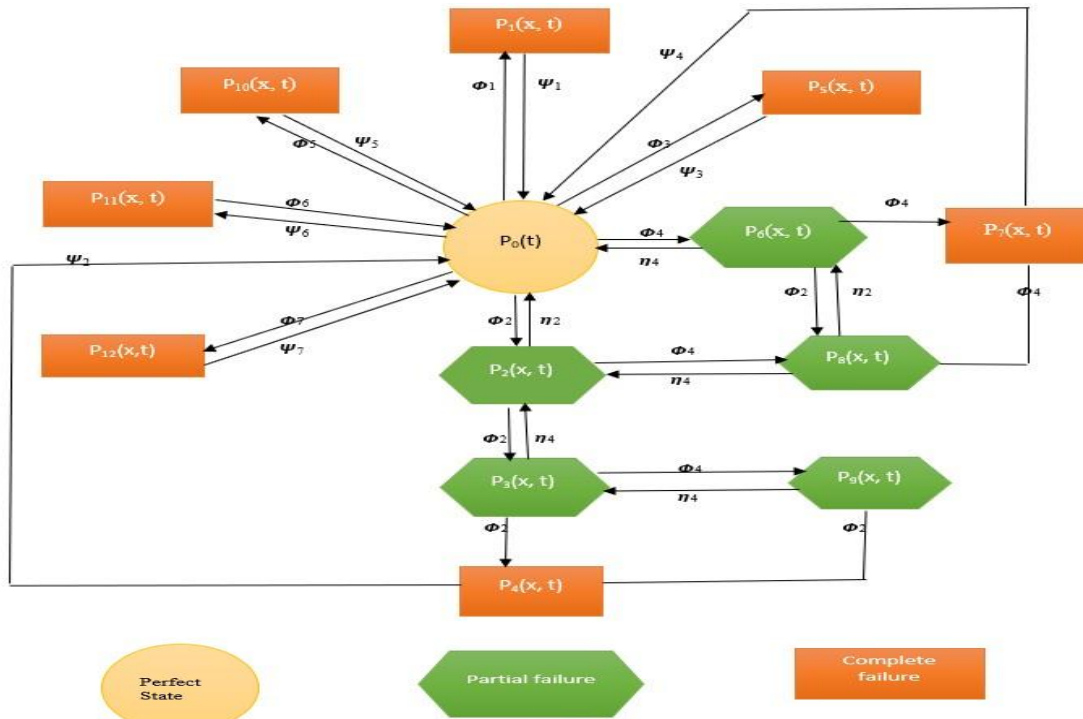


Figure2: State Transition Diagram of Spirulina production plant

## II. Formulation and Solution of model

The probability of considerations and continuity of reasoning relates the following set of difference differential equations to the mathematical model above.

$$\left[ \frac{\partial}{\partial t} + \phi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5 + \phi_6 + \phi_7 \right] P_0(t) = \int_0^\infty \Psi_1 P_1(x, t) dx + \int_0^\infty \eta_2 P_2(x, t) dx + \int_0^\infty \eta_4 P_6(x, t) dx + \int_0^\infty \Psi_2 P_4(x, t) dx + \int_0^\infty \Psi_4 P_7(x, t) dx + \int_0^\infty \Psi_3 P_5(x, t) dx + \int_0^\infty \Psi_5 P_{10}(x, t) dx + \int_0^\infty \Psi_6 P_{11}(x, t) dx + \int_0^\infty \Psi_7 P_{12}(x, t) dx \quad (1)$$

$$\left( \frac{\partial}{\partial t} + \frac{\partial}{\partial x} + \psi_1 \right) P_1(x, t) = 0 \quad (2)$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} + \phi_2 + \phi_4 + \eta_2\right) P_2(x, t) = 0 \tag{3}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} + \phi_2 + \phi_4 + \eta_2\right) P_3(x, t) = 0 \tag{4}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} + \Psi_2\right) P_4(x, t) = 0 \tag{5}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} + \Psi_3\right) P_5(x, t) = 0 \tag{6}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} + \phi_2 + \phi_4 + \eta_2\right) P_6(x, t) = 0 \tag{7}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} + \Psi_4\right) P_7(x, t) = 0 \tag{8}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} + \phi_2 + \eta_2 + \eta_2\right) P_8(x, t) = 0 \tag{9}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} + \phi_2 + \eta_4\right) P_9(x, t) = 0 \tag{10}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} + \Psi_5\right) P_{10}(x, t) = 0 \tag{11}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} + \Psi_6\right) P_{11}(x, t) = 0 \tag{12}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} + \Psi_7\right) P_{12}(x, t) = 0 \tag{13}$$

Boundary conditions:

$$P_3(0, t) = \phi_2^2 P_0(t) \tag{14}$$

$$P_4(0, t) = \phi_2^3 (1 + \phi_4) P_0(t) \tag{15}$$

$$P_i(0, t) = \phi_j P_0(t), \text{ where } i=1,2,5,6,10,11,12 \text{ \& } j=1,2,3,4,5,6,7 \tag{16}$$

$$P_7(0, t) = \phi_4^2 P_0(t) \tag{17}$$

$$P_8(0, t) = (\phi_2 \phi_4 + \phi_2 \phi_4) P_0(t) \tag{18}$$

$$P_9(0, t) = \phi_2^3 \phi_4 P_0(t) \tag{19}$$

$$P_0(0) = 1 \tag{20}$$

Solving (1)-(21),

$$\bar{P}_0(s) = \frac{1}{\epsilon(s)} \tag{21}$$

$$\bar{P}_1(s) = \frac{\phi_1}{\epsilon(s)} \left[ \frac{1 - \bar{S}\psi_1(s)}{s} \right] \tag{22}$$

$$\bar{P}_2(s) = \frac{\phi_2}{\epsilon(s)} \left[ \frac{1 - \bar{S}\eta_2(s + \phi_2 + \phi_4)}{(s + \phi_2 + \phi_4)} \right] \tag{23}$$

$$\bar{P}_3(s) = \frac{\phi_2^2}{\epsilon(s)} \left[ \frac{1 - \bar{S}\eta_2(s + \phi_2 + \phi_4)}{(s + \phi_2 + \phi_4)} \right] \tag{24}$$

$$\bar{P}_4(s) = \frac{\phi_2^3 (1 + \phi_4)}{\epsilon(s)} \left[ \frac{1 - \bar{S}\psi_2(s)}{s} \right] \tag{25}$$

$$\bar{P}_5(s) = \frac{\phi_3}{\epsilon(s)} \left[ \frac{1 - \bar{S}\psi_3(s)}{s} \right] \tag{26}$$

$$\bar{P}_6(s) = \frac{\phi_4}{\epsilon(s)} \left[ \frac{1 - \bar{S}\eta_2(s + \phi_2 + \phi_4)}{(s + \phi_2 + \phi_4)} \right] \tag{27}$$

$$\bar{P}_7(s) = \frac{\phi_4^2}{\epsilon(s)} \left[ \frac{1 - \bar{S}\psi_4(s)}{s} \right] \tag{28}$$

$$\bar{P}_8(s) = \frac{(\phi_4 \phi_2 + \phi_2 \phi_4)}{\epsilon(s)} \left[ \frac{1 - \bar{S}\eta_2(s + \phi_4)}{(s + \phi_4)} \right] \tag{29}$$

$$\bar{P}_9(s) = \frac{\phi_4 \phi_2^3}{\epsilon(s)} \left[ \frac{1 - \bar{S}\eta_2(s + \phi_2)}{(s + \phi_2)} \right] \tag{30}$$

$$\bar{P}_{10}(s) = \frac{\phi_5}{\epsilon(s)} \left[ \frac{1 - \bar{S}\psi_5(s)}{s} \right] \tag{31}$$

$$\bar{P}_{11}(s) = \frac{\phi_6}{\epsilon(s)} \left[ \frac{1 - \bar{S}\psi_6(s)}{s} \right] \tag{32}$$

$$\bar{P}_{12}(s) = \frac{\phi_7}{\epsilon(s)} \left[ \frac{1 - \bar{S}\psi_7(s)}{s} \right] \tag{33}$$

Where,

$$\epsilon(s) = \left[ (s + \phi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5 + \phi_6 + \phi_7) - \phi_1 \bar{S}\psi_1(s) - \phi_2 \bar{S}\eta_2(s + \phi_2 + \phi_4) \phi_2^3 (1 + \phi_4) \bar{S}\psi_2(s) - \phi_3 \bar{S}\psi_3(s) - \phi_4 \bar{S}\eta_2(s + \phi_2 + \phi_4) - \phi_4^2 \bar{S}\psi_4(s) - \phi_5 \bar{S}\psi_5(s) - \phi_6 \bar{S}\psi_6(s) - \phi_7 \bar{S}\psi_7(s) \right] \tag{34}$$

The probability of a system being in an operating mode or a failed state at any given moment are



transformed using a Laplace transform as follows:

$$\bar{P}_{up}(s) = \bar{P}_0(s) + \bar{P}_2(s) + \bar{P}_3(s) + \bar{P}_6(s) + \bar{P}_8(s) + \bar{P}_9(s) \tag{35}$$

$$\bar{P}_{up}(s) = \frac{1}{\epsilon(s)} \left\{ 1 + \phi_2 \left[ \frac{1 - \bar{S}_{\eta_2}(s + \phi_2 + \phi_4)}{(s + \phi_2 + \phi_4)} \right] + \phi_2^2 \left[ \frac{1 - \bar{S}_{\eta_2}(s + \phi_2 + \phi_4)}{(s + \phi_2 + \phi_4)} \right] + \phi_4 \left[ \frac{1 - \bar{S}_{\eta_2}(s + \phi_2 + \phi_4)}{(s + \phi_2 + \phi_4)} \right] + (\phi_2 \phi_4 + \phi_2 \phi_4) \left[ \frac{1 - \bar{S}_{\eta_2}(s + \phi_4)}{(s + \phi_4)} \right] + \phi_2^3 \phi_4 \left[ \frac{1 - \bar{S}_{\eta_4}(s + \phi_2)}{(s + \phi_2)} \right] \right\} \tag{36}$$

$$\bar{P}_{down}(s) = 1 - \bar{P}_{up}(s) \tag{37}$$

### III. Results

#### I. Availability Analysis

Taking,  $S_{\mu_0}(s) = \bar{S}_{\exp [x^\theta + \{\log \eta(x)\}^\theta]^{1/\theta}}(s) = \frac{\exp [x^\theta + \{\log \eta(x)\}^\theta]^{1/\theta}}{s + \exp [x^\theta + \{\log \eta(x)\}^\theta]^{1/\theta}}$ ,  $\bar{S}_\eta(s) = \frac{\eta}{s + \eta}$  and failure rates are  $\phi_1 = .002, \phi_2 = .003, \phi_3 = .004, \phi_4 = .005, \phi_5 = .003, \phi_6 = .007, \phi_7 = .001$  And repair rates  $\eta_1 = \eta_2 = \eta_3 = \eta_4 = \eta_5 = \eta_6 = \eta_7 = 1 = \Psi_1 = \Psi_2 = \Psi_3 = \Psi_4 = \Psi_5 = \Psi_6 = \Psi_7$  in equation [36], One may get the availability expression as: taking the inverse Laplace transform.

$$\text{Availability} = [.01678017142869e^{-1.01717779780318t} - .24872963381646e^{-.06371454763895t} - .37315067813499e^{-.0765367498357t} + .00599206711030e^{-.005t} + .00000002708736e^{-.003t} + 1.5990990379009e^{-.007t} + .000017998275150te^{-.005t} + .000000000135161851te^{-.003t} + .001601056166718te^{-.007t}] \tag{38}$$

Taking time  $t = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$ , We determined several values for availability with equation [38] as shown in Table 1 and graph in Fig. 3

**Table 1:** Availability vs time (t)

Time	A(t)
0	1.00000
1	0.99080
2	0.98550
3	0.97960
4	0.97170
5	0.96130
6	0.94820
7	0.93230
8	0.91370
9	0.89220
10	0.86790

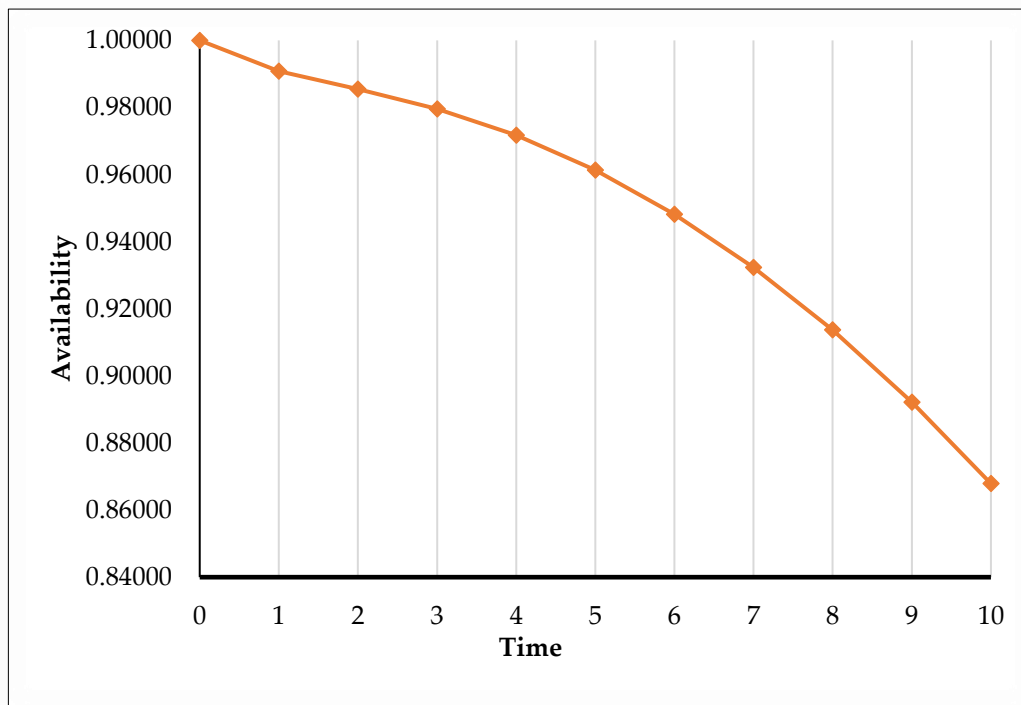


Figure 3: Availability v/s Time

## II. Reliability Analysis

Assuming all repair rates is equal to zero in equation [36] and taking failure rates as  $\phi_1 = .002, \phi_2 = .003, \phi_3 = .004, \phi_4 = .005, \phi_5 = .003, \phi_6 = .007, \phi_7 = .001$  after which, using the Inverse Laplace transform, we obtained Equation [39]. as shown in Table 2 and graph in Fig. 4.

$$R(t) = \{ .0015e^{-.005t} + .00000000061363636e^{-.003t} + .5535555494191e^{-.025t} + .4449444444e^{-.007t} \} \quad (39)$$

Table 2: Reliability v/s Time

Time	R(t)
0	1.00000
1	0.98320
2	0.96680
3	0.95070
4	0.93500
5	0.91960
6	0.90450
7	0.88980
8	0.87540
9	0.86120
10	0.84740

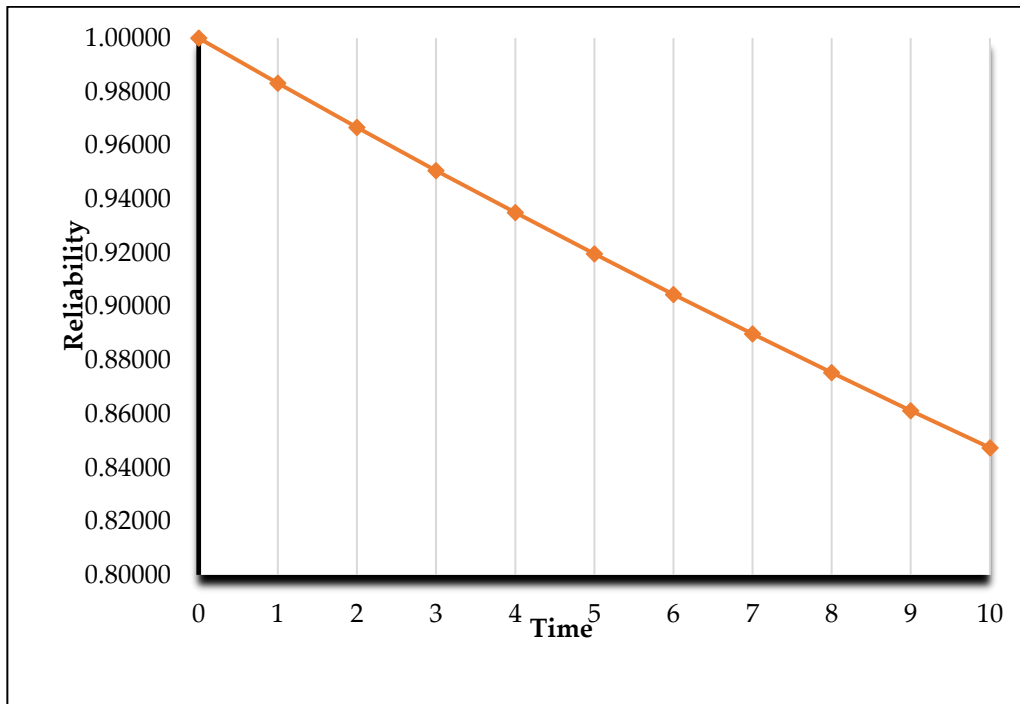


Figure 4: Reliability vs Time

### III. MTTF Analysis

Assuming all repair rates is equal to zero in equation [36], we arrive at the formula for MTTF as  $s$  tends to zero

$$MTTF = \lim_{s \rightarrow 0} \bar{P}_{up}(s) = \frac{1}{\phi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5 + \phi_6 + \phi_7} \left\{ \frac{2\phi_2 + 3\phi_2^2 + 2\phi_4 + 2\phi_2\phi_4 + \phi_2^3\phi_4 + \phi_2^2\phi_4^2}{\phi_2 + \phi_4} \right\} \quad (40)$$

and taking failure rates as  $\phi_1 = .002, \phi_2 = .003, \phi_3 = .004, \phi_4 = .005, \phi_5 = .003, \phi_6 = .007, \phi_7 = .001$  and varying failure rates one by one as .001,.002,.003,.004,.005,.006,.007,.008,.009,.010 in equation [38], and we can get the variation of mean time to failure with respect to failures rates as shown in Table 3 and graph in Fig. 5.

Table 3: MTTF V/S Failure rates

Failure rate	$\phi_1$	$\phi_2$	$\phi_3$	$\phi_4$	$\phi_5$	$\phi_6$	$\phi_7$
0.001	83.630210	87.050725	91.232957	95.630953	87.266306	105.638160	80.285002
0.002	80.285002	83.523810	87.266306	91.263637	83.630210	100.356252	77.197117
0.003	77.197117	80.285002	83.630210	87.282610	80.285002	95.577383	74.337965
0.004	74.337965	77.299148	80.285002	83.636906	77.197117	91.232957	71.683037
0.005	71.683037	74.537042	77.197117	80.285002	74.337965	87.266306	69.211208
0.006	69.211208	71.974032	74.337965	77.192310	71.683037	83.630210	66.904168
0.007	66.904168	69.589089	71.683037	74.329632	69.211208	80.285002	64.745969
0.008	64.745969	67.364113	69.211208	71.672080	66.904168	77.197117	62.722658
0.009	62.722658	65.283423	66.904168	69.198279	64.745969	74.337965	60.821971
0.01	60.821971	63.333349	64.745969	66.889747	62.722658	71.683037	59.033090

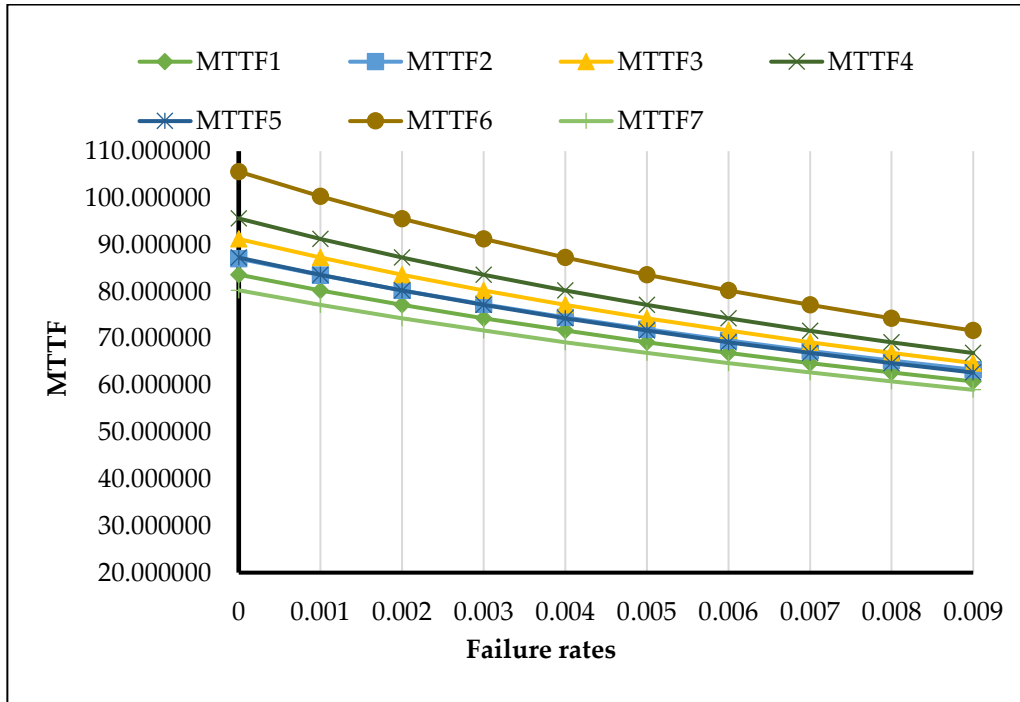


Figure 5: MTTF v/s Failure rates

#### IV. Sensitivity Analysis

With the partial differentiation of MTTF regarding the failure rate of the system, the sensitivity in MTTF of the system can be evaluated. The MTTF sensitivity may be calculated by using the set of parameters  $\phi_1 = .002, \phi_2 = .003, \phi_3 = .004, \phi_4 = .005, \phi_5 = .003, \phi_6 = .007, \phi_7 = .001$  and in the partial differentiation of MTTF, as given in the Table 4 and graphs in Fig. 6

Table 4: Sensitivity of MTTF as a function of failures rates

Failure rates	$\frac{\partial(MTTF)}{\partial\phi_1}$	$\frac{\partial(MTTF)}{\partial\phi_2}$	$\frac{\partial(MTTF)}{\partial\phi_3}$	$\frac{\partial(MTTF)}{\partial\phi_4}$	$\frac{\partial(MTTF)}{\partial\phi_5}$	$\frac{\partial(MTTF)}{\partial\phi_6}$	$\frac{\partial(MTTF)}{\partial\phi_7}$
0.001	-1313.161	-3686.090	-1562.770	-4580.640	-1429.831	-2095.237	-1210.209
0.002	-1210.209	-3376.410	-1429.831	-4164.710	-1313.161	-1890.952	-1118.906
0.003	-1118.906	-3000.000	-1313.161	-3805.760	-1210.209	-1715.149	-1037.559
0.004	-1037.559	-2922.050	-1210.209	-3492.520	-1118.906	-1562.770	-964.771
0.005	-964.771	-2606.310	-1118.906	-3217.000	-1037.559	-1429.831	-899.382
0.006	-899.382	-2424.520	-1037.559	-2973.190	-964.771	-1313.161	-840.423
0.007	-840.423	-2312.060	-964.771	-2756.240	-899.382	-1210.209	-787.077
0.008	-787.077	-2169.620	-899.382	-2562.400	-840.423	-1118.906	-738.653
0.009	-738.653	-2017.450	-840.423	-2388.270	-787.077	-1037.559	-694.564
0.010	-694.564	-1909.720	-787.077	-2231.420	-738.653	-964.771	-654.309

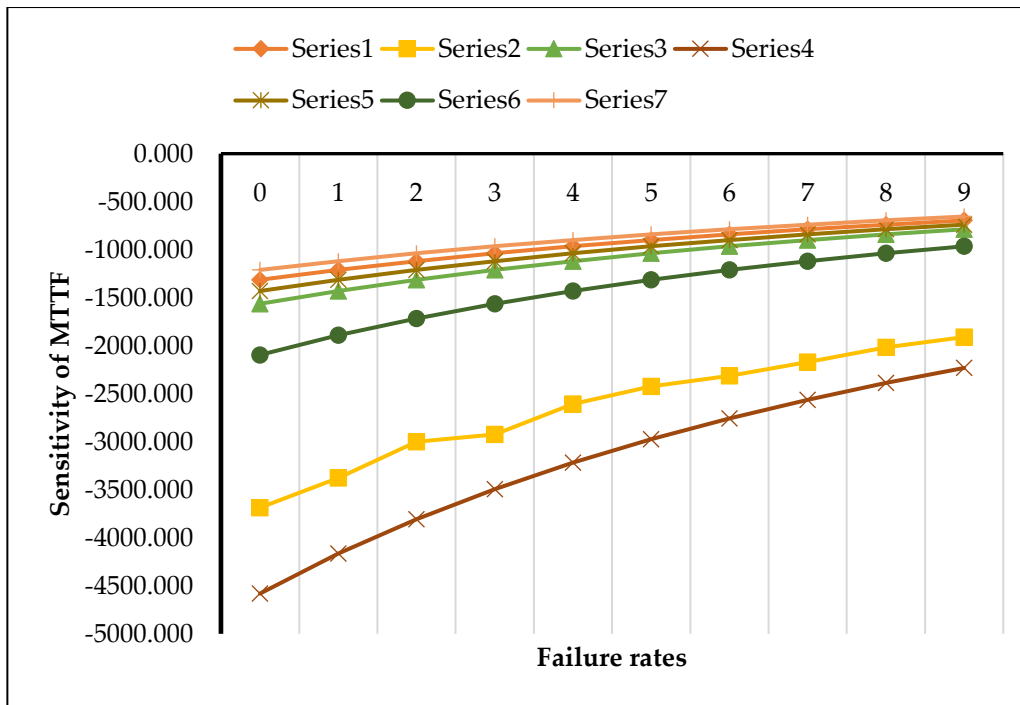


Figure 6: Sensitivity of MTTF V/S failures rates

V. Profit Analysis

Formula presented as follows may be used to compute the expected profit within the period [0, t):

$$E_p(t) = Z_1 \int_0^t P_{up}(t) - Z_2 t$$

Taking  $Z_1 = 1$  and  $Z_2 = .05, .10, .15, .20, .25, .30, .35$  and varying  $t = 0, 1, 2, 3, \dots, 10$ . units of time then the expected profit is

$$E_p(t) = \{0.0164967e^{-1.017177797t} + 3.90381228e^{-0.637145476t} + 4.875444529e^{-0.765367498t} - 1.19841342e^{-0.005t} - 0.000009e^{-0.003t} - 228.4427196e^{-0.007t} - 0.00000033te^{-0.003t} - 0.000011e^{-0.003t} - 0.0035te^{-0.005t} - 0.719928e^{-0.005t} - 0.2287222te^{-0.007t} - 32.674612e^{-0.007t} + 0.232222t + 254.2729352 - Z_2 t\} \tag{41}$$

As given in the Table 5 and graphs in Fig. 7.

Table 5: Expected profit v/s Time

Time	$Z_2 = .05$	$Z_2 = .10$	$Z_2 = .15$	$Z_2 = .20$	$Z_2 = .25$	$Z_2 = .30$	$Z_2 = .35$
0	0	0	0	0	0	0	0
1	1.204884	1.154884	1.104884	1.054884	1.004884	0.954884	0.904884
2	2.415288	2.315288	2.215288	2.115288	2.015288	1.915288	1.815288
3	3.657103	3.507103	3.357103	3.207103	3.057103	2.907103	2.757103
4	4.926165	4.726165	4.526165	4.326165	4.126165	3.926165	3.726165
5	6.219485	5.969485	5.719485	5.469485	5.219485	4.969485	4.719485
6	7.534606	7.234606	6.934606	6.634606	6.334606	6.034606	5.734606
7	8.869361	8.519361	8.169361	7.819361	7.469361	7.119361	6.769361
8	10.22178	9.821781	9.421781	9.021781	8.621781	8.221781	7.821781
9	11.59005	11.14005	10.69005	10.24005	9.790051	9.340051	8.890051
10	14.92179	14.42179	13.92179	13.42179	12.92179	12.42179	11.92179

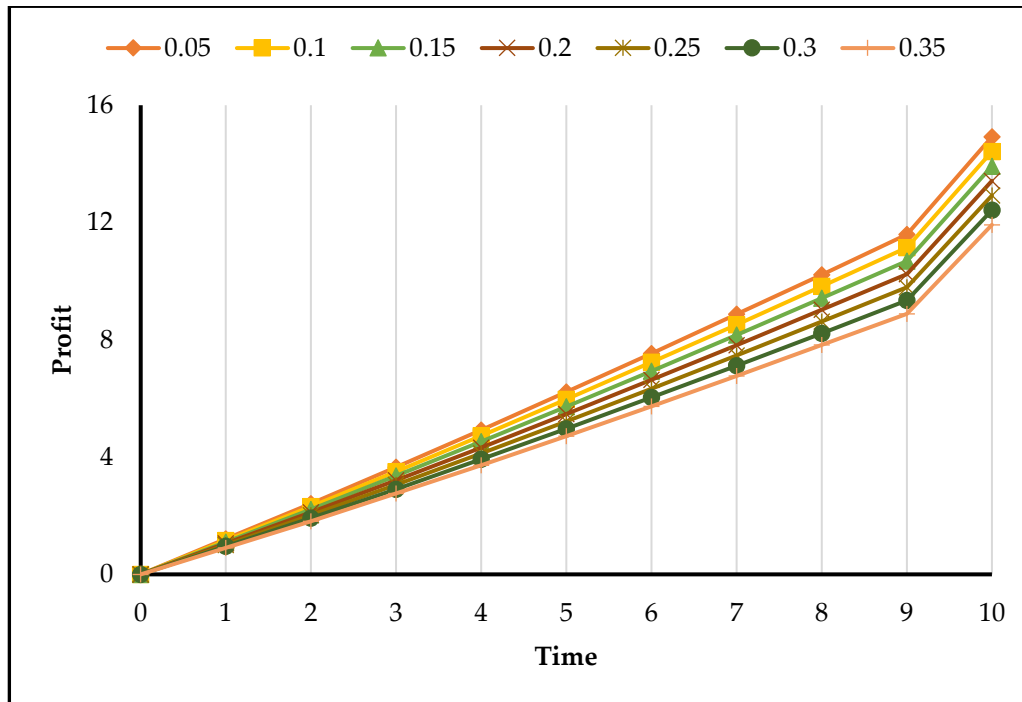


Figure 7: Profit v/s Time

#### IV. Discussion

##### I. Interpretation of the result & Discussion

To analyse and conduct the Spirulina production plant while taking reliability metrics into account for various values of failure and repair rates. When the failure rates are set at various values,  $\phi_1 = .002, \phi_2 = .003, \phi_3 = .004, \phi_4 = .005, \phi_5 = .003, \phi_6 = .007, \phi_7 = .001$  namely, Table.1 shows the information on the availability of the plant repairable system concerning the time variation.

Figure 3's simulation demonstrates how availability declines over time. The graph unequivocally demonstrates that the system's availability is higher when the time span is 5 years or less. A similar way is shown in Figure 4 for the system's reliability over time. The graph shows how reliability decreases as time  $t$  increases from 0 to 10. The time interval, on the other hand, is more reliable.

As shown in Figures 4 and 5, adding more units to standby can increase system availability and reliability by performing a perfect repair in the case of an incomplete failure, replacing the affected subsystem with a new one in the case of a full failure, performing routine inspections and preventative maintenance, hiring more repair equipment, and other methods.

A simulation of the mean time to failure vs the failure rate is shown in Figure 5. The graph demonstrates that the MTTF drops as the failure rate rises. The MTTF decreases as the failure rate rises, lowering the system's duration. To extend the system's MTTF and duration, fault-tolerant components should be used.

One can see from Table 5 and Figure 6 that System MTTF is extremely sensitive to the failure rates of the washing chamber. The MTTF of the spirulina manufacturing facility is significantly impacted when the failure rate of the washing chamber rises. In this case MTTF is much less responsive.

The connection between profit and time  $t$  for  $Z_2 = .05, .10, .15, .20, .25, .30, .35$  is shown Table 5 in Figure 7. The graph shows that the expected profit falls with increasing time for any value of  $Z_2$ . Yet, the anticipated profit increases as the value decline. The anticipated profit will increase by putting the substitution and redundancy concepts into practice.

## II. Conclusion

In this study, the Markov model was used to assess the plant's reliability at the spirulina production plant. From the explanation above, we deduce the following: The MTTF is extremely sensitive to the failure rate of the washing unit; as soon as this number even marginally changes, the MTTF's sensitivity rating increases drastically. So, the engineers of the spirulina production plant should pay more attention to the maintenance of the system's fourth unit (washing unit). This unit mostly affects the plant's functioning. For this unit, reliable equipment should be used to cause the least possible system disruption. Timely preventative maintenance will improve the system's performance. The spirulina production plant will greatly benefit from this study in terms of improving its efficiency and maintenance strategy.

## References

- [1] Agarwal, S. C., & Bansal, S. (2009) 'Reliability analysis of a standby redundant complex system with changing environment under the head of the line repair discipline' *Bulletin of Pure and Applied Sciences*, 28(1), 165-173.
- [2] Xie, L., Lundteigen, M.A and Liu, Y. (2020) 'Reliability and barrier assessment of series-parallel systems subject to cascading failures' *Proceedings of the Institution of Mechanical Engineers Part O Journal of Risk and Reliability* 234(4):1748006X1989923 DOI:10.1177/1748006X19899235
- [3] Agarwal, S. C., Sahani, M., & Bansal, S. (2010) 'Reliability characteristic of a cold-standby redundant system' *International Journal of Research and Review in Applied Sciences (IJRRAS)*, 3(2), 193-199.
- [4] Yusuf, I. and Hussaini, N. (2014) 'A comparative analysis of three-unit redundant systems with three types of failures', *Arabian Journal for Science and Engineering*, Vol. 39, No. 4, pp.3337-3349.
- [5] Zhang A., Wang Z., Bao R., Chengrui Liu, Qiong Wu, Shihao Cao (2023) A novel failure time estimation method for degradation analysis based on general nonlinear Wiener processes. *Reliability Engineering & System Safety* Volume 230108913 <https://doi.org/10.1016/j.res.2022.108913>.
- [6] Agarwal, S. C., & Bansal, S. (2015) 'Cost analysis of solar thermal electric power plant' *International Journal of Advanced Technology in Engineering and Science*, 3(10), 12-22.
- [7] Arora N., Kumar D., (1996) 'Stochastic and maintenance planning of ash handling system in thermal power plant' *International Journal of Microelectronics Reliability*, 37 5 819-834.
- [8] Bansal, S., (2018) 'Availability Analysis of a Repairable Redundant System Under Preemptive-Resume Repair Discipline' *Int. J. Math. And Appl.*, 6(1-D), 665-671.
- [9] Bansal S., Tyagi, S.(2021) 'Mathematical modeling and availability analysis of leaf spring manufacturing plant' *Pertanika journal of science & technology*, Vol.29, No.2, pp. 1041-1051.
- [10] Bansal, S., Tyagi, S., Verma, V.K (2022). 'Performance Modeling and Availability Analysis of Screw Manufacturing Plant' *Materials Today: Proceedings*, Vol.57, No.5, pp.1985-1988.
- [11] Chaudhary, P., and Bansal, S. (2023) 'Assessment of the Reliability Performance of Hydro-Electric Power Station, *IEEE* DOI: 10.1109/ICACITE57410.2023.10182482.
- [12] Chauhan, S.K and Mali, S.C. (2016) 'Reliability Evaluation of Series-Parallel and Parallel-

Series Systems for Arbitrary Values of the Parameters' International Journal of Statistics and Reliability Engineering Vol. 3(1), pp. 10-19.

[13] Fouladirad, M., Giorgio M, Pulcini (2022) 'G. A transformed gamma process for bounded degradation phenomena' *Qual Reliab Eng Int.* (2022); **1**(1): 1– 10. doi: <https://doi.org/10.1002/qre.3167>.

[14] Godara, U., and Bansal, S., (2023) 'Performance of Reliability Factors in Steam Turbine Generator Power Plant Using Boolean Function Technique and Neural Network Approach' *IEEE DOI: 10.1109/ICAECIS58353.2023.10170451*

[15] Kabiru, H., Ibrahim, Singh V.V. and Abulkareem Lado (2017) 'Reliability Assessment of Complex System Consisting Two Subsystems Connected in Series Configuration Using Gumbel-Hougaard Family Copula Distribution' *Journal of Applied Mathematics and Bioinformatics*, Vol.7, no.2, 1- 27.

[16] Uswarman, R., & Rushdi, A. M. (2021) 'Reliability evaluation of rooftop solar photovoltaic using coherent threshold systems' *Journal of Engineering Research and Reports*, 20(2), 32–44. Article no. JERR.64840, ISSN: 2582-2926.

[17] Lai WF, Zwetsloot IM. (2022) 'A data-driven ensemble ranking system of production quality across manufacturers—a case study for risk assessment in the solar industry' *Qual Reliability Eng. Int.* **1**(1): 1– 10. doi: <https://doi.org/10.1002/qre.3190>.

[18] Tyagi, S.L., and Bansal, S., (2023) 'Optimization Model for Wastewater Treatment Process' *IEEE DOI: 10.1109/ICACITE57410.2023.10182482*.

[19] Maihulla, A., Yusuf I., Bala S. (2021) 'Performance and cost assessment of a small solar photovoltaic system using Gumbel-Haugard Family Copula Analysis' *Research square* <https://doi.org/10.21203/rs.3.rs-939621/v1>.

[20] Meynaoui A, Marrel A, Laurent B. (2022) 'Second-level global sensitivity analysis of numerical simulators with application to an accident scene in a sodium-cooled fast reactor' *Qual. Relia. Eng. Int.* **1**(1): 1– 10. doi: <https://doi.org/10.1002/qre.3157>.



# EFFICIENT FRAMEWORK OF SECURITY FOR INTERNET OF THINGS

Dr. Mihir Mehta<sup>1</sup>, Dr. Kajal Patel<sup>2</sup>, Dr. Komal Anadkat<sup>3</sup>

•

<sup>1</sup>Assistant Professor, Computer Engineering department, G.E.C., Gandhinagar, Gandhinagar,  
India. mihir\_mehta@gecg28.ac.in

<sup>2</sup> Associate Professor, Computer Engineering department, Vishwakarma College of Engineering,  
Ahmedabad, India. kajalpatel@vgecg.ac.in

<sup>3</sup>Assistant Professor, Information Technology department, G.E.C., Gandhinagar,  
Gandhinagar, India. komalanadkat@gecg28.ac.in

## Abstract

*IoT security represents a highly compelling subject of research at present. The absence of a viable security solution for IoT applications could render them ineffective across various domains such as healthcare, smart homes, inventory management, smart agriculture, and more. Within the IoT architecture, security services like Confidentiality, Integrity, and Authentication play pivotal roles. In our research, we have concentrated on the Authentication service, which is fundamental for distinguishing users and devices unequivocally within a network. Authentication serves as the initial and crucial step in establishing secure communications among diverse IoT devices and users within the network. A compromised Authentication service could open the door for unauthorized users or devices to infiltrate the network, potentially leading to harmful activities like Masquerade attacks, Man-in-the-Middle (MITM) attacks, and Replay attacks. Currently, Authentication stands as a widely adopted and essential method for granting access to devices within IoT networks. Our contribution involves the development of a Multi-factor IoT Authentication Model, leveraging two key parameters: Device Context Information and Dynamic Key-based authentication. Our proposed approach begins by verifying the origin of information. If the origin is deemed valid, our model proceeds to validate the identity of the device. In the event of an intruder attempting to manipulate the device's origin from its predefined context to an alternative location, our system can swiftly detect this deviation, thereby enabling the rejection of communication requests from compromised devices. Following the verification of context information, we initiate mutual authentication between the IoT device and the server, employing the Challenge-response model. As a result of this second step, individual Session keys are generated at both the device and server sides, facilitating secure communication within a specific time window.*

**Keywords:** Internet of Things, Multi-factor Authentication, Dynamic key based Authentication.

## I. Introduction

The realm of IoT security represents a highly significant area of research in the current era. It has garnered substantial attention from researchers across industry, academia, and various government agencies. A report by CISCO in April 2019 projected a staggering 50 billion devices to be interconnected with the internet by the end of 2020. This exponential growth presents a substantial opportunity for malicious actors to launch diverse cyber-attacks on IoT systems, primarily due to the open architecture inherent in IoT networks. Traditional security approaches are ill-suited for IoT devices, primarily due to their inherent limitations, including constrained storage capacity and computational power. Moreover, IoT devices must function in harsh and unpredictable environments, making them vulnerable to an array of security threats. Consequently, there is an

imperative need to develop security solutions tailored to the resource constraints of IoT devices while providing essential attributes such as Confidentiality, Integrity, and Authentication in IoT networks.

Outlined below are some of the key challenges in IoT security.

**Open Architecture:** In IoT, all devices are interconnected through the internet, adhering to an open framework. This openness amplifies the potential for various security threats.

**System Limitations:** IoT devices face constraints concerning memory, computational power, CPU capacity, and energy. These limitations render traditional security approaches unsuitable for direct deployment in IoT systems.

**Absence of Standards:** The diversity of IoT devices hinders standardization efforts. Each IoT device functions as a standalone system comprising hardware, firmware, and communication interfaces. Ensuring security at the design phase, crafting secure code, and conducting rigorous verification/validation during the manufacturing process are essential. Nevertheless, there is currently no practical means to enforce and standardize these security methods across all devices.

**Deficient Trust and Integrity:** With a multitude of devices connected to the internet, it becomes nearly impossible to verify that each device maintains adequate safeguards and remains up-to-date with the latest security updates. A single vulnerable link in the network can grant intruders access to numerous devices. Ensuring trust and data integrity for every IoT device is of paramount importance.

**Insecure Web Interfaces:** Vulnerable web interfaces in IoT devices are susceptible to various threats, including account enumeration and brute force attacks. For example, attackers may gain unauthorized access to websites by attempting numerous password combinations, potentially compromising administrative policies and sensitive data. Attackers can also manipulate the credentials of legitimate users.

Addressing these challenges is crucial to establishing a robust and secure IoT ecosystem that can withstand the evolving landscape of cyber threats.

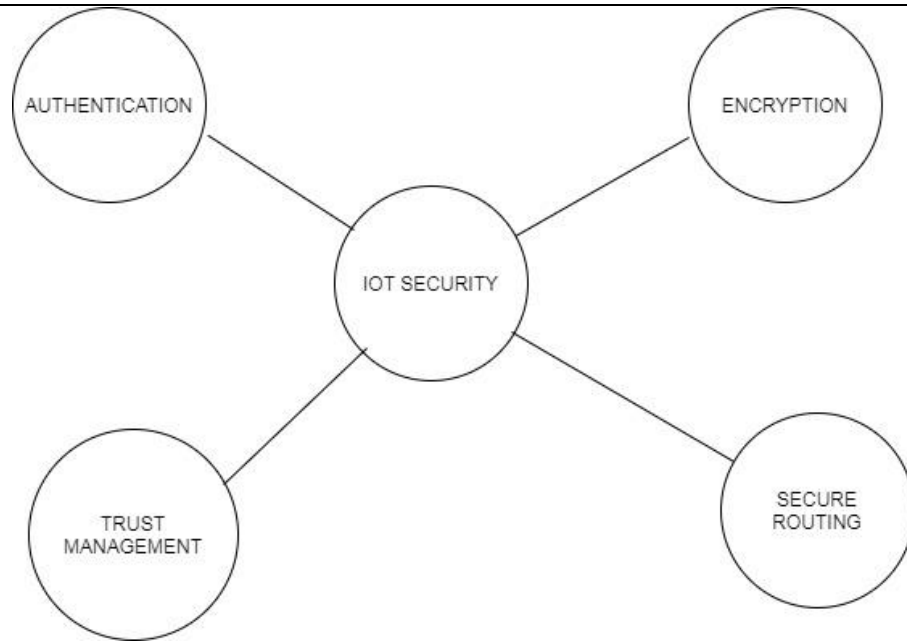
There are certain security issues present in IoT Architecture, they are Authentication, Encryption, Trust Management & Secure Routing.

**Authentication:** Authentication plays a pivotal role in identifying devices and users within an IoT system, granting access exclusively to authorized entities. In IoT systems, authentication can be realized through various methods, including Identity-based authentication, Token-based authentication, PUF-based authentication, and Procedure-based authentication.

**Encryption:** Encryption is essential for achieving end-to-end security in IoT systems. The primary objective of encryption within the IoT ecosystem is to establish effective end-to-end communication through the utilization of symmetric and asymmetric cryptographic algorithms. However, IoT devices face resource limitations, which necessitate a departure from traditional encryption algorithms like AES and DES, as they are not directly suitable for the constraints of IoT networks.

**Trust Management:** IoT trust management is fundamentally geared towards identifying and isolating malicious nodes within the IoT network. The overarching aim is to identify and subsequently remove such nodes from the network, thus enabling secure access control within the IoT environment.

**Secure Routing:** Within the context of data transmission in IoT networks, the presence of malicious nodes poses a significant threat. These malicious nodes have the potential to divert data packets towards them, infiltrating routing and forwarding decision processes for both data and control packets. As such, ensuring secure routing mechanisms becomes imperative in safeguarding the integrity of IoT networks.



**Figure 1:** Security Issues for Internet of Things

## II. Literature Review

In [1], researchers Hokeun Kim and Edward Lee proposed an approach for authentication in IoT devices, emphasizing a provincially centralized and universally distributed method. Trust serves as a fundamental prerequisite for authentication in IoT systems, and the authors discussed the implementation of a certificate-based scenario to establish trust between clients and servers. They identified two key methods for deploying trust within a network: (1) Utilizing a Centralized Trusted Authority and (2) Leveraging Distributed and Trusted stakeholders. The authors developed a network framework named "Auth," which incorporates local authentication and authorization entities. Auth, implemented as open-source software in Java and accessible on GitHub, facilitates authorization for locally registered entities (IoT devices) and manages trust relationships with other Auth instances. The framework securely stores the credentials of endorsed devices and access policies within a database. The authorization process involves the assignment of session keys, cryptographic keys used for specific access activities.

In [2], authors Mohammad Wazid, Ashok Kumar Das, and others discussed a lightweight authentication protocol known as the "User Authenticated Key Management Protocol (UAKMP)" designed for a concept called Hierarchical Internet of Things (HIOT). This protocol utilizes three authentication factors: (1) user smart cards, (2) passwords, and (3) personal biometrics. The method employs a combination of cryptographic message digest functions and symmetric encryption/decryption. UAKMP involves six essential steps: (1) Enrollment of various sensor nodes, (2) Enrollment of users, (3) User sign-up, (4) Authentication and key agreement, (5) Password change, and (6) Integration of newly joined sensor nodes. Gateway nodes store critical information required for authentication in all deployed sensing nodes, including their identity. The protocol assumes that the Gateway node is trustworthy, as a breach of its security could endanger the entire network, potentially leading to node impersonation attacks and denial of service attacks.

In [3], authors Ning Wang, Ting Jiang, and their team presented an authentication approach primarily focused on physical layer attributes. Physical layer authentication involves the examination of various physical attributes, including Received Signal Strength (RSS) and Channel

Impulse Response (CIR). The proposed method incorporates machine learning, specifically a Feedforward Neural Network, for classification tasks. This choice of neural network offers advantages such as rapid learning, ease of construction, and minimal human intervention. Binary hypothesis testing is used to detect spoofing attacks, framing the problem within an Alice-Bob-Eve model, where Alice is the legitimate transmitter, Bob is the legitimate receiver, and Eve is an illegitimate transmitter attempting to impersonate another node with a false address. The challenge addressed in this method is determining whether the second message received by Bob, after the first one confirmed to be from Alice, is still sent by Alice or not.

In [4], authors Muhammad Naveed Aman, Sachin Taneja, and others introduced a token-based authentication method that employs OAuth 2.0, an open authentication and authorization standard. This method aims to mitigate security risks associated with conventional client-server authentication, where clients use resource owners' credentials, potentially leading to password leakage and data breaches. The proposed approach involves three main steps: (1) The client sends an authorization request to the Authorization Server (AS), (2) The AS verifies the client's authenticity and, if verified, issues an access token to the client, and (3) The client uses this access token to authenticate itself to the resource server (RS) and access requested resources. However, the method is susceptible to replay attacks if an intruder captures an access token generated by the Authorization Server, as it could be misused for impersonation attacks.

In [5], authors Prosanta Gope and Biplab Sikdar presented a lightweight two-factor authentication approach for IoT devices, addressing the vulnerabilities of password-based and key-based methods to physical and side-channel attacks. Their approach combines two factors: (1) a secret shared key and (2) a Physical Unclonable Function (PUF). During registration, an IoT device transmits its identity along with a registration request to the server. The server responds by generating a random challenge (C), which it sends back to the client IoT device. The client computes a response to the challenge using its PUF and sends it back to the server for verification. If the response is correct, the server generates an alias identity and session key for the device, storing these details in its database. However, the method does not consider environmental parameters, which can affect PUF output, and is vulnerable to man-in-the-middle attacks, replay attacks, and spoofing attacks.

In [6], authors Muhammad Naveed Aman and Biplab Sikdar presented two-factor authentication algorithms for IoT devices, considering the low-cost nature of IoT devices that makes them susceptible to spoofing and impersonation attacks. Their method combines PUF and device hardware fingerprints for authentication. After device identity verification, the server provides a new challenge to the IoT device, which computes a response using its PUF and the provided challenge. However, this approach is vulnerable to replay attacks, as intruders can intercept Challenge-Response pairs exchanged between the IoT device and the server and use them for predicting other CRPs. Additionally, it does not provide security against man-in-the-middle attacks.

In [7], authors Zahoor Ahmed Alizai, Noquia Fateema Tarin, and others introduced a multifactor authentication approach based on digital signatures and device capabilities. This schema utilizes a secure TLS channel, with a digital signature serving as a second factor for authentication. Device authentication relies on the verification of device capability, involving data processing tasks. However, this approach demands high computational resources due to the involvement of asymmetric cryptography, making it unsuitable for resource-constrained IoT devices. Furthermore, it is vulnerable to impersonation and denial-of-service attacks.

In [8], authors Moritz Loske, Lukas Rothe, and others proposed context-aware authentication methods for IoT devices, addressing the limitations of existing cryptography-based approaches in IoT networks with resource-constrained devices. Context-aware authentication incorporates environmental information, such as temperature, luminosity, radio signals, and device location, to improve the authentication process. While this method reduces computational overhead, it does not provide confidentiality and is susceptible to man-in-the-middle attacks, replay attacks, and spoofing attacks. Therefore, it is best used as one parameter within a multi-factor-based authentication approach to enhance security.

In [9], authors Tarak Nandy, Sananda Bhattacharya, and their team discussed the existing authentication approaches for IoT and emphasized the need for strong and secure authentication methods. In IoT networks, various devices communicate with each other and users, making proper security crucial to prevent credential theft and attacks on the IoT network. The authors identified various attacks on IoT authentication, including masquerade attacks, man-in-the-middle attacks, denial-of-service attacks, forging attacks, guessing attacks, physical attacks, and routing attacks.

**Table 1:** *IoT Attacks & description*

Attacks	Description
Masquerade attack	In this attack, adversary misuses the identity of the legal user to get access to the network.
Man in the Middle attack	In this attack, adversary intercepts the communication between two parties and also can modify the communication contents.
DOS attack	In this attack, adversary floods the network with fake requests so legal user cannot use resources at that time. Network and resources are unavailable for them.
Forging attack	In this attack, adversary emulates a system or legal user to gain access to the network.
Guessing attack	In this attack, adversary predicates credentials of legal user by brute force approach or dictionary approach to gain access of the network.
Physical attack	In this attack, adversary tries to get physical access of the resource and can change physical location of resource to launch the attack.
Routing attack	In this attack, adversary advertises a false route for packet delivery from source to destination.

Problem Statement: Design & Development of Lightweight Multi-factor IoT Authentication approach by considering Context Parameter & Dynamic Key Parameter (Vault, Random Number) for addressing location spoofing attack, Eavesdropping attack, Replay attack & Identity Stolen attack.

Advantages of Context Information Parameter:

**Early Detection of Attackers:** When contextual variables, such as location information, are validated during the login session, it becomes possible to identify and detect request messages from potential attackers at an early stage. This early detection eliminates the need to unnecessarily verify other authentication factors during the session, thereby enhancing the security system's performance and reducing delays.

**Crucial for Decision-Making:** In domains like Military and Industry applications, the context parameter of a device plays a pivotal role in the decision-making process. If a device is legitimate but its context information has been tampered with, it can transmit incorrect or faulty data, which can have adverse effects on system performance. Therefore, validating context information is essential, along with device identity validation, before initiating a communication session.

Advantages of Dynamic Key-Based IoT Authentication:

**Enhanced Security:** In symmetric encryption, both communicating parties share the same pair of keys. However, if a third party gains access to the key or analyzes network traffic, they can infer the communication content. Consequently, long-term use of a fixed session key is insecure in IoT devices.

**"One Time One Cipher" Approach:** To address this security concern, the "One Time One Cipher" approach is employed, where the key used for encryption and decryption differs for each session

and expires after each use. This approach ensures the uniqueness and dynamic nature of the key. Session keys are generated securely and efficiently on both the device and server sides, considering parameters such as the Vault and Random Number Generation. This proactive measure helps prevent Key Stolen and Eavesdropping attacks, enhancing overall security.

### III. Methodology

#### Step 1: Context-Based Authentication

A. During the login request to the server, an IoT device transmits its login request along with contextual information. Specifically, the IoT device sends its location information in the form of Cartesian coordinates to the server.

B. The server proceeds to validate these context parameters by comparing them to the stored records in its database. In this validation process, the server calculates the Angle of Arrival (AoA) for the requested IoT device and matches the result with the stored AoA information for that device in the database. If these physical context parameters match, it provides evidence that the device is legitimate and identified at its original location.

#### Step 2: Dynamic Key-Based Authentication

If the device successfully passes the context-based authentication test, we introduce a second factor to enhance our authentication process, known as Dynamic Key-Based Authentication. In this phase, IoT Device and Server mutually authenticate each other initially by employing a Challenge-Response mechanism. Following a successful mutual authentication, a Session Key is generated for communication within a specific time window.

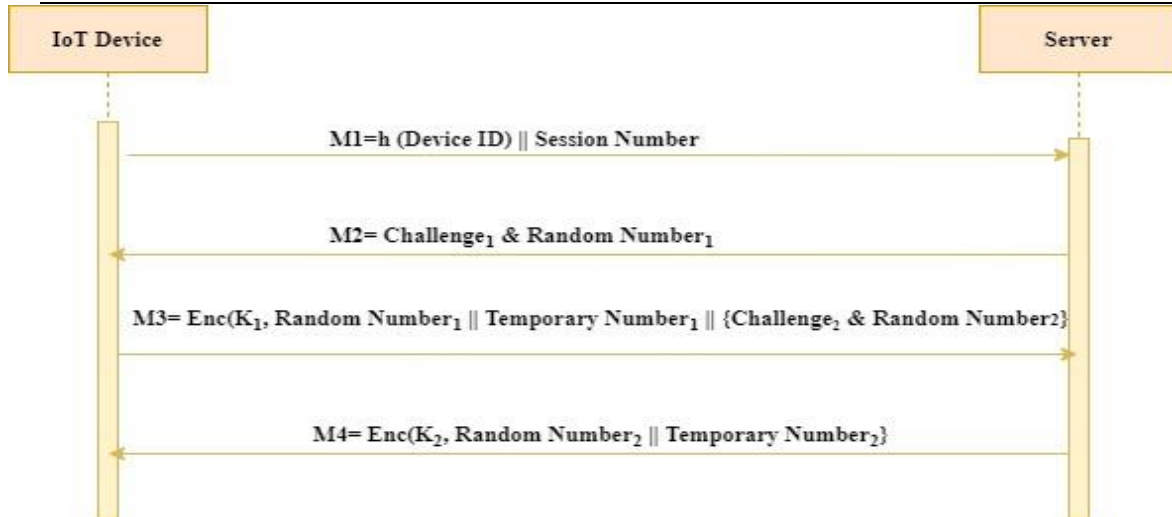
The detailed procedure for Dynamic Key-Based Authentication is as follows:

**Vault:** The Vault consists of 64 keys, with each key being 128 bits in length and represented in hexadecimal format. All of these keys are organized in an 8x8 matrix format, which is stored both on the IoT device and the server. To enhance security, these keys can be stored in an encrypted format at both ends. Each key in this 8x8 matrix can be denoted as  $K[0][0]$ ,  $K[0][1]$ , ...,  $K[7][7]$ . During the initial deployment, this 8x8 matrix is shared between the IoT device and the server.

**Challenge-Response Mechanism:** Our proposed protocol employs a Handshaking concept to achieve mutual authentication between the IoT device and the server. The diagram below illustrates the sequence of messages exchanged between the IoT device and the server to facilitate Mutual Authentication.

**Table 2:** Notations for the proposed Dynamic Key Based Authentication

Notation	Description
$\parallel$	Concatenation Operation
$\oplus$	Ex-OR Operation
$h$	Message Digest Function
Random Number	128-bit Random Number for Mutually Authentication
Temporary Number (Nonce)	Purpose 128-bit Random Number for Session Key Generation Purpose



**Figure 2:** Message Exchange Sequence for the proposed authentication structure

The communication process between an IoT Device and an IoT Server involves several steps to establish a secure authentication session. Below is a description of these steps.

**1. Initiation of Communication Request:**

- The IoT Device initiates communication by sending a request (M1) to the IoT Server.
- Request message M1 includes a message digest of the Device ID and the Session number, which helps maintain the authentication session.
- $M1 = h(\text{Device ID}) \parallel \text{Session Number}$

**2. Challenge Message Generation by Server:**

- The Server verifies the message digest value for the Device ID.
- If valid, the Server generates a Challenge Message (M2) for the IoT Device.
- M2 contains Challenge1 and a Random number1.
- Challenge1 comprises q distinct numbers, each pointing to an index in an 8x8 Matrix stored in a secure vault.
- The value of q must be less than the total number of keys stored in the vault.
- $\text{Challenge1} = \{C1, C2, C3, \dots, C8\}$
- $M2 = \{\text{Challenge1}, \text{Random Number1}\}$

**3. Response Generation by IoT Device:**

- The IoT Device generates a response for the assigned challenge.
- A temporary key of 128 bits (K1) is generated by performing XOR operations on the key values indexed by the challenge message.
- $\text{Temporary Key } K1 \text{ at IoT Device Side} = K[C1] \oplus K[C2] \oplus \dots \oplus K[Cq]$
- The IoT Device creates a response by encrypting Random Number1 || Temporary Number1 using K1 as the encryption key.
- Here, Temporary Number1 is a 128-bit random number generated by the IoT Device for future use in generating a Session key for subsequent communication.
- $M3 = \text{Enc}(K1, \text{Random Number1} \parallel \text{Temporary Number (Nonce)1} \parallel \{\text{Challenge2, Random Number2}\})$ .
- The IoT Device also generates a separate challenge message (Challenge2) for the IoT Server in a similar manner.

**4. Response Generation by Server:**

- Upon receiving the message from the IoT Device, the Server generates a temporary key (K2) using the indexes from Challenge2 stored in its secure vault.
- No key sharing is required between the IoT Device and the Server.
- After obtaining key K2, the Server decrypts message M3.

- If the Server retrieves Random number1 from M3, it indicates that the receiver of the previous challenge message (M2) was a legitimate IoT Device.
  - The Server then generates a response for the IoT Device's challenge (M3).
  - Message M4 from the Server to the IoT Device is encrypted using temporary key K2 and includes Random Number2 and Temporary Number2.
  - Temporary Key K2 at Server Side =  $K[C1] \oplus K[C2] \oplus \dots \oplus K[Cq]$ .
  - $M4 = \text{Enc}(K2, \text{Random Number2} || \text{Temporary Number (Nonce)2})$ .
5. **Authentication by IoT Device:**
- The IoT Device receives message M4 and decrypts it by generating temporary key K2 from its secure vault, using the content of Challenge C2.
  - If the IoT Device obtains Random number2, it signifies that the Server is also authenticated.
6. **Session Key Generation:**
- After mutual authentication between the IoT Device and the Server, they generate a temporary session key using Temporary Number1 and Temporary Number2.
  - Session Key =  $\text{Temporary Number1} \oplus \text{Temporary Number2}$ .

**Contribution of our Research Work:**

1. The proposed work aims to implement light weight mutual authentication approach for IoT devices which can avoid the possibility of Key Stolen attack, Eavesdropping attack and Location Spoofing attack.
2. The proposed work plans to verify contextual information of a device when it initiates a session with reference node. Parameter AoA- Angle of arrival will be utilized for context matching. So, prevention of Location Spoofing attack can be done at initial stage. It will reduce energy consumption, delay and also intrusion activities during session.
3. The proposed work plan to generate the session key as a part of IoT device authentication in a dynamic way. The working principal for dynamic key generation will be "One Session, One Cipher". It will generate session key on both sides –device and server in a secure, efficient way by considering parameters- Vault and Random number generation. So, prevention of Key Stolen attack and Eavesdropping attack will be possible.

## IV. Security Analysis of the Proposed Method

**Protection against Location Spoofing Attack:**

Proof: The distinguishing feature of the proposed protocol lies in its ability to verify the location of the IoT device, ensuring that authentication requests originate from a known location. Consequently, if an adversary seizes an IoT device and attempts authentication from a remote, unauthorized location, their efforts will be in vain. We have implemented a Localization approach, utilizing location-specific attributes such as AoA (Angle of Arrival), to fortify protection against Location Spoofing attacks.

**Protection against Man-in-the-Middle Attack:**

Proof: A Man-in-the-Middle (MitM) attack involves an attacker intercepting communications between two parties with the intention of secretly eavesdropping on or modifying the transmitted data. The significant feature of the proposed method is that adversaries cannot compute the session key due to the reliance on Random number generation in its generation process. Importantly, in our protocol, the session key is not explicitly transmitted between the Server and the device. Instead, it is computed independently by the device and server at their respective locations. Consequently, adversaries are unable to access the session key required to launch a MITM attack.

**Protection against Replay Attack:**

Proof: The initiation of a new session with a device encompasses both the context-based authentication process and the dynamic key-based authentication approach for key establishment.



During this authentication phase, each device shares a nonce and a Session ID. The Session ID is unique for each new session and serves as a timestamp within our protocol. In the event of an attacker attempting to replay previous session authentication messages, these messages will be discarded due to the presence of an old Session ID that has already expired. Furthermore, the attacker cannot manipulate or update the Session ID as it is transmitted in an encrypted form, with only the destination node, i.e., the Server, possessing the knowledge of it after decryption with its key. Even if an adversary were to submit the same authentication message to the server after a certain period of time, they would not succeed. This is because our protocol generates a new Nonce (Random Number) for each session, rendering any previous nonce random number request for session establishment immediately invalid.

#### **Device Anonymity:**

Proof: In relation to Device Anonymity, our proposed approach refrains from transmitting the actual device identity in any message exchange or communication with the server node. Instead of the device ID, a message digest value of the device ID is transmitted along with the session number. Since the message digest function adheres to a one-way property, it becomes computationally infeasible for an intruder to deduce the device ID from the message without knowledge of the specific hash algorithm used.

#### **Brute force attempts Analysis for the proposed Approach:**

We have securely stored a total of 64 keys, each with a length of 128 bits, in both the IoT device and the Server's vaults. Temporary keys are generated through the XOR operation using these stored keys. Let's calculate the efforts required to derive these Temporary keys.

An intruder needs to select 8 keys out of the total 64 keys, resulting in a total possible combination of  ${}^{64}C_8$ , calculated as follows:

$$\begin{aligned} {}^{64}C_8 &= 64! / (64-8)! 8! \\ &= 64! / 56! 8! \\ &= 64 \cdot 63 \cdot 62 \cdot 61 \cdot 60 \cdot 59 \cdot 58 \cdot 57 / 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 \\ &= 17, 84, 62, 98, 76, 37, 760 / 40, 3 20 \end{aligned}$$

Total Possible key combinations at IoT device side = 4,42,61,65,368.

Similarly, total possible key combinations at Server side for selecting 8 different keys from 64 keys vault to generate second temporary key = 4,42,61,65,368.

Total computations required to capture both temporary key from vault = 8,85,23,30,736.

Assuming that an intruder can perform 1 million computations in 1 hour, it would take them a total of 8,852.33 hours or approximately 368 days to recover Temporary Key 1 and Temporary Key 2 from the vault. This is a significant time frame, and since we also update vault values regularly, our suggested schema provides security against Key-stolen attacks.

Even if an adversary possesses knowledge about the dynamic key authentication approach, it remains computationally infeasible for them to directly derive the session key.

## V. Conclusion

The Internet of Things (IoT) encompasses a multitude of physical devices capable of seamless data exchange. These devices connect directly to the web, operating in an open environment, which presents opportunities for intruders to launch various cyber-attacks. IoT security is a critical research domain that engages both academic and industry researchers. Within the realm of IoT security, the CIA Model—Confidentiality, Integrity, and Authentication—is of paramount importance. Authentication, in particular, plays a central role in ensuring the security of IoT networks as it uniquely identifies each device connected to the network. In our investigation, we thoroughly examined the challenges inherent in existing IoT authentication algorithms. We uncovered potential cyber threats, including Replay attacks, Man-in-the-Middle (MITM) attacks, Location Spoofing attacks, and Key Stolen attacks, which can compromise the security of current IoT authentication architectures. Furthermore, we conducted an in-depth review of the work conducted by various

experts in the field of authentication. Through this review, we pinpointed research gaps that still exist in the domain of IoT authentication, highlighting opportunities for researchers to contribute their expertise and develop precise and efficient security solutions. There is a pressing need for the creation of an efficient IoT Authentication Multi-factor algorithm that is lightweight—demanding fewer resources—and is rooted in context verification and dynamic key generation approaches. To substantiate our proposal, we conducted an informal security analysis, demonstrating that our approach effectively safeguards against Key Stolen, MITM, and Replay threats. Furthermore, we established that it is computationally infeasible for an intruder to breach our suggested approach within a finite timeframe and with limited resources.

## References

- [1] Hokeun Kim and Edward A. Lee (2017). Authentication and Authorization for the Internet of Things, *IEEE Internet of Things Journal*, 19: 27-33.
- [2] Mohammad Wazid, Ashok Kumar Das, Vanga Odelu, Neeraj Kumar, Mauro Conti, Minh Jo (2017). Design of Secure User Authenticated Key Management Protocol for Generic IoT Networks, *IEEE Internet of Things Journal*, 5:269-282.
- [3] Ning Wang, Ting Jiang, ShichaoLv and Liang Xiao, Senior Member (2017). Physical-Layer Authentication Based on Extreme Learning Machine. *IEEE Communications*, 21:1557-1560.
- [4] Muhammad Naveed Aman, Sachin Taneja, Biplab Sikdar, Kee Chaing Chua, and Massimo Alioto (2019). Token-Based Security for the Internet of Things With Dynamic Energy-Quality Tradeoff, *IEEE Internet of Things Journal*, 6:2843-2859.
- [5] Vikas Hassija, Vinay Chamola ,Vikas Saxena , Divyansh Jain, Pranav Goyal, And Biplab Sikdar (2019). A Survey on IoT Security: Application Areas, Security Threats, and Solution Architectures, *IEEE Access*, 7:82721-82743.
- [6] Prosanta Gope and Biplab Sikdar (2018). Lightweight and Privacy-Preserving Two-Factor Authentication Scheme for IoT Devices, *IEEE Internet of Things Journal*, 6:580-589.
- [7] Sulabh Bhattarai and Yong Wang (2018). End-to-End Trust and Security for Internet of Things Applications, *IEEE Computer Society*, 51:20-27.
- [8] Muhammad Naveed Aman, Mohamed Haroon Basheer and Biplab Sikdar (2019). Two factor Authentication for IOT with Location Information, *IEEE Internet of Things Journal*, 6(2): 3335-3351.
- [9] Yan Zhao, Shiming Li and Liehui Jiang (2018) Secure and Efficient User Authentication Scheme Based on Password and Smart Card for Multi-server Environment, *WILEY Hindawai Security and Communication Networks*, 18:1-13.
- [10] Majid Alotaibi (2018). An Enhanced Symmetric Cryptosystem and Biometric-Based Anonymous User Authentication and Session Key Establishment Scheme for WSN, *IEEE Access*, 6:70072-70087.
- [11] Zahoor Ahmed Alizai, Noquia Fatima Tareen and Iqura Jadoon (2018). Improved IoT Device Authentication Scheme Using Device Capability and Digital Signatures, *IEEE International Conference on Applied and Engineering Mathematics*, <https://doi.org/10.1109/ICAEM.2018.8536261>.
- [12] Moritz Loske, Lukas Rothe and Dominik Gertler (2019). Context-Aware Authentication: State-of-the-Art Evaluation and Adaption to the IIoT, *IEEE 5th World Forum on Internet of Things (WF-IoT)*, <https://doi.org/10.1109/WF-IoT.2019.8767327>.
- [13] Armin Babaei, Gregor Schiele (2019). Physical Unclonable Functions in the Internet of Things: State of the Art and Open Challenges, *Sensors*, 19 (14):3208 <https://doi.org/10.3390/s19143208>.
- [14] Baibhab Chatterjee, Shovan Maity (2019) RF-PUF: Enhancing IoT Security through Authentication of Wireless Nodes using In-situ Machine Learning, *IEEE Internet of Things Journal*, 6(1): 388-398.
- [15] Tarak Nandy, Norjihan Abdul Ghani and Sananda Bhattacharya (2019). Review on Security of Internet of Things Authentication Mechanism, *IEEE Access*, 7: 151054-151089.

- [16] Santosh Krishna B V and Gnanasekaran T (2017). A Systematic Study of Security Issues in Internet-of-Things (IoT), *IEEE International conference on I-SMAC*, <https://doi.org/10.1109/I-SMAC.2017.8058318>.
- [17] Mardiana binti Mohamad Noor, Wan Haslina Hassan (2019). Current research on Internet of Things (IoT) security: A survey, *ELSEVEIR Computer Networks*, 148: 283-294.
- [18] Chang-le Zhong, Zhen Zhu and Ren-gen Huang (2017). Study on the IOT Architecture and Access Technology, *IEEE 16th International Symposium on Distributed Computing and Applications to Business, Engineering and Science*, <https://doi.org/10.1109/DCABES.2017.32>.
- [19] Jeffrey Voas, Bill Agresti (2018). A Closer Look at the IoT's "Things", *IEEE Computer Society*, 20 (3): 11-14.
- [20] Jyoti Deogirikar and Amarsinh Vidhate (2017). Security Attacks in IoT: A Survey, *IEEE International conference on I-SMAC*, <https://doi.org/10.1109/I-SMAC.2017.8058363>.
- [21] Zhiping Jiang, Kun Zhao and Junzhao Du (2020). PHYAlert: identity spoofing attack detection and prevention for a wireless edge network, *Journal of Cloud Computing*, 9 (5):1-13.

# M/M/ $\infty$ QUEUE WITH IMPATIENT CUSTOMERS

GULAB SINGH BURA



Department of Mathematics and Statistics, Banasthali Vidyapith, Rajasthan, INDIA  
[gulabsingh@banasthali.in](mailto:gulabsingh@banasthali.in)

## Abstract

*In this paper we proposed an M/M/ $\infty$  queue with impatient customers. Generally, customers are impatient due to long waits in queue but in this work, we consider the case when customers are not impatient due to long waits but they are impatient due to the poor quality of service. We model and analyze this queueing system by using continued fraction technique and obtained the probability mass function of the customers present in the system in time dependent form. Also, we calculate the average queue size. Finally, some graphical representations are given to illustrate the model.*

**Keywords:** M/M/ $\infty$  Queue, Transient Solution, Impatient Customers, Laplace Transform.

## 1. INTRODUCTION

We present an M/M/ $\infty$  queue with impatient customers. The impatient behavior of customer is common in many real life queueing situations such as in hospital during emergencies, inventory systems, telecommunications system etc. When the waiting time is sufficiently large or intolerable the customers may become impatient and decide to leave (i.e. balk or renege) the system before being served. The study of queueing models with impatient customers play an important role in many revenue generating queueing system. There is an extensive literature available on queues with impatient customers (see e.g., [4], [5], [14]). First attempt in this field was made by Haight [14]. After that, Al-Seedy et.al. [2] obtained the transient behavior of an M/M/1 queue with balking. The single server Markovian queue with reneging was proposed by Haight [15]. Ancker et.al. ([4], [5]) considered an M/M/1/N queue with both balking and reneging simultaneously. Multi-server queueing model with impatient customers was investigated by Varshney et.al. [23]. Time dependent solution of the M/M/c queueing model was proposed by Al-Seedy et.al. [3]. The concept of balking with heterogeneous servers have been proposed by Abou El-Ata [6] and Singh [19]. Queues with catastrophes and impatient customers have been investigated by various authors. Yechiali [24] consider the case of impatient customers when server is down due to catastrophes. Sudhesh [22] extends the work of Yechiali [24] and obtained the time dependent solution. Altman and Yechiali [1] considered an infinite server queueing system with impatient customers under the situation where servers are free and doing some additional task. A GI/G/1 queue with disaster and customer impatient was studied by Chakravarthy [12]. Customers impatient due to priority has been analyzed by Choi et. al. [11]. Sudhesh et. al. [20] obtained the transient solution of two heterogeneous servers queue with impatient customers when server is down due to the occurrence of breakdown. Vacation queueing model are also analyzed with impatient behavior of customers by various researchers. Ammar [8] obtained the transient solution of a waiting server, vacation queueing model with impatient customers. Sampath et. al. [21] extends the work of Ammar [8] by considering multiple vacation in place of single vacation. Perel and Yechiali [17] give the steady state solution of an M/M/c ( $c=1, 1 < c < \infty, c = \infty$ ) queue with slow server and impatient behavior of

customers in two random phases. Generally it is assumed that customers are impatient due to long waits in queue but in this work authors considered that customers may be impatient due to slow service rate. In our work we also consider that customers may be impatient due to poor quality of service provided by the servers and obtained the time dependent solution of an M/M/∞ queue with impatient customers. Generally, in self service models or an infinite server models there is no question of impatient behavior of customers because the entering customers immediately get service and there is no waiting line in the system. But in our case the customer is impatient not due to long waits but due to the quality of the service provided by the server. It may possible that a customer is impatient due to poor quality of service. The motivation for studying this model comes from the field of telecommunications. Let us consider a university campus which is providing a free Wi-Fi service for their students at their campus. Every student entering in the campus may use the free Wi-Fi service. The service starts as soon as he joins the campus and carries until the end of the campus. A small university campus contains thousands of such self served customers who are using this service. Each customer receives identical quality of signals of Wi-Fi connection. The quality of signals may vary and fluctuate randomly. The poor quality of signals causes the impatient behavior of customers. Whenever a customer enters into the campus and finds poor quality of signals of the Wi-Fi connection, he may decide to leave the system without getting served i.e. customer balk from the system. On other hand, he joins the service but leaves the system due to poor quality of Wi-Fi connection, this also becomes a case of customer renege from the system. Hence, our operating model is a suitable preposition.

## 2. MATHEMATICAL MODEL

M/M/∞ queue with impatient customers is in operation. Arrivals occur one by one in a Poisson stream with mean rate  $\alpha$ . There are infinite servers and service time are exponentially distributed with parameter  $\beta$ . Capacity of the system is infinite. After entering the system, the customers either decide to join the service with probability  $\theta$  or balk with probability  $1 - \theta$ , where  $0 \leq \theta < 1$ . After joining the service, if he finds a poor quality of signals of Wi-Fi connection then, the customer will wait for a certain length of time  $T$ , exponentially distributed with parameter  $\gamma$ , for improving the quality of service. If it has not improved by then, the customer abandons and leave the system without getting complete service. Let  $P_n(t)$  be the probability that the random variable  $N(t)$  assumes the value  $n$  i.e.

$$P_n(t) = P(N(t) = n)$$

## 3. TRANSIENT SOLUTION

In this section, we provide the transient solution of the presented queueing model. For this, the differential- difference equations are given as:

$$P'_0(t) = -(\alpha\theta)P_0(t) + (\beta + \gamma)P_1(t) \tag{1}$$

$$P'_n(t) = -(\alpha\theta + n(\beta + \gamma)) P_n(t) + (n + 1) (\beta + \gamma) P_{n+1}(t) + \theta\alpha P_{n-1}(t), n \geq 1. \tag{2}$$

Initially, at  $t=0$ ,

$$P_n(0) = \begin{cases} 1 & \text{if } n = 0; \\ 0 & \text{otherwise .} \end{cases} \tag{3}$$

Laplace transformation of Eq.(2) with initial condition Eq.(3) results the following equation

$$(s + \theta\alpha + n(\beta + \gamma)) P_n^*(s) = (n + 1) (\beta + \gamma) P_{n+1}^*(s) + \theta\alpha P_{n-1}^*(s) \tag{4}$$

After simplification, Eq.(4), gives

$$\frac{P_n^*(s)}{P_{n-1}^*(s)} = \frac{\theta\alpha}{(s + \theta\alpha + n(\beta + \gamma) - (n + 1)(\beta + \gamma) \frac{P_{n+1}^*(s)}{P_n^*(s)}} \tag{5}$$

$$= \frac{\frac{\theta\alpha}{\beta+\gamma}}{\left(\frac{s}{\beta+\gamma} + \frac{\theta\alpha}{\beta+\gamma} + n\right) - \frac{(n+1)\frac{\theta\alpha}{\beta+\gamma}}{\left(\frac{s}{\beta+\gamma} + \frac{\theta\alpha}{\beta+\gamma} + (n+1)\right) - \frac{(n+2)\frac{\theta\alpha}{\beta+\gamma}}{\left(\frac{s}{\beta+\gamma} + \frac{\theta\alpha}{\beta+\gamma} + (n+2)\right)} - \dots} \quad (6)$$

Now using the identity given by Lorentzen and Waadeland [16]

$$\frac{{}_1F_1(c+1; q+1; z)}{{}_1F_1(c; q; z)} = \frac{q}{q-z} \frac{(c+1)z}{q-z+1} \frac{(c+2)z}{q-z+2} \dots \quad (7)$$

Use of Eq.(7) in Eq.(6), gives

$$\frac{P_n^*(s)}{P_{n-1}^*(s)} = \frac{\theta\alpha}{(\beta+\gamma)} \frac{{}_1F_1(n+1; \frac{s}{\beta+\gamma} + n + 1; \frac{-\theta\alpha}{\beta+\gamma})}{{}_1F_1(n; \frac{s}{\beta+\gamma} + n; \frac{-\theta\alpha}{\beta+\gamma})}, \quad (8)$$

therefore for  $n \geq 1$ , we have

$$P_n^*(s) = \left(\frac{\theta\alpha}{(\beta+\gamma)}\right)^n \frac{{}_1F_1(n+1; \frac{s}{\beta+\gamma} + n + 1; \frac{-\theta\alpha}{\beta+\gamma})}{\prod_{i=1}^n \left(\frac{s}{\beta+\gamma} + i\right) {}_1F_1(1; \frac{s}{\beta+\gamma} + 1; \frac{-\theta\alpha}{\beta+\gamma})} P_0^*(s), \quad (9)$$

$$P_n^*(s) = \zeta_n^*(s) P_0^*(s), \quad (10)$$

where

$$\zeta_n^*(s) = \left(\frac{\theta\alpha}{(\beta+\gamma)}\right)^n \frac{{}_1F_1(n+1; \frac{s}{\beta+\gamma} + n + 1; \frac{-\theta\alpha}{\beta+\gamma})}{\prod_{i=1}^n \left(\frac{s}{\beta+\gamma} + i\right) {}_1F_1(1; \frac{s}{\beta+\gamma} + 1; \frac{-\theta\alpha}{\beta+\gamma})}. \quad (11)$$

It is well known that

$$\sum_{n=0}^{\infty} P_n^*(s) = \frac{1}{s}, \quad (12)$$

by the use of Eq.(10) in Eq.(12), we get

$$P_0^*(s) = \frac{1}{s} \left[ 1 + \sum_{n=1}^{\infty} \zeta_n^*(s) \right]^{-1} \quad (13)$$

$$P_0^*(s) = \frac{1}{s} \left[ \sum_{k=0}^{\infty} \left( \sum_{n=1}^{\infty} \zeta_n^*(s) \right)^k \right], \quad (14)$$

after taking inverse Laplace transform of Eq.(10), we get

$$P_n(t) = \zeta_n(t) * P_0(t), \quad (15)$$

where the symbol \* denotes the convolution and

$$P_0(t) = \int_0^t \sum_{k=0}^{\infty} \left( \sum_{n=1}^{\infty} \zeta_n(y) \right)^k dy. \quad (16)$$

Next we derive the expression for  $\zeta_n(t)$ , where  $\zeta_n(t)$  represents the inverse Laplace transform of  $\zeta_n^*(s)$ .

From Eq.(11)

$$\zeta_n^*(s) = \left(\frac{\theta\alpha}{(\beta+\gamma)}\right)^n \frac{{}_1F_1(n+1; \frac{s}{\beta+\gamma} + n + 1; \frac{-\theta\alpha}{\beta+\gamma})}{\prod_{i=1}^n \left(\frac{s}{\beta+\gamma} + i\right) {}_1F_1(1; \frac{s}{\beta+\gamma} + 1; \frac{-\theta\alpha}{\beta+\gamma})}.$$

It is well known that

$${}_1F_1\left(n+1; \frac{s}{\beta+\gamma} + n+1; \frac{-\theta\alpha}{\beta+\gamma}\right) = \sum_{k=0}^{\infty} \frac{(n+1)_k \left(-\frac{\theta\alpha}{\beta+\gamma}\right)^k}{\left(\frac{s}{\beta+\gamma} + n+1\right)_k k!}$$

where  $(\beta)_k$  represents the Pochhammer symbol, i.e.

$$(\beta)_k = \begin{cases} 1 & \text{if } k = 0; \\ \beta(\beta+1)(\beta+2)\dots(\beta+k-1) & \text{if } k = 1, 2, 3, \dots \end{cases}$$

Therefore

$$\frac{{}_1F_1\left(n+1; \frac{s}{\beta+\gamma} + n+1; \frac{-\theta\alpha}{\beta+\gamma}\right)}{\prod_{i=1}^n \left(\frac{s}{\beta+\gamma} + i\right)} = \sum_{k=0}^{\infty} \frac{\binom{n+k}{k} \left(-\frac{\theta\alpha}{\beta+\gamma}\right)^k}{\prod_{i=1}^{n+k} \left(\frac{s}{\beta+\gamma} + i\right)}$$

Applying partial fraction expansion, the above equation can be written as

$$\frac{{}_1F_1\left(n+1; \frac{s}{\beta+\gamma} + n+1; \frac{-\theta\alpha}{\beta+\gamma}\right)}{\prod_{i=1}^n \left(\frac{s}{\beta+\gamma} + i\right)} = (\beta+\gamma) \sum_{k=0}^{\infty} \binom{n+k}{k} \left(-\frac{\theta\alpha}{\beta+\gamma}\right)^k \sum_{i=1}^{n+k} \frac{(-1)^{i-1}}{(n+k-i)!(i-1)!(s+i(\beta+\gamma))}. \tag{17}$$

Also

$${}_1F_1\left(1; \frac{s}{\beta+\gamma} + 1; \frac{-\theta\alpha}{\beta+\gamma}\right) = \sum_{k=0}^{\infty} (-\theta\alpha)^k d_k^*(s),$$

where

$$d_k^*(s) = \frac{1}{\prod_{i=1}^k (s+i(\beta+\gamma))} \text{ and } d_0^*(s) = 1.$$

By the use of the identity given in Srivastava and Kashyap [18]

$$\frac{1}{{}_1F_1\left(1; \frac{s}{\beta+\gamma} + 1; \frac{-\theta\alpha}{\beta+\gamma}\right)} = \sum_{k=0}^{\infty} (\theta\alpha)^k e_k^*(s), \tag{18}$$

where  $e_0^*(s) = 1$ , and for  $k=1,2,3,\dots$

$$e_k^*(s) = \begin{vmatrix} d_1^*(s) & 1 & & & & & \\ d_2^*(s) & d_1^*(s) & 1 & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ d_{k-1}^*(s) & d_{k-2}^*(s) & d_{k-3}^*(s) & \cdot & \cdot & d_1^*(s) & 1 \\ d_k^*(s) & d_{k-1}^*(s) & d_{k-2}^*(s) & \cdot & \cdot & d_2^*(s) & d_1^*(s) \end{vmatrix}$$

$$= \sum_{j=1}^k (-1)^{j-1} d_j^*(s) e_{k-j}^*(s).$$

By substituting Eq.(17) and Eq.(18) in Eq.(11), we get

$$\zeta_n^*(s) = (\theta\alpha)^n \sum_{j=0}^{\infty} (-\theta\alpha)^j \binom{n+j}{j} d_{n+j}^*(s) \sum_{k=0}^{\infty} (\theta\alpha)^k e_k^*(s).$$

On inversion, we obtain

$$\zeta_n(t) = (\theta\alpha)^n \sum_{j=0}^{\infty} (-\theta\alpha)^j \binom{n+j}{j} d_{n+j}(t) \sum_{k=0}^{\infty} (\theta\alpha)^k e_k(t), \tag{19}$$

where

$$d_k(t) = \frac{1}{(\beta + \gamma)^{k-1}} \sum_{i=1}^k \frac{(-1)^{i-1}}{(k-i)!(i-1)!} e^{-i(\beta+\gamma)t}, k = 1, 2, 3, \dots,$$

and

$$e_k(t) = \sum_{i=1}^k (-1)^{i-1} d_i(t) * e_{k-i}(t), \quad k = 2, 3, 4, \dots; \quad e_1(t) = d_1(t)$$

#### 4. TIME DEPENDENT MOMENTS

##### 4.1. Mean

Let  $A(t)$  represents the average value of the random variable  $N(t)$ , therefore

$$A(t) = E(N(t)) = \sum_{n=1}^{\infty} nP_n(t) \tag{20}$$

Initially, at  $t=0$ , Eq(20) gives

$$A(0) = 0,$$

which implies

$$A'(t) = \sum_{n=1}^{\infty} nP'_n(t), \tag{21}$$

where  $A'(t)$  denote the differentiation of  $A(t)$ .

Application of Eq.(2) in Eq.(21),after some calculation gives

$$A'(t) + (\beta + \gamma)A(t) - \theta\alpha = 0. \tag{22}$$

which is a linear differential equation in  $A(t)$ , whose solution gives

$$A(t) = \frac{\theta\alpha}{\beta + \gamma} [1 - e^{-(\beta+\gamma)t}] \tag{23}$$

##### 4.2. Variance

Let  $\text{Var}(N(t))$  represents the variance of the random variable  $N(t)$ , therefore

$$\text{Var}(N(t)) = E[N(t) - E(N(t))]^2$$

which may be written as

$$\text{Var}(N(t)) = b(t) - [A(t)]^2, \tag{24}$$

where

$$b(t) = E(N^2(t)) = \sum_{n=1}^{\infty} n^2 P_n(t),$$

with

$$b(0) = 0,$$

also

$$b'(t) = \sum_{n=1}^{\infty} n^2 P'_n(t) \tag{25}$$

substitution of  $P'_n(t)$  in Eq.(25), after some calculation results in the form of a linear differential equation in  $b(t)$  i.e.

$$b'(t) = -(2\beta + \gamma)b(t) + (2\theta\alpha + \beta + \gamma)M(t) + \theta\alpha \tag{26}$$



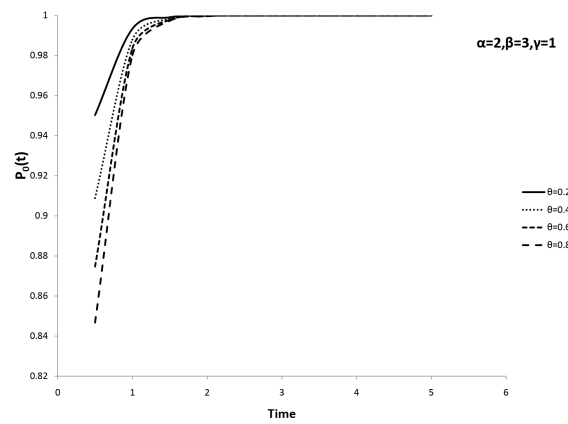
which after integration gives

$$b(t) = \frac{(2\theta\alpha + \beta + \gamma)\theta\alpha(\beta - (\gamma + 3\beta)e^{-(\gamma+2\beta)t} + (\gamma + 2\beta))e^{-(\gamma+\beta)t}}{\beta(\gamma + 2\beta)(\gamma + \beta)} + \frac{\theta\alpha}{(\gamma + 2\beta)}[1 - e^{-(\gamma+2\beta)t}]. \tag{27}$$

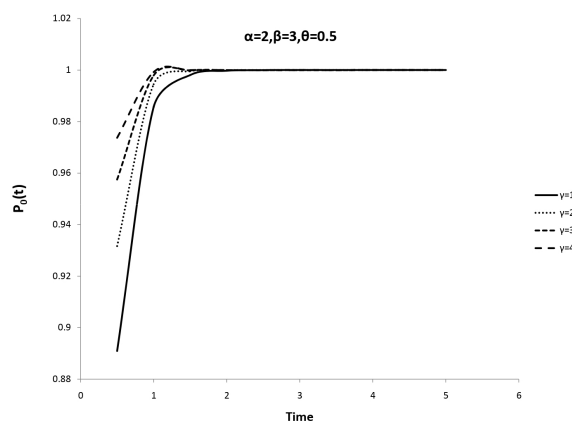
Substitution of Eq.(27) in Eq.(24), gives the expression of Var(N(t)).

### 5. GRAPHICAL ILLUSTRATIONS

In this section, we presents some graphical results to observe the time dependent behavior of various probabilities and average number of customers in the system.



**Figure 1:**  $P_0(t)$  versus Time



**Figure 2:**  $P_0(t)$  versus Time

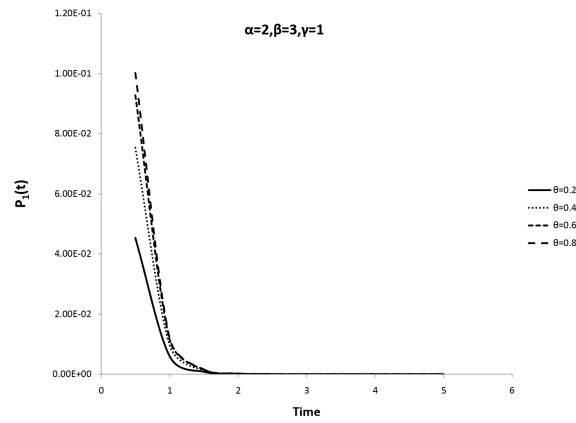


Figure 3:  $P_1(t)$  versus Time

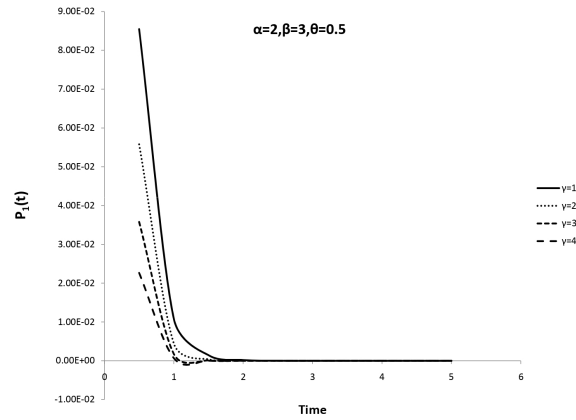


Figure 4:  $P_1(t)$  versus Time

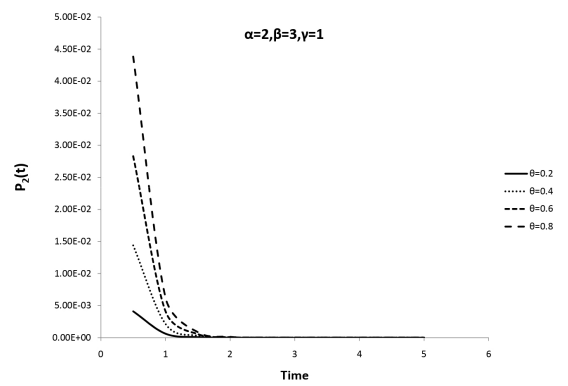


Figure 5:  $P_2(t)$  versus Time

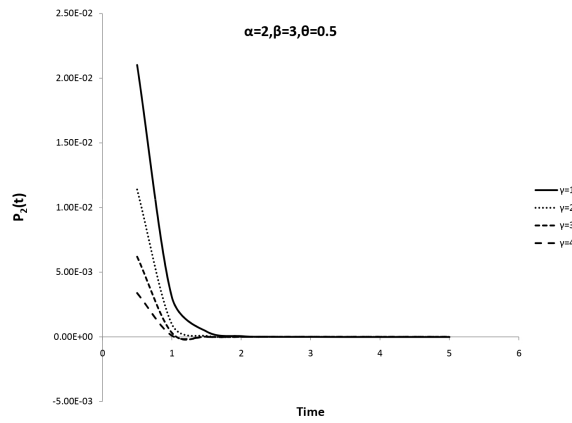


Figure 6:  $P_2(t)$  versus Time

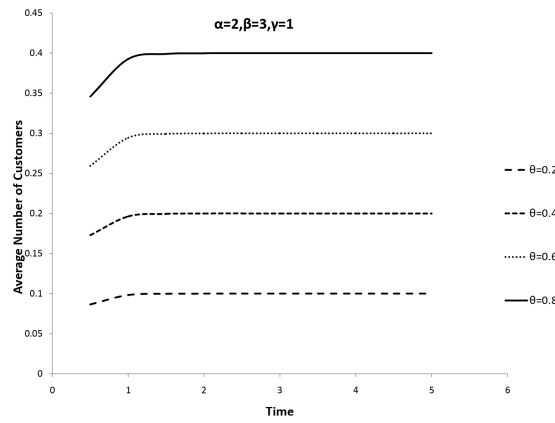


Figure 7: Average Number of Customers versus Time

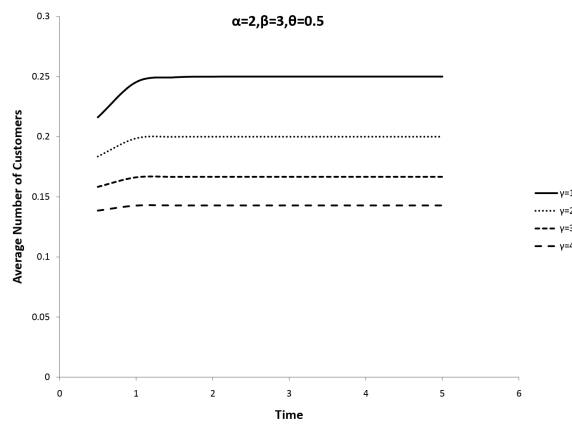


Figure 8: Average Number of Customers versus Time

Fig.(1 to 6) illustrates that as time increases, all the probability curves except  $P_0(t)$  are decreases initially and then attain the steady state after  $t = 2$ . Also we notice that the probabilities  $P_1(t)$  and  $P_2(t)$  increases with increasing  $\theta$  and the probability  $P_0(t)$  decreases while we increase the value of  $\theta$ . Further, we observe that if  $\gamma$  increases the probability of an empty system i.e.  $P_0(t)$  increases while the other probabilities decrease. Fig.(7 and 8 ) explain the situation that the average number of customers in the system increases with time initially and then finally attains steady state. The average number of customers increase with the increasing values of  $\theta$  and decrease with the increasing values of  $\gamma$ .

## 6. CONCLUSION

impatient behavior of customers is common in many real life queueing situations. Approximately, in all previous work available in the literature, it has been assumed that customers are impatient due to long waits in queue. But in the present study we analyze the case in which customers are impatient due to the quality of the service provided by the server. We have obtained the probability mass function of the number of customers in time dependent form. Also, we have determined the transient mean and variance of the number of customers. At the end, for observing the time dependent behavior of various probabilities, we provide some graphical illustrations.

## REFERENCES

- [1] Altman, E and Yechiali, U (1983). Infinite-server queues with system's additional task and impatient customers. *Probability in the Engineering and Informational Sciences*, 2: 292–303.
- [2] Al-Seedy, R. O. and Kotb, K.A. M.(1992). Transient solution of the state dependent queue M/M/1 with balking. *AMSE.*, 35:55–64.
- [3] Al-Seedy, R. O., El-Sherbiny,A. A, El-Shehawy, S. A. and Ammar, S. I(2009). Transient solution of the M/M/c queue with balking and reneging *Computers and Mathematics with applications*,57: 1280–1285.
- [4] Ancker Jr, C. J. and Gafarian, A. V (1963). Some queueing problem with balking and reneging:I *Operations Research*, 11: 88–100.
- [5] Ancker Jr, C. J. and Gafarian, A. V (1963). Some queueing problem with balking and reneging:II *Operations Research*, 11: 928–937.
- [6] Abou El-Ata, M. O. (1983). On Poisson queues with both balking and heterogeneous servers. *Delta J. Sci.*, 7: 292-303.
- [7] Abate J. and Whitt W. (1999). Computing Laplace transforms for numerical inversion via continued fraction. *INFORMS J.Comput*,11:394–405. Andrews, L. W. Special function of Mathematics for engineers, McGraw Hill., Singapore 1992.
- [8] Ammar,S. I.(2017). Transient solution of an M/M/1 vacation queue with a waiting server and impatient customers. *Journal of Egyptian Mathematical society*, 25: 337–342.
- [9] Bura, Gulab Singh (2019). Transient solution of an M/M/ $\infty$  queue with catastrophes. *Communication in Statistics-Theory and Methods*, 48(14): 3439–3450.
- [10] Bura, Gulab Singh (2022). M/M/ $\infty$  queue with catastrophes and repairable servers. *Reliability:Theory and Methoods*, 17:143–153.
- [11] Choi,B. D.,Kim, B. and Chung, J.(2001). M/M/1 Queue with Impatient Customers of Higher Priority *Queueing Systems*, 38: 49–66.
- [12] Chakravarthy, S. R.(2009). A disaster queue with Markovian arrivals and impatient customers. *Applied Mathematics and Computation*, 214: 48–59.
- [13] Gradshteyn,I. S. and Ryzhik,I. M. Table of integrals, Series, and products, Academic Press, 2000.
- [14] Haight, F. A.(1957). Queueing with balking *Biometrika*, 44:360–369.
- [15] Haight, F. A.(1959). Queueing with reneging *Metrika*,2: 186-197.
- [16] lorentzen, L.and Waadeland, H. Continued fraction with application. Studies in computational mathematics, Elsevier,Amsterdam 1992.

- [17] Perel, N., Yechiali, U.(2010). Queues with slow servers and impatient customers. *European Journal of Operational Research.*, 201: 247–258.
- [18] Srivastava H. M. and Kashyap B. R. K. Special functions in queuing theory, Academic Press, New York 1982.
- [19] Singh, V. P. (1969). Two server Markovian queue with balking: heterogeneous vs homogeneous servers. *Operations Research*, 18: 145–159.
- [20] Sudhesh, R., Savithal, P.and Dharmaraja, S. (2017). Transient analysis of a two-heterogeneous servers queue with system disaster, server repair and customers impatience. *TOP* 25: 179–205.
- [21] Sampath,M. I. G. S., Kalidass, K. and Liu, J.(2020). Transient analysis of an M/M/1 queueing system subjected to multiple differentiated vacation, impatient customers and a waiting server with application to IEEE 802.16E power saving mechanism. *Indian Journal of Pure and Applied Mathematics*, 51: 297–320.
- [22] Sudhesh, R. (2010). Transient analysis of a queue with system disaster and customer impatience. *Queueing Syst.*, 66: 95–105.
- [23] Varshney, K., Jain, M. and Sharma, G. C. (1989). A multi-server queueing model with balking and reneging via diffusion approximation. *journal of Physical and natural Science*, 10: 10–15.
- [24] Yechiali, U. (2007). Queues with system disaster and impatient customers when system is down. *Queueing Syst.*, 56: 95–102.

# REGULARITY OF ALTERNATE QUADRA SUBMERGING POLAR FUZZY GRAPH AND ITS APPLICATION

ANTHONI AMALI A<sup>1\*</sup>, J . JESINTHA ROSLINE <sup>2</sup>

•  
Auxilium College(Autonomous), Vellore - 632006,  
Affiliated to Thiruvalluvar University, Serkadu, Tamil Nadu, India,  
anthoniamaliasir@gmail.com, jesi.simple@gmail.com

## Abstract

*Fuzzy soft sets and graphs are invented to solve uncertain problems in the field of Applied mathematics. It is a general mathematical tool introduced with many parameters to model the vagueness of the changing world. The insight learning of the AQSP fuzzy soft graphs paved the way to discover the extension of the AQSP fuzzy soft graph. In this research article we introduce the Regularity of AQSP fuzzy soft graph with definitions, theorems, properties, and real-life applications. The aim of this invention is mainly to obtain the parametric values in submerging level of confidence  $[-0.5, 0.5] \subset [-1,1]$ . The scope of this new AQSP fuzzy soft graph is to solve the imprecise problems in the field of Mathematical Engineering, Bio Mathematics, Economics, Medical Science, Artificial Intelligence and Machine learning. The regularity of AQSP fuzzy soft graph is combined with the concepts of regular, totally regular, and perfectly regular. The application of this new graph is developed for governing of the women safety vehicle network in different spots with membership submerging values. The future extension can be applied in Approximate reasoning, Mathematical psychology, Decision making for medical diagnosis.*

**Keywords:** Regular AQSP fuzzy soft graph, Totally regular, Perfectly regular AQSP fuzzy soft graph, Alternate Quadra Submerging level of confidence.

## 1. INTRODUCTION

The concept of graph theory was introduced by Euler in 1736. He concreted the way to find the solution of Konigsberg bridge problem. In 1965 Zadeh[20] invented Fuzzy set theory as a mathematical fuzzy tool for handling uncertainties like vagueness, ambiguity, and imprecision in linguistic variables. Fuzzy set has resulted as a potential area of interdisciplinary exploration and the fuzzy graph theory is of modern inducement. The first definition of fuzzy graph was determined by Kaufmann[10] in 1973, based on Zadeh's fuzzy relation in 1971. In 1975, Rosenfeld[16] introduced the concept of fuzzy graph. The structure of fuzzy graphs, using fuzzy relations, obtaining contrasts of several graph hypothetical concepts are the masterpiece of Rosenfeld. Operations on fuzzy graphs were exposed by J.N.Moderson[14] and C.S.Peng. A.Nagoorgani[8] and K.Radha[9] invented the concept of regular fuzzy graphs in 2008.

In 1999, D.Molodtsov[12] intended the notion of soft set theory to solve complicated uncertain problems in Applied Mathematics, Engineering and Environmental studies. In 2001, P.K.Maji[11], initiated the concept of fuzzy soft sets. Zou and Xio discussed the application of the fuzzy soft sets in an imprecise scenario. Later, Akram[4] and Nawaz[15] presented new ideas known as fuzzy soft graphs. A.Pouhassani[24] and H.Doostie studied degree, total degree, regularity and total regularity of fuzzy soft graph and its properties. Regular fuzzy soft graphs and its related properties are investigated by B.Akhilandeswari. The concepts of fuzzy bipolar

soft sets and bipolar fuzzy soft sets have been introduced by Naz and Shabir. Aslam et al studied some basic operations on bipolar fuzzy soft sets.

In this article, we portray a new mathematical fuzzy graph model AQSP Fuzzy Soft graph for dealing imprecise information by integrating the concepts of fuzzy graph and fuzzy soft graphs. We estimate the regularities of AQSP fuzzy soft graphs and some of their characteristics and properties. Here Regular AQSP fuzzy soft graphs, and totally regular AQSP fuzzy soft graphs and perfectly regular AQSP fuzzy soft graphs are examined. Total degree of an AQSP fuzzy soft graph is designed. Theorems for regular AQSP fuzzy soft graphs and totally regular AQSP fuzzy soft graphs are presented. A necessary condition under which they are equivalent is provided. Some properties of regular AQSP fuzzy soft graphs, perfectly regular AQSP fuzzy soft graphs are reviewed with real life applications. The perception of AQSP fuzzy soft graph membership values with submerging level of confidence is applicable in Machine learning and medical psychology. We explored the AQSP fuzzy soft graph module in Governing of women safety police vehicle network with membership score functions.

## 2. PRELIMINARIES

### 2.1. Fuzzy Graph [16]

Let  $\tilde{U}$  is a non-empty set. A fuzzy graph is a set of two of functions  $G : (\sigma, \mu)$  where  $\sigma$  is a fuzzy subset of  $\tilde{U}$ ,  $\mu$  is a symmetric fuzzy relation on  $\sigma$ , where  $\sigma : \tilde{U} \rightarrow [0, 1]$  and the edge set  $\mu : \tilde{U} \times \tilde{U} \rightarrow [0, 1]$  such that,  $\mu(x, y) \leq \min(\mu(x), \mu(y)) \forall x, y \in \tilde{U}$ . The underlying crisp graph of fuzzy graph  $G : (\sigma, \mu)$  is with the notion  $G^* : (\sigma^*, \mu^*)$  where  $\sigma^*$  is denoted as the non-empty set  $\tilde{U}$  of vertices and  $\mu^* = E \in V \times V$ .

### 2.2. Fuzzy Soft graph [13]

A fuzzy soft graph  $G = (G^*, F, K, A)$  is a four tuple such that

1.  $G^* = (V, E)$  is a simple graph.
2.  $A$  is a non empty set of parameters.
3.  $(F, A)$  is a fuzzy soft vertex set  $V$ .
4.  $(K, A)$  is a fuzzy soft edge set  $E$ .
5.  $F(a), K(a)$  is a fuzzy soft graph of  $G^* \forall a \in A$ .

Then it satisfies the condition,  $K(a)(x, y) \leq F(a)(x) \wedge F(a)(Y) \forall a \in A$  and  $(x, y) \in V$ .

### 2.3. Fuzzy soft graph degree of a vertex [4]

Let  $G = (G^*, F, K, A)$  be a fuzzy soft graph on  $G^*$ . The fuzzy soft graph degree of a vertex  $a$  is defined as  $deg_G(a) = \sum_{e \in A} \sum_{x \neq y} K(e)(x, y) \forall a \in A$  and  $(x, y) \in V$ .

### 2.4. Regular Fuzzy soft graph [4]

Let  $G = (G^*, F, K, A)$  be a regular fuzzy soft graph if  $(F(e), K(e))$  is regular fuzzy graph of degree  $k$  for all  $e_i \in A$  then  $G$  is a  $k$ - regular fuzzy soft graph.

### 2.5. Order of fuzzy soft graph [4]

Let  $G_{A,V} = ((A, \sigma), (A, \mu))$  be a fuzzy soft graph. Then the order of fuzzy soft graph  $G_{A,V} = \sum_{e \in A} \sum_{x \in A} \sigma_e(x)$ .

### 2.6. AQSP Fuzzy Soft Graph [18]

Let  $V = ((\sigma_1^P(x), \sigma_1^N(x), \rho_1^P(x), \rho_1^N(x)), (\sigma_2^P(x), \sigma_2^N(x), \rho_2^P(x), \rho_2^N(x)) \dots (\sigma_n^P(x), \sigma_n^N(x), \rho_n^P(x), \rho_n^N(x)))$  be a nonempty AQSP fuzzy set.  $E$  (Parameters set) and  $A_{AQSP} \subset E$ . Also let,

- (i)  $\sigma^P : A_{AQSP} \rightarrow F_{AQSP}(V)$  (Collection of all AQSP fuzzy subsets in  $V$ ),  $e \mapsto \sigma_e^P$ , and  $\sigma_e^P : V \rightarrow [0, 1]$ ,  $v_i \mapsto \sigma_e^P$  then  $(A_{AQSP}, \sigma^P) : \text{AQSP fuzzy soft vertex set.}$
- (ii)  $\sigma^N : A_{AQSP} \rightarrow F_{AQSP}(V)$  (Collection of all AQSP fuzzy subsets in  $V$ ),  $e \mapsto \sigma_e^N$ , and  $\sigma_e^N : V \rightarrow [-1, 0]$ ,  $v_i \mapsto \sigma_e^N$  then  $(A_{AQSP}, \sigma^N) : \text{AQSP fuzzy soft vertex set.}$
- (iii)  $\rho^P : A_{AQSP} \rightarrow F_{AQSP}(V)$  (Collection of all AQSP fuzzy submerge subsets in  $V$ ),  $e \mapsto \rho_e^P$ , and  $\rho_e^P : V \rightarrow [0, 0.5]$ ,  $v_i \mapsto \rho_e^P$  then  $(A_{AQSP}, \rho^P) : \text{AQSP fuzzy soft vertex set.}$
- (iv)  $\rho^N : A_{AQSP} \rightarrow F_{AQSP}(V)$  (Collection of all fuzzy submerge subsets in  $V$ ),  $e \mapsto \rho_e^N$ , and  $\rho_e^N : V \rightarrow [-0.5, 0]$ ,  $v_i \mapsto \rho_e^N$  then  $(A_{AQSP}, \rho^N) : \text{AQSP fuzzy soft vertex set.}$
- (v)  $\mu^P : A_{AQSP} \rightarrow F_{AQSP}(V \times V)$  (Collection of all AQSP fuzzy subsets in  $V \times V$ ),  $e \mapsto \mu_e^P$ ,  $\mu_e^P : V \times V \rightarrow [0, 1]$ ,  $(v_i, v_j) \mapsto \mu_e^P(v_i, v_j)$  then  $(A_{AQSP}, \mu^P) : \text{AQSP fuzzy soft membership edge set.}$
- (vi)  $\mu^N : A_{AQSP} \rightarrow F_{AQSP}(V \times V)$  (Collection of all AQSP fuzzy subsets in  $V \times V$ ),  $e \mapsto \mu_e^N$ , and  $\mu_e^N : V \times V \rightarrow [-1, 0]$ ,  $(v_i, v_j) \mapsto \mu_e^N(v_i, v_j)$  then  $(A_{AQSP}, \mu^N) : \text{AQSP fuzzy soft non - membership edge set.}$
- (vii)  $\gamma^P : A_{AQSP} \rightarrow F_{AQSP}(V \times V)$  (Collection of all AQSP fuzzy subsets in  $V \times V$ ),  $e \mapsto \gamma_e^P$ , and  $\gamma_e^P : V \times V \rightarrow [0, 0.5]$ ,  $(v_i, v_j) \mapsto \gamma_e^P(v_i, v_j)$  then  $(A_{AQSP}, \gamma^P) : \text{AQSP fuzzy soft submerge membership edge set.}$
- (viii)  $\gamma^N : A_{AQSP} \rightarrow F_{AQSP}(V \times V)$  (Collection of all AQSP fuzzy subsets in  $V \times V$ ),  $e \mapsto \gamma_e^N$ , and  $\gamma_e^N : V \times V \rightarrow [-0.5, 0]$ ,  $(v_i, v_j) \mapsto \gamma_e^N(v_i, v_j)$  then  $(A_{AQSP}, \gamma^N) : \text{AQSP fuzzy soft submerge membership edge set. Then the AQSP fuzzy soft graph is, } ((A_{AQSP}, (\sigma^P, \sigma^N, \rho^P, \rho^N)), ((A_{AQSP}, (\mu^P, \mu^N, \gamma^P, \gamma^N))) \text{ if the conditions are satisfied}$ 
  - (a)  $\mu_e^P(x, y) \leq \sigma_e^P(x) \wedge \sigma_e^P(y)$ ,      (b)  $\mu_e^N(x, y) \geq \sigma_e^N(x) \vee \sigma_e^N(y)$ ,
  - (c)  $\gamma_e^P(x, y) \leq \rho_e^P(x) \wedge \rho_e^P(y)$ ,      (d)  $\gamma_e^N(x, y) \geq \rho_e^N(x) \vee \rho_e^N(y)$ , for all  $e \in A_{AQSP}$  and for all values of  $x, y = 1, 2, 3, \dots, n$  and this AQSP fuzzy soft graph is denoted as  $G_{AQSP}(A, V)$ .

### 3. METHOD

The essential definition of AQSP fuzzy soft graph method is deliberated with an examples.

#### 3.1. Alternate Quadra Sub - merging Polar(AQSP) Fuzzy Graph

An Alternate Quadra - Submerging Polar (AQSP) Fuzzy Graph  $G = (\sigma_{AQSP}, \mu_{AQSP})$  is a fuzzy graph with crisp graph  $G^* = (\sigma_{AQSP}^*, \mu_{AQSP}^*)$  is given as  $V = (\sigma^P(x), \sigma^N(x), \rho^P(x), \rho^N(x))$  which is the membership value of vertices along with the uncertain membership value of edges is given as,  $E = V \times V = (\mu^P(x, y), \mu^N(x, y), \gamma^P(x, y), \gamma^N(x, y))$ .

Here the vertex set  $V$  is defined with the given condition in a unique method which is an alternate contrast submerging polarized uncertain transformation. Here  $\sigma^P = V \rightarrow [0, 1]$ ,  $\sigma^N = V \rightarrow [-1, 0]$ ,  $\rho^P = d | 0.5, \sigma^P(x) |$  and  $\rho^N = -d | -0.5, \sigma^N(x) |$ . Here  $(-0.5, 0.5)$  is the fixation of uncertain alternate contrast polarized submerging transformation into certain consistent preferable position. And the edge set  $E$  satisfies the following sufficient conditions.

- (i)  $\mu^P(x, y) \leq \min(\sigma^P(x), \sigma^P(y))$ ,      (ii)  $\mu^N(x, y) \geq \max(\sigma^N(x), \sigma^N(y))$
- (iii)  $\gamma^P(x, y) \leq \min(\rho^P(x), \rho^P(y))$       (iv)  $\gamma^N(x, y) \geq \max(\rho^N(x), \rho^N(y))$ ,

$\forall(x, y) \in E$ . By definition,  $\mu^P = V \times V \rightarrow [0, 1] \times [1, 0]$ ,  $\mu^N = V \times V \rightarrow [-1, 0] \times [0, -1]$  and the submerging mappings,  $\gamma^P = V \times V \rightarrow [0, 0.5] \times [0.5, 0]$ ,



$\gamma^N = V \times V \rightarrow [-0.5, 0] \times [0, -0.5]$ , which denotes the impact of the alternate quadrant polarized fuzzy mapping.

The maximum of submerging presumption to be at the level of confidence  $[0, 0.5] \subseteq [0, 1]$  and the minimum of submerging presumption level of confidence is  $[-0.5, 0] \subseteq [-1, 0]$  extension of the graph with its membership and non - membership values portrait the unique level of submerging destination in an AQSP fuzzy graph.

Also it must satisfy the condition,  $-1 \leq \sigma^P(x) + \sigma^N(x) \leq 1$  and  $|\rho^P(x) + \rho^N(x)| \leq 1$  with constrains  $0 \leq \sigma^P(x) + \sigma^N(x) + |\rho^P(x) + \rho^N(x)| \leq 2$  such that the uncertain status of submerging presumption, transform into its precise consistent level with fixation mid - value 0.5, which implies that level of confidence 0.5 in an AQSP as the valuable membership of its position which is real and valid in the fuzzification. The example of AQSP fuzzy graph is given in Figure.1.

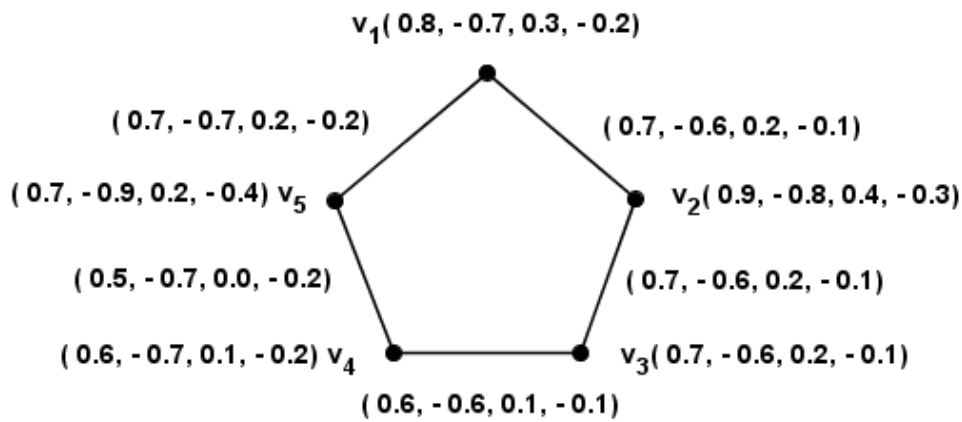


Figure 1: AQSP Fuzzy Graph  $G = (\sigma_{AQSP}, \mu_{AQSP})$

### 3.2. Example of AQSP Fuzzy Soft Graph

Consider an AQSP fuzzy soft graph  $G_{AQSP}(A, V)$ , where  $V = (v_1, v_2, v_3, v_4)$  and  $E = (e_1, e_2, e_3)$ . Here  $G_{AQSP}(A, V)$  is described in Table.1. and  $\mu_e(v_i, v_j) = 0, \forall (v_i, v_j) \in V \times V \setminus \{(v_1, v_2), (v_2, v_3), (v_3, v_4), (v_1, v_4), (v_1, v_3)\}$  for all  $e \in E$ .

Table 1: Tabular representation of AQSP Fuzzy Soft Graph parameter vertex set.

$(\sigma, \rho)$	$v_1$	$v_2$	$v_3$	$v_4$
$e_1$	(0.6, -0.7, 0.1, -0.2)	(0.7, -0.8, 0.2, -0.3)	(0.8, -0.9, 0.3, -0.4)	(0.6, -0.7, 0.1, -0.2)
$e_2$	(0.7, -0.6, 0.2, -0.1)	(0.8, -0.7, 0.3, -0.2)	(0.9, -0.8, 0.4, -0.3)	(0.8, -0.8, 0.3, -0.3)
$e_3$	(0.8, -0.6, 0.3, -0.1)	(0.9, -0.7, 0.4, -0.2)	(0.8, -0.8, 0.3, -0.3)	(0.9, -0.9, 0.4, -0.4)

**Table 2:** Tabular representation of AQSP Fuzzy Soft Graph parameter edge set.

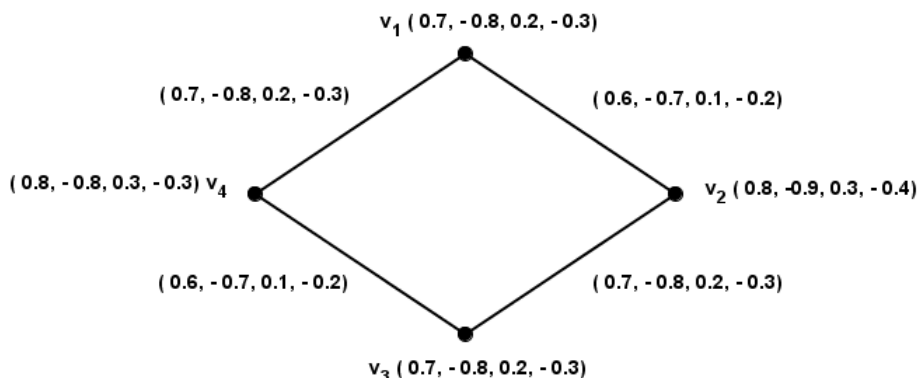
$(\mu, \gamma)$	$v_1, v_2$	$v, v_3$	$v_3, v_4$	$v_4, v_1$	$v_1, v_3$
$e_1$	(0.6, -0.7, 0.1, -0.2)	(0.7, -0.8, 0.2, -0.3)	(0.6, -0.7, 0.1, -0.2)	(0.6, -0.7, 0.1, -0.2)	(0.6, -0.7, 0.1, -0.2)
$e_2$	(0.7, -0.6, 0.2, -0.1)	(0.7, -0.7, 0.2, -0.2)	(0.8, -0.8, 0.3, -0.3)	(0.7, -0.6, 0.2, -0.1)	(0.6, -0.6, 0.1, -0.1)
$e_3$	(0.8, -0.6, 0.3, -0.1)	(0.8, -0.7, 0.3, -0.2)	(0.8, -0.7, 0.3, -0.2)	(0.7, -0.6, 0.2, -0.1)	(0.8, -0.6, 0.3, -0.1)

Table. 2. represents the AQSP fuzzy graph with parametric membership and non - membership with submerge values.

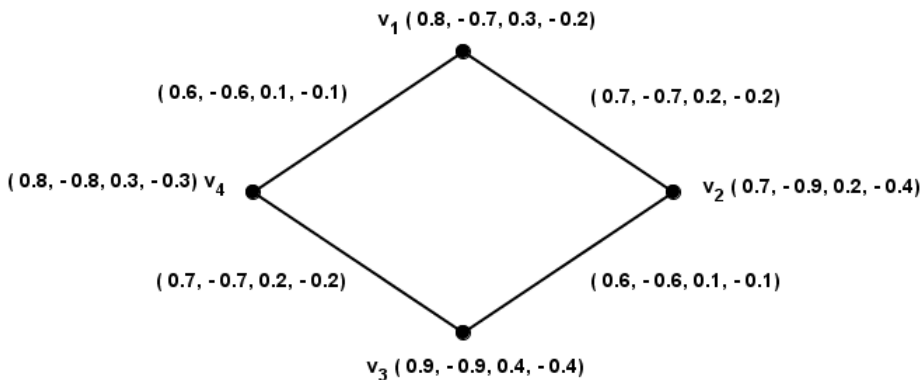
#### 4. DESCRIPTIONS OF THE REGULARITY OF AQSP FUZZY SOFT GRAPH

##### 4.1. Regular AQSP Fuzzy Ssoft Graph

Let  $G^* = (\sigma^*, \mu^*)$  be a crisp graph and  $G_{AQSP}(A, V)$  be an regular AQSP fuzzy soft graph of  $G^*$ . Then  $G_{AQSP}(A, V)$  is said to be an regular AQSP soft graph, if  $R_{AQSP}(e_i)$  is an regular AQSP fuzzy soft graph of degree k for all  $e_i \in A_{AQSP}$ , then  $G_{AQSP}(A, V)$  is a k - regular AQSP fuzzy soft graph.



**Figure 2:**  $G_{AQSP}(A, V)$  - Corresponding to the parameter  $e_1$



**Figure 3:**  $G_{AQSP}(A, V)$  - Corresponding to the parameter  $e_2$

### 4.2. Example of an AQSP Fuzzy Soft Graph

Consider, an AQSP fuzzy soft graph,  $G_{AQSP}(A, V)$ , the vertex set  $V = (v_1, v_2, v_3, v_4)$  and let the corresponding parameters  $E = (e_1, e_2)$ . Here  $G_{AQSP}(A, V) = ((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N)), ((A_{AQSP}), (\mu^P, \mu^N, \gamma^P, \gamma^N))$  is described by Table.3 and Table. 4  $(v_1, v_2, v_3, v_4)$ .

### 4.3. Remark on Regular AQSP Fuzzy Graph

From Figure.2 and Figure.3 we get the result that the regular AQSP fuzzy graph which can not be a totally regular AQSP fuzzy graph. Table 3. represents the AQSP fuzzy soft graph vertex set.

**Table 3:** Tabular representation of AQSP Fuzzy Soft Graph parameter vertex set.

$(\mu, \gamma)$	$v_1$	$v_2$	$v_3$	$v_4$
$e_1$	(0.7, - 0.8, 0.2,- 0.3)	(0.8, - 0.9, 0.3, - 0.4)	(0.7, - 0.7, 0.2, - 0.2 )	(0.8, - 0.8, 0.3,- 0.3)
$e_2$	(0.8, - 0.7, 0.3,- 0.2)	(0.7, - 0.9, 0.2, - 0.4)	(0.9, -0.9, 0.4, - 0.4 )	(0.8, - 0.8, 0.3 ,- 0.3)

**Table 4:** Tabular representation of AQSP Fuzzy Soft Graph parameter edge set.

$(\mu, \gamma)$	$v_1v_2$	$v_2v_3$	$v_3v_4$	$v_4v_1$
$e_1$	(0.6, - 0.7, 0.1,- 0.2)	(0.7, - 0.8, 0.2, - 0.3)	(0.6, - 0.7, 0.1, - 0.2 )	(0.7, - 0.8, 0.2,- 0.3)
$e_2$	(0.7, - 0.7, 0.2,- 0.2)	(0.6, - 0.6, 0.1, - 0.1)	(0.7, - 0.7, 0.2, - 0.2 )	(0.6, - 0.6, 0.1,- 0.1)

Table. 4 represents the corresponding edges,  $(v_1, v_2), (v_2, v_3), (v_3, v_4), (v_4, v_1)$  , for all values of  $e \in A_{AQSP}$ .

### 4.4. Totally Regular AQSP Fuzzy Soft Graph

Let  $G^* = (\sigma, \mu)$  be a simple graph and  $G_{AQSP}(A, V)$  be an AQSP fuzzy soft graph of  $G^*$ . Then  $G_{AQSP}(A, V)$  is said to be a totally regular AQSP fuzzy soft graph if  $R_{AQSP}(A, V)$  is totally regular fuzzy soft graph for all values of  $e_i \in A_{AQSP}$ , then  $G_{AQSP}(A, V)$  is called k totally regular AQSP fuzzy soft graph.

**Theorem 1.** If  $G_{AQSP}(A, V)$  satisfies the condition of regular and totally regular AQSP fuzzy soft graph, then we prove that  $((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N))$  is a constant AQSP fuzzy soft function in  $H_{AQSP}(A, V)$  of  $G^*$  for all values of  $e \in A_{AQSP}$ .

**Proof.** Let  $G_{AQSP}(A, V)$  satisfies the condition of regular and totally regular AQSP fuzzy soft graph. Then we have the degree of vertices as,

- (i)  $deg\sigma_e^P(a) = k_1, deg\sigma_e^N(a) = k_2, deg\rho_e^P(a) = k_3, deg\rho_e^N(a) = k_4$  and
- (ii)  $tdeg\sigma_e^P(a) = l_1, tdeg\sigma_e^N(a) = l_2, tdeg\rho_e^P(a) = l_3, tdeg\rho_e^N(a) = l_4$ .

In AQSP fuzzy subgraphs  $H_{AQSP}(A, V)$  for all values of  $e \in A_{AQSP}, a \in V$ . This implies that,  
 $deg\sigma_e^P(a) + A_{AQSP} \sigma_e^P(a) = l_1,$   
 $deg\sigma_e^N(a) + A_{AQSP} \sigma_e^N(a) = l_2,$   
 $deg\rho_e^P(a) + A_{AQSP} \rho_e^P(a) = l_3,$   
 $deg\rho_e^N(a) + A_{AQSP} \rho_e^N(a) = l_4 \in H_{AQSP}(A, V), \forall e \in A_{AQSP}, a \in V.$

$$\begin{aligned}
 A_{AQSP} \sigma_e^P(a) &= l_1 - k_1 \\
 A_{AQSP} \sigma_e^N(a) &= l_2 - k_2 \\
 A_{AQSP} \rho_e^P(a) &= l_3 - k_3 \\
 A_{AQSP} \rho_e^N(a) &= l_4 - k_4 \in H_{AQSP}(A, V), \\
 \forall e \in A_{AQSP}, a \in V.
 \end{aligned}$$

Hence,  $((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N))$  is a constant AQSP fuzzy soft function in  $H_{AQSP}(A, V)$  of  $G^*$  for all values of  $e \in A_{AQSP}$ . ■

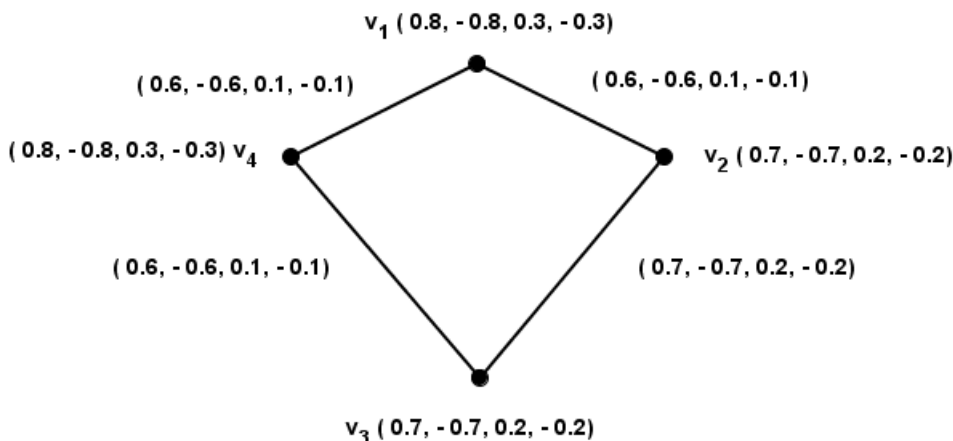


Figure 4:  $G_{AQSP}(A, V)$  - Corresponding to the parameter  $e_1$

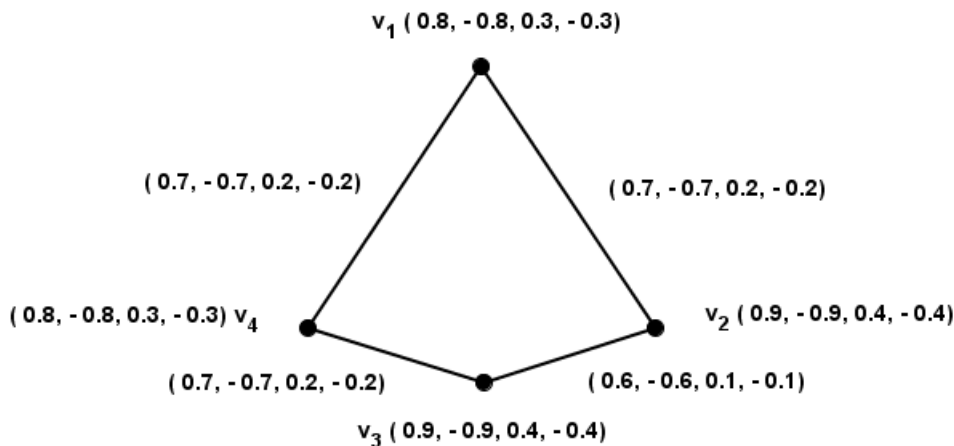


Figure 5:  $G_{AQSP}(A, V)$  - Corresponding to the parameter  $e_2$

#### 4.5. Example Totally Regular AQSP Fuzzy Soft Graph

Consider, an AQSP fuzzy soft graph,  $G_{AQSP}(A, V)$ , the vertex set  $V = (v_1, v_2, v_3, v_4)$  and let the corresponding parameters  $E = (e_1, e_2)$  is shown in the Figure.4 and Figure.5.

Here  $G_{AQSP}(A, V) = ((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N)), ((A_{AQSP}), (\mu^P, \mu^N, \gamma^P, \gamma^N))$  is described Figure.6  $(v_1, v_2, v_3, v_4)$ . Figure. 7 represents the corresponding edges,  $(v_1, v_2), (v_2, v_3), (v_3, v_4), (v_4, v_1)$ , for all values of  $e \in A_{AQSP}$ .

**Table 5:** Tabular representation of AQSP Fuzzy Soft Graph parameter vertex set.

$(\mu, \gamma)$	$v_1$	$v_2$	$v_3$	$v_4$
$e_1$	( 0.8, - 0.8, 0.3,- 0.3)	( 0.7, - 0.7, 0.2, - 0.2)	( 0.7, - 0.7, 0.2, - 0.2 )	( 0.8, - 0.8, 0.3,- 0.3)
$e_2$	( 0.8, - 0.8, 0.3,- 0.3)	( 0.9, - 0.9, 0.4, - 0.4)	( 0.9, -0.9, 0.4, - 0.4 )	( 0.8, - 0.8, 0.3 , - 0.3)

**Table 6:** Tabular representation of AQSP Fuzzy Soft Graph parameter edge set.

$(\mu, \gamma)$	$v_1v_2$	$v_2v_3$	$v_3v_4$	$v_4v_1$
$e_1$	( 0.6, - 0.6, 0.1,- 0.1)	( 0.7, - 0.7, 0.2, - 0.2)	( 0.6, - 0.6, 0.1, - 0.1 )	( 0.6, - 0.6, 0.1,- 0.1)
$e_2$	( 0.7, - 0.7, 0.2,- 0.2)	( 0.6, - 0.6, 0.1, - 0.1)	( 0.7, - 0.7, 0.2, - 0.2 )	( 0.7, - 0.7, 0.2,- 0.2)

#### 4.6. Example of AQSP Fuzzy Soft Graph

Consider, an AQSP fuzzy soft graph,  $G_{AQSP}(A, V)$ , the vertex set  $V = (v_1, v_2, v_3, v_4)$  and let the corresponding parameters  $E = (e_1, e_2)$ .

Here  $G_{AQSP}(A, V) = ((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N)), ((A_{AQSP}), (\mu^P, \mu^N, \gamma^P, \gamma^N))$  is described by Table.5 and Table. 5 such as,  $(v_1, v_2), (v_2, v_3), (v_3, v_4), (v_1, v_3), (v_1, v_4), (v_4, v_1), (v_1, v_1)$  , for all values of  $e \in A_{AQSP}$ .

#### 4.7. Remark on Regular AQSP Fuzzy Soft Graph

From Theorem.5.7. we get the result if  $G_{AQSP}(A, V)$  is a regular AQSP fuzzy soft graph and  $((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N))$  is a constant AQSP fuzzy soft function, then  $G_{AQSP}^C(A, V)$  is a regular AQSP fuzzy soft graph.

#### 4.8. Remark on Totally Regular AQSP Fuzzy Soft Graph

From Theorem.5.7. similarly we get the result if  $G_{AQSP}(A, V)$  is a totally regular AQSP fuzzy soft graph and  $((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N))$  is a constant AQSP fuzzy soft function, then  $G_{AQSP}^C(A, V)$  is a totally regular AQSP fuzzy soft graph.

**Theorem 2.** Let  $G_{AQSP}(A, V) = ((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N)), ((A_{AQSP}), (\mu^P, \mu^N, \gamma^P, \gamma^N))$ , for all values of  $e \in A_{AQSP}$ . be an AQSP fuzzy soft graph with the vertex and edge membership and non - membership submerging values. Then we prove that,

- (i)  $\sum_{a \in A} tdeg_{G_{AQSP}(A, V)} (\sigma_e^P(a) = 2S(G_{AQSP}(A, V)) + O(G_{AQSP}(A, V)))$
- (ii)  $\sum_{a \in A} tdeg_{G_{AQSP}(A, V)} (\sigma_e^N(a) = 2S(G_{AQSP}(A, V)) + O(G_{AQSP}(A, V)))$
- (iii)  $\sum_{a \in A} tdeg_{G_{AQSP}(A, V)} (\rho_e^P(a) = 2S(G_{AQSP}(A, V)) + O(G_{AQSP}(A, V)))$
- (iv)  $\sum_{a \in A} tdeg_{G_{AQSP}(A, V)} (\rho_e^N(a) = 2S(G_{AQSP}(A, V)) + O(G_{AQSP}(A, V)))$

**Proof.** (i)  $tdeg_{G_{AQSP}(A, V)} (\sigma_e^P(a)) = \sum_{e \in A_{AQSP}} (\sum_{a \in V} (\mu_e^P(a, b) + \sigma_e^P(a))$   
 $\implies \sum_{a \in V} tdeg_{G_{AQSP}(A, V)} (\sigma_e^P(a)) = \sum_{a \in V} (\sum_{e \in A_{AQSP}} (\sum_{a \in V} (\mu_e^P(a, b) + \sigma_e^P(a))$

$$= \sum_{a \in A} tdeg_{G_{AQSP}(A,V)} (\sigma_e^P(a)) = 2S(G_{AQSP}(A,V)) + O(G_{AQSP}(A,V)).$$

For non - membership AQSP fuzzy soft graph values are,

$$\begin{aligned} \text{(ii)} \quad tdeg_{G_{AQSP}(A,V)} (\sigma_e^N(a)) &= \sum_{e \in A_{AQSP}} (\sum_{a \in V} (\mu_e^N(a,b) + \sigma_e^N(a)), \\ \implies \sum_{a \in V} tdeg_{G_{AQSP}(A,V)} (\sigma_e^N(a)) &= \sum_{a \in V} (\sum_{e \in A_{AQSP}} (\sum_{a \in V} (\mu_e^N(a,b) + \sigma_e^N(a)), \\ &= \sum_{a \in A} tdeg_{G_{AQSP}(A,V)} (\sigma_e^N(a)) = 2S(G_{AQSP}(A,V)) + O(G_{AQSP}(A,V)). \end{aligned}$$

Now, the Submerging membership values are,

$$\begin{aligned} \text{(iii)} \quad tdeg_{G_{AQSP}(A,V)} (\rho_e^P(a)) &= \sum_{e \in A_{AQSP}} (\sum_{a \in V} (\gamma_e^P(a,b) + \rho_e^P(a)), \\ \implies \sum_{a \in V} tdeg_{G_{AQSP}(A,V)} (\rho_e^P(a)) &= \sum_{a \in V} \sum_{e \in A_{AQSP}} (\sum_{a \in V} (\gamma_e^P(a,b) + \rho_e^P(a)), \\ &= \sum_{a \in A} tdeg_{G_{AQSP}(A,V)} (\rho_e^P(a)) = 2S(G_{AQSP}(A,V)) + O(G_{AQSP}(A,V)). \end{aligned}$$

For the Submerging non - membership values are,

$$\begin{aligned} \text{(iv)} \quad tdeg_{G_{AQSP}(A,V)} (\rho_e^N(a)) &= \sum_{e \in A_{AQSP}} (\sum_{a \in V} (\gamma_e^N(a,b) + \rho_e^N(a)), \\ \implies \sum_{a \in V} tdeg_{G_{AQSP}(A,V)} (\rho_e^N(a)) &= \sum_{a \in V} \sum_{e \in A_{AQSP}} (\sum_{a \in V} (\gamma_e^N(a,b) + \rho_e^N(a)), \\ &= \sum_{a \in A} tdeg_{G_{AQSP}(A,V)} (\rho_e^N(a)) = 2S(G_{AQSP}(A,V)) + O(G_{AQSP}(A,V)). \quad \blacksquare \end{aligned}$$

## 5. PROPERTIES OF REGULAR AND TOTALLY REGULAR AQSP FUZZY SOFT GRAPH

**Theorem 3.** The size of the  $(k_1, k_2, k_3, k_4)$  regular AQSP fuzzy soft graph  $(G_{AQSP}(A, V)$  on  $G^* = (V, E)$  is (i)  $\frac{pk_1}{2}$ , (ii)  $\frac{pk_2}{2}$ , (iii)  $\frac{pk_3}{2}$  and (iv)  $\frac{pk_4}{2}$  where  $p = |V|$  and  $deg \sigma_e^P(a) = k_1$ ,  $deg \sigma_e^N(a) = k_2$ ,  $deg \rho_e^P(a) = k_3$  and  $deg \rho_e^N(a) = k_4$

**Proof.**

$$\begin{aligned} \text{(i)} \quad S(G_{AQSP}(A, V)) &= \sum_{e \in A_{AQSP}} (\sum_{a \neq b} \mu_e^P(a, b)) \\ \text{since } G_{AQSP}(A, V) \text{ is a } k_1 \text{ regular AQSP fuzzy soft graph we get,} \\ deg \sigma_e^P(a) &= k_1, \forall a \in V, \\ \text{Now, } S(G_{AQSP}(A, V)) &= \sum_{e \in A_{AQSP}} (\sum_{a \neq b} \mu_e^P(a, b)) \\ &= \sum_{a \in V} \frac{deg \sigma_e^P(a)}{2}, \\ \sum_{a \in V} \frac{deg \sigma_e^P(a)}{2} &= \sum_{a \in V} \frac{k_1}{2}. \\ \text{(ii)} \quad S(G_{AQSP}(A, V)) &= \sum_{e \in A_{AQSP}} (\sum_{a \neq b} \mu_e^N(a, b)) \\ \text{since } (G_{AQSP}(A, V)) \text{ is a } (k_1, k_2, k_3, k_4) \text{ regular AQSP fuzzy soft graph we get,} \\ deg \sigma_e^N(a) &= k_2, \forall a \in V, \\ \text{Now, } S(G_{AQSP}(A, V)) &= \sum_{e \in A_{AQSP}} (\sum_{a \neq b} \mu_e^N(a, b)) \\ &= \sum_{a \in V} \frac{deg \sigma_e^N(a)}{2} \\ \sum_{a \in V} \frac{deg \sigma_e^N(a)}{2} &= \sum_{a \in V} \frac{k_2}{2} \\ \text{(iii)} \quad S(G_{AQSP}(A, V)) &= \sum_{e \in A_{AQSP}} (\sum_{a \neq b} \gamma_e^P(a, b)) \\ \text{since } (G_{AQSP}(A, V)) \text{ is a } k_3 \text{ regular AQSP fuzzy soft graph we get,} \\ deg \rho_e^P(a) &= k_3, \forall a \in V, \\ \text{Now, } S(G_{AQSP}(A, V)) &= \sum_{e \in A_{AQSP}} (\sum_{a \neq b} \gamma_e^P(a, b)) \\ &= \sum_{a \in V} \frac{deg \rho_e^P(a)}{2}, \end{aligned}$$

$$\sum_{a \in V} \frac{\deg \rho_e^P(a)}{2} = \sum_{a \in V} \frac{k_3}{2}.$$

$$(iv) S(G_{AQSP}(A, V)) = \sum_{e \in A_{AQSP}} (\sum_{a \neq b} \gamma_e^N(a, b))$$

since  $(G_{AQSP}(A, V))$  is a  $k_4$  regular AQSP fuzzy soft graph we get,

$$\deg \rho_e^P(a) = k_4, \forall a \in V,$$

$$\text{Now, } S(G_{AQSP}(A, V)) = \sum_{e \in A_{AQSP}} (\sum_{a \neq b} \gamma_e^N(a, b))$$

$$\sum_{a \in V} \frac{\deg \rho_e^N(a)}{2}$$

$$\sum_{a \in V} \frac{\deg \rho_e^N(a)}{2} = \sum_{a \in V} \frac{k_4}{2}$$

Hence, The size of the  $(k_1, k_2, k_3, k_4)$  regular AQSP fuzzy soft graph  $(G_{AQSP}(A, V))$  on  $G^* = (V, E)$  is  $\frac{pk_1}{2}, \frac{pk_2}{2}, \frac{pk_3}{2}$  and  $\frac{pk_4}{2}$  where  $p = |V|$  ■

**Theorem 4.** If  $G_{AQSP}(A, V) = ((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N)), ((A_{AQSP}), (\mu^P, \mu^N, \gamma^P, \gamma^N))$  be an regular AQSP fuzzy on  $G^* = (\sigma^*, \mu^*)$  is a  $k$ -totally regular AQSP fuzzy soft graph. Then,  $2S(G_{AQSP}(A, V)) + O(G_{AQSP}(A, V)) = (\sigma^P pk, \sigma^N pk, \gamma^P pk, \gamma^N pk)$  where,  $(\sigma^P p, \sigma^N p, \gamma^P p, \gamma^N p) = |V|$ .

**Proof.** Since,  $G_{AQSP}(A, V)$  is a  $k$ -totally regular AQSP fuzzy soft graph,  $tdeg_{G_{AQSP}(A, V)} \sigma^P a = k_1, tdeg_{G_{AQSP}(A, V)} \sigma^N a = k_2, tdeg_{G_{AQSP}(A, V)} \rho^P a = k_3$  and  $tdeg_{G_{AQSP}(A, V)} \rho^N a = k_4, \forall a \in V$ .

$$\implies deg_{G_{AQSP}(A, V)} \sigma^P a + \sum_{e \in A} \sigma_e^P(a), deg_{G_{AQSP}(A, V)} \sigma^N a + \sum_{e \in A} \sigma_e^N(a), deg_{G_{AQSP}(A, V)} \rho^P a + \sum_{e \in A} \rho_e^P(a) \text{ and } deg_{G_{AQSP}(A, V)} \rho^N a + \sum_{e \in A} \rho_e^N(a), \forall a \in V.$$

$$\implies \sum_{a \in V} deg_{G_{AQSP}(A, V)} \sigma^P a + \sum_{a \in V} \sum_{e \in A_{AQSP}} \sigma_e^P a = \sum_{a \in V},$$

$$\sum_{a \in V} deg_{G_{AQSP}(A, V)} \sigma^N a + \sum_{a \in V} \sum_{e \in A_{AQSP}} \sigma_e^N a = \sum_{a \in V}.$$

For submerging AQSP fuzzy soft graph values are,

$$\sum_{a \in V} deg_{G_{AQSP}(A, V)} \rho^P a + \sum_{a \in V} \sum_{e \in A_{AQSP}} \rho_e^P a = \sum_{a \in V},$$

$$\sum_{a \in V} deg_{G_{AQSP}(A, V)} \rho^N a + \sum_{a \in V} \sum_{e \in A_{AQSP}} \rho_e^N a = \sum_{a \in V}.$$

$$\implies tdeg_{G_{AQSP}(A, V)} \sigma^P a = k_1, tdeg_{G_{AQSP}(A, V)} \sigma^N a = k_2,$$

$$tdeg_{G_{AQSP}(A, V)} \rho^P a = k_3 \text{ and } tdeg_{G_{AQSP}(A, V)} \rho^N a = k_4, \forall a \in V.$$

$$\text{Hence, } 2S(G_{A, V}(AQSP)) + O(G_{A, V}(AQSP)) = (\sigma^P pk, \sigma^N pk, \gamma^P pk, \gamma^N pk). \quad \blacksquare$$

**Theorem 5.** If  $(G_{AQSP}(A, V)) = ((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N)), ((A_{AQSP}), (\mu^P, \mu^N, \gamma^P, \gamma^N))$  be an AQSP fuzzy soft graph on  $G^* = (\sigma^*, \mu^*)$  is a  $k$ -regular AQSP fuzzy soft graph. Then

$$(i) O(G_{AQSP}(A, V)) = n(l_1 - k_1), (ii) O(G_{AQSP}(A, V)) = n(l_2 - k_2),$$

$$(iii) O(G_{AQSP}(A, V)) = n(l_3 - k_3) \text{ and } (iv) O(G_{AQSP}(A, V)) = n(l_4 - k_4) \text{ where } n = |V|.$$

**Proof.** Since  $(G_{AQSP}(A, V))$  is an  $k$ -regular AQSP fuzzy soft graph, then we have

$$deg_{G_{AQSP}(A, V)} \sigma^P a = k_1, deg_{G_{AQSP}(A, V)} \sigma^N a = k_2, deg_{G_{AQSP}(A, V)} \rho^P a = k_3 \text{ and}$$

$$deg_{G_{AQSP}(A, V)} \rho^N a = k_4, \forall a \in V. \text{ Here, } (G_{AQSP}(A, V)) \text{ is totally regular}$$

AQSP fuzzy soft graph, then we consider,

$$tdeg_{G_{AQSP}(A, V)} \sigma^P a = l_1, tdeg_{G_{AQSP}(A, V)} \sigma^N a = l_2, tdeg_{G_{AQSP}(A, V)} \rho^P a = l_3$$

$$\text{and } tdeg_{G_{AQSP}(A, V)} \rho^N a = l_4, \forall a \in V. \text{ Now we have,}$$

$$\sum_{a \in A} tdeg_{G_{AQSP}(A, V)} \sigma^P a = \sigma^P pk_1,$$

$$\sum_{a \in A} tdeg_{G_{AQSP}(A, V)} \sigma^N a = \sigma^N pk_2,$$

$$\sum_{a \in A} tdeg_{G_{AQSP}(A, V)} \rho^P a = \rho^P pk_3,$$

$$\sum_{a \in A} tdeg_{G_{AQSP}(A, V)} \rho^N a = \rho^N pk_4.$$

(i) The AQSP fuzzy soft graph membership value is,

$$\implies \sum_{a \in V} l_1 = \sum_{a \in V} deg_{G_{AQSP}(A, V)} \sigma^P a + O(G_{AQSP}(A, V))$$

$$\implies \sum_{a \in V} l_1 = \sum_{a \in V} k_1 + O(G_{AQSP}(A, V))$$

$$\implies nl_1 = nk_1 + O(G_{AQSP}(A, V))$$

$$\implies O(G_{AQSP}(A, V)) = nk_1 - nl_1$$

$$\implies O(G_{AQSP}(A, V)) = n(k_1 - l_1)$$

$$O(G_{AQSP}(A, V)) = n(l_1 - k_1).$$

(ii) The AQSP fuzzy soft graph non-membership value is,

$$\implies \sum_{a \in V} l_2 = \sum_{a \in V} \text{deg}_{G_{AQSP}(A, V)} \sigma^N a + O(G_{AQSP}(A, V))$$

$$\implies \sum_{a \in V} l_2 = \sum_{a \in V} k_2 + O(G_{AQSP}(A, V))$$

$$\implies nl_2 = nk_2 + O(G_{AQSP}(A, V))$$

$$\implies O(G_{AQSP}(A, V)) = nk_2 - nl_2$$

$$\implies O(G_{AQSP}(A, V)) = n(k_2 - l_2)$$

$$O(G_{AQSP}(A, V)) = n(l_2 - k_2).$$

(iii) The AQSP fuzzy soft graph submerrging membership value is,

$$\implies \sum_{a \in V} l_3 = \sum_{a \in V} \text{deg}_{G_{AQSP}(A, V)} \rho^P a + O(G_{AQSP}(A, V))$$

$$\implies \sum_{a \in V} l_3 = \sum_{a \in V} k_3 + O(G_{AQSP}(A, V))$$

$$\implies nl_3 = nk_3 + O(G_{AQSP}(A, V))$$

$$\implies O(G_{AQSP}(A, V)) = nk_3 - nl_3$$

$$\implies O(G_{AQSP}(A, V)) = n(k_3 - l_3)$$

$$O(G_{AQSP}(A, V)) = n(l_3 - k_3).$$

(iv) The AQSP fuzzy soft graph submerrging non- -membership value is,

$$\implies \sum_{a \in V} l_3 = \sum_{a \in V} \text{deg}_{G_{AQSP}(A, V)} \rho^N a + O(G_{AQSP}(A, V))$$

$$\implies \sum_{a \in V} l_3 = \sum_{a \in V} k_3 + O(G_{AQSP}(A, V))$$

$$\implies nl_3 = nk_3 + O(G_{AQSP}(A, V))$$

$$\implies O(G_{AQSP}(A, V)) = nk_3 - nl_3$$

$$\implies O(G_{AQSP}(A, V)) = n(k_3 - l_3)$$

$$O(G_{AQSP}(A, V)) = n(l_3 - k_3). \text{ Hence the result.}$$

■

## 6. PERFECTLY REGULAR AQSP FUZZY SOFT GRAPH

Let  $G_{AQSP}(A, V)$  be an AQSP fuzzy soft graph on  $V$ . Then  $G_{AQSP}(A, V)$  is called as perfectly regular AQSP fuzzy soft graph if  $G_{AQSP}(A, V) = ((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N)), ((A_{AQSP}), (\mu^P, \mu^N, \gamma^P, \gamma^N))$  is a regular and totally regular AQSP fuzzy soft graph  $\forall e_i \in A_{AQSP}$ .

**Table 7:** Tabular representation of AQSP Fuzzy Soft Graph parameter vertex set.

$(\mu, \gamma)$	$v_1$	$v_2$	$v_3$	$v_4$
$e_1$	(0.8, - 0.8, 0.3,- 0.3)	(0.8, - 0.8, 0.3,- 0.3)	(0.8, - 0.8, 0.3,- 0.3)	(0.8, - 0.8, 0.3,- 0.3)
$e_2$	(0.9, - 0.9, 0.4,- 0.4)	(0.9, - 0.9, 0.4,- 0.4)	(0.9, - 0.9, 0.4,- 0.4)	(0.9, - 0.9, 0.4,- 0.4)

Table 7. represent the AQSP Fuzzy Soft Graph corresponding parametric vertex set

**Table 8:** Tabular representation of AQSP Fuzzy Soft Graph parameter edge set.

$(\mu, \gamma)$	$v_1v_2$	$v_2v_3$	$v_3v_4$	$v_4v_1$
$e_1$	(0.7, - 0.7, 0.2, - 0.2)	(0.7, - 0.7, 0.2, - 0.2)	(0.7, - 0.7, 0.2, - 0.2)	(0.7, - 0.7, 0.2, - 0.2)
$e_2$	(0.8, - 0.7, 0.3,- 0.2)	(0.8, - 0.7, 0.3, - 0.2)	(0.8, - 0.7, 0.3, - 0.2)	(0.8, - 0.7, 0.3,- 0.2)

Table 8. explains the AQSP Fuzzy Soft Graph corresponding parametric edge set



### 6.1. Example of AQSP Fuzzy Soft Graph

From the Figure.8 we get the result of AQSP fuzzy soft graph with the condition,  $tdeg_{G_{AQSP}}(A, V) = 2S(G_{AQSP}(A, V) + O(G_{AQSP}(A, V))) : 5.6 + 3.2 = 8.8$  where,  $2S(G_{AQSP}(A, V) = 5.6$  and  $O(G_{AQSP}(A, V)) = 3.2$ , then,  $tdeg_{G_{AQSP}}(A, V) = 8.8$ . Using Figure.6 and Figure.7 we can get the same result of AQSP fuzzy soft graph.

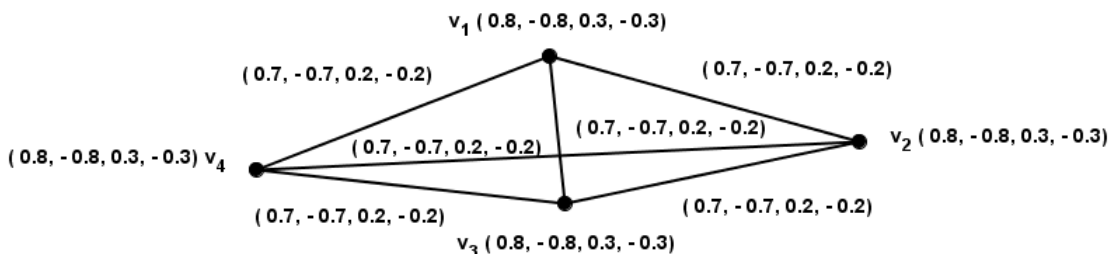


Figure 6: Perfectly regular AQSP fuzzy soft graph Corresponding to the parameter  $e_1$

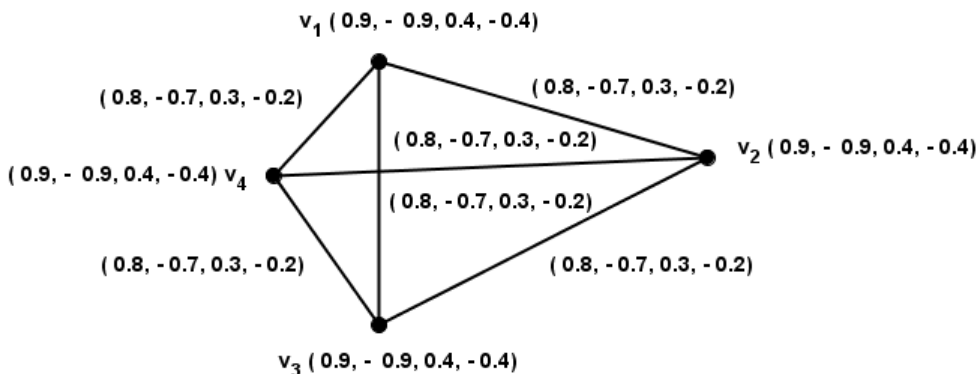


Figure 7: Perfectly regular AQSP fuzzy soft graph Corresponding to the parameter  $e_1$

**Theorem 6.** For a perfectly regular AQSP fuzzy soft graph  $G_{AQSP}(A, V)$  we have  $((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N))$  is a constant function.

**Proof.** From Theorem .4 and Theorem. 5 we prove that  $G_{AQSP}(A, V) = ((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N)), ((A_{AQSP}), (\mu^P, \mu^N, \gamma^P, \gamma^N))$  is perfectly regular AQSP fuzzy soft graph. ■

**Theorem 7.** Let  $G_{AQSP}(A, V)$  be an AQSP fuzzy soft graph. Then we prove that  $G_{AQSP}(A, V)$  is perfectly regular AQSP fuzzy soft graph if and only if the given conditions are satisfied for edges and vertices with membership values.

- (i)  $\sum_{x \neq y} \mu_e^P(x, y) = \sum_{z \neq y} \mu_e^P(z, y)$
- (ii)  $\sum_{x \neq y} \mu_e^N(x, y) = \sum_{z \neq y} \mu_e^N(z, y)$
- (iii)  $\sum_{x \neq y} \gamma_e^P(x, y) = \sum_{z \neq y} \gamma_e^P(z, y)$
- (iv)  $\sum_{x \neq y} \gamma_e^N(x, y) = \sum_{z \neq y} \gamma_e^N(z, y) \quad \forall x, y \in V, e_i \in A_{AQSP}$ .
- (v)  $\sigma_e^P(x) = \sigma_e^P(z)$ , (vi)  $\sigma_e^N(x) = \sigma_e^N(z)$
- (vii)  $\rho_e^P(x) = \rho_e^P(z)$ , (viii)  $\rho_e^N(x) = \rho_e^N(z)$ ,  $\forall x, y \in V, e_i \in A_{AQSP}$ .

**Proof.** Consider,  $G_{AQSP}(A, V)$  is perfectly regular AQSP fuzzy soft graph. By definition  $G_{AQSP}(A, V)$  is regular AQSP fuzzy soft graph, hence it trivially satisfies (i), (ii), (iii) and (iv). Therefore we have the following,

$$\begin{aligned} \text{deg}_{G_{AQSP}(A,V)}\sigma^P(x) &= \text{deg}_{G_{AQSP}(A,V)}\sigma^P(z), \\ \text{deg}_{G_{AQSP}(A,V)}\sigma^N(x) &= \text{deg}_{G_{AQSP}(A,V)}\sigma^N(z), \\ \text{deg}_{G_{AQSP}(A,V)}\rho^P(x) &= \text{deg}_{G_{AQSP}(A,V)}\rho^P(z), \\ \text{deg}_{G_{AQSP}(A,V)}\rho^N(x) &= \text{deg}_{G_{AQSP}(A,V)}\rho^N(z), \forall x, z \in V, e_i \in A_{AQSP}. \end{aligned}$$

Thus implies the results by proposition 8.2.in the following,

$$\begin{aligned} \sum_{x \neq y} \mu_e^P(x, y) &= \sum_{z \neq y} \mu_e^P(z, y) \\ \sum_{x \neq y} \mu_e^N(x, y) &= \sum_{z \neq y} \mu_e^N(z, y) \\ \sum_{x \neq y} \gamma_e^P(x, y) &= \sum_{z \neq y} \gamma_e^P(z, y) \\ \sum_{x \neq y} \gamma_e^N(x, y) &= \sum_{z \neq y} \gamma_e^N(z, y) \quad \forall x, y \in V, e_i \in A_{AQSP}. \end{aligned}$$

by Theorem.6, (v), (vi), (vii) and (viii) also holds.

Conversely, suppose that  $G_{AQSP}(A, V)$  is an AQSP fuzzy soft graph such that it satisfies the conditions from (i), (ii), (iii) and (iv).

$$\begin{aligned} \text{deg}_{G_{AQSP}(A,V)}\sigma^P(x) &= \text{deg}_{G_{AQSP}(A,V)}\sigma^P(z) = r_1, \\ \text{deg}_{G_{AQSP}(A,V)}\sigma^N(x) &= \text{deg}_{G_{AQSP}(A,V)}\sigma^N(z), r_2 \\ \text{deg}_{G_{AQSP}(A,V)}\rho^P(x) &= \text{deg}_{G_{AQSP}(A,V)}\rho^P(z), r_3 \\ \text{deg}_{G_{AQSP}(A,V)}\rho^N(x) &= \text{deg}_{G_{AQSP}(A,V)}\rho^N(z) = r_4, \forall x, z \in V, e_i \in A_{AQSP}. \end{aligned}$$

This implies that  $G_{AQSP}(A, V)$  is a regular AQSP fuzzy soft graph.

From, (v), (vi), (vii) and (viii) we get the result,

$$\begin{aligned} \text{(v)} \quad \sigma_e^P(x) &= \sigma_e^P(z) = k_1, \\ \text{(vi)} \quad \sigma_e^N(x) &= \sigma_e^N(z) = k_2 \\ \text{(vii)} \quad \rho_e^P(x) &= \rho_e^P(z) = k_3, \\ \text{(viii)} \quad \rho_e^N(x) &= \rho_e^N(z) = k_4, \forall x, z \in V, e_i \in A_{AQSP}. \end{aligned}$$

Thus,  $((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N))$  is a constant AQSP fuzzy soft function.

$$\begin{aligned} \text{tdeg}_{G_{AQSP}(A,V)}\sigma^P(z) &= \text{deg}_{G_{AQSP}(A,V)}\sigma^P(z) + \sigma^P(z) = r_1 + k_1, \\ \text{tdeg}_{G_{AQSP}(A,V)}\sigma^P(w) &= \text{deg}_{G_{AQSP}(A,V)}\sigma^P(w) + \sigma^P(w) = r_1 + k_1, \\ \text{tdeg}_{G_{AQSP}(A,V)}\sigma^N(z) &= \text{deg}_{G_{AQSP}(A,V)}\sigma^N(z) + \sigma^N(z) = r_2 + k_2, \\ \text{tdeg}_{G_{AQSP}(A,V)}\sigma^N(w) &= \text{deg}_{G_{AQSP}(A,V)}\sigma^N(w) + \sigma^N(w) = r_2 + k_2, \\ \text{tdeg}_{G_{AQSP}(A,V)}\rho^P(z) &= \text{deg}_{G_{AQSP}(A,V)}\rho^P(z) + \rho^P(z) = r_3 + k_3, \\ \text{tdeg}_{G_{AQSP}(A,V)}\rho^P(w) &= \text{deg}_{G_{AQSP}(A,V)}\rho^P(w) + \rho^P(w) = r_3 + k_3, \\ \text{tdeg}_{G_{AQSP}(A,V)}\rho^N(z) &= \text{deg}_{G_{AQSP}(A,V)}\rho^N(z) + \rho^N(z) = r_4 + k_4, \\ \text{tdeg}_{G_{AQSP}(A,V)}\rho^N(w) &= \text{deg}_{G_{AQSP}(A,V)}\rho^N(w) + \rho^N(w) = r_4 + k_4, \forall x, z \in V, e_i \in A_{AQSP}. \end{aligned}$$

The toally regular AQSP fuzzy soft graph is,

$$\begin{aligned} \text{tdeg}_{G_{AQSP}(A,V)}\sigma^P(z) &= \text{tdeg}_{G_{AQSP}(A,V)}\sigma^P(w) = k_1, \\ \text{tdeg}_{G_{AQSP}(A,V)}\sigma^N(z) &= \text{tdeg}_{G_{AQSP}(A,V)}\sigma^N(w) = k_2, \\ \text{tdeg}_{G_{AQSP}(A,V)}\rho^P(z) &= \text{tdeg}_{G_{AQSP}(A,V)}\rho^P(w) = k_3, \\ \text{tdeg}_{G_{AQSP}(A,V)}\rho^N(z) &= \text{tdeg}_{G_{AQSP}(A,V)}\rho^N(w) = k_4 \quad \forall x, z \in V, e_i \in A_{AQSP}. \end{aligned}$$

Hence  $G_{AQSP}(A, V)$  is toally regular AQSP fuzzy soft graph. This implies that

$G_{AQSP}(A, V) = ((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N)), ((A_{AQSP}), (\mu^P, \mu^N, \gamma^P, \gamma^N))$  is perfectly regular AQSP fuzzy soft graph and  $((A_{AQSP}), (\sigma^P, \sigma^N, \rho^P, \rho^N))$  is a constant function. therefore,  
 $\text{tdeg}_{G_{AQSP}(A,V)}\rho^P(z) = \text{tdeg}_{G_{AQSP}(A,V)}\rho^P(w) = k_1, k_2, k_3, \text{ and } k_4, \forall x, z \in V, e_i \in A_{AQSP}.$



## 7. APPLICATION OF AQSP FUZZY SOFT GRAPH

AQSP fuzzy soft graph can be used in the governing of women safety police network (WSPN) of a city or a district or any Non safety area region. The WSPN can be utilized using AQSP fuzzy soft graph, where the police vehicle depots are the vertices  $(v_1, v_2, v_3, \dots, v_n)$  and the route connecting two police vehicles are considered as corresponding edges.

For women safety police inspectors are positioned and the objective of the Governing problem is to find the minimum number of women inspectors required who will inspect the police vehicle for a particular time and particular bus stop or any region. The following description of AQSP fuzzy soft graph will help to find the solution of Patrolling of Police vehicle Network.

### 7.1. Method of AQSP Fuzzy Soft Graph Women Safety Police Vehicle Network

1. Let  $V = (v_1, v_2, v_3 \dots v_n)$  be the vertices of AQSP fuzzy soft graph police vehicle depots in a particular women safety vehicle network corresponding to the women Institutions, Companies, Colleges and Working places especially in bus stops.
2. We consider the edges as women working regions  $E = (v_1 v_2, v_1 v_3, v_2 v_4, \dots v_m v_n)$ . The vertices membership and non-membership values of the police vehicle  $V_i$  is determine as  $V_i \in A_{AQSP}$  for  $i = 1, 2, \dots n$ .
3. Now, define a term safety of women work is satisfied, which is the minimum number of women saved from particular people who distubs them while they stay or travel or work in different places. It is denoted as S vertices. The vehicle route is denoted by edges  $R = v_i v_j$  in Alternate quadra submerging polar fuzzy soft graph.
4. Find the membership values of the women safety vehicle route  $v_i v_j$  between the range [-1,1] using AQSP fuzzy graph soft graph with the given conditions if
  - (i)  $S > R$ , for AQSP membership values (ii)  $S < R$ , for AQSP non-membership values.
5. (a)  $\mu_e^P(x, y) \leq \sigma_e^P(x) \wedge \sigma_e^P(y)$ ,      (b)  $\mu_e^N(x, y) \geq \sigma_e^N(x) \vee \sigma_e^N(y)$ ,  
 (c)  $\gamma_e^P(x, y) \leq \rho_e^P(x) \wedge \rho_e^P(y)$ ,      (d)  $\gamma_e^N(x, y) \geq \rho_e^N(x) \vee \rho_e^N(y)$ , for all  $e \in A_{AQSP}$  and for all values of  $x, y = 1, 2, 3, \dots, n$ .
6. Let the capacity of five women police vehicle depots as vertices  $v_1 = 4, v_2 = 3, v_3 = 5, v_4 = 4$ , number of women exist in the spot facing dangerous situation denoted as edges,  $v_1, v_2 = 55, v_2, v_3 = 95, v_3, v_4 = 100, v_4, v_1 = 92$  are tabulated below.
7. The score values are measured by the AQSP score formula which gives the result of low and high self-esteem influential person,  $\frac{1}{n} \left( \frac{1}{I^P} \sum \varphi_x^P - \frac{1}{I^N} \sum \varphi_x^N \right)$

**Table 9:** Tabular representation of AQSP Fuzzy Soft Graph parameter vertex set.

$(\sigma, \rho)$	$v_1$	$v_2$	$v_3$	$v_4$
$e_1$	( 0.6, - 0.8, 0.1,- 0.3)	( 0.7, - 0.7, 0.2, -0.2)	( 0.8, - 0.9, 0.3, - 0.4 )	( 0.6, - 0.9, 0.1,- 0.4)
$e_2$	( 0.7, - 0.9, 0.2,- 0.4)	( 0.8, - 0.6, 0.3, -0.1)	( 0.9, - 0.8, 0.4, - 0.3 )	( 0.8, - 0.8, 0.3,- 0.3)
Score	0.500	0.900	0.925	0.900

The score values of the women needed safety in different spots are given with membership and non membership values of the edges are  $v_1, v_2 = 0.550, v_2, v_3 = 0.950, v_3, v_4 = 1.000, v_4, v_1 = 0.925$ . The police vehicle  $v_3 = 3$  is the important vehicle to be in the spot  $v_3, v_4 = 1.000$  where women in that area need safety. The bar diagram given below shows the result.

**Table 10:** Tabular representation of AQSP Fuzzy Soft Graph parameter edge set.

$(\mu, \gamma)$	$v_1v_2$	$v_2v_3$	$v_3v_4$	$v_4v_1$
$e_1$	( 0.7, - 0.7, 0.2, - 0.2)	( 0.7, - 0.7, 0.2, - 0.2)	(1.0, - 0.7, 0.5, - 0.2)	( 0.7, - 0.7, 0.2, - 0.2)
$e_2$	( 0.8, - 0.7, 0.3,- 0.2)	( 0.8, - 0.7, 0.3, - 0.2)	( 0.9, - 1.0, 0.4, - 0.5 )	( 0.8, - 0.7, 0.3,- 0.2)
Score	0.550	0.950	1.000	0.925

## 8. CONCLUSION

The Alternate Quadra Submerging Polar (AQSP) fuzzy graph is introduced with the basic perception of Fuzzy soft graphs. In this article, we introduce the new module AQSP fuzzy soft graphs with suitable definitions, theorems, examples, and properties. The membership and non-membership values of AQSP fuzzy soft graph is introduced with submerging level of confidence [-0.5,0.5]. The introduction of this new module AQSP fuzzy soft graph is an indispensable concept that can be rather developed into interdisciplinary subjects. The main purpose of this new graph is to find the reliable corresponding parametric membership values. The regular, totally regular, and perfectly regular AQSP fuzzy soft graph combinatoric concepts and properties can be applied in Combinatoric subjects, Applied Mathematics, Statistics, Probability, Artificial intelligence, Approximate reasoning, Teaching learning projects and Mathematical psychology. Different types of AQSP fuzzy soft graphs and the Network method of Governing the women safety vehicle in different spots are presented specifically. Finding the important police vehicle, connected routes and spots are the extent of the AQSP fuzzy soft graph. In future the extension of the AQSP fuzzy soft graph can be developed in Decision making analysis, medical diagnosis, and machine learning. The regularity of AQSP fuzzy soft sets and graphs are applicable in real life situations which are uncertain. The combinatoric membership and non-membership submerging values can be found using corresponding parameters in different fuzzy fields.

### Declarations

### Acknowledgements

The authors do thankful to the editor for giving an opportunity to submit our research article in this esteemed journal. And grateful to the Institution for providing SEED Money and MATLAB software for the research purpose.

### Conflict of interest

The authors declared that they have no conflict of interest regarding the publication of the research article.

### Contributions

The authors worked equally regarding the publication of the research article.

## REFERENCES

- [1] Akram, Muhammad. *Bipolar fuzzy graphs* Information sciences 181, no. 24 (2011): 5548-5564.
- [2] Akram, Muhammad, and Saira Nawaz. *Fuzzy soft graphs with applications* Journal of Intelligent and Fuzzy Systems 30, no. 6 (2016): 3619-3632.
- [3] Akram, Muhammad, and Sundas Shahzadi. *Novel intuitionistic fuzzy soft multiple-attribute decision-making methods* Neural Computing and Applications 29 (2018): 435-447.
- [4] Akram, Muhammad, and Saira Nawaz. *On fuzzy soft graphs* Italian journal of pure and applied mathematics 34 (2015): 497-514.
- [5] Atanassov, Krassimir T., *Intuitionistic fuzzy sets* Fuzzy sets and Systems 20, no. 1 (1986): 87-96.
- [6] Bhattacharya, Prabir. *Some remarks on fuzzy graphs* Pattern recognition letters 6, no. 5 (1987): 297-302.

- [7] Chellamani, P., D. Ajay, Said Broumi, and T. Antony Alphonse Ligori. *An approach to decision-making via picture fuzzy soft graphs* Granular Computing (2021): 1-22.
- [8] Gani, A. Nagoor, R. Jahir Hussain, and S. Yahya Mohamed. *Irregular Intuitionistic fuzzy graph* IOSR Journal of Mathematics (IOSR-JM) 9 (2014): 47-51.
- [9] Gani, A. Nagoor, and K. Radha. *Regular property of fuzzy graphs* Bulletin of Pure and Applied Science 27, no. 2 (2008): 415-423.
- [10] Kauffman, A. *Introduction a la Theorie Des Sous - Ensembles Flous*, Paris, 1973." no. 1 (1975): 120-120.
- [11] Maji, Pabitra Kumar. *More on intuitionistic fuzzy soft sets*. In Rough Sets, Fuzzy Sets, Data Mining and Granular Computing: 12th International Conference, RSFDGrC 2009, Delhi, India, December 15-18, 2009. Proceedings 12, pp. 231-240. Springer Berlin Heidelberg, 2009.
- [12] Molodtsov, Dmitriy. *Soft set theory?first results* Computers and mathematics with applications 37, no. 4-5 (1999): 19-31.
- [13] Mohinta, Sumit, and T. K. Samanta. *An introduction to fuzzy soft graph* Mathematica Moravica 19, no. 2 (2015): 35-48.
- [14] Mordeson, John N., and Premchand S. Nair. *Fuzzy graphs and fuzzy hypergraphs*, Vol. 46. Physica, 2012.
- [15] Nawaz, Hafiza Saba, and Muhammad Akram. *Oligopolistic competition among the wireless internet service providers of Malaysia using fuzzy soft graphs* Journal of Applied Mathematics and Computing (2021): 1-36.
- [16] Rosenfeld, Azriel. *Fuzzy graphs*, In Fuzzy sets and their applications to cognitive and decision processes, pp. 77-95. Academic press, 1975.
- [17] Rosline, J. Jesintha, and T. Pathinathan. *Triple layered fuzzy graph* International Journal of Fuzzy Mathematical Archive 8, no. 1 (2015): 36-42.
- [18] Anthoni Amali A and Rosline, J. Jesintha. *Alternate Quadra Submerging Polar Fuzzy Soft Graph Reliability: Theory and Applications* 18, no. 3 (74) (2023): 97-112.
- [19] Shahzadi, Gulfam, Muhammad Akram, and Bijan Davvaz. *Pythagorean fuzzy soft graphs with applications* Journal of Intelligent and Fuzzy Systems 38, no. 4 (2020): 4977-4991.
- [20] Sunitha, M. S., and A. Vijayakumar. *Studies on fuzzy graphs* PhD diss., Department of Mathematics, 2001.
- [21] Zadeh, Lotfi A. *Fuzzy sets* Information and control, 8, no. 3 (1965): 338-353.
- [22] Zhang, Wen-Ran, and Lulu Zhang. *YinYang bipolar logic and bipolar fuzzy logic* Information Sciences 165, no. 3-4 (2004): 265-287.
- [23] Mathew, Sunil, John N. Mordeson, and Davender S. Malik., *Fuzzy graph theory*, Vol. 363. Berlin, Germany: Springer International Publishing, 2018.
- [24] Rosline, J. Jesintha, and T. Pathinathan., *Structural core graph of double layered fuzzy graph*, Intern. J. Fuzzy Mathematical Archive 8, no. 2 (2015), 59-67.
- [25] Pouhassani, A., and H. Doostie. *On the properties of fuzzy soft graphs* Journal of Information and Optimization Sciences 38, no. 3-4 (2017): 541-557.

## STATISTICAL ANALYSIS OF SPLIT-PLOT DESIGN USING SPECIAL TYPE OF GRAPHS

V. Saranya<sup>1</sup>, S. Kavitha<sup>2\*</sup>, M. Pachamuthu<sup>3</sup> and S. Vijayan<sup>4</sup>



<sup>1</sup>Research Scholar, Department of Statistics, Periyar University, Salem, Tamil Nadu, India  
saranya88stat@gmail.com

<sup>2\*</sup>Assistant Professor, Department of Statistics, Periyar University, Salem, Tamil Nadu, India  
pustatkavitha@gmail.com

<sup>3</sup>Assistant Professor, Department of Statistics, Periyar University, Salem, Tamil Nadu, India  
pachamuthu@periyaruniversity.ac.in

<sup>4</sup>Research Scholar, Department of Statistics, Periyar University, Salem, Tamil Nadu, India  
vijaystatistician1210@gmail.com

### Abstract

*When all experimental runs cannot be done under homogeneous conditions, blocking can be utilized to increase the power for testing treatment effects. In many real-life environments, there is at least one factor that is hard to change, leading to a split-plot structure. This paper demonstrates how to generate certain graphs using main-plot and sub-plot analyses, as well as providing a catalog. As a result, during situations where the candidate set is too huge to be tractable, the design of split-plot experiments becomes computationally feasible. The designs were considered ideal because they were capable and efficient in estimating the fixed effects of the suitable statistical model given the split-plot design structure. The Split-Plot Design (SPD) is the complete block design which plays an important role in the fields of agriculture, medicine, and industries. This SPD is specifically suited for a two-factor experiment that has more treatments than can be accommodated by a complete block design. In an SPD, one factor is assigned to the main-plot. The assigned first factor is called the main-plot factor. The main-plot is then divided into subplots and the second factor is called the sub-plot factor. SPD is most used for (i) few experimental materials may be rare while the other experimental materials may be available in large quantity, (ii) the levels of one or more treatment factor or easy to change and the alteration of levels of other treatment factors are costly or time-consuming. Given the extensive study done in graph theory, it has developed to be a very broad subject in mathematics. Graphs are important because they are a visual way of expressing information. A graph shows data that is equivalent to many words. A graph can convey information that is difficult to express in words. A bipartite graph is a type of graph in which the entire graph may be divided into two bipartite sets, with edges connecting vertices in one set to vertices in the other. Vertex coloring is the procedure of assigning labels or colors to each vertex in a graph. The data set was also manually analyzed to validate the software-analyzed outcomes. R gave the same results as the manual analysis, showing that they were both correct. R is mainly command-based. The proposed approach is demonstrated using agricultural and industrial examples.*

**Key Word:** Split-plot design, complete bipartite graph, colored graph.

## 1. Introduction

Tamil Nadu is one of the leading rice-growing states in India and has been successfully cultivating rice since ancient times as the state has all the favorable climatic conditions suitable for rice cultivation. Rice research was initiated in Madras State to increase rice production and productivity (mixed). 1902 at Samalkota in East Godavari district was extended to 12 more places. This study uses primary data to determine the rice production of the Salem district of Tamil Nadu. The present study aims to analyze rice cultivation using different levels of nitrogen. The constraint analysis is applied to find out the problems of paddy cultivation.

A split plot design and some graphs were adopted for this present study. The split-plot design originated in the field of agriculture. Experimenters applied one treatment to a large area of land, called a whole plot, and other treatments to smaller areas of land within the whole plot called a subplot. Split plots have two types of factors Hard-to-change (HTC) applied to the whole plots and Easy-to-change (ETC) applied to the subplots. In such case different sizes of plots are required and the resulting design is known as Split Plot Design (*SPD*). In 1925, Fisher developed these designs for the purpose of agricultural experiments.

One of the fastest-expanding sciences in modern technology is graph theory. Graphs are commonly used in applications of many fields to represent different objects and their relationships. The declaration of an object vertex serves as the graph's visual representation, while an edge represents the relationships between objects. Graph theory has recently become established as a significant mathematical tool in a wide range of fields, including functional research, chemistry, genetics, and linguistics, as well as electrical engineering, geography, sociology, and architecture of themselves.

Wooding W M [1] has discussed split-plot designs characteristics and applications. To design the first section, models and least squares are reviewed. The main part shows how a fundamental split-plot design is created through a process of "evolution," starting with a completely random model and progressing through a randomized blocks design to a split-plot while using the same set of runs. George Box and Stephen Jones [2] have evaluated the applicability of split-plot designs for the experimental setting and have concentrated on the use of statistical experimental designs in designing goods that are robust to environmental factors. They conclude that the split-plot and strip-block designs are valuable for creating strong products. Peter Goos and Martina Vandebroek [3] have developed an exchange algorithm for constructing D-optimal split-plot designs and the resulting designs are analyzed. Natalino Calegario et al. [4] have analyzed the split-split-plot design and established the impact of fertilizer concentration on the establishment of Begonia and Petunia. Then they draw the conclusion that the pH values declined with fertilizer concentration over time and the EC values increased over time, resulting in values that limited nutrient availability and plant growth.

Bradley Jones and Peter Goos [5] have suggested a fresh technique for producing ideal split-plot designs. These split-plot designs are best when they are effective at estimating the fixed effects of the proper statistical model, given the structure of the design. Pwasong A D and Choji D N [6] have analyzed the rabbit feeds data obtained from the Department of Agricultural Science, Federal College of Education Pankshin and determined that there is any significant variation in the categories of feeds given. The result illustrates that there was no significant difference between the various types of feed utilized to feed the rabbits.

Bradley Jones [7] has suggested for the use of split-plot designs in industrial applications are provided after an examination of current developments. Johannes Ledolter [8] has reviewed the factorial split-plot design and fractional factorial split-plot designs experiments and uses several illustrative examples to illustrate why they frequently occur in industrial investigations. Abhishek K. Shrivastava [9] has presented an effective method for constructing split-plot design catalogues by transforming the design isomorphism problem to a graph isomorphism problem utilizing a new graph

representation. Derya Dogan and Pinar Dundar [10] have introduced the new concept of average covering number of a graph and establish the brief relationship between the average covering number and some other graph parameters.

David I J and Adehi M U [11] have utilized a 21×52 split-plot experiment with three replicates for comparison. Here they review the sorghum thresher's improved threshing efficiency. Vahide Hajihassani and Yadollah Rajaei [12] have used five agricultural machinery companies that have existed accepted in the Tehran Stock Exchange since (1388-1390) and a sample that is representative of society, conduct a split-plot design model study on the factors impacting liquidity accepted in stock exchange Agricultural Machinery companies. David I J et al. [13] have presented the steps for the estimated generalized least square (EGLS) technique, which estimates the parameters of a nonlinear split-plot design (SPD) model utilizing theoretical iterative Gauss Newton via Taylor Series expansion. Yoshimi Egawaa et al. [14] have discussed the 4-connected graph in triangles and let  $G$  be a 4-connected graph, and let  $E^{\sim}(G)$  denote the set of those edges of  $G$  which are not contained in a triangle, and let  $E_c(G)$  denote the set of 4-contractible edges of  $G$ . We show that if  $3 \leq |E^{\sim}(G)| \leq 4$  or  $|E^{\sim}(G)| \geq 7$ , then  $|E_c(G)| \geq (|E^{\sim}(G)| + 8)/4$  unless  $G$  has one of the three specified configurations.

**Table 1:** Background of this research

Review	The related articles of Split-plot design, application of split-plot design, graph theory, vertex coloring and split-plot design with colored graph are given.
Example 1	Yield of paddy in different level of nitrogen and the given data are collected from the agriculture filed of salem district.
Example 2	Application method of paint in different mixing level and the given data are collected from different hardware's in salem district.

## 2. Preliminaries

### 2.1 Split-plot Design

A randomized complete block design with two factors is no longer a randomized complete block design because the order of experiments is controlled to obtain observations in each treatment under each block. Splitting the randomization of an experiment to obtain observations under the treatment of one factor is called a split-plot design.

### 2.2 Complete Bipartite Graph

A complete bipartite graph is a graph whose vertex set  $V$  can be divided into two subsets  $V_1$  and  $V_2$  such that no edge has both endpoints in the same subset and every edge is connected to every vertex of the first subset and every vertex of the second subset.

### 2.3 Colored Graph

In a graph, the procedure for assigning the labels (colors) to the nodes or edges or areas is known as graph coloring. In this assignment no two adjacent vertices or adjacent edges or adjacent areas are getting the same color.

## 3. Statistical Analysis of Split Plot Design

The liner model for Split Plot Design is.



$$Y_{ijk} = \mu + r_i + t_j + s_k + ts_{jk} + \partial_{ij} + \varepsilon_{ijk}, \forall i=1,2,\dots,r; j=1,2,\dots,v; k=1,2,\dots,n.$$

Where,  $Y_{ijk}$  is the observation corresponding to  $k^{th}$  level of sub plot factor (B),  $j^{th}$  level of main plot factor (A) and  $i^{th}$  replication.

$\mu$  is general mean effect.

$r_i$  is  $i^{th}$  replication effect.

$t_j$  is  $j^{th}$  main - plot treatment effect.

$s_k$  is  $k^{th}$  sub - plot treatment effect.

$ts_{jk}$  is interaction effect.

The error components  $\partial_{ij}$  and  $\varepsilon_{ijk}$  are independently and normally distributed with mean zero and respective variance  $\sigma^2 \partial$  and  $\sigma^2 \varepsilon$ .

### 3.1 Main-Plot Analysis

This analysis part is based on the comparisons of main plot totals:

The levels of A are assigned to the main plots within blocks based on RBD and the sum of squares are given below,

- Correction factor (CF) =  $\frac{G^2}{rvn}$
- Total sum of square (SST) =  $\Sigma X^2 - CF$
- Replication sum of square (SSR) =  $\frac{\Sigma R^2}{vn} - CF$
- Main-plot sum of square (SS (MP)) =  $\frac{\Sigma A^2}{rn} - CF$
- Main-plot error sum of square (SSE1) =  $\frac{\Sigma (AR)^2}{n} - CF - SSR - SS(MP)$

### 3.2 Sub-Plot Analysis

This analysis part is based on the comparisons of sub plot totals:

- Sub-plot sum of square (SS (SP)) =  $\frac{\Sigma B^2}{rv} - CF$
- Interaction effect sum of square (A×B) =  $\frac{\Sigma (AB)^2}{r} - CF - SSA - SSB$
- Sub-plot error sum of square (SSE2) =  $SST - (SSR + SS (MP) + SSE1 + SS (SP) + (A \times B))$

The analysis of the variance table is outlined as follows

**Table 2:** ANOVA for split-plot designs

Sv	Df	Ss	Mss	F-Ratio
Replication	(r-1)	SSR	$S_R^2 = \frac{SSR}{(r-1)}$	$F_R = \frac{S_R^2}{S_{E_1}^2} \sim F_{(r-1), (r-1)(v-1)}$
MP(A)	(v-1)	SSA	$S_A^2 = \frac{SSA}{(v-1)}$	$F_A = \frac{S_A^2}{S_{E_1}^2} \sim F_{(v-1), (r-1)(v-1)}$
MPE (E1)	(r-1)(v-1)	SSE1	$S_E^2 = \frac{SSE_1}{(r-1)(v-1)}$	-
SP (B)	(n-1)	SSB	$S_B^2 = \frac{SSB}{(n-1)}$	$F_B = \frac{S_B^2}{S_{E_2}^2} \sim F_{(v-1), (r-1)(n-1)}$
IE(AB)	(v-1)(n-1)	SSAB	$S_{AB}^2 = \frac{SSAB}{(v-1)(n-1)}$	$F_{AB} = \frac{S_{AB}^2}{S_{E_2}^2} \sim F_{(v-1), (r-1)(n-1)}$

SPE( $E_2$ )	$v(r-1)(n-1)$	SSE <sub>2</sub>	$S_A^2 = \frac{SSE_2}{v(r-1)(n-1)}$	-
Total	$rvn-1$	SST	-	-

### 3.3 Flow – Chart

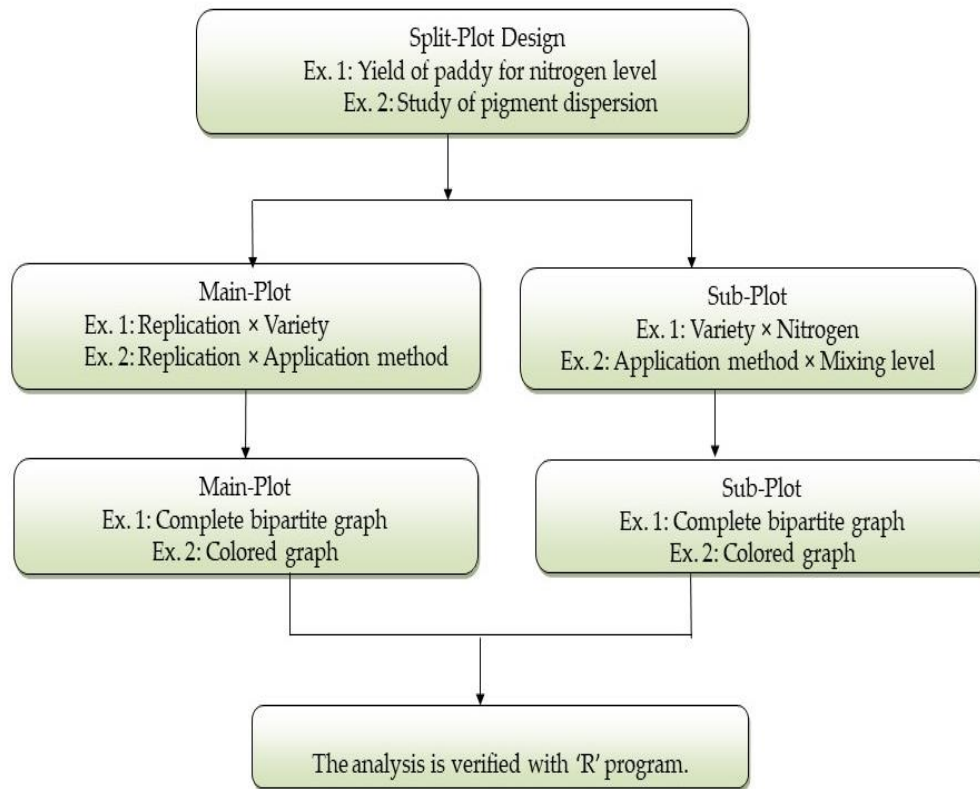


Figure 1: Flow chart

## 4. Construction of Split-Plot Design using Complete Bipartite Graph

### 4.1 Method for Construction of Complete Bipartite Graphs

- Let us consider the main-plot and sub-plot as vertex set  $S$ . This vertex set can be divided into subsets of  $S_1$  and  $S_2$ .
- In main-plot, the replication is considered as first subset  $S_1$  and variety as second subset  $S_2$ .
- Now consider the first vertex of first subset and then  $R_1$  is connected to all the vertices of second subset through edges.
- Next consider the second vertex and it is connected to all the vertices of the second subset through the edges.
- Similarly, all the remaining vertices of the first subset are connected to all the vertices of second subset through the corresponding edges.
- Finally, we get the complete bipartite graph for main-plot and sub-plot.

#### 4.1.1 Application

This example is to determine the yield response in N fertilization between different paddy varieties, three varieties of Paddy ( $V_1 = \text{ADT 36}$ ,  $V_2 = \text{ASD 16}$ ,  $V_3 = \text{IR50}$ ) are the treatments of main plot, nitrogen rates such as 0, 30 and 60 Kg/ha are the sub-plot treatments. The study was replicated four times and the primary data gathered for this experiment from the agricultural field of salem district of tamil nadu in India and shown in table 3.

**Table 3:** Replication wise data for yield of paddy (Kg/ha)

Replication	$R_1$	$R_2$	$R_3$	$R_4$
Variety	Nitrogen ( $N_1$ )			
$V_1$	15.8	19.2	13.2	13.2
$V_2$	20.8	15.3	20.5	13.8
$V_3$	15.9	16.3	16.2	12.8
	Nitrogen ( $N_2$ )			
$V_1$	17.8	20.5	14.8	13.8
$V_2$	24.8	20.8	18.8	17.8
$V_3$	18.5	16.1	20.8	12.2
	Nitrogen ( $N_3$ )			
$V_1$	21.1	24.8	13.8	18.8
$V_2$	30.5	19.2	25.7	15.2
$V_3$	18.3	18.2	22.8	10.8

**Table 4:** Replication  $\times$  variety ( $R \times V$ ) for main – plot

	$V_1$	$V_2$	$V_3$	Replication Total
$R_1$	54.7	76.1	52.7	183.5
$R_2$	64.5	55.3	51	170.8
$R_3$	41.8	65	59.8	166.6
$R_4$	45.8	46.8	35.8	128.4
Variety Total	206.8	243.2	199.3	649.3

The procedure for constructing the complete bipartite graph mentioned in section 4.1 is followed for the main-plot and sub-plot for the above experiments and then the finalized complete bipartite graph.

From the above table 4 as vertex is fixed as  $S$ , which is divided into two subsets, figure 2 shows that  $S_1$  (replication) and  $S_2$  (variety). Figure 3 shows that the first replication vertex ( $R_1$ ) and it is connected to all the vertices of variety ( $V_1, V_2$  and  $V_3$ ) through the edge values  $54.7(Y_1)$ ,  $76.1(Y_2)$  and  $54.7(Y_3)$ .

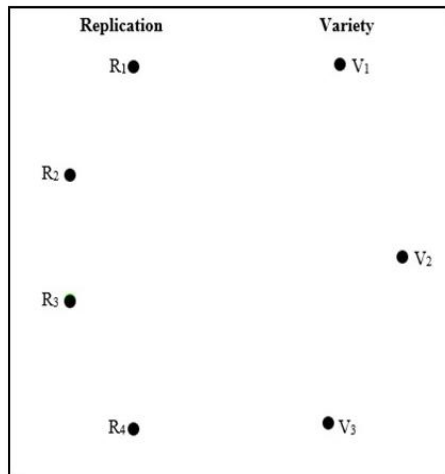


Figure 2: Graph of subsets

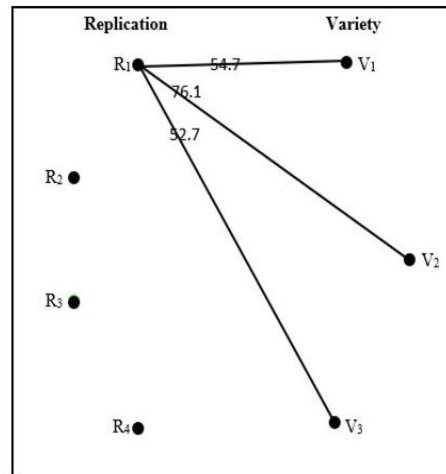


Figure 3: Graph for first replication ( $R_1$ )

Figure 4 shows that the second replication vertex ( $R_2$ ) and it is connected to all the vertices of variety ( $V_1, V_2$  and  $V_3$ ) through the edge values  $64.5(Y_1)$   $55.3(Y_2)$  and  $51(Y_3)$ . Similarly, figure 5 shows that the third and fourth replication vertices  $R_3$  and  $R_4$  are connected to all the vertices of variety ( $V_1, V_2$  and  $V_3$ ) through the corresponding edge values ( $Y_1, Y_2$  and  $Y_3$ )  $41.8, 65$ , and  $59.8 (Y_1, Y_2$  and  $Y_3)$   $45.8, 46.8$  and  $35.8$ .

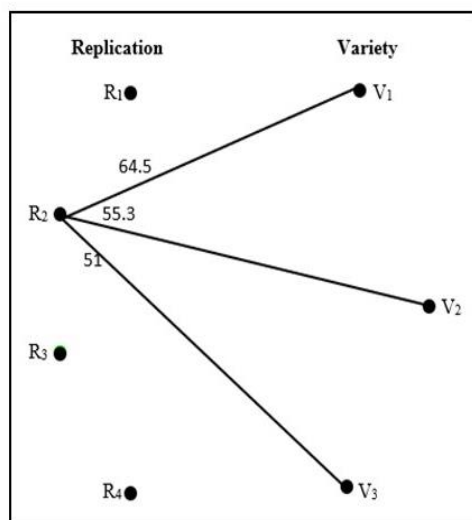


Figure 4: Graph for second replication ( $R_2$ )

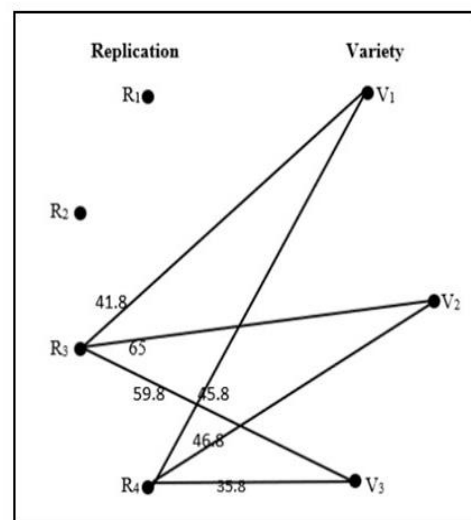


Figure 5: Graph for third and fourth replication ( $R_3$  and  $R_4$ )

Finally, figure 6 shows that the complete bipartite graph of variety and replication for main - plot.

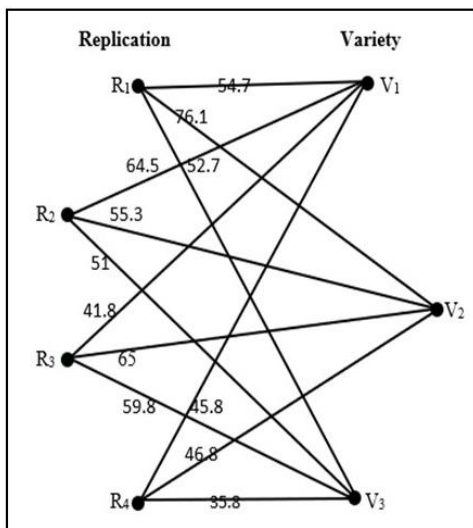


Fig. 6: Graph for complete bipartite graph of main - plot

Table 5: Variety  $\times$  nitrogen ( $V \times N$ ) for sub-plot

	$N_1$	$N_2$	$N_3$	Variety Total
$V_1$	61.4	66.9	78.5	206.8
$V_2$	70.4	82.2	90.6	243.2
$V_3$	61.2	67.6	70.5	199.2
Nitrogen Total	193	216.7	239.6	649.3

The construction of complete bipartite graph for the sub-plot are given below.

From the above table 5 as vertex is fixed as  $G$ , which is divided into two subsets, figure 7 shows that  $G_1$ (variety) and  $G_2$ (nitrogen). Figure 8 shows that the first variety ( $V_1$ ) is connected to all nitrogen ( $N_1, N_2$  and  $N_3$ ) through the values 61.4( $Y_1$ ), 66.9( $Y_1$ ) and 78.5( $Y_1$ ).

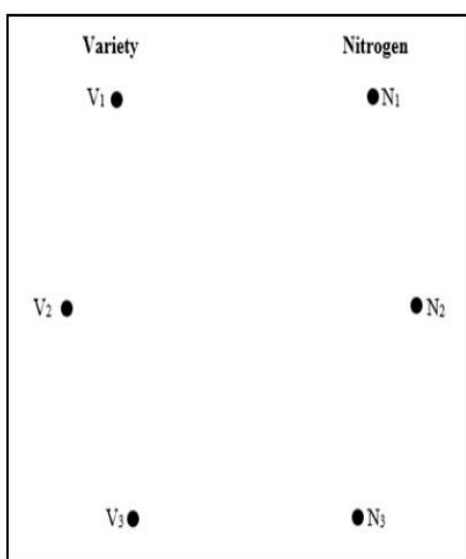


Figure 7: Graph for vertex subset

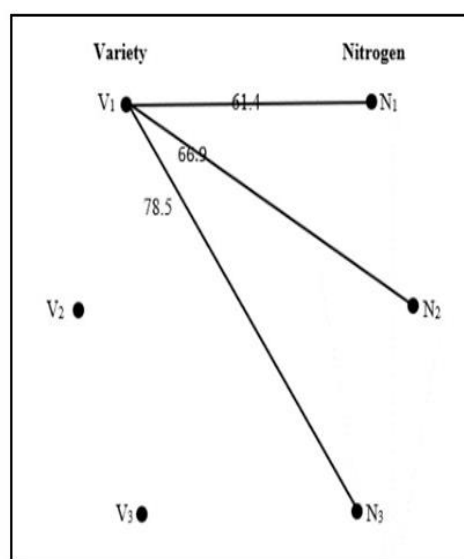


Figure 8: Graph for first variety ( $V_1$ )

Similarly, figure shows that the second and third variety  $V_2$  and  $V_3$  is connected to all the nitrogen ( $N_1, N_2$  and  $N_3$ ) through the corresponding values 70.4, 82.2 and 90.6 ( $Y_1, Y_2$  and  $Y_3$ ), 61.2, 67.6 and 70.7 ( $Y_1, Y_2$  and  $Y_3$ ). Finally, figure 10 shows the complete bipartite graph for variety and nitrogen.

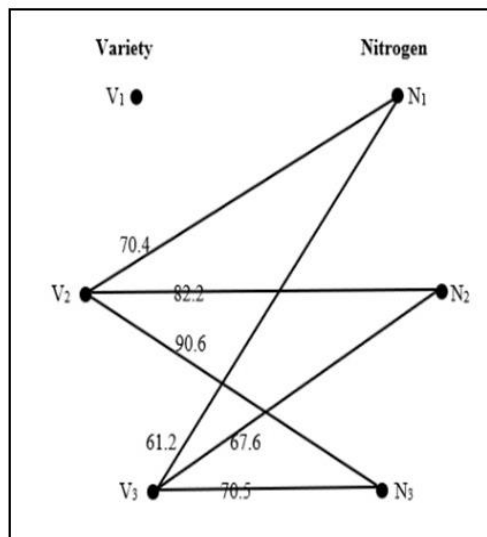


Figure 9: Graph for second and third variety ( $V_2$  and  $V_3$ )

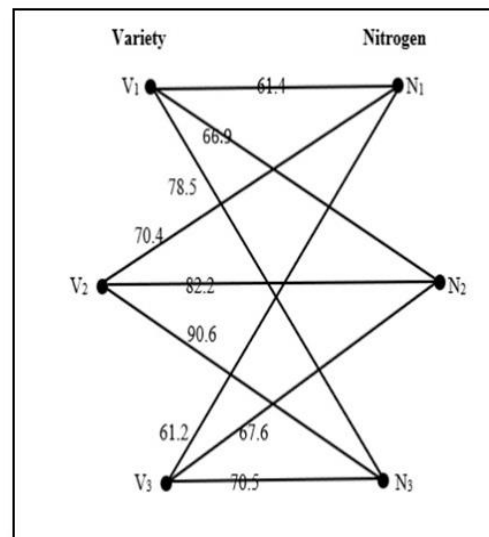


Figure 10: Complete bipartite graph for sub – plot

The sum of squares for main- plot:

- Correction factor (CF) = 17910.8469
- Total sum of square (SST) = 637.46
- Replication sum of square (SSR) = 187.709
- Variety sum of square (SSV) = 91.9006
- Main-plot error sum of square ( $SSE_1$ ) = 175.4527

The sum of squares for sub-plot:

- Nitrogen sum of square (SSN) = 90.4906
- Interaction effect sum of square ( $V \times N$ ) = 10.4194
- Sub-plot error sum of square ( $SSE_2$ ) = 81.4869

The ANOVA table for split-plot design is shown in below table:

Table 6: ANOVA table for split-plot design

Sv	Df	Ss	Mss	F-Ratio	P-Value
Replication	3	187.7098	62.570	2.1397	0.196473
Variety (A)	2	91.9006	45.950	1.5714	0.282634
Main - plot error( $E_1$ )	6	175.4527	29.242	-	-
Nitrogen(B)	2	90.4906	45.245	9.9945	0.001204**
Interaction (AB)	4	10.4194	2.605	0.5754	0.684090
Sub - plot error( $E_2$ )	18	81.49	4.527	-	-
Total	35	-	-	-	-

The table value of replication and variety are greater than the calculated values. So, the null hypothesis is accepted. There is no significant difference between the four replications and three varieties. The table value of nitrogen level is greater than the calculated value. So, the null hypothesis is accepted. There is no significant difference between the three nitrogen levels. The table value of the interaction effect is

also greater than the calculated value. So, the null hypothesis is accepted.

There is no significant difference between the interaction effects. The P-value of the above experiment is greater than the 5% level of significant. Therefore, the null hypothesis is accepted. There is no significant difference that occurred in the above experiment.

## 4.2 Method for Construction of Colored Graph

- Let us consider the main-plot and sub-plot factors as the vertex set  $S_1$  and  $S_2$ . Here the common factor assigned as  $S_1$  and the other factor assigned  $S_2$ .
- The vertices of set  $S_1$  and  $S_2$  are colored using the vertex coloring and the vertices are differentiate with different colors. Now consider the first vertex of  $S_1$  and it is connected to the corresponding vertices of  $S_2$  through edges.
- Next the second vertex of  $S_1$  which is connected to the corresponding vertices of  $S_2$  through edges.
- Similarly, all the remaining vertices of the first set are connected to the corresponding vertices of second set through the corresponding edges.
- Finally, we get the colored graph (vertex coloring graph) for main-plot and sub-plot.

### 4.2.1 Application

The test is designed to examine pigment dispersion in paint. Three different mixing levels of a particular pigment are studied. The procedure consists of three application methods (brushing, sparing, and rolling) and measured the percentage reflectance of a pigment. Four days required running the experiment from hardware shops in salem district and the data obtained below.

**Table 7:** Replication wise data form mixes level and application method of paint

Replication	$R_1$	$R_2$	$R_3$	$R_4$
Application Method	Mixing level ( $M_1$ )			
$A_1$	65.8	70.2	65.2	69.2
$A_2$	69.8	65.3	70.5	63.8
$A_3$	70.8	67.3	68.2	69.8
	Mixing level ( $M_2$ )			
$A_1$	68.7	73.5	69.9	66.8
$A_2$	74.8	70.8	68.8	67.8
$A_3$	50.8	69.1	71.8	63.2
	Mixing level ( $M_3$ )			
$A_1$	72.2	77.8	71.6	70.8
$A_2$	81.5	69.2	75.7	65.2
$A_3$	69.3	71.6	77.8	60.8

**Table 8:** Replication  $\times$  application method ( $R \times A$ ) for main - plot

	$A_1$	$A_2$	$A_3$	Replication Total
$R_1$	206.8	226.1	190.9	623.8

$R_2$	221.5	205.3	208	634.8
$R_3$	206.7	215	217.8	639.5
$R_4$	206.8	196.8	193.8	597.4
Application Total	841.8	843.2	810.5	2495.5

**Table 9:** Application method  $\times$  mixing ( $A \times M$ ) for sub-plot

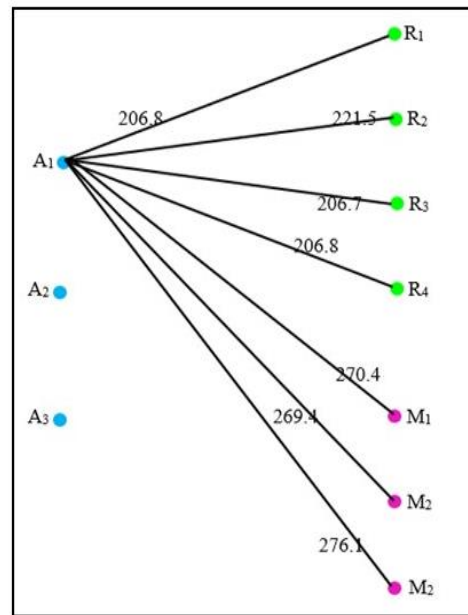
	$M_1$	$M_2$	$M_3$	Application Total
$A_1$	270.4	219	292.4	841.8
$A_2$	269.4	282.2	291.6	843.2
$A_3$	276.1	254.9	279.5	810.5
Mixing Total	815.9	816.1	863.5	2495.5

The procedure for constructing the colored graph mentioned in section 4.2 for main-plot and sub-plot for the above experiments and then the finalize.

Here we take main-plot factors (replication and application) and sub-plot factors (application method and mixing level) as the set  $S_1$  and  $S_2$ . Figure 11 shows that the Here  $S_1$  consists of the common factor which is application and  $S_2$  consists of factors such as replication and mixing level. Next, figure 12 shows that the first vertex  $A_1$  of first set and then  $A_1$  is connected to the corresponding (replication and mixing) vertices of the second set through edges.



**Figure 11:** Colored graph of vertex set



**Figure 12:** Colored graph for first application method

Next figure 13 shows that the second vertex  $A_2$  of first set and it is connected to the corresponding vertices of second set. Similarly, figure 14 shows that the third vertex  $A_3$  of first set are connected to the corresponding vertices of second set and finally, we get the colored graph (vertex coloring graph) for main-plot and sub-plot.



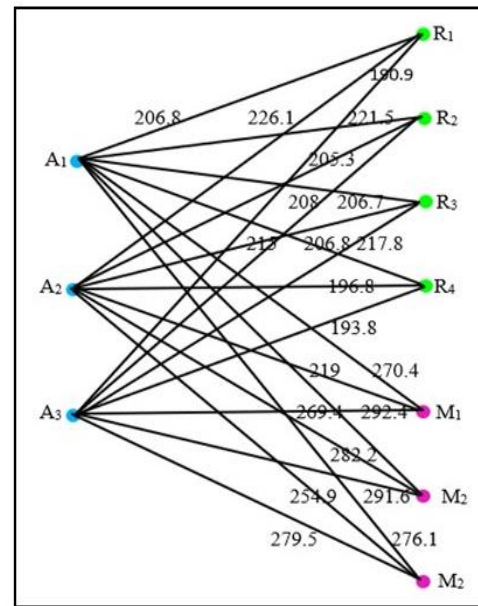
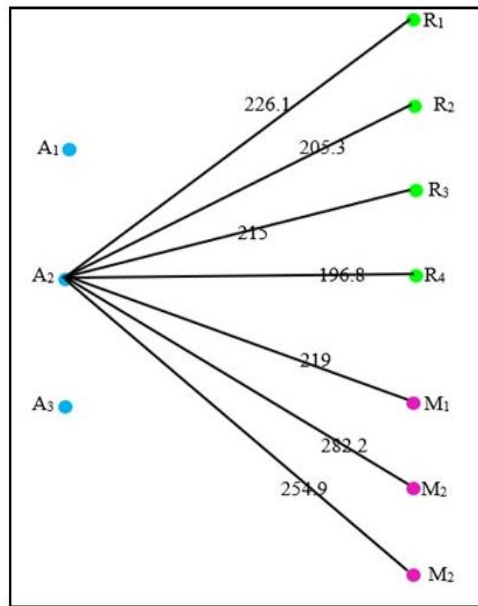


Figure 13: Colored graph for second application method

Figure 14: Colored graph for main – plot and sub - plot

The sum of squares for main-plot:

- Correction factor (CF) = 172986.6738
- Total sum of square (SST) = 975.3762
- Replication sum of square (SSR) = 118.269533
- Application method sum of square (SSA) = 56.970367
- Main-plot error sum of square (SSE<sub>1</sub>) = 173161.9137

The sum of squares for main-plot:

- Sub-plot sum of square (SSM) = 125.348
- Interaction effect sum of square (A×M) = 87.0449
- Sub-plot error sum of square (SSE<sub>2</sub>) = 334.2397

Table 10: ANOVA table for split- plot design

Sv	Df	Ss	Mss	F-Ratio	P-Value
Replication	3	118.2695	39.4217	0.9330	0.48062
Application method(A)	2	56.97037	28.4852	0.6742	0.54435
Main Plot Error(E <sub>1</sub> )	6	253.503	42.2505	-	-
Mixing(M)	2	125.3487	62.674	3.3752	0.05691
Interaction (AM)	4	87.0449	21.7612	1.1719	0.35616
Sub-plot Error(E <sub>2</sub> )	18	334.2397	18.5688	-	-
Total	35	-	-	-	-

The table values of replication and application method are greater than the calculated values. So, the null hypothesis is accepted. There is no significant difference between the four replications and three application methods. The table value of mixing level is greater than the calculated value. So, the null hypothesis is accepted. There is no significant difference between the three application methods. The table value of the interaction effect is also greater than the calculated value. So, the null hypothesis is accepted. There is no significant difference between the interaction effects.

The P-value of the above experiment is greater than the 5% level of significant. Therefore, the null hypothesis is accepted. There is no significant difference that occurred in the above experiment.

**Table 11:** Table for comparative study

ANOVA	Example 1	Example 2
Traditional method	Null hypothesis is accepted	Null hypothesis is accepted
R-Software method	Null hypothesis is accepted	Null hypothesis is accepted

## 5. Conclusion

Many of these real-world agricultural and industrial experiments involve factors called HTC. In these situations, experimenters have realized that the most efficient way to conduct an experiment is to fix the level of the hard-to-change factor and then run all or some combination of the easily changeable factors. This is repeated a few times. As we have seen, this leads to a split-plot design. Accounting for the split-plot nature of the design is equally important in the analysis of the data because the split-plot test contains two error terms. The present paper is classified into three parts namely rice production, nitrogen level and variety of rice in salem district. Constraint analysis is used to increase rice production. To construct and analyze the SPD using some special type of graphs through numerical examples from different field and the hypothesis testing is compared by the split-plot ANOVA method with software using method. When comparing the results of these methods, they produce the same results. Here some special type of graphs is used to construct the SPD. In future, there is an idea to expanding this procedure to other experimental designs such as strip-plot design and incomplete block designs etc.

### List of Abbreviation

Sv - Sources of variance  
 Df - Degrees of freedom  
 Ss - Sum of squares  
 Mss - Mean sum of squares  
 MP(A) – Main -Plot(A)  
 MPE(E<sub>1</sub>) – Main – Plot Error (E<sub>1</sub>)  
 SP(B) – Sup -Plot(B)  
 SPE(E<sub>2</sub>) - Sup – Plot Error (E<sub>2</sub>)  
 IE(AB) - Interaction Effect (AB)

### References

- [1] George Box and Stephen Jones. (1992). Split-Plot Designs for Robust Product Experimentation, *Journal of Applied Statistics*, (1):3-26.
- [2] Robin J. Wilson. Introduction to Graph Theory, Fourth Edition, Addison Wesley Longman Limited, England, 1996.
- [3] Peter Goos, and Martina Vandebroek. (2001). Optimal Split-Plot Designs, *Journal of Quality Technology*, 33(4): 436-450.

- [4] Calegario, C. L. L., Calegario, N., Lemos, P. C., and de Mello, A. A. (2006). Effect of different fertilization practices on the growth of two plant species: Begonia sp. AND Petunia sp. *Cerne*, 12(2):107-112.
- [5] Bradley Jones, and Peter Goos. (2007). A Candidate-Set-Free Algorithm for Generating D-Optimal Split-Plot Designs, *Applied Statistics*, 56(3):347–364.
- [6] Pwasong, A D and Choji D N. (2008). The Application of the Split-Plot Design in the Analysis of Experimental Rabbit Feeds Data, *Science World Journal*, 3(2):73-77.
- [7] Bkggradley Jones. (2009). Split-Plot Designs: What, Why, and How, *Journal of Quality Technology*, 41(4): 340-361.
- [8] Johannes Ledolter. (2010). Split-Plot Designs: Discussion and Examples", *International Journal Quality Engineering and Technology*, 1(4):441-457.
- [9] Abhishek, K. Shrivastava. (2013). Efficient Construction of Split-Plot Design Catalogs using Graphs, *IIE Transactions*,45, 1137–1152.
- [10] Derya Dogan, and Pinar Dundar. (2013). The Average Covering Number of a Graph Hindawi Publishing Corporation, *Journal of Applied Mathematics*, 1-4.
- [11] A. Bondy, U.S.R. Murty. Graph theory, *Springer International Edition*, 2013.
- [12] David, I J and Adehi, M U. (2014). Effectiveness of Split-Plot Design over Randomized Complete Block Design in Some Experiments", *Journal of Biology, Agriculture and Healthcare*, 4(19): 75-80.
- [13] Vahide Hajihassani, and Yadollah Rajaei. (2015). Using Model of Split-Plot Design in the Study of Factors Affecting on Accepted in Stock Exchange Agricultural Machinery Companies Liquidity, *Indian Journal of Science and Technology*, 8:526–529.
- [14] David I J, Asiribo, O E and Dikko, H G. (2018). On Parameters Estimation of Nonlinear Split-Plot Design Model with EGLS-MLE, *Professional Statisticians Society of Nigeria*, 2:381-386.
- [15] Douglas, C. Montgomery. Design and analysis of experiments, Fifth Edition, *International Student Version*, Arizona State University, 2016.
- [16] Egawa, Y., Kotani, K., and Nakamura, S. (2018). Structure of edges in a 4-connected graph not contained in triangles and the number of contractible edges. *AKCE International Journal of Graphs and Combinatorics*, 15(2):202-210.
- [17] David, I. J., Asiribo, O. E., and Dikko, H. G. (2019). Parameter estimation of nonlinear split-plot design models: a theoretical framework. *Journal of Reliability and Statistical Studies*, 117-129.
- [18] Santoso, K. A., Agustin, I. H., Prihandini, R. M. and Alfarisi, R. (2019, April). Vertex colouring using the adjacency matrix. In *Journal of Physics: Conference Series* 1211(1): 012019.

# INFERENCE ON THE TIME-DEPENDENT STRESS-STRENGTH RELIABILITY MODELS BASED ON FINITE MIXTURE MODELS

KRISHNENDU K. <sup>1</sup>, ANNIE SABITHA PAUL <sup>2</sup>, DRISYA M. <sup>3</sup> & JOBY K. JOSE <sup>4</sup>

•

<sup>1,4</sup> Dept. of Statistical Sciences, Kannur University, Kerala 670567, India.

<sup>2</sup> Dept. of Applied Sciences, Govt. College of Engineering Kannur, Kerala 670563, India.

<sup>3</sup> Dept. of Statistics, Govt. Victoria College Palakkad, Kerala 678001, India.

<sup>1</sup> krishnendu61195@gmail.com , <sup>2</sup> anniesabithapaul@gmail.com,

<sup>3</sup> drisyam.m@gmail.com, <sup>4</sup> jobyk@kannuruniv.ac.in.

## Abstract

*Time-dependent stress-strength reliability engages with the chance of survival for systems with dynamic strength and/or dynamic stress. When a system is allowed to run continuously, each run will cause a change in the strength of the system. The repeated occurrence of stress on the system over each run will affect the survival capacity of the system. In this paper, we consider the distribution of time taken for the completion of a run by the system follows gamma and the stress or strength of the system follows a finite mixture of lifetime probability models. Here we consider two cases in which the first case deals with stress and strength following a finite mixture of Weibull distribution and in the second case the stress and strength is assumed to follow a finite mixture of the power-transformed half-logistic distribution. Moreover, the strength of the system is assumed to decrease by a constant and the stress acting on the system is assumed to increase by a constant over each run. We obtained the expression of the stress-strength reliability function and explained the ML and Bayesian methods for the estimation of the reliability at various time points.*

**Keywords:** Time-dependent Stress-strength reliability, Gamma Renewal process, Finite mixture distribution, Expectation Maximization algorithm, Markov Chain Monte Carlo method.

## 1. INTRODUCTION

In reliability theory, stress-strength reliability measures the chance of the strength of a system to overcome the stress acting on it. Every object or individual has its own strength for survival. When they are subject to any kind of stress, they will survive only if their strength surpasses the stress. Stress-strength reliability model can be used to compare the effectiveness of two treatments, to compare the life length of two equipment, etc. Let  $Y$  denotes the random strength of the system under consideration and  $X$  is the stress acting on that system. Then the stress-strength reliability of the system is denoted by  $R$  and is defined as  $R = P[X < Y]$ .

The concept of stress-strength reliability theory was originated by Birnbaum [2]. Kotz et.al. [11] discussed point and interval estimation of stress-strength models using different approaches. Baklizi and Eidous [1] proposed an estimator of  $R$  based on kernel estimators of the densities of  $X$  and  $Y$ . Zhou [20] illustrated the estimation of  $R$  using the bootstrap method. Recently many authors discussed classical and Bayesian methods of estimating  $R$  for different probability models, see Pakdaman et al. [12] Xavier and Jose [15,16], Xavier *et al.* [17, 18] and Jose et.al. [7,10].

Nowadays, research on stress-strength reliability estimation focuses on the case where the stress, strength or both changes with respect to time, and hence the term time-dependent stress-strength reliability. Let  $Y(t)$  represent the strength of a system at time  $t$  and  $X(t)$  be the stress on the system at  $t$ . Under the time-dependent stress-strength reliability model, we are interested in the estimation of the stress-strength reliability function

$$R(t) = P[X(t) < Y(t)], \tag{1}$$

which gives the chance of survival of the system at time  $t$ . For example, quite often we have to download files to mobile phones. The downloaded files consume the memory space of the phone corresponding to the size of that file. It will cause a reduction in the speed of functioning of the phone. So each time we download a new file, the number of files piled up in the phone memory which will reduce the functioning speed of the phone. Time dependent stress-strength reliability models were studied in Yadav [19], Gopalan and Venkateswarlu [5, 6], Eryilmaz [4] and Siju and Kumar [13, 14], Jose and Drisya [8, 9] and Drisya et al. [3].

Time-dependent stress-strength reliability engages with the chance of survival for systems with dynamic strength and/or dynamic stress. When a system is allowed to run continuously, each run will cause a change in the strength of the system. The repeated occurrence of stress on the system over each run will affect the survival capacity of the system. In this paper, we consider the distribution of time taken for the completion of a run by the system follows gamma and the stress or strength of the system follows a finite mixture of lifetime probability models. Here we consider two cases in which the first case deals with stress and strength following a finite mixture of Weibull distribution and in the second case the stress and strength are assumed to follow a finite mixture of the power-transformed half-logistic distribution. Moreover, the strength of the system is assumed to decrease by a constant and the stress acting on the system is assumed to increase by a constant over each run.

This paper is organized as follows. Estimation of stress-strength reliability function with gamma cycle times under random fixed stress and strength is discussed in Section 2. The expressions for stress-strength reliability function under a finite mixture of Weibull and a finite mixture of power-transformed half-logistic distributions are also derived. A brief description of the EM algorithm for estimating  $R(t)$  is given in Section 3 with numerical illustrations based on simulated data. Computation of the Bayes estimate of  $R(t)$  using the Markov Chain Monte Carlo method is illustrated in Section 4 with a numerical illustration based on simulated data.

## 2. ESTIMATION OF $R(t)$ BASED ON FINITE MIXTURE DISTRIBUTION

Consider a system that is allowed to work continuously. The system executes several runs during the time period of observation say  $(0, t)$ . The time taken for completion of a run by the system is a random variable and we call it cycle time. In this paper, we assume that the cycle times are gamma-distributed. Hence the total number of runs within the entire time period will have a renewal process. Let the cycle time  $Z$  follows gamma distribution with p.d.f.,

$$f(z) = \frac{a^k z^{k-1} e^{-az}}{(k-1)!}; z \geq 0. \tag{2}$$

Then the number of runs during the time interval  $(0, t)$ , say  $N(t)$  has the following distribution.

$$\begin{aligned} P_n(t) &= p[N(t) = n] \\ &= e^{-at} \sum_{r=nk}^{(n+1)k-1} \frac{(at)^r}{r!}; n = 0, 1, 2, \dots \end{aligned} \tag{3}$$

Let  $X_j$  be the stress imposed on the system during  $j^{th}$  cycle time and the corresponding strength of the system be  $Y_j$ . Also let the initial strength of the system, say  $Y_0$  be a continuous random variable with density function  $h(y_0)$  and the initial stress on the system  $X_0$  also be a

continuous random variable with p.d.f  $g(x_0)$ . The system is allowed to run continuously and when it runs, its strength decreases by  $a_0$  and the stress increases by  $b_0$  on completion of each run. Hence, the probability that the system works after  $n$  runs is given by

$$\begin{aligned} R_n &= P((X_1 < Y_1) \cap (X_2 < Y_2) \cap \dots \cap (X_n < Y_n)) \\ &= P((x_0 + b_0 < y_0 - a_0) \cap (x_0 + 2b_0 < y_0 - 2a_0) \cap \dots \cap (x_0 + nb_0 < y_0 - na_0)) \\ &= P(x_0 + n(a_0 + b_0) < y_0) \\ &= \int_0^\infty \int_{x_0+n(a_0+b_0)}^\infty h(y_0)g(x_0)dy_0dx_0 \end{aligned} \tag{4}$$

Therefore the reliability of the system at time  $t$  is

$$\begin{aligned} R(t) &= \sum_{n=0}^\infty P_n(t)R_n \\ &= \sum_{n=0}^\infty P_n(t) \int_0^\infty \int_{x_0+n(a_0+b_0)}^\infty h(y_0)g(x_0)dy_0dx_0 \end{aligned} \tag{5}$$

$$= \sum_{n=0}^\infty e^{-at} \sum_{r=nk}^{(n+1)k-1} \frac{(at)^r}{r!} \int_0^\infty \int_{x_0+n(a_0+b_0)}^\infty h(y_0)g(x_0)dy_0dx_0 \tag{6}$$

In particular, consider the case that stress acting on the system do not vary throughout the observation period as well as the strength of the system decreases by a constant say,  $a_0$ . Then the probability of functioning of the system after  $n$  runs is given by

$$\begin{aligned} R_n &= P[(X_1 < Y_1) \cap (X_2 < Y_2) \cap \dots \cap (X_n < Y_n)] \\ &= P[(x_0 < Y_0 - a_0) \cap (x_0 < Y_0 - 2a_0) \cap \dots \cap (x_0 < Y_0 - na_0)] \\ &= P[(x_0 + na_0 < Y_0)] \\ &= \int_{x_0+na_0}^\infty h(y_0)dy_0 \end{aligned} \tag{7}$$

Therefore, the value of  $R(t)$  can be obtained as

$$\begin{aligned} R(t) &= \sum_{n=0}^\infty P_n(t)R_n \\ &= \sum_{n=0}^\infty e^{-at} \sum_{r=nk}^{(n+1)k-1} \frac{(at)^r}{r!} \int_{x_0+na_0}^\infty h(y_0)dy_0. \end{aligned}$$

### 2.1. R(t) based on finite mixture Weibull distribution

Let the initial strength of the system follow a mixture of Weibull distributions with p.d.f.

$$h(y_0) = \sum_{i=1}^{m_1} \pi_i \frac{\alpha}{\beta_i} y_0^{\alpha-1} e^{-y_0^\alpha/\beta_i}, y_0 \geq 0, \alpha > 0, 0 < \pi_i < 1, \beta_i > 0; i = 1, 2, \dots, m_1. \tag{8}$$

and initial stress on the system follows a mixture of Weibull distribution with p.d.f.

$$g(x_0) = \sum_{j=1}^{m_2} p_j \frac{\alpha}{\theta_j} x_0^{\alpha-1} e^{-x_0^\alpha/\theta_j}, x_0 \geq 0, \alpha > 0, 0 < p_j < 1, \theta_j > 0; j = 1, 2, \dots, m_2. \tag{9}$$

When the system runs, its strength decreases by  $a_0$  and the stress increases by  $b_0$  on completion of each run. The time taken for completion of a run is assumed to be a gamma variate. Then the chance for survival of the system after  $n$  runs is

$$R_n = \sum_{i=1}^{m_1} \pi_i \sum_{j=1}^{m_2} p_j e^{-(n(a_0+b_0))^\alpha/\beta_i}; n = 1, 2, \dots \tag{10}$$

with

$$R_0 = \sum_{i=1}^{m_1} \pi_i \sum_{j=1}^{m_2} p_j \frac{\beta_i}{\beta_i + \theta_j} \tag{11}$$

Then the corresponding stress-strength reliability function is obtained as

$$R(t) = e^{-at} \sum_{r=0}^{k-1} \frac{(at)^r}{r!} \sum_{i=1}^{m_1} \pi_i \sum_{j=1}^{m_2} p_j \frac{\beta_i}{\beta_i + \theta_j} + \sum_{n=1}^{\infty} e^{-at} \sum_{r=nk}^{(n+1)k-1} \frac{(at)^r}{r!} \sum_{i=1}^{m_1} \pi_i \sum_{j=1}^{m_2} p_j e^{-(n(a_0+b_0))^\alpha / \beta_i} \tag{12}$$

Change in  $R(t)$  corresponding to change in different parameters stress and strength distributions are given in Figure 1. From the figure, it is clear that the value of  $R(t)$  increases with an

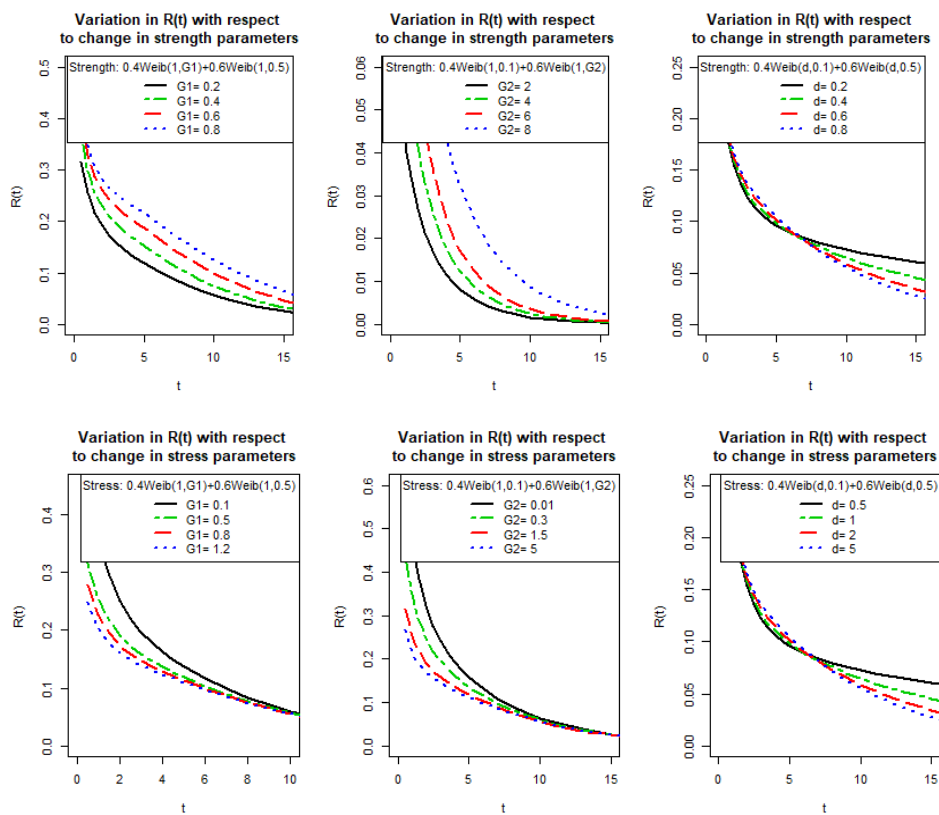


Figure 1: Variation in  $R(t)$  corresponding to change in parameters

increase in shape parameter values and decreases with an increase in scale parameter values of strength when the initial strength of the the system is Weibull-distributed. Also  $R(t)$  increases with an increase in shape parameter values of stress distribution.

As a particular case assume that the strength of the system has a mixture Weibull distribution with parameters  $(\alpha, \beta_i); i = 1, 2, \dots, m_1$ , and the stress is fixed. Then the chance of the system working after the completion of n runs is,

$$R_n = e^{-(x_0+na_0)^\alpha / \beta} \tag{13}$$

and the corresponding stress strength reliability is obtained as

$$R(t) = \sum_{n=0}^{\infty} e^{-at} \sum_{r=nk}^{(n+1)k-1} \frac{(at)^r}{r!} \sum_{i=1}^{m_1} \pi_i e^{-(x_0+na_0)^\alpha / \beta_i} \tag{14}$$

$$= \sum_{i=1}^{m_1} \pi_i \sum_{n=0}^{\infty} e^{-at} \sum_{r=nk}^{(n+1)k-1} \frac{(at)^r}{r!} e^{-(x_0+na_0)^\alpha / \beta_i} \tag{15}$$

## 2.2. R(t) based on finite a mixture of power transformed half logistic distribution

The p.d.f of the power transformed half-logistic distribution (Xavier and Jose (2020)) is given by

$$f(y) = \begin{cases} 2\delta\gamma y^{\gamma-1} e^{-\delta y^\gamma} (1 + e^{-\delta y^\gamma})^{-2}, & 0 \leq y < \infty; \delta > 0; \gamma > 0. \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

Now, let us assume that initial strength (  $Y_0$  ) of the system follows a mixture of power transformed half logistic distribution with p.d.f

$$h(y_0) = \sum_{i=1}^{m_1} \pi_i 2\delta_{1i} \gamma_{1i} y_0^{\gamma_{1i}-1} e^{-\delta_{1i} y_0^{\gamma_{1i}}} (1 + e^{-\delta_{1i} y_0^{\gamma_{1i}}})^{-2}, \tag{17}$$

$0 \leq y_0 < \infty, \delta_{1i} > 0, 0 < \pi_i < 1, \gamma_{1i} > 0, ; i = 1, 2, \dots, m_1$ . It is also assumed that initial stress on the system (  $X_0$  ) follows the mixture of power transformed half logistic distribution with p.d.f

$$g(x_0) = \sum_{j=1}^{m_2} p_j 2\delta_{2j} \gamma_{2j} x_0^{\gamma_{2j}-1} e^{-\delta_{2j} x_0^{\gamma_{2j}}} (1 + e^{-\delta_{2j} x_0^{\gamma_{2j}}})^{-2}, \tag{18}$$

$0 \leq x_0 < \infty, \delta_{2j} > 0, 0 < p_j < 1, \gamma_{2j} > 0, ; j = 1, 2, \dots, m_2$ .

Hence,  $R_n$  is given by

$$R_n = 4 \sum_{i=1}^{m_1} \pi_i \sum_{j=1}^{m_2} p_j \delta_{2j} \gamma_{2j} \times \int_0^\infty [1 - (1 + e^{-\delta_{1i}(x_0+n(a_0+b_0))^\gamma})^{-1}] x_0^{\gamma_{2j}-1} e^{-\delta_{2j} x_0^{\gamma_{2j}}} (1 + e^{-\delta_{2j} x_0^{\gamma_{2j}}})^{-2} dx_0. \tag{19}$$

Then, the stress-strength reliability is given by

$$R(t) = 4 \sum_{n=0}^{\infty} e^{-at} \sum_{r=nk}^{(n+1)k-1} \frac{(at)^r}{r!} \sum_{i=1}^{m_1} \pi_i \sum_{j=1}^{m_2} p_j \delta_{2j} \gamma_{2j} \times \int_0^\infty [1 - (1 + e^{-\delta_{1i}(x_0+n(a_0+b_0))^\gamma})^{-1}] x_0^{\gamma_{2j}-1} e^{-\delta_{2j} x_0^{\gamma_{2j}}} (1 + e^{-\delta_{2j} x_0^{\gamma_{2j}}})^{-2} dx_0 \tag{20}$$

Change in R(t) corresponding to change in different parameters stress and strength distributions are given in Figure 2. From the graph, when the stress and strength parameters follow a mixture of power transformed half logistic distribution, the increase in the parameters results in a decrease in the R(t) and after a point, they converge. Particularly when stress is fixed and strength of the system has a mixture of power transformed half logistic distribution with parameters  $(\delta_i, \gamma_i) : i = 1, 2, \dots, m_1$ , the chance of the system working after the completion of n runs is,

$$R_n = \sum_{i=1}^{m_1} 2\pi_i [1 - (1 + e^{-\delta_i(x_0+na_0)^\gamma})^{-1}]. \tag{21}$$

Then, corresponding R(t) is given by

$$R(t) = \sum_{n=0}^{\infty} e^{-at} \sum_{r=nk}^{(n+1)k-1} \frac{(at)^r}{r!} \sum_{i=1}^{m_1} 2\pi_i [1 - (1 + e^{-\delta_i(x_0+na_0)^\gamma})^{-1}]. \\ = \sum_{i=1}^{m_1} \pi_i \sum_{n=0}^{\infty} e^{-at} \sum_{r=nk}^{(n+1)k-1} \frac{(at)^r}{r!} 2[1 - (1 + e^{-\delta_i(x_0+na_0)^\gamma})^{-1}]. \tag{22}$$



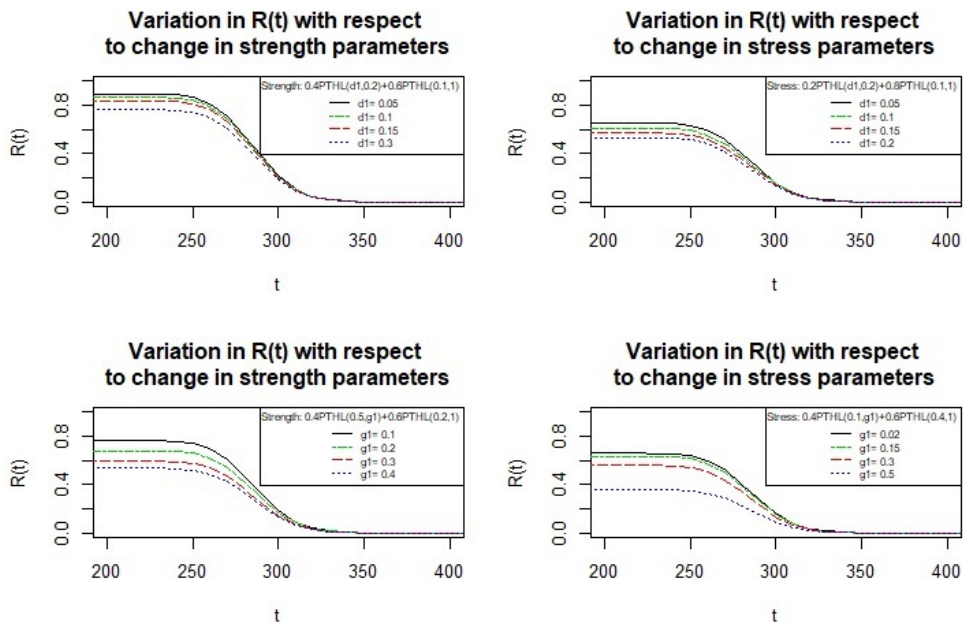


Figure 2: Variation in  $R(t)$  corresponding to change in stress and strength parameters

### 3. ML ESTIMATION OF $R(T)$ USING EM ALGORITHM

In this section, we describe the ML estimation of the reliability function. Assume that the strength and the stress follow a finite mixture distribution with densities  $h(y)$  and  $g(x)$  respectively. where

$$h(y) = \sum_{i=1}^{m_1} \pi_i h_i(y), \quad 0 < \pi_i < 1, \quad \sum_{i=1}^{m_1} \pi_i = 1. \quad (23)$$

and

$$g(x) = \sum_{j=1}^{m_2} p_j g_j(x), \quad 0 < p_j < 1, \quad \sum_{j=1}^{m_2} p_j = 1. \quad (24)$$

The cycle time follows a gamma distribution with p.d.f.

$$f(z) = \frac{a^k z^{k-1} e^{-az}}{(k-1)!}, \quad z \geq 0. \quad (25)$$

Let  $(x_1, x_2, \dots, x_n)$  and  $(y_1, y_2, \dots, y_m)$  and  $(z_1, z_2, \dots, z_r)$  be random samples on stress, strength and cycle time respectively. Then the joint likelihood function is

$$L = \prod_{i=1}^n g(x_i) \prod_{j=1}^m h(y_j) \prod_{t=1}^r f(z_t) \quad (26)$$

and the corresponding log-likelihood function

$$\begin{aligned} l &= \sum_{i=1}^n \log g(x_i) + \sum_{j=1}^m \log h(y_j) + \sum_{t=1}^r \log f(z_t) \\ &= l_1 + l_2 + l_3 \end{aligned} \quad (27)$$

As the log-likelihood function is the sum of log-likelihoods corresponding to the random samples of stress, strength as well as cycle time respectively and since the parameters are independent the stress, strength, and cycle time parameters can be obtained by maximizing corresponding log-likelihood function. The ML estimates of stress and strength parameters can be computed by using Expectation - Maximization algorithm.

### 3.1. ML Estimation of R(t) based on a finite mixture of Weibull distribution

Assuming that the cycle time distribution follows gamma distribution and the initial stress follows a finite mixture of Weibull distribution with parameters  $(\alpha, \beta_i), i = 1, 2, \dots, m_1$  and strength follows finite mixture of Weibull distribution with parameters  $(\alpha, \theta_j), j = 1, 2, \dots, m_2$  expression for the stress strength reliability function is derived in the previous section. We can estimate the stress, strength and cycle time parameters separately. The ML estimates of stress and strength parameters can be computed by using EM algorithm. Here we summarize the EM algorithm for computing the parameters of a finite mixture of Weibull distribution. Consider the strength data consists of n independent and identically distributed observations  $(y_1, y_2, \dots, y_n)$  from a finite mixture of Weibull distribution with p.d.f.

$$h(y, \alpha, \beta) = \sum_{i=1}^{m_1} \pi_i h_i(y, \alpha, \beta_i), \quad \beta = (\beta_i; i = 1, 2, \dots, m_1)$$

Where

$$h_i(y, \alpha, \beta_i) = \frac{\alpha}{\beta_i} y^{\alpha-1} e^{-\frac{y^\alpha}{\beta_i}}; y > 0, \alpha > 0, \beta_i > 0; i = 1, 2, \dots, m_1$$

The associated log-likelihood function is

$$L(y, \alpha, \beta) = \sum_{j=1}^n \log h(y, \alpha, \beta). \tag{28}$$

The MLE of  $\hat{\alpha}, \hat{\beta}$  is determined such that

$$L(y, \hat{\alpha}, \hat{\beta}) = \sup_{\alpha, \beta} L(y, \alpha, \beta). \tag{29}$$

Define a variable  $z_{ij}$  such that  $z_{ij} = 1$  if  $j^{th}$  unit of the sample comes from the  $i^{th}$  component and  $z_{ij} = 0$  otherwise. Since each component comes from exactly one component, we have  $\sum_{i=1}^k z_{ij} = 1, \pi_i = P[z_{ij} = 1]$ .

$$Y_i | z_{ij=1} \sim \text{Weibull}(\alpha, \beta_i), i = 1, 2, \dots, m_1.$$

In missing data setup y can be considered as incomplete data and  $x = (x_1, x_2, \dots, x_n)$  where  $x_j = (y_j, z_j)$  and  $z_j = (z_{ij}, i = 1, 2, \dots, m_1)$  as a complete data set. The density function corresponding to the observations in the complete data set is

$$h_c(x_j, \alpha, \beta) = h_c(y_j, z_j, \alpha, \beta) = \sum_{i=1}^{m_1} \pi_i I_{z_{ij}} h_i(y_j, \alpha, \beta_i). \tag{30}$$

and the likelihood function is

$$L_c(x, \alpha, \beta_i) = \sum_{j=1}^n \log h_c(x_j, \alpha, \beta). \tag{31}$$

The EM algorithm iteratively maximizes  $Q(\alpha, \beta | \alpha, \beta^{(t)}) = E(L_c(x, \alpha, \beta | y, \alpha, \beta^{(t)}))$  instead of maximizing  $L(y, \alpha, \beta)$ , where  $\alpha, \beta^{(t)}$  is the current value at  $t$  and then compute the expectation

$$E_{\alpha, \beta^{(t)}}(L_c(x, \alpha, \beta) | y) = \sum_{j=1}^n \sum_{i=1}^{m_1} E_{\alpha, \beta_i^{(t)}}(z_{ij} | y) (\log \pi_i + \log h_i(y_j, \alpha, \beta_i)) \tag{32}$$

$$\begin{aligned} E_{\alpha, \beta_i^{(t)}}(z_{ij} | y) &= P_{\alpha, \beta_i^{(t)}}(z_{ij} = 1 | y) \\ &= \frac{\pi_i^{(t)} h_i(y_j, \alpha, \beta_i)}{\sum_{i=1}^{m_1} \pi_i^{(t)} h_i(y_j, \alpha, \beta_i)}, j = 1, 2, \dots, n; i = 1, 2, \dots, m_1 \end{aligned} \tag{33}$$

$$= \tau_{ij}(y_j, \alpha, \beta_i) \tag{34}$$

It is the posterior probability that  $j^{th}$  observation belongs to the  $i^{th}$  component in the  $t^{th}$  iteration. Thus we have

$$Q(\alpha, \beta_i | \alpha, \beta^{(t)}) = \sum_{j=1}^n \sum_{i=1}^{m_1} \tau_{ij}(\alpha, \beta_i^{(t)}) (\log \pi_i + \log h_i(y_j, \alpha, \beta_i)). \tag{35}$$

Hence the EM algorithm consists of the following two steps.

**Step1.E-step:** Compute  $Q(\alpha, \beta | \alpha, \beta^{(t)})$

**Step2.M-step:** Compute the value of  $\alpha, \beta^{(t+1)}$  that maximizes  $Q(\alpha, \beta | \alpha, \beta^{(t)})$ .

If  $\tau_{ij}$  where observable posterior probabilities, then MLE of  $\pi$  is simply given by

$$\hat{\pi}_i = \sum_{j=1}^n \frac{\tau_{ij}}{n}, \quad i = 1, 2, \dots, m_1,$$

which is the proportion of the sample having arisen from the  $i^{th}$  component of the mixture.

For the  $(t + 1)^{th}$  update other parameters  $\alpha$  and  $(\beta_1, \beta_2, \dots, \beta_{m_1})$ , we have to obtain the solution of

$$\sum_{j=1}^n \sum_{i=1}^{m_1} \tau_{ij}(\pi^{(t)}) \frac{\partial}{\partial \alpha, \beta_i} \log h_i(y_j, \alpha, \beta_i) = 0 \tag{36}$$

We repeat the procedure until the desired accuracy is obtained. Hence we get the estimates of the strength parameters as:

$$\hat{\beta}_i(t + 1) = \left[ \frac{\sum_{j=1}^n \tau_{ij} y_j^{\alpha(t+1)}}{\sum_{j=1}^n \tau_{ij}} \right] \tag{37}$$

$$\hat{\alpha}(t + 1) = n \left[ \sum_{i=1}^{m_1} \frac{1}{\hat{\beta}_i(t)} \sum_{j=1}^n \tau_{ij} y_j^{\hat{\alpha}(t)} \log(y_j) - \sum_{i=1}^{m_1} \sum_{j=1}^n \tau_{ij} \log(y_j) \right]^{-1} \tag{38}$$

Similarly, we can estimate the stress parameters. The ML estimates of gamma cycle time parameters can be obtained by standard procedures. Using the ML estimates of the stress, strength, and cycle time parameters and applying the invariance property of the ML estimators we can find the value of  $R(t)$ .

We use the Monte Carlo simulation technique to estimate  $R(t)$  for systems with initial strength and initial stress following Weibull mixture and cycle times following gamma distribution. We have done the entire numerical analysis using R. The numerical illustration of ML of  $R(t)$  with gamma cycle time with Weibull mixture initial stress and strength for different time values is given in Table 1. In which  $y_0$  represent initial strength and  $x_0$  represent initial stress of the system. For a fixed time interval, we draw samples for cycle time and the number of cycles based on the distributional assumption of cycle times. The maximum number of cycles up to which the total cycle time does not exceed the length of the time interval under consideration is taken as the number of runs during the time interval. The cycle time observed during each run constitutes the simulated sample of cycle times. The command `rweibull` helps in simulating samples from the Weibull distribution. Samples to represent initial stress and initial strength distributions, when both are mixtures of Weibull distributions are generated using this command. We repeat the entire simulation experiment 1,000 times.

From the table, it is clear that  $R(t)$  decreases as the time increases, when the initial stress and strength of the system is distributed as a mixture of Weibull distribution with gamma cycle time.

**Table 1:** ML Estimation of  $R(t)$  with Weibull mixture initial stress and strength

Cycle time		Stress and Strength		$a_0$	$x_0$	$t$	$R(t)$
Parameters	Estimated	Parameters	Estimated				
G(0.5,2)	$a = 0.4881$ $k = 2.0840$	$Y_0 : 0.8W(0.3,0.6)+$	$\alpha = 0.2890$	1	0.02	10	0.2230
		$0.2W(0.3,2)$	$\theta = (0.5892, 1.9454)$			25	0.1539
		$X_0 : 0.3W(0.3,1)+$	$\alpha = 0.2954$			50	0.1175
		$0.7W(0.3,0.3)$	$\beta = (1.1603, 0.2957)$			75	0.0989
						100	0.0869
G(0.5,4)	$a = 0.5231$ $k = 3.1780$	$Y_0 : 0.6W(5,0.3)+$	$\alpha = 5.0982$	0.001	0.08	10	0.9155
		$0.4W(5,2)$	$\theta = (0.2992, 2.1114)$			25	0.9154
		$X_0 : 0.3W(5,0.1)+$	$\alpha = 5.1089$			50	0.6168
		$0.7W(5,0.2)$	$\beta = (0.1008, 0.2292)$			75	0.2228
						100	0.0498
G(1,2)	$a = 0.5231$ $k = 3.1780$	$Y_0 : 0.4W(2,1.2)+$	$\alpha = 2.0008$	0.02	0.05	10	0.8699
		$0.6W(2,4)$	$\theta = (1.3100, 4.1216)$			25	0.8938
		$X_0 : 0.7W(2,4)+$	$\alpha = 1.9279$			50	0.7896
		$0.3W(2,2.5)$	$\beta = (4.3704, 2.6518)$			75	0.6559
						100	0.5188
G(1,4)	$a = 0.9910$ $k = 3.9622$	$Y_0 : 0.2W(1.2,0.8)+$	$\alpha = 1.2056$	0.1	0.05	10	0.5913
		$0.8W(1.2,2.4)$	$\theta = (0.7759, 2.3253)$			25	0.4231
		$X_0 : 0.3W(1.2,4)+$	$\alpha = 1.1128$			50	0.2357
		$0.7W(1.2,3.2)$	$\beta = (3.9000, 3.9901)$			75	0.1292
						100	0.0693

### 3.2. ML Estimation of $R(t)$ based on a finite mixture of power-transformed half-logistic distribution

By assuming that the cycle time follows gamma distribution and the initial stress and strength follow the mixture of power transformed half logistic distribution with parameters  $(\delta_{1i}, \gamma_{1i}), i = 1, 2, \dots, m_1$  and  $(\delta_{2j}, \gamma_{2j}), j = 1, 2, \dots, m_2$  respectively, the corresponding stress-strength reliability is given in the previous section. Now, consider independent and identically distributed strength observations  $y = (y_1, y_2, \dots, y_n)$  from a finite a mixture of power transformed half logistic mixture with p.d.f.

$$h(y, \delta_1, \gamma_1) = \sum_{i=1}^{m_1} \pi_i h_i(y, \delta_{1i}, \gamma_{1i}).$$

Where

$$h_i(y) = \begin{cases} 2\delta_{1i}\gamma_{1i}y^{\gamma_{1i}-1}e^{-\delta_{1i}y^{\gamma_{1i}}} \left(1 + e^{-\delta_{1i}y^{\gamma_{1i}}}\right)^{-2}, & 0 \leq y < \infty; \delta_{1i} > 0; \gamma_{1i} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (39)$$

$i = 1, 2, \dots, m_1$ . Using the EM algorithm explained earlier, we get the ML estimates of the strength parameters as

$$\pi_i = \frac{\sum_{j=1}^n \tau_{ij}}{n}, i = 1, 2, \dots, m_1. \quad (40)$$

$$\hat{\delta}_{1i} = \frac{\sum_{j=1}^n \tau_i(y_j; \delta_{1i}, \gamma_{1i})}{\sum_{j=1}^n (\tau_i(y_j; \delta_{1i}, \gamma_{1i}) y_j^{\gamma_{1i}} [1 - \frac{2e^{-\delta_{1i}y_j^{\gamma_{1i}}}}{1 + e^{-\delta_{1i}y_j^{\gamma_{1i}}}}])}; i = 1, 2, \dots, m_1. \quad (41)$$

$$\hat{\gamma}_{1i} = \frac{\sum_{j=1}^n \tau_i(y_j; \delta_{1i}, \gamma_{1i})}{\sum_{j=1}^n (\tau_i(y_j; \delta_{1i}, \gamma_{1i}) \log(y_j) [\delta_{1i} y_j^{\gamma_{1i}} (1 - \frac{2e^{-\delta_{1i}y_j^{\gamma_{1i}}}}{1 + e^{-\delta_{1i}y_j^{\gamma_{1i}}}}) - 1])}; i = 1, 2, \dots, m_1. \quad (42)$$

Similarly, we can find the stress estimates and hence we can find  $R(t)$  by the estimated parameters.

We use the Monte Carlo simulation technique to estimate  $R(t)$  for systems with initial strength and initial stress following a finite mixture of power-transformed half logistic distribution and cycle times following Gamma distribution. Table 2 gives the estimated value of  $R(t)$  with gamma cycle time with power transformed half logistic mixture initial stress and strength for different time values. The package bayesmeta available in R software allows sampling from half-logistic distribution. Then sample from power transformed half logistic distribution is simulated using simple conversion techniques.

**Table 2:** ML Estimation of  $R(t)$  with PTHL mixture initial stress and strength

Cycle time		Stress and Strength		$a_0$	$x_0$	$t$	$R(t)$
Parameters	Estimated	Parameters	Estimated				
G(0.5,1)	$a = 0.4885$ $k = 0.9687$	$Y_0 : 0.6\text{PTHL}(5,0.3)$	$\delta_2 = (5.0927, 4.9764)$	0.001	0.008	10	0.00756
		$+0.4\text{PTHL}(0.4,2)$	$\gamma_2 = (0.1067, 0.2025)$			50	0.00755
		$X_0 : 0.3\text{PTHL}(5,0.1)$	$\delta_1 = (9.2026, 2.0041)$			100	0.00755
		$+0.7\text{PTHL}(5,0.2)$	$\gamma_1 = (1.1615, 6.0581)$			150	0.00753
						200	0.00367
G(0.5,2)	$a = 0.9522$ $k = 1.9138$	$Y_0 : 0.7\text{PTHL}(8,0.5)$	$\delta_2 = (7.3349, 1.1391)$	0.001	0.08	10	0.0629
		$+0.3\text{PTHL}(0.5,2.5)$	$\gamma_2 = (0.5245, 1.8925)$			50	0.0629
		$X_0 : 0.2\text{PTHL}(4,0.5)$	$\delta_1 = (3.8315, 5.0771)$			125	0.0626
		$+0.8\text{PTHL}(5,0.2)$	$\gamma_1 = (0.1008, 0.2292)$			140	0.0525
						150	0.0313
G(1,1)	$a = 1.0097$ $k = 1.0029$	$Y_0 : 0.2\text{PTHL}(2,0.2)$	$\delta_2 = (1.9903, 3.9043)$	0.002	0.005	10	0.1067
		$+0.8\text{PTHL}(4,2.4)$	$\gamma_2 = (0.1961, 2.4533)$			50	0.1067
		$X_0 : 0.6\text{PTHL}(4,2)$	$\delta_1 = (3.9133, 5.1319)$			75	0.1063
		$+0.4\text{PTHL}(4,6,2)$	$\gamma_1 = (2.1198, 1.9368)$			100	0.0490
						125	0.0008
G(1,2)	$a = 1.0099$ $k = 1.9968$	$Y_0 : 0.1\text{PTHL}(3,2.4)$	$\delta_2 = (2.6875, 3.0715)$	0.002	0.005	10	0.1038
		$+0.9\text{PTHL}(3,1.2)$	$\gamma_2 = (2.2066, 1.1926)$			50	0.1038
		$X_0 : 0.8\text{PTHL}(2.5,1.1)$	$\delta_1 = (2.4427, 5.3425)$			75	0.1034
		$+0.2\text{PTHL}(5,2)$	$\gamma_1 = (1.1208, 2.0481)$			100	0.0477
						125	0.0008

From this table, we can see that,  $R(t)$  decreases as time increases, when the initial stress and strength of the system is distributed as a mixture of power transformed half logistic distribution with gamma cycle time.

#### 4. BAYESIAN ESTIMATION OF $R(T)$ USING MCMC METHOD

In this section, we describe the Bayesian estimation of the reliability function. The stress and strength follow a finite mixture distribution with densities  $g(x)$  and  $h(y)$  respectively and the cycle time follows a gamma distribution. Let  $(x_1, x_2, \dots, x_n)$ ,  $(y_1, y_2, \dots, y_m)$  and  $(z_1, z_2, \dots, z_r)$  be random samples on stress, strength, and cycle time respectively. Then, the joint likelihood function is

$$L = \prod_{i=1}^n g(x_i) \prod_{j=1}^m h(y_j) \prod_{t=1}^r f(z_t) \tag{43}$$

where

$$g(x) = \sum_{j=1}^{m_2} p_j g_j(x), \quad 0 < p_j < 1, \quad \sum_{j=1}^{m_2} p_j = 1. \tag{44}$$

and

$$h(y) = \sum_{i=1}^{m_1} \pi_i h_i(y), \quad 0 < \pi_i < 1, \quad \sum_{i=1}^{m_1} \pi_i = 1. \tag{45}$$

The cycle time follows a gamma distribution with p.d.f.

$$f(z) = \frac{a^k z^{k-1} e^{-az}}{(k-1)!}, \quad z \geq 0. \tag{46}$$

We assume prior probabilities corresponding to each parameter to get a Bayesian estimate of the reliability function.

#### 4.1. Bayesian Estimation of R(t) based on a finite mixture of Weibull distribution

Let the cycle time follow a gamma distribution with parameters  $(a, k)$  and stress and strength of the system follow a mixture of Weibull distribution with parameters  $(\alpha, \beta_j), j = 1, 2, \dots, m_2$  and  $(\delta, \gamma_i), i = 1, 2, \dots, m_1$  respectively. The expression for stress-strength reliability is given in section 2. Here we discuss the estimation of the parameters by the Bayesian estimation method. Treating  $Z_i$  as the auxiliary variable, such that

$$X_j | Z_j = i \sim g_i(x, \alpha, \beta_i) \text{ and } p(Z_j = i) = p_i, j = 1, 2, \dots, n, i = 1, 2, \dots, m_2.$$

$$Y_j | Z_j = i \sim h_i(y, \delta, \gamma_i) \text{ and } p(Z_j = i) = \pi_i, j = 1, 2, \dots, m, i = 1, 2, \dots, m_1.$$

Where

$$g_i(x) = \frac{\alpha}{\beta_i} x^{\alpha-1} e^{-x^\alpha / \beta_i}, x \geq 0, \alpha > 0, \beta_i > 0; i = 1, 2, \dots, m_1. \tag{47}$$

and

$$h_i(y) = \frac{\alpha}{\theta_j} y^{\alpha-1} e^{y^\alpha / \theta_j}, y \geq 0, \alpha > 0, \theta_j > 0; j = 1, 2, \dots, m_2. \tag{48}$$

We can simplify the likelihood function into the form,

$$L = \prod_{t=1}^r \frac{a^k z_t^{k-1} e^{-az_t}}{(k-1)!} \prod_{i=1}^{m_2} \left\{ \pi_i^{n_{1i}} \left( \frac{\alpha}{\beta_i} \right)^{n_{1i}} \left( \prod_{j=1}^n x_j^{z_{ij}} \right)^{\alpha-1} e^{\frac{1}{\beta_i} \sum_{j=1}^n (z_{ij} x_j^\alpha)} \right\} \prod_{k=1}^{m_1} \left\{ p_k^{n_{2k}} \left( \frac{\delta}{\gamma_k} \right)^{n_{2k}} \left( \prod_{l=1}^n y_l^{z_{kl}} \right)^{\delta-1} e^{\frac{1}{\gamma_k} \sum_{l=1}^n (z_{kl} y_l^\delta)} \right\} \tag{49}$$

We fix the Dirichlet prior distribution for  $\pi = (\pi_1, \pi_2, \dots, \pi_{m_1})$  and  $p = (p_1, p_2, \dots, p_{m_2})$ , gamma prior for  $\beta_i, \gamma_k; i = 1, 2, \dots, m_2, k = 1, 2, \dots, m_1$  non-informative prior for  $\alpha, \delta, a$  and  $k$ . The variable  $z_{ij}$  is such that  $z_{ij} = 1$  if  $j^{th}$  unit of the sample comes from the  $i^{th}$  component and  $z_{ij} = 0$  otherwise. Also  $n_{1i} = \sum_{j=1}^{m_2} z_{ij}$  and  $n_{2k} = \sum_{j=1}^{m_1} z_{kj}$ . Hence

$$\begin{aligned} \pi &\sim \text{Dirichlet}(\mu_{11}, \mu_{12}, \dots, \mu_{1m_2}) \\ p &\sim \text{Dirichlet}(\mu_{21}, \mu_{22}, \dots, \mu_{2m_1}) \\ \pi_{3i}(\beta_i) &\propto \beta_i^{a_{1i}-1} e^{-b_{1i}\beta_i}; i = 1, 2, \dots, m_2 \\ \pi_{4k}(\gamma_k) &\propto \gamma_k^{a_{2k}-1} e^{-b_{2k}\gamma_k}; k = 1, 2, \dots, m_1 \\ \pi_5(\alpha), \pi_6(\delta), \pi_7(a), \pi_8(k) &\propto 1 \end{aligned}$$

where  $\mu_1 = (\mu_{11}, \mu_{12}, \dots, \mu_{1m_2}), \mu_2 = (\mu_{21}, \mu_{22}, \dots, \mu_{2m_2}), (a_{1i}, a_{2i}); i = 1, 2, \dots, m_2$  and  $(b_{1j}, b_{2j}); j = 1, 2, \dots, m_1$  are the hyper-parameters. Since the cycle time parameters have a non-informative prior, their estimates coincide with the ML estimates. The joint prior distribution of  $\pi, p, \beta, \gamma, \alpha$ , and  $\delta$  can be written as,

$$g(\pi, p, \beta, \gamma, \alpha, \delta) \propto \prod_{i=1}^{m_2} p_i^{\mu_{1i}-1} \beta_i^{(a_{1i}-1)} e^{b_{1i}\beta_i} \prod_{j=1}^{m_1} \pi_j^{\mu_{2j}-1} \gamma_j^{(a_{2j}-1)} e^{b_{2j}\gamma_j} \tag{50}$$

Where  $\beta = (\beta_1, \beta_2, \dots, \beta_{m_2})$  and  $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_{m_1})$ . The posterior probability is given by,

$$\begin{aligned}
 h(\pi, p, \beta, \gamma, \alpha, \delta | x, y, z_{ij}) &\propto \prod_{i=1}^{m_2} p_i^{\mu_i-1} \beta_i^{(a_i-1)} e^{b_1 \beta_i} \prod_{j=1}^{m_1} \pi_j^{\mu_j-1} \gamma_j^{(a_j-1)} e^{b_2 \gamma_j} \\
 &\prod_{t=1}^r \frac{a^k z_t^{k-1} e^{-a z_t}}{(k-1)!} \prod_{i=1}^{m_2} \left\{ \pi_i^{n_{1i}} \left( \frac{\alpha}{\beta_i} \right)^{n_{1i}} \left( \prod_{j=1}^n x_j^{z_{ij}} \right)^{\alpha-1} e^{\frac{1}{\beta_i} \sum_{j=1}^n (z_{ij} x_j^\alpha)} \right\} \\
 &\prod_{k=1}^{m_1} \left\{ p_k^{n_{2k}} \left( \frac{\delta}{\gamma_k} \right)^{n_{2k}} \left( \prod_{l=1}^n y_l^{z_{kl}} \right)^{\delta-1} e^{\frac{1}{\gamma_k} \sum_{l=1}^n (z_{kl} y_l^\delta)} \right\} \quad (51)
 \end{aligned}$$

Then, the conditional posterior distributions of  $\pi, p, \beta, \gamma, \alpha$ , and  $\delta$  are:

$$\pi \sim \text{Dirichlet}(\mu_{11} + n_{11}, \mu_{12} + n_{12}, \dots, \mu_{1m_2} + n_{1m_2}) \quad (52)$$

$$p \sim \text{Dirichlet}(\mu_{21} + n_{21}, \mu_{22} + n_{22}, \dots, \mu_{2m_1} + n_{2m_1}) \quad (53)$$

$$\pi_1(\alpha | \beta, x, z) \propto \prod_{i=1}^{m_2} \left\{ \alpha^{n_{1i}} \left( \prod_{j=1}^n x_j^{z_{ij}} \right)^{\alpha-1} e^{-\frac{1}{\beta_i} \sum_{j=1}^n x_j^\alpha} \right\} \quad (54)$$

$$\pi_2(\delta | \gamma, y, z) \propto \prod_{i=1}^{m_1} \left\{ \delta^{n_{2i}} \left( \prod_{j=1}^n y_j^{z_{ij}} \right)^{\delta-1} e^{-\frac{1}{\gamma_i} \sum_{j=1}^n y_j^\delta} \right\} \quad (55)$$

$$\pi_{3i}(\beta_i | \alpha, \beta_i^*, x, z) \propto \beta_i^{-n_{1i} + a_i - 1} e^{-\frac{1}{\beta_i} \sum_{j=1}^n x_j^\alpha} e^{-b_1 \beta_i}; i = 1, 2, \dots, m_2 \quad (56)$$

$$\pi_{4i}(\gamma_i | \delta, \gamma_i^*, y, z) \propto \gamma_i^{-n_{2i} + a_i - 1} e^{-\frac{1}{\gamma_i} \sum_{j=1}^n y_j^\delta} e^{-b_2 \gamma_i}; i = 1, 2, \dots, m_1. \quad (57)$$

Where  $\beta_i^* = \{\beta_i, i = 1, 2, i - 1, i + 1, \dots, m_2\}$  and  $\gamma_i^* = \{\gamma_i, i = 1, 2, i - 1, i + 1, \dots, m_1\}$ .

The posterior distributions of  $\alpha, \beta_i, \delta$ , and  $\gamma_i$  cannot be reduced analytically to a well-known distribution. So we use the Markov chain Monte Carlo method with Gibbs sampling under Metropolis-Hastings algorithm for computing Bayes estimate using the statistical software, R. The Metropolis-Hastings algorithm with chi-square proposal density is used for generating samples from  $(\pi, p, \alpha, \beta, \delta, \gamma)$ , where  $\pi = (\pi_1, \pi_2, \dots, \pi_{m_1})$ ,  $p = (p_1, p_2, \dots, p_{m_2})$   $\beta = \{\beta_i, i = 1, 2, \dots, m_2\}$ , and  $\gamma = \{\gamma_i, i = 1, 2, \dots, m_1\}$  is given as follows.

**ALGORITHM – 1 :**

**Step1.** Set the initial values  $(\pi^0, p^0, \alpha^0, \beta^0, \delta^0, \gamma^0)$

**Step2.** Generate  $z_{ij}$  values using sample  $x$

**Step3.** Generate  $\pi^t$

**Step4.** Using the proposal density  $g(\alpha) \sim \chi_{(x)}^2$  where  $x$  is the d.f and choose  $x = \alpha^{t-1}$  Generate another random variable  $y$  from the chi-square density  $g$ . Generate  $u$  from Uniform(0,1). If  $u < \frac{\pi_1(y)g(x)}{\pi_1(x)g(y)}$  accept  $y$  and set  $\alpha^t = y$ ; otherwise set  $\alpha^t = x$

**Step5.** Using the proposal density  $g(\beta_i) \sim \chi_{(x)}^2$  where  $x$  is the d.f and choose  $x = \beta_i^{t-1}$  Generate another random variable  $y$  from the chi-square density  $g$ . Generate  $u$  from Uniform(0,1). If  $u < \frac{\pi_{3i}(y)g(x)}{\pi_{3i}(x)g(y)}$  accept  $y$  and set  $\beta_i^t = y$ ; otherwise set  $\beta_i^t = x$ . Repeat the procedure and generate  $\beta_i^t, i = 1, 2, \dots, m_2$

**Step6.** Generate  $z_{ij}$  values using sample  $y$

**Step7.** Generate  $p^t$

**Step8.** Using the proposal density  $g(\delta) \sim \chi_{(x)}^2$  where  $x$  is the d.f and choose  $x = \delta^{t-1}$  Generate random variable  $y$  from the chi-square density  $g$ . Generate  $u$  from Uniform(0,1). If  $u < \frac{\pi_2(y)g(x)}{\pi_2(x)g(y)}$  accept  $y$  and set  $\delta^t = y$ ; otherwise set  $\delta^t = x$

**Step9.** Using the proposal density  $g(\gamma_i) \sim \chi_{(x)}^2$  where  $x$  is the d.f and choose  $x = \gamma_i^{t-1}$  Generate random variable  $y$  from the chi-square density  $g$ . Generate  $u$  from Uniform(0,1). If  $u < \frac{\pi_{4i}(y)g(x)}{\pi_{4i}(x)g(y)}$  accept  $y$  and set  $\beta_i^t = y$ ; otherwise set  $\beta_i^t = x$ . Repeat the procedure and generate  $\gamma_i^t, i = 1, 2, \dots, m_1$   
**Step10** Compute  $R(t)$ .  
**Step11** Increment  $t$ .

Table 3 provides the estimated values of  $R(t)$  by the Bayesian estimation method when the stress and strength of the system follow a mixture of two Weibull distributions with gamma cycle time. We assume that the mixture proportions  $\pi$  and  $p$  are known and the component parameters  $(\alpha, \beta, \delta, \gamma)$  are unknown and are following gamma prior distributions. Also, we assume that cycle time parameters  $a$  and  $k$  follow non-informative prior. Since the cycle time parameters have a non-informative prior, their estimates coincide with the ML estimates. The table shows Bayes estimates of the parameters  $(\alpha, \beta, \delta, \gamma)$  and Bayes estimate of the reliability function  $R(t)$  for different time values corresponding to various sets of hyperparameter values. The table shows that  $R(t)$  decreases as time increases.

**Table 3:** Bayesian Estimation of  $R(t)$  with Weibull mixture initial stress strength

Cycle time		Stress and Strength		$a_0$	$x_0$	$t$	$R(t)$
Parameters	Estimated	Parameters	Estimated				
G(0.5,2)	$a = 0.4881$	$Y_0:0.8W(0.3,0.6)+$	$\delta = 0.2890$	1	0.02	10	0.2230
	$k = 2.0840$	$0.8W(0.3,2)$	$\gamma = (0.5892, 1.9454)$			25	0.1539
		$X_0:0.3W(0.3,1)+$	$\alpha = 0.2954$			50	0.1175
		$0.7W(0.3,0.3)$	$\beta = (1.1603, 0.2957)$			75	0.0989
						100	0.0869
G(0.5,4)	$a = 0.5231$	$Y_0:0.6W(5,0.3)+$	$\delta = 5.0982$	0.001	0.08	10	0.9155
	$k = 3.1780$	$0.4W(5,2)$	$\gamma = (0.2992, 2.1114)$			25	0.9154
		$X_0:0.3W(5,0.1)+$	$\alpha = 5.1089$			50	0.6168
		$0.7W(5,0.2)$	$\beta = (0.1008, 0.2292)$			75	0.2228
						100	0.0498
G(1,2)	$a = 0.9975$	$Y_0:0.3W(0.2,2)+$	$\delta = 0.2005$	0.01	0.02	10	0.7598
	$k = 1.9738$	$0.7W(0.2,5)$	$\gamma = (0.8536, 8.6029)$			25	0.7208
		$X_0:0.4W(0.2,0.9)+$	$\alpha = 5.1089$			50	0.6868
		$0.6W(0.2,8)$	$\beta = (0.1008, 0.2292)$			75	0.6652
						100	0.6492
G(1,4)	$a = 0.9652$	$Y_0:0.5W(2,0.2)+$	$\delta = 2.0344$	0.001	0.05	10	0.7455
	$k = 3.7469$	$0.5W(2,6)$	$\gamma = (0.2093, 5.9099)$			25	0.6070
		$X_0:0.5W(2,1)+$	$\alpha = 1.9980$			50	0.4140
		$0.5W(2,10)$	$\beta = (1.0200, 9.8969)$			75	0.3370
						100	0.2908

#### 4.2. Bayesian Estimation of $R(t)$ based on a finite mixture of power-transformed half-logistic distribution

Let the cycle time follows a gamma distribution with parameters  $(a, k)$  and stress and strength of the system follow a mixture of power transformed half logistic distribution with parameters  $(\delta_j, \gamma_j), j = 1, 2, \dots, m_2$  and  $(\alpha_i, \theta_i), i = 1, 2, \dots, m_1$  respectively. The expression for stress-strength reliability is given in section 2. Here we discuss the estimation of the parameters by the Bayesian estimation method. Consider the auxiliary variable  $Z_j$ , such that

$$X_j|Z_j = i \sim g_i(x, \delta_i, \gamma_i) \text{ and } p(Z_j = i) = p_i, j = 1, 2, \dots, n, i = 1, 2, \dots, m_2$$

$$Y_j|Z_j = i \sim h_i(y, \alpha_i, \theta_i) \text{ and } p(Z_j = i) = \pi_i, j = 1, 2, \dots, m, i = 1, 2, \dots, m_1$$



Where

$$g_i(x) = \begin{cases} 2\delta_i\gamma_i x^{\gamma_i-1} e^{-\delta_i x^{\gamma_i}} (1 + e^{-\delta_i x^{\gamma_i}})^{-2}, & 0 \leq x < \infty; \delta_i > 0; \gamma_i > 0. \\ 0 & \text{otherwise.} \end{cases} \quad (58)$$

$$f(y) = \begin{cases} 2\alpha_i\theta_i y^{\theta_i-1} e^{-\alpha_i y^{\theta_i}} (1 + e^{-\alpha_i y^{\theta_i}})^{-2}, & 0 \leq y < \infty; \alpha_i > 0; \theta_i > 0. \\ 0 & \text{otherwise.} \end{cases} \quad (59)$$

Then likelihood function  $L$  is

$$L = \prod_{i=1}^n \frac{a^k y_i^{k-1} e^{-a y_i}}{(k-1)!} \prod_{i=1}^{m_2} \left\{ \pi_i^{n_{1i}} 2^{n_{1i}} \gamma_i^{n_{1i}} \delta_i^{n_{1i}} \left( \prod_{j=1}^n x_j^{z_{ij}} \right)^{\gamma_i-1} e^{-\delta_i \sum_{j=1}^n (z_{ij} x_j^{\gamma_i})} \prod_{j=1}^n \left[ (1 + e^{-\delta_i x_j^{\gamma_i}})^{-2z_{ij}} \right] \right\} \prod_{k=1}^{m_1} \left\{ p_k^{n_{2k}} 2^{n_{2k}} \theta_k^{n_{2k}} \alpha_k^{n_{2k}} \left( \prod_{j=1}^m y_j^{z_{kj}} \right)^{\theta_k-1} e^{-\alpha_k \sum_{j=1}^m (z_{kj} y_j^{\theta_k})} \prod_{j=1}^m \left( 1 + e^{-\alpha_k y_j^{\theta_k}} \right)^{-2z_{kj}} \right\} \quad (60)$$

We fix the Dirichlet prior distribution for  $\pi = (\pi_1, \pi_2, \dots, \pi_{m_1})$  and  $p = (p_1, p_2, \dots, p_{m_2})$ , gamma prior for  $\delta_i, \alpha_k; i = 1, 2, \dots, m_2, k = 1, 2, \dots, m_1$  non-informative prior for  $\gamma, \theta, a$  and  $k$ . Since the cycle time parameters have a non-informative prior, their estimates coincide with the ML estimates. The variable  $z_{ij}$  is such that  $z_{ij} = 1$  if  $j^{th}$  unit of the sample comes from the  $i^{th}$  component and  $z_{ij} = 0$  otherwise. Also  $n_{1i} = \sum_{j=1}^{m_2} z_{ij}$  and  $n_{2k} = \sum_{j=1}^{m_1} z_{kj}$ . Hence

$$\begin{aligned} \pi &\sim \text{Dirichlet}(\mu_{11}, \mu_{12}, \dots, \mu_{1m_2}) \\ p &\sim \text{Dirichlet}(\mu_{21}, \mu_{22}, \dots, \mu_{2m_1}) \\ \pi_{3i}(\delta_i) &\propto \delta_i^{a_{1i}-1} e^{-b_{1i}\delta_i}; \\ \pi_{4k}(\alpha_k) &\propto \alpha_k^{a_{2k}-1} e^{-b_{2k}\alpha_k}; k = 1, 2, \dots, m_1 \\ \pi_5(\gamma_i), \pi_6(\theta_k), \pi_7(a), \pi_8(k) &\propto 1; i = 1, 2, \dots, m_2; k = 1, 2, \dots, m_1 \end{aligned}$$

where  $\mu_1 = (\mu_{11}, \mu_{12}, \dots, \mu_{1m_2}), \mu_2 = (\mu_{21}, \mu_{22}, \dots, \mu_{2m_1}), (a_{1i}, a_{2i}); i = 1, 2, \dots, m_2$  and  $(b_{1j}, b_{2j}); j = 1, 2, \dots, m_1$  are the hyper-parameters. Since the cycle time parameters have a non-informative prior, their estimates coincide with the ML estimates.

Now proceeding as in the case of Bayesian estimation of  $R(t)$  based on the finite mixture of Weibull distribution discussed in the previous section we can easily obtain the conditional marginal distributions  $\pi, p, \delta, \alpha, \gamma$ , and  $\theta$ . The conditional posterior distributions of  $\pi, p, \delta, \alpha, \gamma$ , and  $\theta$  are:

$$\pi \sim \text{Dirichlet}(\mu_{11} + n_{11}, \mu_{12} + n_{12}, \dots, \mu_{1m_2} + n_{1m_2}) \quad (61)$$

$$p \sim \text{Dirichlet}(\mu_{21} + n_{21}, \mu_{22} + n_{22}, \dots, \mu_{2m_1} + n_{2m_1}) \quad (62)$$

$$\pi_{3i}(\delta_i | \gamma, \delta_i^*, x, z_{ij}) \propto \delta_i^{a_{1i} + n_{1i} - 1} e^{-(\delta_i \sum_{j=1}^n (z_{ij} x_j^{\gamma_i}) + b_{1i}\delta_i)} \prod_{j=1}^n \left[ 1 + e^{-\delta_i x_j^{\gamma_i}} \right]^{-2z_{ij}} \quad (63)$$

$i = 1, 2, \dots, m_2.$

$$\pi_{4k}(\alpha_k | \theta, \alpha_k^*, y, z_{kj}) \propto \alpha_k^{a_{2k} + n_{2k} - 1} e^{-(\alpha_k \sum_{j=1}^m (z_{kj} y_j^{\theta_k}) + b_{2k}\alpha_k)} \prod_{j=1}^m \left[ 1 + e^{-\alpha_k y_j^{\theta_k}} \right]^{-2z_{kj}} \quad (64)$$

$k = 1, 2, \dots, m_1.$

$$\pi_{5i}(\gamma_i|\delta, \gamma_i^*, x, z_{ij}) \propto \gamma_i^{n1_i} \left( \prod_{j=1}^n x_j^{z_{ij}} \right)^{\gamma_i-1} e^{-\left(\delta_i \sum_{j=1}^n (z_{ij} x_j^{\gamma_i})\right)} \prod_{j=1}^n \left[ 1 + e^{-\delta_i x_j^{\gamma_i}} \right]^{-2z_{ij}} : \\ i = 1, 2, \dots, m_2. \tag{65}$$

$$\pi_{6k}(\theta_k|\alpha, \theta_k^*, y, z_{kj}) \propto \theta_k^{n2_k} \left( \prod_{j=1}^m y_k^{z_{kj}} \right)^{\theta_k-1} e^{-\left(\alpha_k \sum_{j=1}^m (z_{kj} y_k^{\theta_k})\right)} \prod_{j=1}^m \left[ 1 + e^{-\alpha_k y_k^{\theta_k}} \right]^{-2z_{kj}} : \\ k = 1, 2, \dots, m_1. \tag{66}$$

Where  $\delta_i^* = \{\delta_j, j = 1, 2, \dots, i - 1, i + 1, \dots, m_2\}$ ,  $\gamma_i^* = \{\gamma_j, j = 1, 2, \dots, i - 1, i + 1, \dots, m_2\}$ ,  $\alpha_k^* = \{\alpha_i, i = 1, 2, \dots, k - 1, k + 1, \dots, m_1\}$  and  $\theta_k^* = \{\theta_i, i = 1, 2, \dots, k - 1, k + 1, \dots, m_1\}$ .

Since the posterior distributions of  $\alpha_k, \theta_k, \delta_i$ , and  $\gamma_i$  cannot be reduced analytically to a well-known distribution, as done in the previous section, we use the Markov chain Monte Carlo method with Gibbs sampling under Metropolis-Hastings algorithm for computing Bayes estimates. We fix the proposal density as the chi-square distribution. The Metropolis-Hastings algorithm with chi-square proposal density is used for generating samples from  $(\pi, p, \alpha, \theta, \delta, \gamma)$ , where  $\pi = (\pi_i, i = 1, 2, \dots, m_1)$ ,  $p = (p_k, k = 1, 2, \dots, m_1)$   $\alpha = \{\alpha_k, k = 1, 2, \dots, m_1\}$ ,  $\theta = \{\theta_k, k = 1, 2, \dots, m_1\}$ ,  $\delta = \{\delta_i, i = 1, 2, \dots, m_2\}$  and  $\gamma = \{\gamma_i, i = 1, 2, \dots, m_1\}$  is given as follows.

**ALGORITHM – 2 :**

**Step1.** Set the initial values  $(\pi^0, p^0, \alpha^0, \theta^0, \delta^0, \gamma^0)$ .

**Step2.** Generate  $z_{ij}$  values using sample  $x$

**Step3.** Generate  $\pi^t$ .

**Step4.** Using the proposal density  $g(\delta_i) \sim \chi_{(x)}^2$ , where  $x$  is the d.f and choose  $x = \delta_i^{t-1}$ . Generate another random variable  $y$  from the chi-square density  $g$ . Generate  $u$  from Uniform(0,1). If  $u < \frac{\pi_{5i}(y)g(x)}{\pi_{5i}(x)g(y)}$  accept  $y$  and set  $\delta_i^t = y$ ; otherwise set  $\delta_i^t = x$ . Repeat the procedure and generate  $\delta_i^t, i = 1, 2, \dots, m_2$ .

**Step5.** Using the proposal density  $g(\gamma_i) \sim \chi_{(x)}^2$ , where  $x$  is the d.f and choose  $x = \gamma_i^{t-1}$  Generate another random variable  $y$  from the chi-square density  $g$ . Generate  $u$  from Uniform(0,1). If  $u < \frac{\pi_{5i}(y)g(x)}{\pi_{5i}(x)g(y)}$  accept  $y$  and set  $\gamma_i^t = y$ ; otherwise set  $\gamma_i^t = x$ . Repeat the procedure and generate  $\beta_i^t, i = 1, 2, \dots, m_2$ .

**Step6.** Generate  $z_{ij}$  values using sample  $y$ .

**Step7.** Generate  $p^t$ .

**Step8.** Using the proposal density  $g(\alpha_k) \sim \chi_{(x)}^2$ , where  $x$  is the d.f and choose  $x = \alpha_k^{t-1}$  Generate random variable  $y$  from the chi-square density  $g$ . Generate  $u$  from Uniform(0,1). If  $u < \frac{\pi_{4k}(y)g(x)}{\pi_{4k}(x)g(y)}$  accept  $y$  and set  $\alpha_k^t = y$ ; otherwise set  $\alpha_k^t = x$ . Repeat the procedure and generate  $\alpha_k^t, k = 1, 2, \dots, m_1$ .

**Step9.** Using the proposal density  $g(\theta_k) \sim \chi_{(x)}^2$ , where  $x$  is the d.f and choose  $x = \theta_k^{t-1}$  Generate random variable  $y$  from the chi-square density  $g$ . Generate  $u$  from Uniform(0,1). If  $u < \frac{\pi_{6k}(y)g(x)}{\pi_{6k}(x)g(y)}$  accept  $y$  and set  $\theta_k^t = y$ ; otherwise set  $\theta_k^t = x$ . Repeat the procedure and generate  $\theta_k^t, k = 1, 2, \dots, m_1$ .

**Step10** Compute  $R(t)$ .

**Step11** Increment  $t$ .

Table 4 provides the estimated values of R(t) by the Bayesian estimation method when the stress and strength of the system follow a mixture of two power-transformed half-logistic distributions with gamma cycle time. We assume that the mixture proportions  $\pi$  and  $p$  are known and

the component parameters  $(\alpha, \theta, \delta, \gamma)$  are unknown and are following gamma prior distributions. Also, we assume that cycle time parameters  $a$  and  $k$  follow non-informative prior. Since the cycle time parameters have a non-informative prior, their estimates coincide with the ML estimates. The table shows Bayes estimates of the parameters  $(\alpha, \theta, \delta, \gamma)$  and Bayes estimate of the reliability function  $R(t)$  for different time values corresponding to various sets of hyperparameter values. The table shows that  $R(t)$  decreases as time increases, as we expected.

**Table 4:** Bayesian Estimation of  $R(t)$  with PTHL mixture initial stress and strength

Cycle time		Stress and Strength		$a_0$	$x_0$	$t$	$R(t)$
Parameters	Estimated	Parameters	Estimated				
G(0.5,1)	$a = 0.5110$ $k = 0.9955$	$Y_0:0.3\text{PTHL}(0.2,5)$	$\alpha = (0.2942, 1.7069)$	0.001	0.005	10	0.0184
		$+0.7\text{PTHL}(2,6)$	$\theta = (4.5235, 5.4221)$			75	0.0185
		$X_0: 0.3\text{PTHL}(1,3)$	$\delta = (1.8201, 2.2808)$			150	0.0184
		$+0.7\text{PTHL}(2.2,4.5)$	$\gamma = (6.1306, 4.6754)$			200	0.0166
						225	0.0090
G(0.5,1)	$a = 0.4759$ $k = 0.9559$	$Y_0:0.6\text{PTHL}(2,6)$	$\alpha = (1.8055, 6.7181)$	0.002	0.005	10	0.1851
		$+0.4\text{PTHL}(4,2)$	$\theta = (5.9235, 2.1754)$			75	0.1864
		$X_0:0.4\text{PTHL}(1,3)$	$\delta = (1.1301, 1.0738)$			150	0.1858
		$+0.6\text{PTHL}(0.5,2)$	$\gamma = (2.4231, 0.6529)$			200	0.0907
						225	0.0202
G(1,4)	$a = 0.9694$ $k = 3.9365$	$Y_0:0.7\text{PTHL}(8,0.5)$	$\alpha = (0.6985, 3.4123)$	0.001	0.005	10	0.2369
		$+0.3\text{PTHL}(0.5,2.5)$	$\theta = (1.8173, 4.9566)$			75	0.2394
		$X_0:0.2\text{PTHL}(4,0.5)$	$\delta = (0.5839, 0.2675)$			125	0.2335
		$+0.8\text{PTHL}(5,0.2)$	$\gamma = (2.2849, 1.6391)$			140	0.1487
						150	0.0589
G(1,4)	$a = 0.9674$ $k = 3.8930$	$\text{Strength}:0.6\text{PTHL}(1.5,5)$	$\alpha = (1.1917, 6.8528)$	0.002	0.08	10	0.0400
		$+0.4\text{PTHL}(5,4)$	$\theta = (6.2016, 3.7667)$			75	0.0404
		$\text{Stress}:0.5\text{PTHL}(2.4,6)$	$\delta = (1.92620, 2.490)$			125	0.0394
		$+0.5\text{PTHL}(0.3,4)$	$\gamma = (5.5248, 3.8446)$			140	0.0251
						150	0.0099

## 5. CONCLUSION

In this paper, we investigated the stress-strength reliability of a system. Here we considered a scenario where the stress and strength of the system follow a finite mixture distribution with gamma cycle time. Specifically, we examined the performance of the system under two types of finite mixture models: a finite mixture of Weibull distribution and a finite mixture of power-transformed half-logistic distribution. To estimate the reliability function  $R(t)$ , we employed two methods: maximum likelihood (ML) estimation using the expectation-maximization (EM) algorithm and Bayesian estimation using the Markov Chain Monte Carlo (MCMC) method. We computed the estimates of  $R(t)$  for different time points corresponding to various sets of parameter values. Based on the graphs and tables presented in the paper, it can be observed that as time increases, the reliability function  $R(t)$  decreases when the stress and strength of the system follow a finite mixture of Weibull or power-transformed half-logistic distribution with gamma cycle time. This suggests that the system becomes less reliable or more prone to failure as time progresses.

## REFERENCES

- [1] Baklizi, A. and Eidous, O. M.(2006) Nonparametric estimation of  $P(X < Y)$  using kernel methods, *Metron* LXIV , 47–60.
- [2] Birnbaum, Z. M. (1956) On a use of the Mann-Whitney Statistic, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 1, 13–17.

- [3] Drisya, M., Jose J. K. and Krishnendu, K.(2022) Time-dependent stress-strength reliability model with phase-type cycle time based on finite mixture models, *American Journal of Mathematical and Management Sciences*, 41(2) 128-147.
- [4] Eryilmaz, S. (2018).Phase-type stress-strength models with reliability applications.*Communications in Statistics-Simulation and Computation*, 47(4), 954-963.
- [5] Gopalan, M. N. and Venkateswarlu, P.(1982) Reliability analysis of time-dependent cascade system with deterministic cycle times, *Microelectronics Reliability* 22(4), 841–872.
- [6] Gopalan, M. N. and Venkateswarlu, P.(1983). Reliability analysis of time-dependent cascade system with random cycle times, *Microelectronics Reliability*, 23(2), 355–366.
- [7] Jose J. K., Xavier, T., and Drisya, M. (2019) Estimation of stress-strength reliability Using Kumaraswamy half-logistic distribution, *Journal of Probability and Statistical Science* 17(2), 141–154.
- [8] Jose J. K. and Drisya, M. (2020). Time-Dependent Stress-Strength Reliability Models based on Phase Type Distribution, *Statistics*, 35, 1345-1371.
- [9] Jose J. K. and Drisya, M. (2021). Stress-Strength Reliability Estimation of Time-Dependent Models with Fixed Stress and Phase Type Strength Distribution. *Revista Colombiana de Estadística*, 44(1), 201-224.
- [10] Jose J. K., Drisya, M. and Manoharan, M. (2021). Estimation of Stress-Strength Reliability using Discrete Phase Type Distribution. *Communications in Statistics- Theory and Methods*, 51(2), 368-386.
- [11] Kotz, S., Lumelskii, Y., and Pensky, M. The stress-strength model and its generalizations: theory and applications, *World Scientific Publishing Co. Pvt. Ltd.*, Singapore, 2003.
- [12] Pakdaman, Z., Ahmadi, J., and Doostparast, M. (2019). Signature-based approach for stress-strength systems. *Statistical Papers*, 60(5), 1631–1647.
- [13] Siju, K. C. and Kumar, M. (2016).Reliability analysis of time-dependent stress strength model with random cycle times, *Perspectives in Science*, 8,654-657.
- [14] Siju, K. C. and Kumar, M. (2017), Reliability computation of a dynamic stress strength model with random cycle times, *International Journal of Pure and Applied Mathematics*, 117, 309-316.
- [15] Xavier, T. and Jose J. K., (2021).A study of stress-strength reliability using a generalization of power-transformed half-logistic distribution, *Communications in Statistics- Theory and Methods*, 50(18), 4335-4351.
- [16] Xavier, T. and Jose J. K., (2021). Estimation of reliability in a multicomponent stress-strength model based on power transformed half-logistic distribution. *Int J Reliab Qual Saf Eng*, 28(2):04.
- [17] Xavier, T., Jose J. K., and Nadarajah, S.(2022). An additive power-transformed half-logistic model and its applications in reliability. *Quality and Reliability Engineering International*, <https://doi.org/10.1002/qre.3119>.
- [18] Xavier, T., Jose J. K., Subhash C. Bagui (2023). Stress Strength Reliability Estimation of a Series System with Cold Standby Redundancy Based on Kumaraswamy Half-Logistic Distribution. *American Journal of Mathematical and Management Sciences*, DOI: 10.1080/01966324.2023.2213835.
- [19] Yadav,R. P. S.(1973) A reliability model for stress strength problem, *Microelectronics Reliability*, 12(2), 119–123.
- [20] Zhou, W. (2008) Statistical inference for  $P(X < Y)$ , *Statistics in Medicine*, 27, 257–279.

## DUAL EXPONENTIAL RATIO ESTIMATOR IN PRESENCE OF NON-RESPONSE

Rafia Jan<sup>1</sup>, T. R. Jan<sup>2</sup> and Faizan Danish<sup>3\*</sup>

<sup>1</sup>Department of Statistics, Government Degree College Bejbehara Anantnag, J&K, India

<sup>2</sup>Department of Statistics, University of Kashmir, J&K, India

<sup>3</sup>Department of Mathematics, School of Advanced Sciences, VIT-AP University, Inavolu,  
Beside AP Secretariat, Amravati, Andhra Pradesh-522237, India.

<sup>1</sup>rafiajan836@gmail.com, <sup>2</sup>drtrjan@gmail.com, <sup>3</sup>danishstat@gmail.com

### Abstract

*The manuscript under consideration delves into a comprehensive exploration of the dual exponential ratio estimator, particularly in the context of non-response scenarios. In the following discourse, we will embark on an intricate journey through this research, emphasizing the pivotal aspects and findings that unravel the significance of this estimator in the realm of statistical estimation. The crux of this investigation revolves around evaluating the Mean Squared Error (MSE) and the Predictive Relative Efficiency (PRE) of the dual exponential ratio estimator. These two performance metrics serve as essential benchmarks for assessing the accuracy and effectiveness of the estimator. Notably, they play a crucial role in determining the estimator's suitability for practical applications, especially in situations where non-response is prevalent. To begin our exploration, it is imperative to understand the fundamental concept of the dual exponential ratio estimator. This estimator is a statistical tool employed in situations where traditional estimators may falter due to non-response, a phenomenon frequently encountered in surveys and data collection. It leverages a dual exponential model to address this challenge, making it a valuable addition to the toolkit of statisticians and researchers. The manuscript embarks on a rigorous theoretical analysis of the dual exponential ratio estimator's MSE and PRE. Through a series of mathematical derivations and proofs, the authors elucidate the underlying principles governing its performance. This theoretical foundation is crucial, as it not only establishes a solid framework for evaluating the estimator but also provides insights into its behavior under different conditions. However, theory alone can only take us so far. To validate the theoretical findings and assess the estimator's practical utility, numerical experiments are conducted. These experiments involve simulations and real-world data scenarios, allowing the authors to draw comparisons between the dual exponential ratio estimator and traditional estimators. The numerical results serve as a bridge between theory and application, offering empirical evidence of the estimator's prowess. In essence, this manuscript fills a critical gap in the field of statistical estimation by thoroughly investigating the dual exponential ratio estimator's performance in the presence of non-response. By juxtaposing its MSE and PRE with those of traditional estimators, it provides valuable insights into the potential advantages of adopting this novel approach. Moreover, the combination of rigorous theory and practical validation ensures that the findings are both intellectually sound and operationally relevant. The dual exponential ratio estimator, as explored and analyzed within these pages, emerges as a promising solution, backed by both theoretical rigor and empirical support. This research contributes not only to the theoretical foundations of statistics but also to its real-world applications, underscoring the estimator's potential to enhance the accuracy and reliability of estimation in the face of non-response complexities.*

**Keywords:** Non-Response (NR), Exponential Estimator, Dual to Ratio Estimator, Mean Square Error and Percent Relative Efficiency.

## I. Introduction

In recent years, the use of sample surveys has gained popularity due to the practicality of overcoming logistical challenges associated with conducting comprehensive census surveys. This trend has led to the widespread adoption of estimators like the ratio, product, and regression estimators for efficiently estimating population parameters, particularly the mean of the variable of interest. These estimators capitalize on the inherent correlation between the study variable and auxiliary variables, either during the survey design or at the estimation stage, to yield accurate results while optimizing resources. The central focus of this research is to develop a novel modified exponential ratio estimator for the population mean. This estimator aims to address potential limitations of existing estimators and enhance the precision of estimates, as evaluated through mean squared error comparisons. By exploring alternative approaches and incorporating adjustments, the researchers anticipate achieving more reliable and efficient estimates of the population mean.

Over the years, several scholars have made significant contributions to the field of survey estimation. Various authors have made numerous work for the estimation of population variance from time to time including [14],[9] , [13], [8] [1], [5], [11],[12],[15] and [10] have made important studies on this topic in the literature. Notably, [17] made pioneering strides by explicitly utilizing auxiliary information for estimation purposes, laying the foundation for the ratio estimator. Subsequently, [18] further advanced this concept by employing auxiliary information to refine estimations.

When dealing with scenarios where the coefficient of correlation is negative between the study variable and auxiliary variables, [19] introduced the product-type estimator, which has proven to be valuable in specific contexts. Additionally, [20] proposed an innovative approach by combining multiple ratio estimators based on individual auxiliary variables positively correlated with the study variable. This technique allowed for greater accuracy in estimation. The product estimator was formalized by [21], providing a well-defined framework for its application. Furthermore, [22] delved into the complexities of ratio estimators involving two or more correlated variables, shedding light on new possibilities for refining estimation methods. The exponential type estimators of population mean were thoroughly investigated by [23] using auxiliary data, resulting in a comprehensive analysis of their performance and potential improvements. [24] took a unique approach by incorporating transformed auxiliary variables, which led to promising results in estimating the mean of the study character. The literature offers an array of other contributions in this area, including the works of [25], [26], [27], and [28], who introduced their respective estimators and demonstrated their efficacy in diverse sampling scenarios. Moreover, [29] and [30] took on the challenge of developing superior exponential type estimators by considering information from two altered auxiliary variables, further expanding the range of available estimation techniques. To gain a more comprehensive understanding of this topic, interested readers can refer to [31], which offers an in-depth exploration of various aspects of survey estimation. In recent times, [32], [33], and [34] have made notable contributions to this area of study, introducing novel ideas and methodologies that hold promise for advancing the field of survey estimation even further.

In conclusion, this research endeavors to create a Generalized Ratio-cum-product estimator of population variance that builds upon the knowledge and advancements made by previous scholars. By harnessing the power of auxiliary information and exploring innovative avenues, the researchers aim to provide an enhanced and efficient approach to estimating the population mean and contributing to the growing body of knowledge in survey estimation techniques.

## II. Notations

Let  $N$  and  $n$  be population and sample of size respectively. Out of  $n$  units ' $n_1$ ' responds and ' $n_2$ ' do not respond accordingly, the population is distributed in ' $N_1$ ' (those who respond) and ' $N_2$ ' (the non-respondents), such that  $N_1 + N_2 = N$ . From sample of ' $n_2$ ' a sub-sample of size  $k$  where  $\left(k = \frac{n_2}{h}, h > 1\right)$  is taken and data is obtained. Further, we define

$$W_1 = \frac{N_1}{N}, W_2 = \frac{N_2}{N}, \lambda = \frac{1}{n} - \frac{1}{N}, \theta = \frac{W_2(h-1)}{n}, w_1 = \frac{n_1}{n}, w_2 = \frac{n_2}{n}, \bar{Y} = \frac{1}{N} \sum_{i=1}^N y_i,$$

$$\bar{Y}_2 = \frac{1}{N_2} \sum_{i=1}^{N_2} y_i, \bar{X} = \frac{1}{N} \sum_{i=1}^N x_i, \bar{X}_2 = \frac{1}{N_2} \sum_{i=1}^{N_2} x_i, C_y = \frac{S_y}{\bar{Y}}, C_{y(2)} = \frac{S_{y(2)}}{\bar{Y}}, C_x = \frac{S_x}{\bar{X}}, C_{x(2)} = \frac{S_{x(2)}}{\bar{X}},$$

$$S_y^2 = \frac{\sum_{i=1}^N (y_i - \bar{Y})^2}{N-1}, S_{y(2)}^2 = \frac{\sum_{i=1}^{N_2} (y_i - \bar{Y}_2)^2}{N_2-1}, S_x^2 = \frac{\sum_{i=1}^N (x_i - \bar{X})^2}{N-1}, S_{x(2)}^2 = \frac{\sum_{i=1}^{N_2} (x_i - \bar{X}_2)^2}{N_2-1},$$

$$C = \rho_{yx} \left( \frac{C_y}{C_x} \right), C_{(2)} = \rho_{yx(2)} \left( \frac{C_{y(2)}}{C_{x(2)}} \right)$$

## III. Existing Estimators

Hansen and Hurwitz proposed an unbiased estimator of  $\bar{Y}$  in case of non-response,

$$\bar{y}^* = w_1 \bar{y}_1 + w_2 \bar{y}_2,$$

where  $\bar{y}_1$  is the sample mean of respondents and  $\bar{y}_2$  is the mean of sub-sample of non-respondents,

The variance is,

$$V(\bar{y}^*) = \bar{Y}^2 (\lambda C_y^2 + \theta C_{y(2)}^2),$$

The unbiased estimator of  $\bar{X}$  in case of non-response is given as,

$$\bar{x}^* = w_1 \bar{x}_1 + w_2 \bar{x}_2,$$

where  $\bar{x}_1$  the sample is mean of the respondents and similarly  $\bar{x}_2$  is the mean of sub-sample.

The variance is

$$V(\bar{x}^*) = \bar{X}^2 (\lambda C_x^2 + \theta C_{x(2)}^2).$$

### 3.1 Case I: Non-response on $y$ only

Ratio estimator of  $\bar{Y}$  in case I is,

$$t_{RI} = \frac{\bar{y}^*}{\bar{x}} \bar{X},$$

$$MSE(t_{RI}) = \bar{Y}^2 \left[ \lambda (C_y^2 + C_x^2 - 2\rho_{yx} C_y C_x) + \theta C_{y(2)}^2 \right].$$

The dual to ratio estimator given by Srivenkentrama (1980) is,

$$t_D = \bar{y} \left( \frac{\bar{x}^\beta}{\bar{X}} \right),$$

where  $\bar{x}^\beta = \left( \frac{\bar{x}-}{\bar{X}} \right), i = 1, 2, \dots, N.$

and the MSE is,

$$MSE(t_D) = \lambda \bar{Y}^2 (C_y^2 + C_x^2 g (g - 2C))$$

The dual of ratio estimator in case of non-response is,

$$t_{D1} = \bar{y}^* \left( \frac{\bar{x}^\beta}{\bar{X}} \right),$$

The MSE is given by

$$MSE(t_{D1}) = \bar{Y}^2 \left[ \lambda (C_y^2 + C_x^2 g (g - 2C)) + \theta C_{y(2)}^2 \right]$$

Singh and Kumar (2009) considered the exponential estimators of  $\bar{Y}$  in case of non-response.

$$t_{ER1} = \bar{y}^* \exp \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$$

The MSE is,

$$MSE(t_{ER1}) = \bar{Y}^2 \left[ \lambda \left( C_y^2 + \frac{1}{4} C_x^2 - \rho_{xy} C_x C_y \right) + \theta (C_{y(2)}^2) \right]$$

The dual exponential estimator for non-response is

$$t_{ED1} = \bar{y}^* \exp \left( \frac{\bar{x}^\beta - \bar{X}}{\bar{x}^\beta + \bar{X}} \right),$$

The MSE is,

$$MSE(t_{ED1}) = \bar{Y}^2 \left[ \lambda \left( C_y^2 + \frac{g^2 C_x^2}{4} - g \rho_{yx} C_x C_y \right) + \theta C_{y(2)}^2 \right]$$

### 3.2 Case II: Non-response on both y and x

The ratio estimator of  $\bar{Y}$  for case II along with MSE is given as,

$$t_{R2} = \frac{\bar{y}^*}{\bar{x}^*} \bar{X},$$

$$MSE(t_{R2}) = \left[ \lambda \bar{Y}^2 (C_x^2 + C_y^2 - 2\rho_{xy} C_x C_y) + \theta \bar{Y}^2 (C_{x(2)}^2 + C_{y(2)}^2 - 2\rho_{xy(2)} C_{x(2)} C_{y(2)}) \right],$$

where  $\rho_{yx} = \frac{S_{yx}}{S_y S_x}$  is the correlation for the overall population, while  $\rho_{yx(2)} = \frac{S_{yx(2)}}{S_{y(2)} S_{x(2)}}$  is the

case of non-respondent group.

Dual of ratio estimator for case II is,

$$t_{D2} = \bar{y}^* \frac{\bar{x}^{*\beta}}{\bar{X}}$$

and the MSE is,

$$MSE(t_{D2}) = \bar{Y}^2 \left[ \lambda (C_y^2 + C_x^2 g (g - 2C)) + \theta (C_{y(2)}^2 + C_{x(2)}^2 g (g - C_{(2)})) \right]$$

The exponential ratio estimator is given as,

$$t_{ER2} = \bar{y}^* \exp \left( \frac{\bar{X} - \bar{x}^*}{\bar{X} + \bar{x}^*} \right)$$

$$MSE(t_{ER2}) = \bar{Y}^2 \left[ \lambda \left( C_y^2 + \frac{1}{4} C_x^2 - \rho_{xy} C_x C_y \right) + \theta \left( C_{y(2)}^2 + \frac{C_{x(2)}^2}{4} - \rho_{xy(2)} C_{x(2)} C_{y(2)} \right) \right]$$



The dual of exponential ratio estimator is

$$t_{ED2} = \bar{y}^* \exp\left(\frac{\bar{x}^{*\beta} - \bar{X}}{\bar{x}^{*\beta} + \bar{X}}\right)$$

And the MSE is given by

$$MSE(t_{ED2}) = \bar{Y}^2 \left[ \lambda \left( C_y^2 + \frac{g^2 C_x^2}{4} - g\rho_{yx} C_x C_y \right) + \theta \left( C_{y(2)}^2 + \frac{g^2 C_{x(2)}^2}{4} - g\rho_{xy(2)} C_{x(2)} C_{y(2)} \right) \right]$$

### 3.3 Proposed Estimator: Case I

The proposed ratio-cum-dual of exponential ratio estimator of  $\bar{Y}$  is given as,

$$t^* = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}}\right)^\alpha \exp\left[\frac{\delta(\bar{x}^\beta - \bar{X})}{(\bar{x}^\beta + \bar{X})}\right], \tag{1}$$

here  $x^\beta = \frac{N\bar{X} - n\bar{x}}{N - n}$ ,  $\alpha$  and  $\delta$  are constants.

here  $x^\beta = \frac{N\bar{X} - n\bar{x}}{N - n}$ ,  $\alpha$  and  $\delta$  are constants.

**Table 1:** Some members of the proposed class of estimator

S. No.	Estimator	Values of Constants	
		$\alpha$	$\delta$
1.	$t^* = \bar{y}^*$	0	0
2.	$t^* \rightarrow t_{R1} = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}}\right)$	1	0
3.	$t^* \rightarrow t_{P1} = \bar{y}^* \left(\frac{\bar{x}}{\bar{X}}\right)$	-1	0
4.	$t^* = \bar{y}^* \left(\frac{\bar{X}^2}{\bar{x}^2}\right)$	2	0
5.	$t^* = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}}\right)^{\frac{1}{2}}$	$\frac{1}{2}$	0
6.	$t^* = \bar{y}^* \exp\left(\frac{\bar{x}^\beta - \bar{X}}{\bar{x}^\beta + \bar{X}}\right)$	0	1
7.	$t^* = \bar{y}^* \exp\left(\frac{(\bar{x}^\beta - \bar{X})}{2(\bar{x}^\beta + \bar{X})}\right)$	0	$\frac{1}{2}$
8.	$t^* = \bar{y}^* \exp\left(\frac{2(\bar{x}^\beta - \bar{X})}{(\bar{x}^\beta + \bar{X})}\right)$	0	2
9.	$t^* = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}}\right) \exp\left(\frac{\bar{x}^\beta - \bar{X}}{\bar{x}^\beta + \bar{X}}\right)$	1	1

10.	$t^* = \bar{y}^* \left( \frac{\bar{X}}{\bar{x}} \right)^{\frac{1}{2}} \exp \left( \frac{1(\bar{x}^\beta - \bar{X})}{2(\bar{x}^\beta + \bar{X})} \right)$	$\frac{1}{2}$	$\frac{1}{2}$
-----	--	---------------	---------------

The associated sample mean is obtained as

$$\bar{x}^\beta = (1 + g)\bar{X} - g\bar{x} \text{ and } g = \frac{n}{N - n}$$

To acquire the MSE, we write

$$\bar{y}^* = \bar{Y}(1 + e_0^*) \text{ and } \bar{x} = \bar{X}(1 + e_1),$$

Such that,

$$E(e_0^{*2}) = (\lambda C_y^2 + \theta C_{y(2)}^2), E(e_1^2) = \lambda C_x^2, E(e_0^* e_1) = \lambda C C_x^2$$

Expressing (1) in  $e$ 's we have

$$\begin{aligned} t^* &= \bar{Y}(1 + e_0) \left( \frac{\bar{X}}{\bar{X}(1 + e_1)} \right)^\alpha \exp \left[ \delta \left\{ \frac{(1 + g)\bar{X} - g(\bar{X}(1 + e_1)) - \bar{X}}{(1 + g)\bar{X} - g(\bar{X}(1 + e_1)) + \bar{X}} \right\} \right] \\ &= \bar{Y}(1 + e_0) (1 + e_1)^{-\alpha} \exp \left[ \delta \left( \frac{-g e_1}{2 - g e_1} \right) \right] \\ &= \bar{Y}(1 + e_0) (1 + e_1)^{-\alpha} \exp \left( -\frac{\delta g e_1}{2} \left( 1 - \frac{g e_1}{2} \right)^{-1} \right) \\ &= \bar{Y}(1 + e_0) \left( 1 - \alpha e_1 + \frac{\alpha(\alpha + 1)}{2} e_1^2 - \dots \right) \left( 1 - \frac{\delta g e_1}{2} \left( 1 - \frac{g e_1}{2} \right)^{-1} + \dots \right) \end{aligned}$$

Ignoring higher order terms,

$$t^* = \bar{Y}(1 + e_0) (1 - \alpha e_1) \left( 1 - \frac{\delta g e_1}{2} \right)$$

$$(t^* - \bar{Y}) = \bar{Y} \left( e_0 - \left( \frac{\delta g + 2\alpha}{2} \right) e_1 \right)$$

Squaring both sides, we get,

$$(t^* - \bar{Y})^2 = \bar{Y}^2 \left( e_0^2 + \left( \frac{\delta g + 2\alpha}{2} \right) e_1^2 - (2\alpha + \delta g) e_0 e_1 \right)$$

Taking expectation, we get the MSE as,

$$MSE(t^*) = \bar{Y}^2 \left[ (\lambda C_y^2 + \theta C_{y(2)}^2) + \lambda C_x^2 \left\{ \left( \frac{\delta g + 2\alpha}{2} \right)^2 - (\delta g + 2\alpha) C \right\} \right] \tag{2}$$

Differentiate (2) w.r.t.  $\alpha$  and equate it to zero,

$$\frac{\partial}{\partial \alpha} MSE(t^*) = \frac{\partial}{\partial \alpha} \left\{ \bar{Y}^2 \left[ (\lambda C_y^2 + \theta C_{y(2)}^2) + \lambda C_x^2 \left\{ \left( \frac{\delta g + 2\alpha}{2} \right)^2 - (\delta g + 2\alpha) C \right\} \right] \right\} = 0$$

$$\frac{\partial}{\partial \alpha} MSE(t^*) = 2\alpha = \delta g + 2C$$

$$\alpha = \frac{\delta g + 2C}{2}$$

We can write

$$(\delta g + 2\alpha) = 2C \tag{3}$$

Using (3) in (2) we get,

$$MSE(t^*)_{\min} = \bar{Y}^2 \left[ \lambda C_y^2 (1 - \rho_{xy}^2) + \theta C_{y(2)}^2 \right] \tag{4}$$

### 3.4 Proposed Estimator: Case II

The proposed estimator of  $\bar{Y}$  is

$$t^{**} = \bar{y}^* \left( \frac{\bar{X}}{\bar{x}^*} \right)^\alpha \exp \left[ \frac{\delta (\bar{x}^{*\beta} - \bar{X})}{(\bar{x}^{*\beta} + \bar{X})} \right] \tag{5}$$

where  $x^{*\beta} = \frac{N\bar{X} - n\bar{x}^*}{N-n}$ ,  $\alpha$  and  $\delta$  are suitably chosen constant.

The associated sample mean is obtained as

$$\bar{x}^\beta = (1 + g)\bar{X} - g\bar{x} \text{ and } g = \frac{n}{N-n}$$

**Table 2:** Some members of the proposed class of estimator

S. No.	Estimator	Values of Constants	
		$\alpha$	$\delta$
1.	$t^* = \bar{y}^*$	0	0
2.	$t^{**} \rightarrow t_{R2} = \bar{y}^* \left( \frac{\bar{X}}{\bar{x}^*} \right)$	1	0
3.	$t^{**} \rightarrow t_{P2} = \bar{y}^* \left( \frac{\bar{x}^*}{\bar{X}} \right)$	-1	0
4.	$t^{**} = \bar{y}^* \left( \frac{\bar{X}^2}{\bar{x}^{*2}} \right)$	2	0
5.	$t^{**} = \bar{y}^* \left( \frac{\bar{X}}{\bar{x}^*} \right)^{\frac{1}{2}}$	$\frac{1}{2}$	0
6.	$t^{**} = \bar{y}^* \exp \left( \frac{\bar{x}^{*\beta} - \bar{X}}{\bar{x}^{*\beta} + \bar{X}} \right)$	0	1
7.	$t^{**} = \bar{y}^* \exp \left( \frac{(\bar{x}^{*\beta} - \bar{X})}{2(\bar{x}^{*\beta} + \bar{X})} \right)$	0	$\frac{1}{2}$
8.	$t^{**} = \bar{y}^* \exp \left( \frac{2(\bar{x}^{*\beta} - \bar{X})}{(\bar{x}^{*\beta} + \bar{X})} \right)$	0	2
9.	$t^{**} = \bar{y}^* \left( \frac{\bar{X}}{\bar{x}} \right) \exp \left( \frac{\bar{x}^{*\beta} - \bar{X}}{\bar{x}^{*\beta} + \bar{X}} \right)$	1	1
10.	$t^{**} = \bar{y}^* \left( \frac{\bar{X}}{\bar{x}} \right)^{\frac{1}{2}} \exp \left( \frac{1(\bar{x}^{*\beta} - \bar{X})}{2(\bar{x}^{*\beta} + \bar{X})} \right)$	$\frac{1}{2}$	$\frac{1}{2}$

To acquire the MSE,

$$\bar{x}^* = \bar{X} (1 + e_1^*),$$

$$E(e_0^{*2}) = (\lambda C_y^2 + \theta C_{y(2)}^2) \text{ and } E(e_0^* e_1^*) = (\lambda C C_x^2 + \theta C_{(2)} C_{x(2)}^2)$$

Ignoring the higher terms, we get,

$$t^{**} = \bar{Y} (1 + e_0) (1 - \alpha e_1) \left( 1 - \frac{\delta g e_1}{2} \right)$$

$$(t^{**} - \bar{Y}) = \bar{Y} \left( e_0 - \frac{(\delta g + 2\alpha)}{2} e_1 \right)$$

Squaring the above equation, we get

$$(t^{**} - \bar{Y})^2 = \bar{Y}^2 \left( e_0^2 + \left\{ \frac{(\delta g + 2\alpha)}{2} \right\}^2 e_1^2 - (\delta g + 2\alpha) e_0 e_1 \right)$$

Taking expectation, we get

$$MSE(t^{**}) = \bar{Y}^2 \left[ \left\{ \left( \lambda C_y^2 + \theta C_{y(2)}^2 \right) + \left( \lambda C_x^2 + \theta C_{x(2)}^2 \right) \right\} \left\{ \left( \frac{\delta g + 2\alpha}{2} \right)^2 - (\delta g + 2\alpha) R \right\} \right], \tag{6}$$

where  $R = \frac{(\lambda C C_x^2 + \theta C_{(2)} C_{x(2)}^2)}{(\lambda C_x^2 + \theta C_{x(2)}^2)}$ .

Differentiate (6) w.r.t  $\alpha$  we get,

$$\frac{\delta}{\delta \alpha} MSE(t^{**}) = \frac{\delta}{\delta \alpha} \left\{ \bar{Y}^2 \left[ \left\{ \left( \lambda C_y^2 + \theta C_{y(2)}^2 \right) + \left( \lambda C_x^2 + \theta C_{x(2)}^2 \right) \right\} \left\{ \left( \frac{\delta g + 2\alpha}{2} \right)^2 - (\delta g + 2\alpha) R \right\} \right] \right\} = 0$$

$$\frac{\delta}{\delta \alpha} MSE(t^{**}) = 2\alpha = \delta g + 2R$$

$$\alpha = \frac{\delta g + 2R}{2}$$

We can write,

$$(\delta g + 2\alpha) = 2R \tag{7}$$

Substituting (7) in (6) we have,

$$MSE(t^{**})_{\min} = \bar{Y}^2 \left[ (\lambda C_y^2 + \theta C_{y(2)}^2) - \frac{(\lambda C C_x^2 + \theta C_{(2)} C_{x(2)}^2)^2}{(\lambda C_x^2 + \theta C_{x(2)}^2)} \right]$$

$$MSE(t^{**})_{\min} = \bar{Y}^2 [(\lambda C_y^2 + \theta C_{y(2)}^2) (1 - \rho^{*2})]. \tag{8}$$

Where  $\rho^* = \frac{Cov(\bar{y}^*, \bar{x}^*)}{\sqrt{V(\bar{y}^*)V(\bar{x}^*)}} = \frac{(\lambda C C_x^2 + \theta C_{(2)} C_{x(2)}^2)}{\sqrt{(\lambda C_y^2 + \theta C_{y(2)}^2)(\lambda C_x^2 + \theta C_{x(2)}^2)}}$

#### IV. Theoretical Efficiency Comparison

##### 4.1 Case I

$$V(t_0) - MSE(t^*) \geq 0.$$

$$\begin{aligned}
 &= \bar{Y}^2(\lambda C_y^2 + \theta C_{y(2)}^2) - \bar{Y}^2[\lambda C_y^2(1 - \rho_{xy}^2) + \theta C_{y(2)}^2] \geq 0 \\
 &= \bar{Y}^2 \lambda C_x^2 C_x^2 \geq 0 \\
 \text{MSE}(t_{R1}) - \text{MSE}(t^*) &\geq 0. \\
 &= \bar{Y}^2[\lambda(C_y^2 + C_x^2 - 2\rho_{yx} C_y C_x) + \theta C_{y(2)}^2] - \bar{Y}^2[\lambda C_y^2(1 - \rho_{xy}^2) + \theta C_{y(2)}^2] \geq 0 \\
 &= \bar{Y}^2 \lambda (C_x - C C_x)^2 \geq 0 \\
 \text{MSE}(t_{ER1}) - \text{MSE}(t^*) &\geq 0 \\
 &= \bar{Y}^2 \left[ \lambda \left( C_y^2 + \frac{1}{4} C_x^2 - \rho_{xy} C_x C_y \right) + \theta (C_{y(2)}^2) \right] - \bar{Y}^2 [\lambda C_y^2(1 - \rho_{xy}^2) + \theta C_{y(2)}^2] \geq 0 \\
 &= \bar{Y}^2 \lambda \frac{C_x^2}{4} (1 - 2C)^2 \geq 0
 \end{aligned}$$

#### 4.2 Case II

$$\begin{aligned}
 V(t_0) - \text{MSE}(t^{**}) &\geq 0. \\
 &= \bar{Y}^2(\lambda C_y^2 + \theta C_{y(2)}^2) - \bar{Y}^2 \left[ (\lambda C_y^2 + \theta C_{y(2)}^2) - \frac{(\lambda C C_x^2 + \theta C_{(2)} C_{x(2)}^2)^2}{(\lambda C_x^2 + \theta C_{x(2)}^2)} \right] \geq 0 \\
 &= \frac{(\lambda C C_x^2 + \theta C_{(2)} C_{x(2)}^2)}{(\lambda C_x^2 + \theta C_{x(2)}^2)} \geq 0 \\
 \text{MSE}(t_{R2}) - \text{MSE}(t^{**}) &\geq 0. \\
 &= \left\{ \begin{aligned} &\lambda \bar{Y}^2 (C_x^2 + C_y^2 - 2\rho_{xy} C_x C_y) + \theta \bar{Y}^2 (C_{x(2)}^2 + C_{y(2)}^2 - 2\rho_{xy(2)} C_{x(2)} C_{y(2)}) \\ &- \bar{Y}^2 \left[ (\lambda C_y^2 + \theta C_{y(2)}^2) - \frac{(\lambda C C_x^2 + \theta C_{(2)} C_{x(2)}^2)^2}{(\lambda C_x^2 + \theta C_{x(2)}^2)} \right] \end{aligned} \right\} \geq 0 \\
 &= \bar{Y}^2 [(\lambda C_x^2 + \theta C_{x(2)}^2) - (\lambda C C_x^2 + \theta C_{(2)} C_{x(2)}^2)]^2 \geq 0 \\
 \text{MSE}(t_{ER2}) - \text{MSE}(t^{**}) &\geq 0. \\
 &= \left\{ \begin{aligned} &\bar{Y}^2 \left[ \lambda \left( C_y^2 + \frac{1}{4} C_x^2 - \rho_{xy} C_x C_y \right) + \theta \left( C_{y(2)}^2 + \frac{C_{x(2)}^2}{4} - \rho_{xy(2)} C_{x(2)} C_{y(2)} \right) \right] \\ &- \bar{Y}^2 \left[ (\lambda C_y^2 + \theta C_{y(2)}^2) - \frac{(\lambda C C_x^2 + \theta C_{(2)} C_{x(2)}^2)^2}{(\lambda C_x^2 + \theta C_{x(2)}^2)} \right] \end{aligned} \right\} \geq 0 \\
 &= \bar{Y}^2 \left[ \left( \lambda \frac{C_x^2}{2} + \theta \frac{C_{x(2)}^2}{2} \right) - (\lambda C C_x^2 + \theta C_{(2)} C_{x(2)}^2) \right]^2 \geq 0
 \end{aligned}$$

### V. Empirical Study

We have used two data sets.  
 Population I: Khare and Sinha (2004).  
 y: weight in kg of children,

x: chest circumference in cm of children  
 Population II: Satici and Kadilar (2011).  
 y: number of successful students,  
 x: number of teachers.

**Table 3: Data Statistics**

Population I [Khare and Sinha (2004)]		Population II[Satici and Kadilar (2011)]	
$N = 95$	$N_1 = 71$	$N = 261$	$N_1 = 196$
$n = 35$	$N_2 = 24$	$n = 90$	$N_2 = 65$
$\bar{X} = 55.86$	$C_{xy(2)} = 0.00395$	$S_y = 415.1944$	$C_x = 1.7595$
$\bar{Y} = 19.5$	$\rho_{xy(2)} = 0.729$	$C_y = 1.8654$	$\rho_{xy} = 0.9705$
$C_x = 0.05860$	$C_{x(2)} = 0.05402$	$\bar{Y} = 222.57$	$S_{x(2)} = 376.48$
$\rho_{xy} = 0.85$	$C_{xy} = 0.00776$	$\bar{X} = 306.43$	$C_{x(2)} = 1.2285$
$C_y = 0.15613$	$C_{y(2)} = 0.12075$	$S_x = 539.1722$	$\rho_{xy(2)} = 0.9733$

**Table 4: The MSE and PRE's of the population I w.r.t unbiased estimator for case I**

Estimators	$h = 2$		$h = 3$		$h = 4$	
	MSE	PRE	MSE	PRE	MSE	PRE
$t_0$	0.2015	100	0.2395	100	0.2813	100
$t_{R1}$	0.1293	155.81	0.1593	150.31	0.1960	143.52
$t_{ER1}$	0.1594	126.41	0.1914	125.13	0.2264	124.21
$t^*_{(prop)}$	<b>0.0791</b>	<b>254.43</b>	<b>0.1172</b>	<b>204.27</b>	<b>0.1574</b>	<b>178.70</b>

**Table 5: The MSE and PRE's of population I w.r.t unbiased estimator for case II**

Estima tors	$h = 2$		$h = 3$		$h = 4$	
	MSE	PRE	MSE	PRE	MSE	PRE
$t_0$	0.2015	100	0.2395	100	0.2813	100
$t_{R1}$	0.1363	147.79	0.1743	137.35	0.2152	130.66
$t_{ER1}$	0.1654	121.81	0.2023	117.73	0.2447	114.93
$t^*_{(prop)}$	0.0832	242.16	0.1022	234.31	0.1228	228.92

**Table 6: The MSE and PRE's population II w.r.t unbiased estimator for case I**

Estimator s	$h = 2$		$h = 3$		$h = 4$	
	MSE	PRE	MSE	PRE	MSE	PRE
$t_0$	1459.59	100	1664.28	100	1868.98	100
$t_{R1}$	278.57	523.95	483.27	344.37	687.96	271.67
$t_{ER1}$	589.96	247.40	796.66	209.43	999.36	187.02
$t^*_{(prop)}$	275.92	529.00	480.92	346.28	685.32	272.72

*Table 7: The MSE and PRE's of population II w.r.t unbiased estimator for case II*

Estimator s	h = 2		h = 3		h = 4	
	MSE	PRE	MSE	PRE	MSE	PRE
$t_0$	1459.59	100	1664.28	100	1868.98	100
$t_{R2}$	84.90	1719.15	95.92	1735.01	106.94	1747.60
$t_{ER2}$	441.20	330.82	497.13	334.38	553.06	337.93
$t^{**}$	82.77	1763.31	94.07	1769.19	105.21	1776.46

## V. Conclusion

In the context of survey sampling and estimation, the ratio-cum-dual of exponential ratio estimator has been proposed as a valuable approach, particularly in cases involving non-response. This innovative method combines elements of the traditional ratio estimator and dual to improve estimation accuracy. To assess the performance of this novel estimator, an essential step is to compute the Mean Squared Error (MSE) expression. This metric provides insights into the estimator's precision and reliability in estimating population parameters. To further evaluate the efficacy of the suggested estimator, both theoretical and empirical analyses have been conducted. Theoretical assessments involve rigorous mathematical proofs and calculations, while empirical evaluations utilize real-world data to validate the estimator's practical utility. The synergy of these two evaluation approaches ensures a comprehensive understanding of the estimator's competence. Upon scrutinizing the results presented in the accompanying table, a compelling conclusion emerges. It is evident that the proposed estimator surpasses the existing estimators found in the literature in terms of efficiency. This conclusion is drawn from a careful consideration of the MSE values, which indicate that the proposed estimator consistently provides more accurate and precise estimates, even in the presence of non-response. Therefore, this study contributes to the field by introducing a superior estimator for survey sampling, offering improved accuracy and reliability in estimating population parameters.

## References

- [1] Chanu, W. W. and Singh B. K. (2015). Improved Exponential Ratio cum Exponential Dual to Ratio Estimator of Finite Population Mean in Presence of Non-Response. *J. Stat. Appl. Pro.*, 4(1), 103-111.
- [2] Hansen, M.H. and Hurwitz, W.N. (1946). The problem of non-response in sample surveys. *J. Amer. Statist. Assoc.* 41, 517-529.
- [3] Ismail, M., Shahbaz, M.Q., and Hanif, M. (2011). A General class of estimator of population mean in the presence of non-response. *Pak. J. Statist.* 27(4), 467-476.
- [4] Khare, B. B and Sinha, R. (2004). Estimation of population ratio using two-phase sampling scheme in the presence of non-response. *Aligarh Journal of Statistics*, 24:43-46.
- [5] Kumar, S. and Bhougal, S. (2011). Estimation of the Population Mean in Presence of Non-Response. *Communications of the Korean Statistical Society*, 18(4), 537-548.
- [6] Pal, S. K. and Singh, H. P. (2016). Finite Population Mean Estimation through a Two-Parameter Ratio Estimator Using Auxiliary Information in Presence of Non-Response. *JAMSI*, 12(2), 5-39.
- [7] Rao, P.S.R.S. (1986). Ratio estimation with sub sampling the non-respondents. *Surv. Methodology*, 12(2), 217-230.
- [8] Rao, P.S.R.S. (1990). Regression estimators with sub sampling of non-respondents, *In-Data Quality Control. Theory and Pragmatics*, (Gunar E. Liepins and V.R.R. Uppuluri, eds) Marcel Dekker, New York, 191-208.

[9] Satici, E., and Kadilar, C. (2011). Ratio estimator for the population mean at the current occasion in the presence of non-response in successive sampling. *Hacettepe Journal of Mathematics and Statistics*, 40(1), 115–24.

[10] Singh, H.P. and Kumar, S. (2009). A general procedure for estimating the population mean in presence of non-response under double sampling using auxiliary information. *SORT*, 33(1), 71-83.



# PROCESS CAPABILITY ANALYSIS FOR NON NORMAL DATA BASED ON BOX-COX TRANSFORMATION THROUGH TESTS OF GOODNESS OF FIT

J. Krishnan

•

Department of Mathematics, Sri Krishna Adithya College of Arts and Science  
Coimbatore – 641042, Tamil Nadu, INDIA  
krrishme92@gmail.com

R. Vijayaraghavan

•

Department of Statistics, Bharathiar University  
Coimbatore 641 046, Tamil Nadu, INDIA  
vijaystatbu@gmail.com

## Abstract

Process capability analysis is an effective and efficient tool for quality assurance. When the distribution of the underlying quality characteristics is not normal, modifications of the basic process capability indices are required. Literature in process control provides avenues to resolve the issue of non-normality and data transformation is one of the approaches frequently applied in practice. Primarily the Box – Cox transformation (BCT) is employed to transform the non normal data into normal data which originally utilizes the method of maximum likelihood estimation (MLE) to find the single transformation parameter  $\lambda$ . There are alternative methods to estimate the optimal parametric value  $\lambda$  using goodness of fit tests rather using MLE method. In order to bring improved estimates, this paper makes a fresh attempt to estimate process capability analysis (PCA) using transformed data through different goodness of fit tests. The simulation study uses variety of asymmetric behaviors from a Weibull distribution generating a random sample of 100 data points to find the best goodness of fit test for better process capability estimates that are compared to the standard of six sigma results for non-normal data. Final result shows that Shapiro-Wilk's (SW) and Artificial Covariate (AC) methods are performing well when compared to the method of MLE. Minitab software and R programming language were utilized for data simulation and analysis.

**Keywords:** Goodness of fit tests, Box-Cox Transformation, Asymmetric, MLE, Weibull distribution, Six sigma.

## 1. Introduction

Process capability indices (PCIs), the statistical tools in quality control, are widely used to meet the required targets set in most of the manufacturing industries. Process capability analysis (PCA) addresses the issues relating to how well a manufacturing process meets the required specification. PCIs defined from normality assumptions cannot be used to accurately measure the performance of non-normal processes. Data transformation for preserving a somewhat normal distribution has been recommended in [5]. The empirical study made in [4] has demonstrated that the findings of transformed data are much superior to the results of the original data. The literature surveys demonstrate that for non-normal distributions such as Lognormal, Weibull, etc., the transformation methods perform well when compared to non-transformation (NT) methods and are considered as consistently superior to NT methods. Further,

NT methods are found to be inadequate in capturing the capability of the process unless the underlying distribution is close to or approximately normal. NT methods are unsatisfactory because the distribution deviates significantly from normal. See, [15].

In PCA, the process variation is defined based on the measure ‘standard deviation’. The short-term and long-term variability may be addressed by the estimated standard deviation obtained from the random sample observations and such an estimate is used while computing the process capability. The short-term variability is considered for computing the process capability indices whereas long-term variability is taken for calculating process performance indices. Hence, capability indices are calculated using samples of data based on short-term or within group variation, whereas performance indices are calculated using all the data points using long-term or overall variation. The process capability indices are denoted by  $C_p$  and  $C_{pk}$ , and process performance indices are denoted by  $P_p$  and  $P_{pk}$ . A detailed review on various methods that are chosen for performance comparison in their ability to handle non-normality in the computation of process capability indices is presented in [13]. The most common and traditional indices being applied by manufacturing industry are process capability index  $C_p$  and process capability ratio  $C_{pk}$  which are given below in Table 1 along with the respective performance indices, where  $\bar{x}$  is the sample mean, USL is the upper specification limit and LSL is the lower specification limit.

**Table 1:** Process Capability and Process Performance Indices

Process capability indices	Process performance indices
$C_p = \frac{USL - LSL}{6\sigma_w}$	$P_p = \frac{USL - LSL}{6\sigma_{overall}}$
$C_{pk} = \text{Min}(\text{CPU}, \text{CPL})$	$P_{pk} = \text{Min}(\text{PPU}, \text{PPL})$
$\text{CPU} = \frac{USL - \bar{x}}{3\sigma_w}, \text{CPL} = \frac{\bar{x} - LSL}{3\sigma_w}$	$\text{PPU} = \frac{USL - \bar{x}}{3\sigma_{overall}}, \text{PPL} = \frac{\bar{x} - LSL}{3\sigma_{overall}}$

According to [15], a better understanding is required about Box - Cox transformation (BCT) and its parameter estimation approach utilizing a search method to estimate the process capabilities. In [17], a method of converting non-normal data into normal data to analyze the data using the process capability indices and an improved Box-Cox transformation model have been proposed to deal with non-normal data and to calculate its process capability indices. In [1], the method of maximum likelihood estimation (MLE) was utilized for finding the ideal parameter  $\lambda$  in Box-Cox transformation. Alternative methods to MLE approach utilizing goodness of fit tests (normality tests) were developed in [3], [10] and [11]. By examining the effect of conversion of non-normal data into normal data with the use of different goodness of fit tests, it is demonstrated in [3] that the method of MLE in estimating the BCT parameter  $\lambda$  could be biased and ineffective. The competence of the different goodness of fit test was also determined in [3] by various measures of errors, estimates of PCI, PPI and defective parts per million (PPM) products.

In order to get improvised estimates of PCI and the result within the standard of six sigma level, a new attempt is made in this paper to estimate process capability analysis implementing different goodness of fit tests in BCT. The results of different goodness of fits tests are recorded and presented to help the practitioner to choose the method which will produce the improvised results in various asymmetric situations, *viz.*, low, moderate and high. Thus, the objectives of this paper is to examine the effectiveness of the different goodness of fit tests involving transformation of non-normal data into normal data using BCT and to recommend a superior test that will produce higher values of process capability with minimum of error and PPM values. It also verifies whether the proposed method produce the results within the standard of six sigma level.

## 2. Methodology

Transforming non-normal data into normal data is one of the frequently used approaches in practice when the observed data do not satisfy the normality assumption. A few approaches which are applied in practice to transform the non-normal data into normal include Johnson's system of transformation (JST), Box-Cox transformation (BCT) and Rosenblatt transformation (RT). Though JST and BCT approaches are equally efficient, the latter would be preferred over the first one for handling non-normal data when computer assisted analysis is available and it also outperforms the other methods. See, [12]. Further, when compared with the JST method, BCT method is more accurate and precise. BCT provides a family of power transformations that will optimally normalize a particular variable. As stated in [2], the BCT method transforms non-normal data into normal data on the positive response variable  $x$  as shown in the below expressions:

$$x^\lambda = \begin{cases} \frac{x^\lambda - 1}{\lambda}, & \text{for } \lambda \neq 0 \\ \log x, & \text{for } \lambda = 0 \end{cases} \quad (1)$$

It may be noted that since an analysis of variance is unchanged by a linear transformation, the expressions given (1) is equivalent to

$$x^\lambda = \begin{cases} x^\lambda, & \text{for } \lambda \neq 0 \\ \log x, & \text{for } \lambda = 0 \end{cases} \quad (2)$$

The estimation of  $\lambda$  is done through various goodness of tests for normality, that are available in the literature, which includes tests, such as Shapiro - Wilk (SW), Anderson Darling (AD), Cramer Von Mises (CVM), Pearson Chi-square (PC), Shapiro - Francia (SF), Lillefors (Kolmogorov - Simirnov) (LT / KS), Jarque - Bera (JB), and artificial covariate method (AC). The BCT approach given in [2] involves the method of maximum likelihood estimation (MLE). Two alternative approaches proposed in [10] and [11], respectively, considered Box - Cox power transformation using maximization of the Shapiro - Wilk  $W$  statistics which forces the data to get closer to normal as much as possible and Anderson - Darling test. In these approaches, Newton - Rapson algorithm has been used to obtain  $\lambda$ . A method is proposed in [3] to simulate a single artificial and non-informative covariate and to find  $\lambda$  minimizing the sum of squares of errors among several simple linear regression models.

The results of the earlier studies presented in the literature, particularly in [1], [7], [10], [14], [16] and [18], would be useful to understand the significance of tests of goodness of fit while transforming non-normal data into normal data. [10] Shows that the test based on SW statistic is a powerful test of normality for a variety of non-normal distributions, the SW statistic is reliable for small samples and in regression applications, the statistic would yield higher  $R^2$ . It is asserted in [7] that the test based on SW statistic is the most powerful test for non-normal distributions.

According to [14], JB test is preferable to the Shapiro-Wilk test when the data exhibit a symmetric distribution with medium or long tails, or a slightly skewed distribution with long tails. [18] Ascertained that the test based on SW statistic is the best one for asymmetric distributions and powerful for symmetric short tailed distributions and has good power qualities throughout a wide variety of asymmetric distributions. Based on the results of a simulation study provided in [1], it is found that all of the transforming approaches performed similarly to one another. One may refer to [9] and [19] for the details on the concepts of six-sigma tools and process capability analysis for non-normal data, respectively.

### 3. Weibull Distribution

Weibull distribution is applicable to a wide range of non-normal processes because it is capable of generating a variety of distinct curves based on its parameters. It exhibits a significant tail behavior, showing a significant effect on the capability of the process. It is frequently utilized in applications that focus on quality and reliability to analyze failure data and to comprehend how failures take place or how often products fail.

The probability density function of a Weibull random variable is given by the following form:

$$f(x) = \begin{cases} \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha}, & x \geq 0 \\ 0, & x < 0 \end{cases},$$

where  $\alpha > 0$  and  $\beta > 0$  are the shape and scale parameters, respectively.

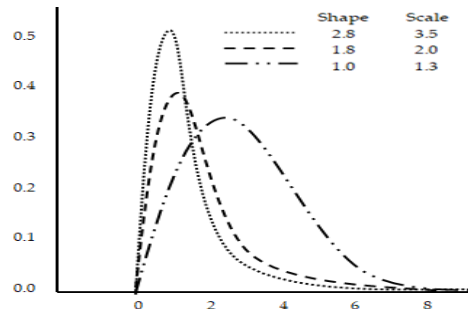
The mean, the variance and the measure of skweness of the Weibull distribution are, respectively, given as follows:

$$E(X) = \mu = \beta \Gamma(1+1/\alpha)$$

$$V(X) = \sigma^2 = \beta^2 [\Gamma(1+2/\alpha) - (\Gamma(1+1/\alpha))^2]$$

$$S_k = \gamma_1 = \frac{1}{\sigma^3} [\beta^3 \Gamma(1+3/\alpha) - 3\mu\sigma^2 - \mu^3]$$

The Weibull distribution with three sets of shape and scale parameters, say (2.8, 3.5), (1.8, 2.0), and (1.0, 1.3) is considered in [6]. The sets of parameters are categorized for the purpose of assessing the effectiveness of low, moderate, and high asymmetric behaviors during the transformation of non-normal data into normal data and carrying out the process capability analysis. The shapes of the density function of Weibull distribution for these sets of parameters are shown in Figure 1.



**Figure 1:** Asymmetric Behavior of Weibull Distribution

### 4. Numerical Illustrations

For a simulation set-up, the data set of size 100 is generated using different asymmetric levels of Weibull distribution. Minitab and R programming were utilized for data simulation and analysis purpose. As given in [6], the lower and upper specification limits are taken as 0.0 and 10. A combination of the box plot, descriptive statistics, measures of errors, like bias, percentage bias, median absolute error (MdAE), root mean square error (RMSE) and radar chart can be used to assess the effectiveness of the method.

This paper considers only the measures of errors and radar plots. In particular, bias, MdAE and RMSE are taken while transforming non-normal data into normal data using different goodness of fit tests in Box - Cox transformation. Once the transformation has been completed, the data have been further utilized to estimate process capability and process performance index and to choose the most effective approach among different goodness of fit tests. According to [8], a process is categorized as inadequate, if  $PCI < 1.00$ ; capable, if  $1.00 \leq PCI \leq 1.33$ ; satisfactory, if  $1.33 \leq PCI \leq 1.50$ ; excellent, if  $1.50 \leq PCI \leq 2.00$ ; and super, if  $\geq 2.00$ . Automotive industry uses  $C_{pk} = 1.33$  as a benchmark in assessing the capability of the process. If  $C_p$  and  $C_{pk}$  are more than or equal to 2 and 1.5, respectively, a process is said to be under six-sigma controls. Similarly,  $P_p$  and  $P_{pk}$  must be more than 2 and 1.5, respectively, for a process to generate six-sigma results. See, [8].

In order to guarantee the quality of the final product and reduce the number of faulty items, quality practitioners will also focus on PPM values. Table 2 lists the process fallout in defective parts per million products in relation to the proportion of good items and PPM values for various sigma levels. The main goal of all quality and industry practitioners is to reach  $6\sigma$  limits and a defect rate of 3.4 PPM has been associated with the process using these indices. On the other hand, the process performance indices, namely  $P_p$  and  $P_{pk}$  are utilized in the industries, particularly in the automobile sector, as the second sorts of estimators.

**Table 2:** Process Fallout in Defective Parts per Million with Respect to Different Sigma Levels

Sigma Level	Percentage	PPM Values
6	99.9997%	3.4
5	99.98%	233
4	99.4%	6,210
3	93.3%	66,807
2	69.1%	308,537
1	30.9%	691,462

**4.1 Low Asymmetric Distribution**

In this sub-section, low asymmetric Weibull distribution with the skewness of 0.13 and 0.31 for the combination of shape and scale parameters 2.8 and 3.5, respectively, has been taken for simulation study. From the error point of view, Bias, MdAE and RMSE values are very less for AD, CVM, SF, LT and PC goodness of fit tests and this ensures that the transformed values are very closer to normal data with minimum error values. For more information, Table 3 and Figure 2 may be referred. On the other hand, from estimation point of view, the transformed data are further taken for the estimation of process capability and process performance. The transformed data sets from SW, LT, AC, and MLE tests show the closeness to the standard normal and produce better results when compared to other methods. The PPM values are recorded as a minimum of 656 and a maximum of 1939 corresponding to the above said methods and are better than the results of  $3\sigma$  and  $4\sigma$  limits and closer to the result of  $5\sigma$  standards. For more information, Table 4 and 5 may be referred.

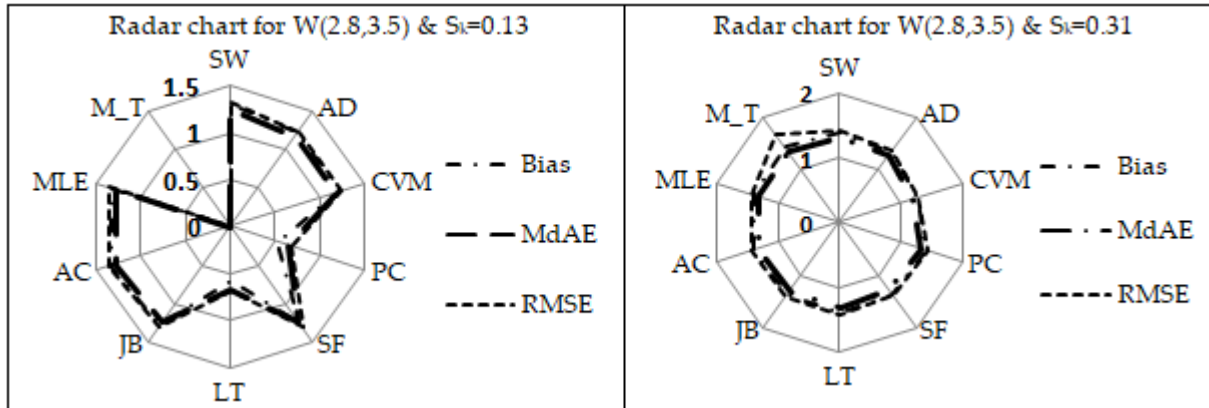
**4.2 Moderate Asymmetric Distribution**

A Weibull distribution with the shape and scale parameters fixed as 1.80 and 2.0, respectively, will represent the moderate asymmetrical non-normal data with skewness 0.64 and 0.94. In the simulation study, Minitab (M\_T) transforms non-normal data into much closer normal data with minimum Bias, MdAE and RMSE values compared to other methods and the corresponding estimate of PC is smaller but with higher PPM values compared to the benchmark result. Thus, the method of transformation using Minitab cannot be taken as a competent method. One may refer to Table 6 and Figure 3.

**Table 3:** Various Measures of Error Values for Low Asymmetric Data After Data Transformation

Methods	Low Asymmetry (SK=0.13)			Low Asymmetry (SK=0.31)		
	Weibull distribution ( $\alpha=2.8, \beta=3.5$ )			Weibull distribution ( $\alpha=2.8, \beta=3.5$ )		
	Bias	MdAE	RMSE	Bias	MdAE	RMSE
SW	1.300	1.245	1.322	1.391	1.320	1.428
AD	1.226	1.184	1.240	1.335	1.273	1.363
CVM	1.226	1.184	1.240	1.246	1.200	1.263
PC	0.527	0.646	0.663	1.391	1.320	1.428
SF	1.271	1.221	1.289	1.363	1.297	1.396
LT	0.571	0.665	0.677	1.391	1.320	1.428
JB	1.285	1.233	1.306	1.377	1.309	1.412
AC	1.343	1.281	1.371	1.392	1.321	1.429
MLE	1.342	1.280	1.370	1.391	1.320	1.428
M_T	*	*	*	1.434	1.345	1.706

\* Transformation not done



**Figure 2:** Radar Chart for Various Measures of Errors After Normalization of Low Asymmetric Distribution

**Table 4:** Estimates of Process Capability and Process Performance Indices for W(2.8, 3.5) Distribution Having  $S_k = 0.13$  After Normalization via Goodness of Fit Tests

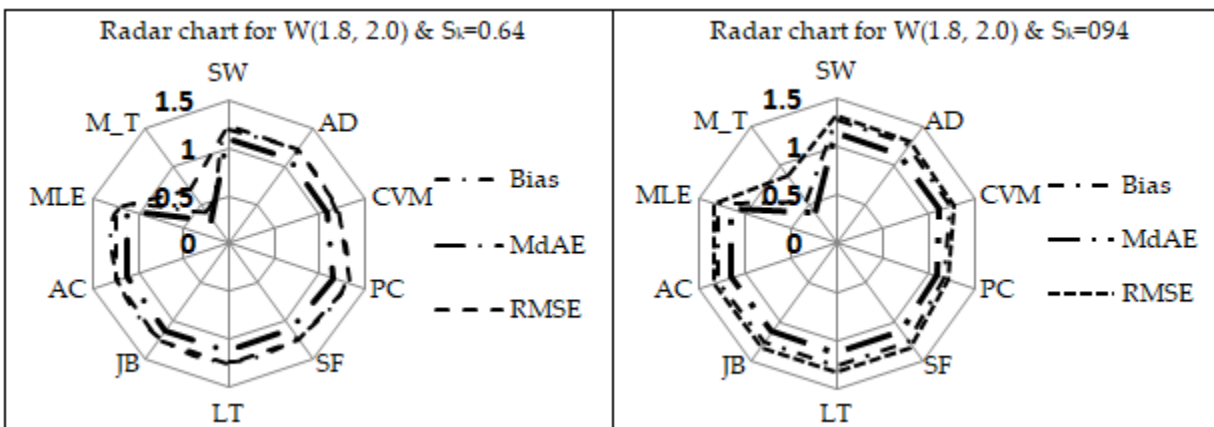
Method	$\lambda$ Value	LSL	USL	PCI (Within Capability)			PPI (Overall Capability)		
				Cp	Cpk	PPM	Pp	Ppk	PPM
W(2.8, 3.5)	-	0	10	1.30	0.82	6828	1.27	0.81	7667
SW	0.75	-1.33	6.16	1.29	1.07	656	1.25	1.04	904
AD	0.79	-1.27	6.54	1.29	1.02	1066	1.25	1.00	1402
CVM	0.85	-1.18	7.15	1.28	0.96	2051	1.25	0.93	2543
PC	0.75	-1.33	6.16	1.29	1.07	656	1.25	1.04	904
SF	0.77	-1.30	6.35	1.29	1.05	841	1.25	1.02	1130
LT	0.75	-1.33	6.16	1.29	1.07	656	1.25	1.04	904
JB	0.76	-1.32	6.26	1.29	1.06	731	1.25	1.03	995
AC	0.75	-1.33	6.16	1.29	1.07	656	1.25	1.04	904
MLE	0.75	-1.33	6.16	1.29	1.07	656	1.25	1.04	904
M_T	0.50	0.00	3.16	1.42	1.28	66	1.36	1.22	127

**Table 5:** Estimates of Process Capability and Process Performance Indices for W(2.8, 3.5) Distribution Having  $S_k = 0.31$  After Normalization via Goodness of Fit Tests

Method	$\lambda$ Value	LSL	USL	PCI (Within Capability)			PPI (Overall Capability)		
				Cp	Cpk	PPM	Pp	Ppk	PPM
W(2.8, 3.5)	-	0	10	1.27	0.80	8026	1.32	0.83	6362
SW	0.81	-1.23	6.74	1.24	0.96	1939	1.29	1.00	1389
AD	0.86	-1.16	7.26	1.25	0.91	3051	1.29	0.95	2259
CVM	0.86	-1.16	7.26	1.25	0.91	3051	1.29	0.95	2259
PC	1.24	-0.81	13.21	1.36	0.67	22553	1.41	0.69	19197
SF	0.83	-1.20	6.94	1.24	0.94	2351	1.29	0.98	1708
LT	1.22	-0.82	12.78	1.35	0.86	21189	1.40	0.70	17959
JB	0.82	-1.22	6.84	1.24	0.95	2106	1.29	0.99	1518
AC	0.78	-1.28	6.44	1.24	1.00	1402	1.29	1.03	981
MLE	0.78	-1.28	6.44	1.24	1.00	1407	1.29	1.03	985
M_T	-	0	10	1.27	0.80	8026	1.32	0.83	6362

**Table 6:** Various Measures of Error Values for Moderate Asymmetric Data After Data Transformation

Methods	Moderate Asymmetry (SK=0.64)			Moderate Asymmetry (SK=0.94)		
	Weibull distribution ( $\alpha=1.8, \beta=2.0$ )			Weibull distribution ( $\alpha=1.8, \beta=2.0$ )		
	Bias	MdAE	RMSE	Bias	MdAE	RMSE
SW	1.204	1.108	1.231	1.271	1.137	1.321
AD	1.195	1.102	1.219	1.255	1.127	1.301
CVM	1.175	1.090	1.195	1.247	1.122	1.290
PC	1.282	1.156	1.326	1.192	1.091	1.221
SF	1.201	1.106	1.227	1.271	1.137	1.321
LT	1.223	1.118	1.253	1.271	1.137	1.321
JB	1.211	1.111	1.238	1.271	1.137	1.321
AC	1.207	1.110	1.234	1.282	1.143	1.335
MLE	1.207	1.110	1.234	1.283	1.143	1.336
M_T	0.420	0.304	0.703	0.524	0.383	0.863



**Figure 3:** Radar Chart for Various Measures of Errors After Normalization of Moderate Asymmetric Distribution

Besides M\_T transformation, the CVM, AD, AF, AC and SW methods of transformation produce less errors and the PC, LT, JB, AC, MLE and SW methods of transformation yield the target results during the estimation of process capability and process performance indices along with the minimum PPM values. For the moderate asymmetric situations, the minimum and maximum PPM values were recorded as 81 and 241, respectively. The goodness of fit tests in the estimation of process capability for moderate asymmetric distribution shows the better results than  $3\sigma$ ,  $4\sigma$  and  $5\sigma$  limits and approach towards the standard of  $6\sigma$ . One may also refer to Table 7 and 8 for more information.

**Table 7:** Estimates of Process Capability and Process Performance Indices for W(1.8, 2.0) Distribution Having  $S_k = 0.64$  After Normalization via Goodness of Fit Tests

Method	$\lambda$ Value	LSL	USL	PCI (Within Capability)			PPI (Overall Capability)		
				Cp	Cpk	PPM	Pp	Ppk	PPM
W(1.8, 2.0)	-	0	10	1.79	0.59	37568	1.80	0.60	36938
SW	0.45	-2.22	4.04	1.44	1.23	110	1.44	1.23	114
AD	0.48	-2.08	4.21	1.44	1.16	252	1.43	1.16	259
CVM	0.54	-1.85	4.57	1.43	1.04	900	1.43	1.04	914
PC	0.19	-5.26	2.89	2.02	1.25	92	2.01	1.24	99
SF	0.46	-2.17	4.10	1.44	1.21	149	1.44	1.20	154
LT	0.39	-2.56	3.73	1.48	1.41	14	1.48	1.40	15
JB	0.43	-2.33	3.93	1.45	1.29	56	1.45	1.28	59
AC	0.44	-2.27	3.99	1.45	1.26	81	1.45	1.25	84
MLE	0.44	-2.27	3.99	1.45	1.26	81	1.45	1.25	84
M_T	0.50	0	3.16	1.43	1.12	398	1.43	1.12	408

**Table 8:** Estimates of Process Capability and Process Performance Indices for W(1.8, 2.0) Distribution Having  $S_k = 0.94$  After Normalization via Goodness of Fit Tests

Method	$\lambda$ Value	LSL	USL	PCI (Within Capability)			PPI (Overall Capability)		
				Cp	Cpk	PPM	Pp	Ppk	PPM
W(1.8, 2.0)	-	0	10	1.50	0.54	51629	1.54	0.56	47940
SW	0.43	-2.33	3.93	1.28	1.17	241	1.32	1.21	151
AD	0.47	-2.13	4.15	1.26	1.08	623	1.30	1.11	428
CVM	0.49	-2.04	4.27	1.26	1.04	949	1.30	1.07	674
PC	0.62	-1.61	5.11	1.26	0.84	6101	1.30	0.86	4922
SF	0.43	-2.33	3.93	1.28	1.17	241	1.32	1.21	154
LT	0.43	-2.33	3.93	1.28	1.17	241	1.32	1.21	154
JB	0.43	-2.33	3.93	1.28	1.17	241	1.32	1.21	154
AC	0.40	-2.50	3.78	1.30	1.25	118	1.34	1.29	70
MLE	0.40	-2.50	3.78	1.30	1.25	118	1.34	1.29	70
M_T	0.50	0	3.16	1.26	1.02	1143	1.30	1.05	822

### 4.3. High Asymmetric Distribution

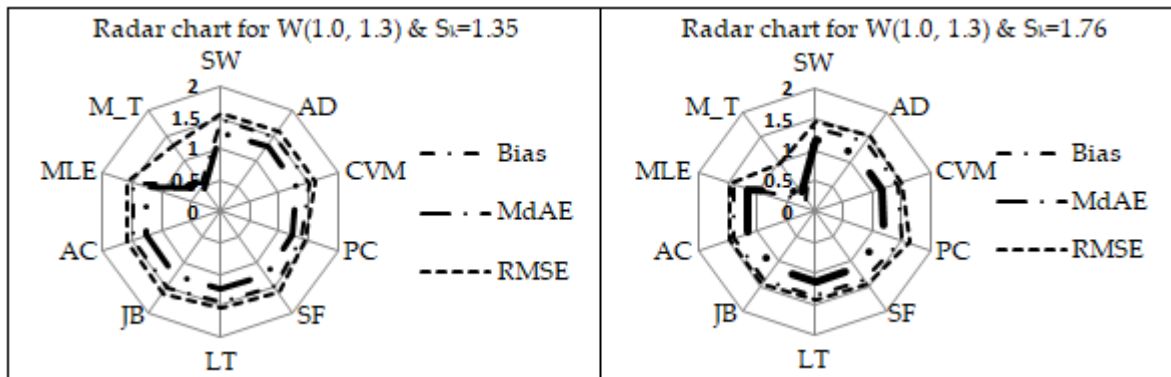
A Weibull distribution with the shape and scale parameters fixed as 1.0 and 1.3, respectively, will represent the high asymmetrical non-normal data with skewness 1.35 and 1.76. Among the different methods, Minitab (M\_T) transforms non-normal data into much closer normal data with minimum Bias, MdAE and RMSE values when compared to other methods, but the corresponding estimate of PCA



shows smaller and more PPM values compared to the standard requirements. Therefore, the method of transformation using Minitab (M\_T) cannot be taken as an effective method. One may refer to Table 9 and Figure 4 for more information. From the point of view of errors, after transforming non normal data into normal data using different goodness of fit tests, the LT, SF, AC and SW, PC and AD methods produce fewer errors. Moreover, the methods such as AC, JB, SW, AD and MLE yield better estimates of process capability and process performance along with lesser PPM values. In this case, the minimum and maximum PPM values are recorded as 740 and 3075, respectively. The goodness of fit tests in the estimation of process capability for moderate asymmetric distribution shows that the process is better than  $3\sigma$  and  $4\sigma$  and approach towards the standard of  $5\sigma$ . One may refer to Table 10 and 11 for more information.

**Table 9:** Various Measures of Error Values for High Symmetric Data After Data Transformation

Methods	High Asymmetry (SK = 1.35)			High Asymmetry (SK = 1.76)		
	Weibull distribution ( $\alpha = 1.0, \beta=1.3$ )			Weibull distribution ( $\alpha=1.0, \beta=1.3$ )		
	Bias	MdAE	RMSE	Bias	MdAE	RMSE
SW	1.473	1.261	1.584	1.382	1.165	1.474
AD	1.480	1.265	1.593	1.414	1.174	1.519
CVM	1.486	1.269	1.602	1.414	1.174	1.519
PC	1.363	1.196	1.442	1.490	1.198	1.641
SF	1.466	1.257	1.576	1.376	1.163	1.465
LT	1.440	1.241	1.542	1.364	1.159	1.448
JB	1.493	1.273	1.611	1.382	1.165	1.474
AC	1.479	1.265	1.593	1.382	1.164	1.472
MLE	1.480	1.265	1.593	1.382	1.165	1.474
M_T	0.536	0.466	1.308	0.237	0.369	0.966



**Figure 4:** Radar Chart for Various Measures of Errors after Normalization of High Asymmetric Distribution

### 5. Results and Discussion

Data transformation and estimation of process capability analysis are the two aspects considered in this section. The effectiveness of different goodness of fit tests is determined by various measures of errors such as Bias, MdAE and RMSE. Based on the numerical illustrations provided in the previous section, it is found that the methods of AD and CVM tests produce lesser errors in low and moderate asymmetric situations, the methods of SW and SF tests yield considerably lesser errors in the case of moderate and high asymmetric behaviors, and the methods of LT and AC tests perform better only on high asymmetric situations. Similarly, the methods of PC, LT, JB, DME, and M\_T tests yield better estimates, but provide

greater PPM values while estimating process capability and process performance indices.

**Table 10:** Estimates of Process Capability and Process Performance Indices for W(1.0, 1.3) Distribution Having  $S_k = 1.35$  After Normalization via Goodness of Fit Tests

Method	$\lambda$ Value	LSL	USL	PCI (Within Capability)			PPI (Overall Capability)		
				Cp	Cpk	PPM	Pp	Ppk	PPM
W(1.0, 1.3)	-	0	10	1.42	0.34	156902	1.39	0.33	160815
SW	0.26	-3.85	3.15	1.14	1.09	744	1.13	1.08	821
AD	0.21	-4.76	2.96	1.22	1.01	1248	1.21	1.00	1316
CVM	0.21	-4.76	2.96	1.22	1.01	1248	1.21	1.00	1316
PC	0.1	-10.0	2.59	1.82	0.84	6000	1.83	0.84	5871
SF	0.27	-3.70	3.19	1.12	1.10	770	1.11	1.09	856
LT	0.29	-3.45	3.28	1.11	1.07	956	1.09	1.06	1071
JB	0.26	-3.85	3.15	1.14	1.09	744	1.13	1.08	821
AC	0.26	-3.85	3.15	1.14	1.09	740	1.13	1.08	817
MLE	0.26	-3.85	3.15	1.14	1.09	744	1.13	1.08	821
M_T	0.28	0	1.90	1.11	1.11	834	1.10	1.10	932

**Table 11:** Estimates of Process Capability and Process Performance Indices for W(1.0, 1.3) Distribution Having  $S_k = 1.76$  After Normalization via Goodness of Fit Tests

Method	$\lambda$ Value	LSL	USL	PCI (Within Capability)			PPI (Overall Capability)		
				Cp	Cpk	PPM	Pp	Ppk	PPM
W(1.0, 1.3)	-	0	10	1.12	0.35	148540	1.15	0.36	142686
SW	0.29	-3.45	3.28	1.00	0.95	3033	0.99	0.94	3397
AD	0.28	-3.57	3.23	1.01	0.94	3075	0.99	0.93	3459
CVM	0.27	-3.70	3.19	1.02	0.93	3136	1.01	0.91	3539
PC	0.46	-2.17	4.10	0.92	0.69	19756	0.92	0.69	19840
SF	0.30	-3.33	3.32	0.99	0.96	3173	0.98	0.95	3533
LT	0.34	-2.94	3.49	0.95	0.90	4652	0.95	0.90	5012
JB	0.26	-3.85	3.15	1.04	0.92	3265	1.02	0.90	3694
AC	0.28	-3.57	3.23	1.01	0.94	3075	0.99	0.93	3458
MLE	0.28	-3.57	3.23	1.01	0.94	3075	0.99	0.93	3458
M_T	0.24	0.00	1.74	1.06	0.90	3639	1.05	0.88	4132

Thus, as a result, it will not be thought of as a useful way to evaluate the capability or a performance of the process, though the methods of SW, AC, SF and MLE tests produce superior results with better estimates and lesser PPM values when compared to other and traditional methods. A small PPM value generally assures that fewer items will be rejected, and it must be lower than the benchmark values to obtain six sigma results. On the basis of the numerical illustrations, it can be observed that the different tests of goodness of fit would guarantee better performance (656 as the minimum and 1939 as the maximum PPM values) in comparison to the typical PPM values of the  $3\sigma$  and  $4\sigma$  limits, and are very close to the outcome of the  $5\sigma$  limits only in low asymmetric behaviors.

The PPM values for moderately asymmetric conditions are found to be 81 and 241 as minimum and maximum values, respectively. This outcome surpasses the  $3\sigma$ ,  $4\sigma$ , and  $5\sigma$  limits and is getting closer to

the benchmark of  $6\sigma$  outcomes. The minimum and maximum PPM values of 740 and 3075 would ensure that the procedure is better than the  $3\sigma$  and  $4\sigma$  limits only under high asymmetrical circumstances. One may refer to Table 12 for the better understanding of the efficiency of different normality tests under various asymmetric behaviors while dealing with non-normal quality characteristics based on the numerical examples, results and discussion.

**Table 12:** Efficiency of Various Tests of Goodness of Fit in Data Transformation and Estimation of Process Capability and Process Performance Indices for Weibull Distribution

Different Asymmetric Levels	Efficiency in data transformation						Efficiency in estimation of PCI/PPI					
	Low Asymmetric		Moderate Asymmetric		High Asymmetric		Low Asymmetric		Moderate Asymmetric		High Asymmetric	
Skewness	0.13	0.31	0.64	0.94	1.35	1.76	0.13	0.31	0.64	0.94	1.35	1.76
SW			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AD	✓	✓	✓	✓								✓
CVM	✓	✓	✓	✓								
PC	✓			✓	✓			✓*	✓*		✓*	
SF		✓	✓	✓	✓	✓				✓	✓	
LT	✓	✓			✓	✓	✓	✓*	✓	✓		
JB						✓			✓	✓	✓	✓*
AC					✓	✓	✓	✓	✓	✓	✓	✓
MLE							✓	✓	✓	✓	✓	✓
M_T	@		✓	✓			✓	✓*				✓*
DME							✓\$	✓\$	✓\$	✓\$	✓\$	✓\$

DME – Direct Minitab Estimation | @ - No transformation done | ✓ - less errors and/or better estimates and less PPM values | ✓\* - Produces less error but higher PPM values | ✓\$ - Produces Better estimates but higher PPM values.

## 6. Conclusion

Process capability analysis is important for any production process and useful for its continuous improvement. This study attempts to compare the ability of various tests of goodness of fit over the method of maximum likelihood in the estimation of the parameter involved in Box - Cox transformation. Primarily, the effectiveness of the tests of goodness of fit in transforming non-normal data into normal data is assessed through various measures of errors along with a radar chart. Based on the numerical example, the solutions to the research problem are turned out and it is observed that, regardless of using different formulas, the estimates of process capability and process performance indices approximately match. It is to be noted that the performance of process capability analysis for non-normal data purely depends on the choices of variation taken into account. Further, the transformed data is extended towards estimating process capability and process performance in order to identify the effective methods for non-normal quality characteristics. As per the results and discussion, one may observe that the measures of errors, and estimates of PCI, PPI and PPM values from SW, AC, SF and MLE methods of goodness of fit tests have higher accuracy in data transformation, greater power in estimating process capability or process performance and leaves smaller PPM values in all asymmetric situations.

By taking into account of the research problem, the SW test outperforms the other tests while

transforming non-normal data into normal data and estimating process capability / performance with smaller PPM values in all the asymmetric situations. However, other methods of tests such as AC and MLE methods can also be considered for handling non-normal quality characteristics and producing considerably good results. Application of different goodness of fit tests to estimate PCA yields smaller PPM values and obviously better results than  $3\sigma$ ,  $4\sigma$  and  $5\sigma$  limits. Implementing goodness of fit tests further helps to obtain the results that are closer to the six sigma standards than the traditional MLE method. Thus, the current MLE technique could be effectively substituted by using goodness of fits tests in Box-Cox transformation to achieve desired results in estimating process capability.

## References

- [1] Asar, O., Ilk, O., and Dag, O. (2017). Estimating Box-Cox Power Transformation Parameter via Goodness-of-Fit Tests, *Communications in Statistics - Simulation and Computation*, **46**, 91 – 105.
- [2] Box, G. E. P., and Cox, D. R. (1964). An analysis of Transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, **26**, 211 - 243.
- [3] Dag O., Asar, O., and Ilk, O. (2014). A Methodology to Implement Box-Cox Transformation When No Covariate is Available, *Communications in Statistics – Simulation and Computation*, **43**, 1740 – 1759.
- [4] Gunter, B. H. (1989). The Use and Abuse of Cpk, *Quality Progress*, **22**, 108 – 109.
- [5] Kane V E (1986), Process Capability Indices, *Journal of Quality Technology*, **18**, 41 – 52.
- [6] Kashif, M., Aslam, M., Al Marshadi, A. H., Jun, C-H. (2017), Evaluation of Modified Non-normal Process Capability Index and Its Bootstrap Confidence Intervals, *IEEE Access*, **5**, 12135 – 12142.
- [7] Oztuna, D., Elhan, A., and Tuccar, E. (2006). Investigation of Four Different Normality Tests In Terms of Type 1 Error Rate and Power under Different Distributions, *Turkish Journal of Medical Sciences*, **36**, 171 - 176.
- [8] Pearn, W. L., and Chen, K. -S. (2002). One-sided Capability Indices CPU and CPL: Decision Making with Sample Information, *International Journal of Quality & Reliability Management*, **19**, 221 – 245.
- [9] Pyzdek, T. (2003). *Pyzdek Six Sigma Handbook: A Complete Guide for Green Belts, Black Belts, and Managers at All Levels*, McGraw-Hill Inc., New York.
- [10] Rahman, M. (1999). Estimating the Box-Cox Transformation via Shapiro-Wilk W Statistic, *Communications in Statistics – Simulation and Computation*, **28**, 223 – 241.
- [11] Rahman, M., and Pearson, L. M. (2008). Anderson-Darling statistic in Estimating the Box-Cox Transformation Parameter, *Journal of Applied Probability & Statistics*, **3**, 23 – 35.
- [12] Sennaroglu, B., and Senvar, O. (2015). Performance Comparison of Box-Cox Transformation and Weighted Variance Methods with Weibull Distribution, *Journal of Aeronautics and Space Technologies*, **8**, 49 - 55.
- [13] Tang, L. C., Than, S. E. (1999). Computing Process Capability Indices for Non-normal Data: A Review and Comparative Study, *Quality and Reliability Engineering International*, **15**, 339 – 353.
- [14] Thadewald, T., and Buning, H. (2007). Jarque-Bera Test and its Competitors for Testing Normality - A Power Comparison, *Journal of Applied Statistics*, **34**, 87 - 105.
- [15] Swamy, D. R., Nagesh, P., and Wooluru, Y. (2016). Process Capability Indices for Non-normal Distribution – A Review, *Proceedings of the International Conference on Operations Research and Management*, January 21 – 22, Mysuru, India.
- [16] Wooluru, Y., Swamy, D. R., and Nagesh, P. (2016). Process Capability Estimation for Non-normally Distributed Data using Robust Methods – A Comparative Study, *International Journal of Quality Research*, **10**, 407 – 420.

- [17] Yang Y and Zhu H (2018). A Study on Non-normal Process Capability Analysis based on Box-Cox Transformation, *Proceedings of the 3rd International Conference on Computational Intelligence and Applications (ICCIA)*, Hong Kong, China, IEEE, 240 – 243.
- [18] Yap, B. W., and Sim, C. H. (2011). Comparisons of Various Types of Normality Tests, *Journal of Statistical Computation and Simulation*, **18**, 2141 – 2155.
- [19] Yoap, T. (2006). Process Capability Analysis for Non-normal Data with Minitab, *In: Six Sigma: Advances Tools for Black Belts and Master Black Belts*, Eds. Tang, L. C., Goh, T. N., Yam, H. S., & Yoap, T, 131 – 149, John Wiley & Sons Ltd., The Atrian, England.

# THE USE OF EXPERIMENTAL MODELLING IN THE PREDICTION OF PRODUCT RELIABILITY

Alena Breznická<sup>1</sup> Pavol Mikuš<sup>2</sup>

•

Faculty of Special Technology, Alexander Dubček University of Trenčín,  
Ku kyselke 469, 911 06, Trenčín, Slovakia<sup>1,2</sup>

[alena.breznicka@tuni.sk](mailto:alena.breznicka@tuni.sk)

[pavol.mikus@tuni.sk](mailto:pavol.mikus@tuni.sk)

## Abstract

*When designing new systems and components, it is very important to correctly determine the degree and ability of the joint to withstand stress and load. Every new product that is intended for the market must meet the requirements for high safety and reliability during the entire life cycle. The presented article deals with the possibility of modelling the ability to withstand such a load, the principle of the interference method was used in the experimental modelling. The interference theory of reliability is based on the analysis of regularities and properties of two random variables that characterize reliability. Among these elementary properties from the point of view of reliability assessment, we can successfully use dependability and lifetime analysis. It originates from the concept of "safe life", which is deterministic, based on determining and respecting the values of reliability factors. The described approach assumes that a malfunction or a faulty function occurs when the strength limit of the object is exceeded, i.e., ability to withstand stress.*

**Keywords:** Reliability, Interference theory, Dependability, Load, Strength

## 1. Introduction

The interference theory of reliability is based on the analysis of the regularities and properties of two random variables that characterize the elementary properties of dependability and lifetime. Interference reliability theory offers reliability prediction in new product design because it can simulate the various loads and stresses that are applied to the product during its life cycle. The method uses the assessment of reliability properties in interesting interactions, which ultimately affect the resulting reliability. Such analyses are important precisely in the first stages of the product life cycle and are therefore successfully included in the process of creation and production of parts. The basic step of the analysis is the observation of two random variables, which we will describe in the following text.

Distribution of random variables

- The first random variable characterizes the operating mode and the resulting operating stress  $L$  (Load stress). Operating stress is caused by the sum of external stress and the conditions of the selected modes of use.

- The second random variable quantifies the strength S (carrying capacity). Strength to load S (Strength) is the ability to withstand physical, or chemical and biological loads, which, because of their action, result in changes causing element failures.

Both parameters of the model are random variables, characterized by random variables or processes. The form of their expression can be expressed by a histogram, or after statistical processing by probability distribution functions [1]. The literature presents many models of analytical quantification of dependability interference for the cases of exponential, normal, Weibull or gamma and log-normal distributions of load probability densities  $f_L(L)$  and strength  $f_S(S)$  [2].

For the combination of different distributions of load and strength, the method of calculating the integrals of the two-dimensional joint function is complex, and the calculation of fault-freeness is difficult. Then it is advantageous to use mathematical or simulation modelling [3]. Today, the reliability of products is successfully predicted already when designing new systems and can effectively use mathematical modelling and simulation. From the point of view of partial reliability properties, in the presented article, the authors will focus on the prediction of dependability modelling. Therefore, we will deal with the calculation method of the interference theory of reliability in the present paper.

## 2. Definition of the model

The assessed system or object  $M_k$ , which is exposed to the load during the monitored time, will be reliable if the given operating stress  $L$  together with a certain probability does not exceed the strength  $S$ .

$$M_k = \Pr(S > L) \tag{1}$$

Where:

$M_k$ ... System or object,

$S$ ...Stress [%/MPa],

$L$ ...Load Stress [%/MPa].

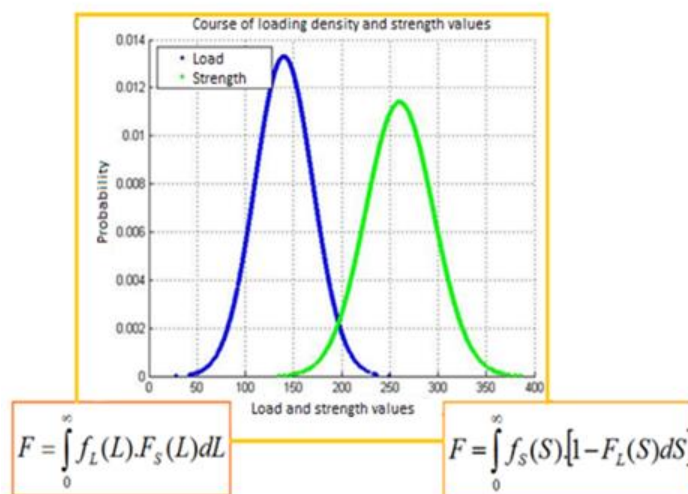


Figure 1: Representation of the range of interference

This area (Figure.1) is proportional to the probability of a malfunction. It expresses the fact that due to the random properties of quantities (mainly their dispersion) there is a certain degree of possibility - the probability that a state will occur when the stress will be greater than the resistance to failure in each case. As a result, a malfunction will occur. When calculating the probability of trouble-free operation  $R$ , assuming that the random variables  $L$ ,  $S$  are independent of each other, we can start from the well-known fact that the probability of the simultaneous occurrence of two independent phenomena is equal to the product of their probabilities.

The quantities characterizing the operating stress  $L$  (load) and the strength of the structure  $S$  are expressed by distribution functions and probability densities. Let us denote the probability density for the random stress variable  $L$  by  $f_L(L)$  and the probability density for the random variable  $S$  against failure by  $f_S(S)$ . Let's denote the distribution function for the random stress variable  $L$  by  $F_L(L)$  and the distribution function for the random stress strength variable  $S$  by  $F_S(S)$ . The quantities  $L$  and  $S$  are random, they have a specific probability distribution law, most often continuous or discrete. They can influence each other, which means to interfere, and this property can therefore be successfully used when assessing the reliability of a technical system or object in general. The extreme points of penetration, which arise during the analysis itself, define the area of mutual influence of both quantities his area is proportional to the probability of a malfunction [4]. The overlapping area defines the area of mutual influence of both quantities. It is proportional to the probability of failure and expresses the fact that due to the random properties of quantities (primarily their dispersion) there is a certain degree (probability) of the possibility that a state will occur when the stress will be greater than the strength to failure in each case and as a result a failure will occur. The area expresses the fact that due to the random properties of the quantities (especially their dispersion) there is a certain degree of possibility - the probability that a state will occur where the stress will be greater than the strength to failure in the given case. As a result, a malfunction will occur [5]. The curves are shown in Fig. 1. When calculating the probability of trouble-free operation  $R$ , we can assume that the random variables  $L$ ,  $S$  are independent of each other, based on the known fact that the probability of the simultaneous occurrence of two independent phenomena is equal to the product of their probabilities. In accordance with the introduction of labels for the probability densities of quantities  $L$  and  $S$ , the following will apply to the probability of dependability operation  $R$ :

$$F = \int_0^{\infty} f_L(L) \cdot F_S(L) dL \quad (2)$$

Or

$$F = \int_0^{\infty} f_S(S) \cdot [1 - F_L(S)] dS \quad (3)$$

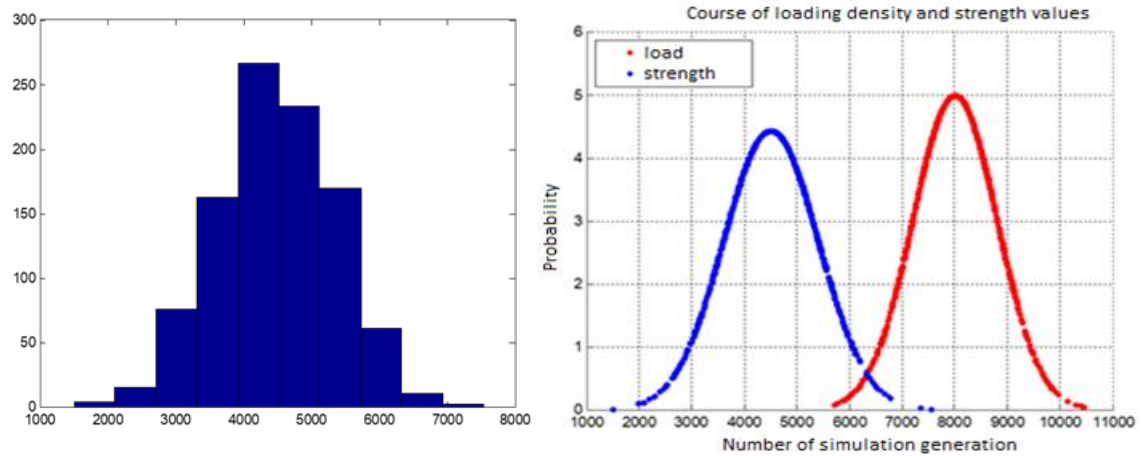
The mentioned relationships are the methodological basis for modelling the failure rate or failure-freeness of elements using the SSI interference method [6,7].

### 3. SSI simulation model

The input quantities of load  $L$  and strength  $S$  have a random character obtained from experimental measurement. The result is a non-parametric distribution of the obtained data, which we can statistically process in the form of a histogram or convert to a usable parametric distribution, as illustrated in Fig. 2. [8]. Both cases provide us with the possibility of generating input quantities and

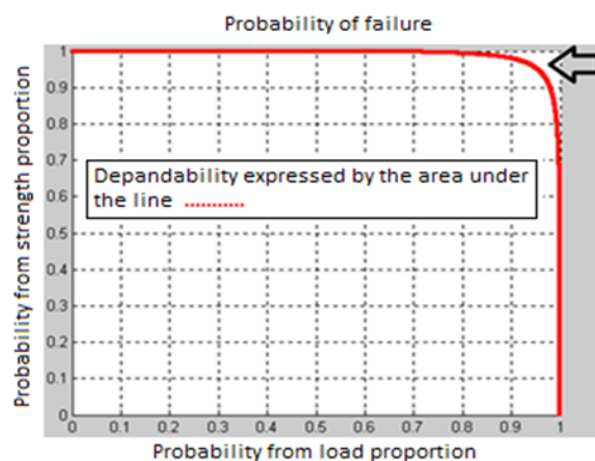


assessing the occurrence of decisive events for the statistical expression of failure rate or failure-freeness of elements using the interference method.



**Figure 2:** Expression of random variables  $L$  and  $S$  by histogram of relative abundances and probability distribution density

For the range of experimental or generated values, we determine the size of the values of the distribution functions  $F_L(S)$  and  $F_S(S)$  for different strength values  $S$ . For the range of values of both functions, stress  $L$  and  $S$  contribute to failure. If we plot the values of  $L$  and  $S$  in the interdependence graph, the intersection represents the product of two independent phenomena. The area below the line of the graph represents the probability of fault-free operation expressed, and the area above the line represents the probability of the occurrence of a fault. A probability distribution model is characterized by a density function and a distribution function based on precisely specified parameters that need to be estimated from the data using a likelihood function. We also test hypotheses in statistical models, which often represent models of causal dependence of dependent variables on predictors. In the experiment, graphic tools are used and serve for a quick and illustrative presentation of the results, especially when it comes to more comprehensive data and mutual comparison of several files. By graphically representing the frequency distribution, we get a clear idea of the nature of the frequency distribution of the observed character.



**Figure 3:** Probability of failure

The modelling procedure was designed as follows. The first step is to obtain the input data of histogram parameters, or the probability distribution of operating load  $L$  and strength  $S$  of the investigated system element. Subsequently, a random level of operating load and strength is generated, thus creating a point of realization of the phenomenon. The next analysis will assess

which area it falls into and show it graphically. The boundary between fault-free and fault-free areas is given by the condition  $S \geq L$  and expressed by the red line in the fig. 3. If the condition  $S \geq L$  is not met, this is a fault condition. We record the number of simulation steps  $N$  and the number of failure states  $n$ . We will statistically process the generated data into the form of values of probability density functions and distribution functions, and by plotting them we will get an idea of their interference. In the proposed procedure, we will successfully use the MATLAB simulation language, because its graphics allow the creation of interactive programs, the environment of which allows the user to dialogically change the parameters of the distributions and judge what load and strength values are acceptable for the structural design application. The program in the basic window offers the option of choosing the type of load distribution and strength of the investigated element, distribution parameters and the number of simulations. If, from the input data used, the simulation results indicate that the required fault-free parameters do not meet, the simulations can be carried out by changing the load and strength parameters until an acceptable level of interference is reached.

#### 4. Steps of the experimental simulation model

The first steps of the analysis require the determination of the number of simulations and the loading of the necessary input data of the parameters of the distributions of the probability density functions of the operating load and the strength of the investigated element. The verification analysis is shown in fig.4.

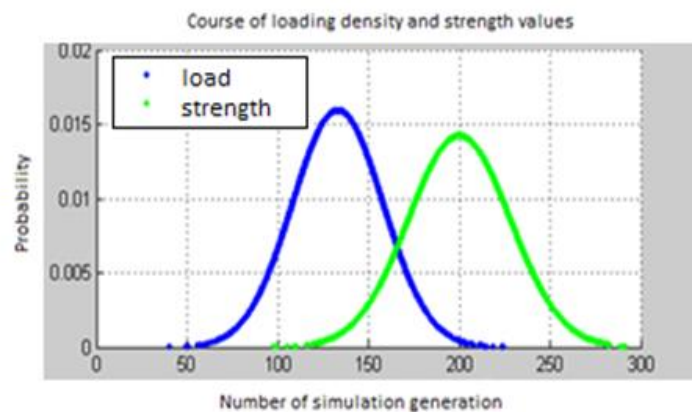


Figure 4: Determination of the number of simulations in the mathematical model

The level of operational load and strength is generated and statistically processed into values of probability density functions and distribution functions. The curves of the distribution functions are shown in fig. 5.

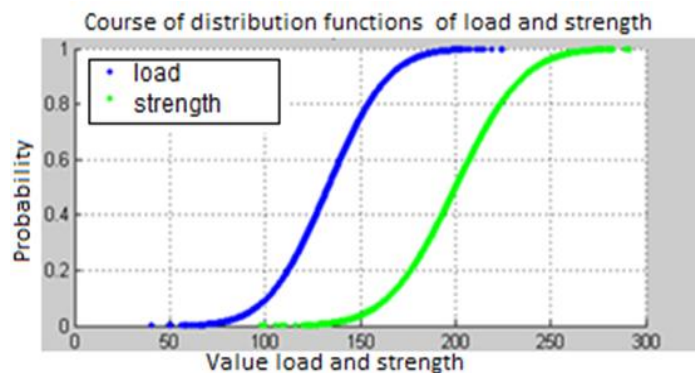


Figure 5: Determining the operating load and strength of the part

The next step of the analysis is to plot the curve of probability density functions and distribution functions, shown in Fig. 6. Determines the minimum value of the strength function and the maximum value of the load function. Plots the interdependence of L and S values. Calculates the size of the area under the graph line.

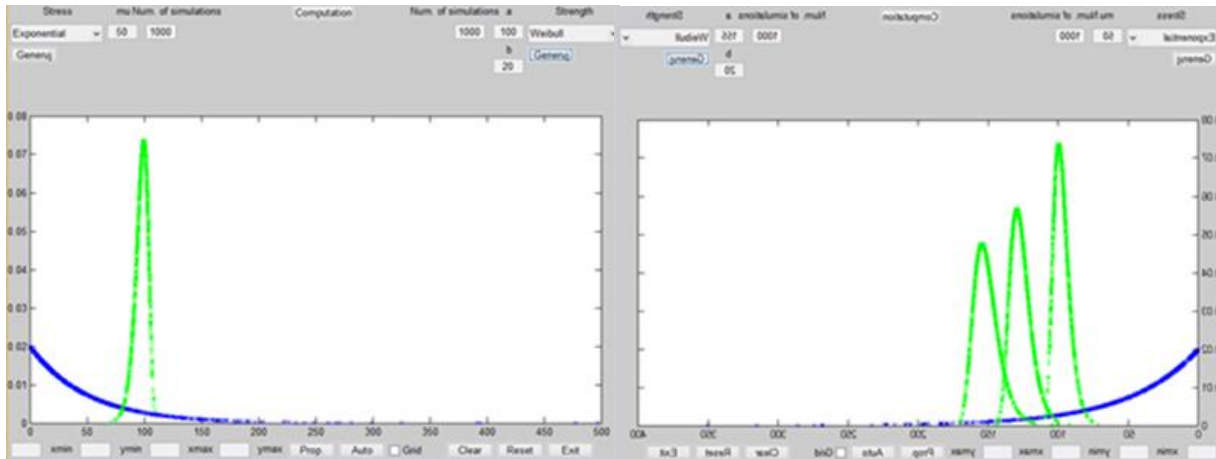


Figure 6: Probability density functions and distribution functions

The next step of the experiment was to analyse the impact of changes in the values of standard deviations.

## 5. Outputs of experimental simulation of parameters of dependability

Load L is a stochastic quantity with properties as in the previous case. It has its distribution of the probability of occurrence at individual levels, which do not change its character (type and parameters of the distribution) with time (period of operation). The resistance of the structure to failure S with time does not change its type (law) of distribution but changes its position relative to the origin of the coordinates. A change in position occurs when the stress repeatedly exceeds a certain threshold limit  $S_c$  of the sensitivity (resistance) of the structure. The application of the dynamic model requires the clarification of some important concepts and properties of the random variables used in the model. Above all, the clarification of the stochastic nature of the quantities S and L, especially their possible change with the time of stress exposure, and further the concept of "accumulation of damage". The possibilities of variations in how the system will react to different strength need to be verified by repeated modelling. ongoing analyses can be evaluated in Fig.7.

Procedure for processing the analysis experiment:

- In the first step, we summarize the input data that evaluates the parameters of histograms, distribution probabilities
- Simulation of random variable operating load and resistance.
- Modelling the point of realization of the phenomenon, assessing which area it falls into and graphically representing it.
- Graphically determine the boundary between the fault-free and fault-free areas, represented by a red line in the picture. 3.
- If the condition  $S \geq L$  is not met, this is a fault condition.
- Control of the number of simulation steps N to the ratio of the number of failure states n.
- We statistically process the generated data into the form of values of probability density functions and distribution functions.
- Generation of mutual interference of phenomena.
- We calculate probability of failure.

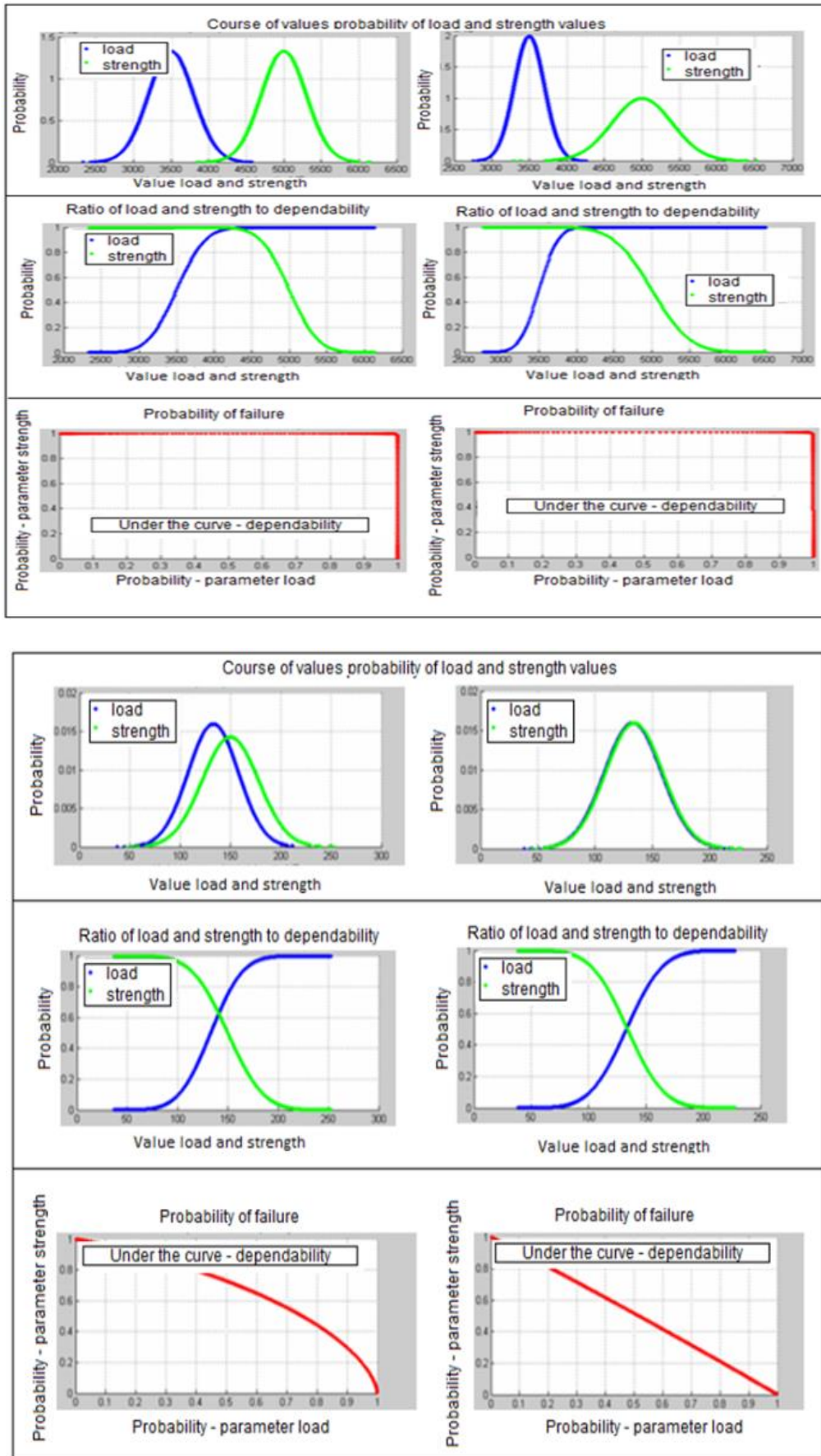
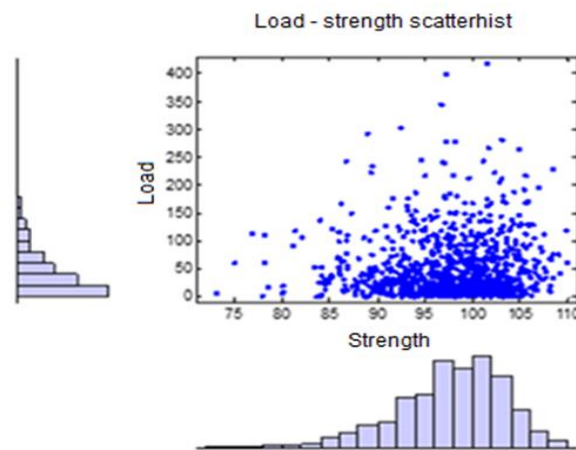


Figure 7: Principle of the impact of changes in the strength parameter

The method of creating a statistical probability model and graphical representation of the share of load and strength in the form of histograms of the generated values can be expressed with the scaatherhist function. The next step is the analysis of the probability of failure achieved in a simulation experiment with parameters according to Table 1. Again, we choose a smaller number of simulations, for indicative results. With a higher number of simulations, the accuracy of the result increases, but the graphical representation of the results deteriorates.

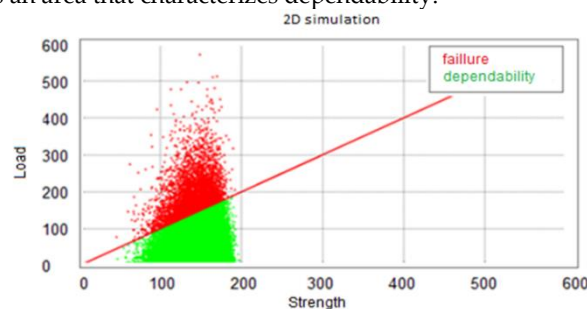
**Table 1:** Input and output parameters of the simulation experiment

Input:		Output:	
Number of simulations	6      1000	Probability of dependability	0.919
Load parameters, Exponential	Median 50	Probability of failure	0.081
	-		
Strength parameters, Weibull	Median 130		
	Standard deviation 20		



**Figure 8:** Graphic representation of load and strength ratio

The results of the simulation and their graphic representation point to a high degree of influence of individual parameters of strength and load. The principle of the approach is shown in Fig. 8 Graphical analysis of the impact of changes in the strength parameter. When assessing the results of the simulation, it must be remembered that the engineering object (element) has the structural and material properties to withstand stress. Load and strength are expressed by quantities that can be characterized as dynamic and stochastic. During operation, the engineering object (element) is stressed by combined effects, namely: Operational stress: operational load, environmental effects, and the human factor. This is represented by the quantity L - load. And resistance to physical stress, chemical stress, and biological stress, represented by the component S – strength. In Fig. 9. an analysis is shown, which provides a graphical output describing the state when we can identify failure. The red area represents a high load that the system is no longer able to withstand. Below the critical line is the permissible area. It is an area that characterizes dependability.



**Figure 9:** Graphical representation of the 2D failure and failure-free set of the realization of 100,000 simulations

## 6. Conclusion

An equivalent stress is a stress that has one constant level (level, amplitude) which, if applied to a component with frequency Sometimes, will cause a failure after the same lifetime of the component as would cause the complete spectrum of stress acting in service at all levels. So, the damage after a certain period of operation (life) caused by this equivalent stress is the same as the damage caused during the same period of operation by the complete spectrum of operational stress [9]. Thus, we can assume that any arbitrary operating stress spectrum can be "converted" to a single level equivalent spectrum of the described properties. The dynamic model is applied primarily to such processes when the strength  $S$  against failure due to repeated exposure to Load  $L$  of different (randomly variable) levels changes with the duration of operation (time). These are e.g., typical cases of element damage due to phenomena associated with material fatigue, exceeding the set parameters limits [10,11].

The results of experimental simulation using the reliability interference method can be summarized in the following advantages:

- The construction, component will be reliable if the operating load  $L$  does not exceed the strength  $S$  with a certain probability.
- The quantities  $L$  and  $S$  are random, and we assume that they have a specific probability distribution law.
- The operational load and strength of the structure will be expressed by probability densities and distribution functions.
- Load and strength are quantities that can influence each other (interfere).
- The extreme points of penetration delimit the area of mutual interference, which is proportional to the probability of the occurrence of a fault.

The simulation model makes it possible to eliminate the shortcomings of classical calculation methods and to use the results of few experimental measurements, to determine the interference of different probability density distributions of randomly variable functions of permitted operating loads and strength, to apply the results to determine the reliability of elements of diverse systems and, last but not least, to use graphic outputs for didactic support of the method explanation SSI and the behaviour of random variables of different probability distributions.

## References

- [1] CÉROVÁ, E. (2023). *Economical and Statistical Optimization of the Maintenance in the Production Process*. In: Manufacturing Technology 2023. Vol. 23, No.1, pp. 32-39.
- [2] RAUSAND, M., ROYLAND, A. (2003). *System Reliability Theory: Models, Statistical Methods, and Applications*, 2nd Edition (Wiley Series in Probability and Statistics), ISBN-13: 978-0471471332.
- [3] VALIŠ, D., POKORA O., KOLÁČEK J. (2020). *System failure estimation based on field data and semi-parametric modelling*, In: Engineering Failure Analysis, Volume 101, pp. 473 – 484, ISSN 13506307.
- [4] VALIŠ, D., ŽÁK, L., VINTR, Z. (2019). *System Condition Assessment Based on Mathematical Analysis*, Conference Proceedings. In: IEEE International Conference on Industrial Engineering and Engineering Management. Volume 2019. pp. 222 - 2269. ISBN 978-153866786-6.
- [5] LEITNER, B., LUSKOVÁ, M. (2019). *Quantified probability estimation of traffic congestion as source of societal risks*. In: Transport Means - Proceedings of the International Conference, Volume 2019. pp. 118 – 123, ISSN 1822296X.

- [6] SAILER, J., HLADÍK, T. (2019). Consistent Maintenance Management Model: Results of changes of maintenance organisation structure and processes". In: *Manufacturing Technology*, Vol. 21, No. 1, pp. 124-131. ISSN 12132489, 2019
- [7] MICHALKOVÁ, P., LEGÁT, V., ALEŠ, Z. (2018). *Dependability analysis of the injection press using Weibull distribution*. In: *Manufacturing Technology*, Vol. 18, No. 4, pp. 625-629. ISSN 12132489.
- [8] DVOŘÁK, Z., LEITNER, B. (2018). *Software tool for railway traffic modelling under conditions of limited permeability*. In: *Transport Means - Proceedings of the International Conference*, pp. 448-454, ISSN 1822296X.
- [9] FABIÁNEK, D., LEGÁT, V., ALEŠ, Z. (2021). *Weibull's analysis of the dependability of critical components of selected agricultural machinery*. In: *Manufacturing Technology*. 2021. Vol. 21, No 5. pp. 605-615.
- [10] TERINGL, A., ALEŠ, Z., LEGÁT, V. (2015). *Dependability Characteristics - Indicators for Maintenance Performance Measurement of Manufacturing Technology*. In: *Manufacturing Technology*. Vol. 15, No. 3. pp. 456-461.
- [11] STAVEK, M., ALEŠ, Z., LEGÁT, V., TERINGL, A. (2015). *Operational Risk Management and Treatment at Technical Systems with Maintenance Support*. In: *Manufacturing Technology*. Vol. 15. No. 3. pp. 429-435.

# A STUDY ON PARTIALLY ACCELERATED LIFE TEST MODEL FOR GENERALIZED INVERSE RAYLEIGH DISTRIBUTION UNDER ADAPTIVE TYPE-II PROGRESSIVE HYBRID CENSORING

Intekhab Alam<sup>1</sup>, Trapty Agarwal<sup>2</sup>, Awakash Mishra<sup>3\*</sup>, Aanchal Gaba<sup>4</sup>

<sup>1,2,3</sup>School of Engineering & Technology, Maharishi University of Information Technology, Noida,  
India

<sup>4</sup>Department of Mathematics, Noida International University, Greater Noida, India

<sup>1</sup>Email- intekhab.pasha54@gmail.com

<sup>2</sup>Email- trapty@gmail.com

<sup>3</sup>Email- awakashmishra@gmail.com

<sup>4</sup>Email- aanchal.gaba54gmail.com

\*-corresponding author

## Abstract

*Modeling and examination of lifetime phenomena are the main aspects of statistical work in a wide variety of scientific and industrial areas. The area of lifetime information analysis has developed and extended quickly with respect to methodology, theory, and fields of applications. The point and interval maximum-likelihood estimations of generalized inverse Rayleigh distribution (GIRD) parameters and the acceleration factor are considered in this work. The estimation procedure is carried out for a partially accelerated step-stress model under adaptive Type-II progressive hybrid censored data. The biases and the mean square errors of the maximum-likelihood estimators are computed to assess their performances in the occurrence of censoring developed in this study through a Monte Carlo simulation study.*

**Keywords:** Partially accelerated life test, Generalized inverse Rayleigh distribution, Newton Raphson method, Adaptive Type-II Progressive Hybrid Censoring, Simulation study.

## I. Introduction

The Partially accelerated life tests (PALTs) are applied by reliability practitioners profitably to calculate approximately the acceleration factor and thus gathering the accelerated information to ordinary surroundings. In a PALT, objects are experience in both regular and accelerated circumstances. Progressive-stress, step-stress, and constant-stress are the three types of PALTs. The assessment performed under these kinds of stress is called accelerated life test (ALT) or partially accelerated life test (PALT). In ALT, the components are placed under stress to obtain additional failures in a tiny time. The key postulation in ALT is that the mathematical model connecting the life span of the component and the stress is acknowledged or can be assumed. In various situations, such a model is neither identified nor assumed. That is, ALT information can't be



gathered to ordinary use circumstances. So, in such situations, PALT is a more appropriate choice to be applied to calculate the statistical model’s parameters. There are three types of PALT i.e. Constant stress PALT (CSPALT), step stress PALT (SSPALT) and progressive stress PALT (PSPALT).

In SSPALT, the test component initiates at ordinary use circumstances for a particular period. If it works successfully at that period, it is placed in stress. Stress continually increases until the examination components are unsuccessful or the examination is ended based on a confident censoring scheme. Rao [1] indicates that the step-stress technique can reduce the investigating period and save many human resources, substances, sources, and cash. In particular, SSPALT should be applied for a trustworthiness study to save time and wealth mainly, when the trial components are of superior reliability and have significant models.

In the present work, we combine an adaptive Type-II progressive hybrid censoring scheme with the step-stress PALT to obtain a step-stress PALT under adaptive Type-II progressive hybrid censored scheme with the GIRD as a lifetime model.

As pointed out by Lin et al. [2], many conditions in existence analysis and reliability research are available, in which components are lost or removed in the investigation prior to failure. The practitioner may not gainful idea about the failure times for all the elements under study. The information detected from this research is called censored information, and the scheme is called censoring scheme. The frequently applied censoring schemes are the Type-I and Type-II censoring scheme, for more details one may refer to Balakrishnan and Ng [3]. Many studies have discussed the hybrid censoring plan, which is a combination of Type-I and Type-II censoring schemes, with the associated statistical inference, see for example, Epstein [4], Balakrishnan and Kundu [5] Childs et al. [6], Gupta and Kundu [7], Kundu [8], Deyand Pradhan [9], and Salah el al. [10] among others. Due to the less flexibility of removing the components from the testing at any position other than the starting point, another censoring scheme was applied, which is called progressively Type-II hybrid censoring schemes. Table 1 summarizes a recent literature review of the different censoring schemes.

**Table 1:** *Related work to the proposed problem*

<b>Author(s) Name</b>	<b>Method</b>	<b>Scheme</b>	<b>Failure Model</b>	<b>Strategy</b>
Abdel-Ghaly et al. [11]	SSALT	Type-II	Pareto distribution	-
Alam et al. [12], Alam and Aquil [13]	CSPALT, SSPALT	Progressive censoring, Adaptive Type-II progressive hybrid censoring	Generalized inverted exponential distribution, Exponentiated Pareto distribution	Maintenance service policy
Abd El-Raheem [14, 15]	CSALT, CSALT	Complete sampling, Type-I censoring	Extension of the exponential distribution	-
Balakrishnan et al. [16]	SSALT	Type-II censoring	Exponential distribution	-
Xiaolin et al. [17]	SSPALT	Progressive Type-II hybrid censoring	Modified Weibull distribution	-
Alam and Aquil [18]	SSPALT	Progressive censoring	Generalized inverted exponential distribution	Maintenance service policy
Alam et al. [19]	SSPALT	Progressive censoring	Power function distribution	Maintenance service policy

Author(s) Name	Method	Scheme	Failure Model	Strategy
Ismail [20, 21]	SSPALT, SSPALT	Adaptive Type-II progressively hybrid censoring, Adaptive Type-I progressively hybrid censoring	Weibull distribution, Weibull distribution	-
Ismail [22]	SSPALT	Type-I progressive hybrid censoring	Weibull distribution	-
El-Sagheer and Ahsanullah [23]	SSPALT	Type-II-Progressive censoring	Lomax distribution	-
Zhou et al. [24]	SSALT	-	Copula function	Competing risk
Srivastava and Mittal [25]	SSPALT	Type-I and Type-II censorings	truncated logistic distribution	-
<b>Proposed Work</b>	SSPALT	Adaptive Type-II progressive hybrid censoring	Generalized Inverse Rayleigh distribution	-

The proposed study is motivated by two factors. The first aims to establish explicit formulas for the likelihood and log-likelihood functions under an adaptive Type-II progressive hybrid censoring scheme. The second is to apply a Monte Carlo simulation study to estimate the performance of the model parameter estimators with an adaptive Type-II progressive hybrid censoring scheme in terms of biases and mean squared errors. The authors presented a study on SSPALT utilizing adaptive Type-II progressive hybrid censoring where the lifespan of test items follows the two parameters GIRD in this work.

The uniqueness of this work stems from the fact that no earlier research has been conducted in this area using the proposed censoring scheme for two parameters GIRD.

The present paper is arranged as; the model illustration and test procedure are presented in section 2. The point and interval estimation is presented in section 3. A simulation study is carried out in section 4 to check the performance of model parameters. The result based on the proposed problem and conclusion is provided in section 5. The real-life implementation of the proposed work is shown in section 6.

## II. Model Illustration and Test Process

The GIRD is one of the most beneficial and important distribution within the inverted scale distributions. It has been considered as an appropriate failure model in life testing and reliability analysis, for more details about GIRD one may refer to Fatima et al. [26]. The GIRD has lots of uses in the area of reliability theories. The Probability density function (pdf) of GIRD presents by the following equation (1);

$$f(y, \lambda, \theta) = \frac{2\lambda}{\theta^2 y^3} e^{-(\theta y)^{-2}} \left[ 1 - e^{-(\theta y)^{-2}} \right]^{\lambda-1}; y, \lambda, \theta > 0 \quad (1)$$

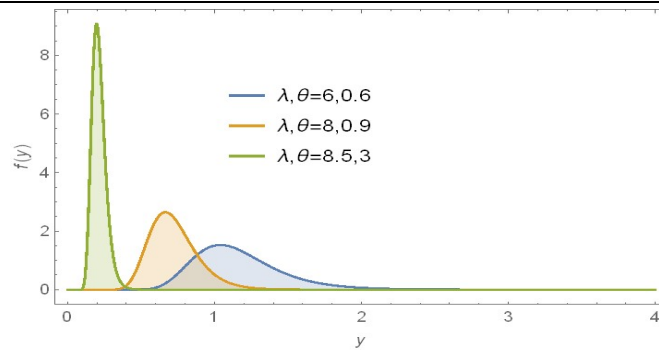


Figure 1: Pdf pattern of GIRD

The cumulative density function (cdf) of GIRD presents by the following equation (2);

$$F(y, \lambda, \theta) = 1 - \left[ 1 - e^{-(\theta y)^{-2}} \right]^\lambda ; y, \lambda, \theta > 0 \quad (2)$$

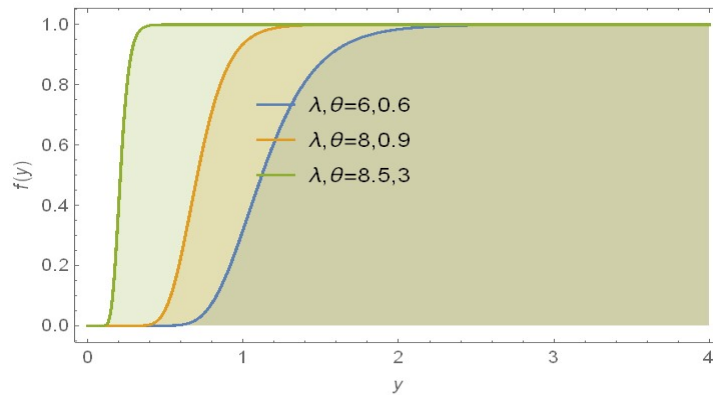


Figure 2: Cdf pattern of GIRD

The reliability function of GIRD is given by

$$S(y, \lambda, \theta) = \left[ 1 - e^{-(\theta y)^{-2}} \right]^\lambda \quad (3)$$

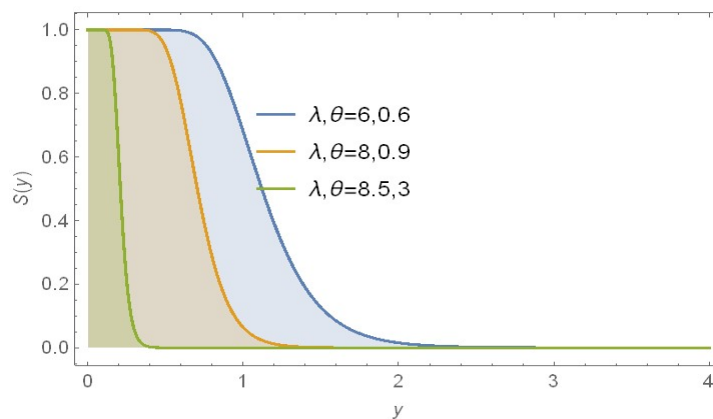


Figure 3: Reliability pattern of GIRD

The hazard rate function of GIRD is presented by the following equation:

$$h(y, \lambda, \theta) = \frac{2\lambda e^{-(\theta y)^{-2}}}{\theta^2 y^3 \left[ 1 - e^{-(\theta y)^{-2}} \right]} \quad (4)$$

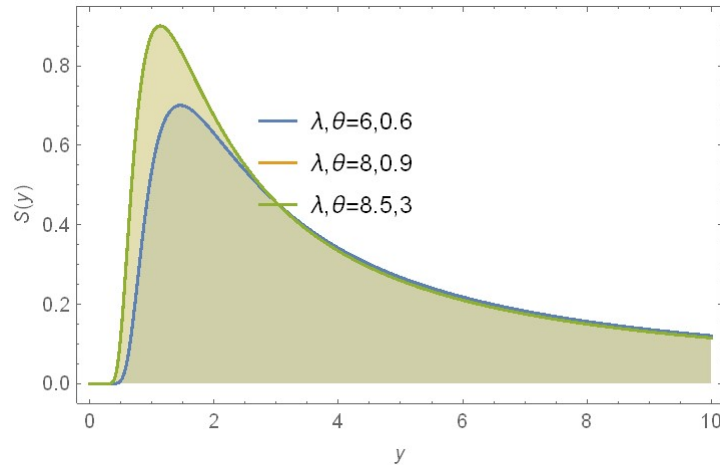


Figure 4: Hazard pattern of GIRD

Figure 1 shows that the Pdf of GIRD is positively skewed, while the shape of Cdf is increasing as shown in Figure 2. Figure 3 and 4 show the Reliability and Hazard shapes of GIRD for different values of  $\lambda$  and  $\theta$ . The figure 3 shows that the reliability function of GIRD is downward skewed for different values of  $\lambda, \theta$ , it becomes flatter and flatter as the shape parameter is increased. The behavior of instantaneous failure rate of the GIRD has an upside-down bathtub shape curve.

The unimodel hazard rate function shows the possibility of decreasing failures as soon as the product has passed a particular moment, during some kind of stress on that product. Thus, the GIRD shows excellent statistical performance and can be a better model to fit real data in many scientific fields.

Kumar and Garg [27] handled an estimation of parameters of GIRD based on randomly censored trials. Bakoban and Abubaker [28] presented the assumption of GIRD with real information applications. Bakoban and Abubaker [29] also proposed a study on the estimation of parameters of GIRD using progressive Type-II censoring.

Under SSPALT the pdf of  $Y$  can be written as:

$$f(y) = \begin{cases} 0, & y \leq 0 \\ f_1(y) = f(y, \lambda, \theta), & 0 < y \leq \tau \\ f_2(y), & y > \tau \end{cases} \quad (5)$$

where,  $f_2(y) = \frac{2\lambda\beta}{\theta^2 y^3} e^{-(\theta(\tau+\beta(y-\tau)))^{-2}} \left[ 1 - e^{-(\theta(\tau+\beta(y-\tau)))^{-2}} \right]^{\lambda-1}; y, \beta, \lambda, \theta > 0$

$f_2(y)$  is attained by applying variable transformation that is projected by DeGroot and Goel [30] and the procedure is given in the following equation:

$$Y = \begin{cases} T & \text{if } T \leq \tau \\ \tau + \beta^{-1}(T - \tau) & \text{if } T > \tau \end{cases} \quad (6)$$

In equation (6),  $T$  is the life span of the article in normal operating circumstances, while  $\tau$  is the time at which stress is changed (i.e., stress change time) and  $\beta$  is the acceleration factor.

In life testing analysis, the Type-I and Type-II censoring ideas are the mainly well-liked and widespread plans. These schemes explain as follows, suppose there are  $n$  apparatus situate on the investigation, then under Type-I censoring, the investigation carries on until reached a pre-specified point  $\kappa$ . But in Type-II censoring, the investigation carries on until a reached a pre-specified quantity of components  $m(\leq n)$ . We cannot take out components from the live experimentation at any moment and any position except the opening point in these two schemes. This is the key weakness of these two schemes. To remove this weakness, progressive Type-II censoring or progressively Type-II hybrid censoring another censoring comes in light. Hybrid censoring is the mixture of two censorings, i.e., Type-I and Type-II censorings. So, the progressive Type-II censoring scheme is described as follows;

In the progressive Type-II censoring scheme environment, the reliability practitioner presets the number of components to be unsuccessful (say  $m$ ) out of the total number of components  $n$ , placed under analysis. At the moment, when initial failure happens,  $R_1$  components among  $n - 1$  leftover (surviving) components are randomly taken off from the life analysis. Similarly,  $R_2$  of the leftover  $n - 2 - R_1$  examination components are eliminated from the analysis at the moment of the second failure. This practice continues until the  $m$ th failure is reached. All the leftover  $R_m = n - m - R_1 - R_2 - \dots - R_{m-1} - 1$  surviving examination components are eliminated from the examination at this point. The  $R_i$  units are situating before the work. The assumption related to progressive censoring and progressively is proposed by many authors such as Balakrishnan [31], Balakrishnan and Agrawala [32], etc.

If a life examination experiment stops randomly at a moment  $\min(Y_{m:m:n}, \varphi)$ , where  $1 \leq m \leq n, 0 < \varphi < \infty$  are determined prior to the experiment, and  $Y_{1:m:n} \leq Y_{2:m:n} \leq \dots \leq Y_{m:m:n}$  are the ordered lifetimes consequential from the study, then  $(R_1, R_2, \dots, R_m)$  are called progressively hybrid censoring (PHC) scheme. If the  $m$ th progressive censored unit occurs before the point  $\varphi$  ( $\varphi > Y_{m:m:n}$ ), then the investigation ends at the moment  $Y_{m:m:n}$ . Else, the examination will end at the moment  $\varphi$ , where  $Y_{j:m:n} < \varphi < Y_{j+1:m:n}$ , hence all the leftover  $(n - \sum_{i=1}^j R_i - j)$  existing units are censored at  $\varphi$ . Here  $j$  is a random variable and denotes the number of unsuccessful units up to  $\varphi$ . The reliability engineer comes with the problem of different censoring schemes, and the practitioner may observe a tiny test size (even it is equal to zero). So, this is not possible to happen with standard suggestion procedures to obtain good results. To remove such type of drawback, another censoring comes in light called adaptive censored samples. This was commenced by Ng et al. [33].

In this scheme, the observed quantity of failed units  $m$  is prefixed and the investigation moment is unlocked to run over the moment  $\kappa$ . The investigation will continue along with pre-determined progressive censoring schemes  $(R_1, R_2, R_3, \dots, R_m)$  if  $Y_{m:m:n} < \varphi$ , otherwise, the ongoing units (on work units), which following the  $(j + 1)h$  to  $(m - 1)h$  experimental failures, are

not uninvolved from the analysis. All the surviving units  $R_m = n - m - \sum_{i=1}^j R_i$  are taken back from the test at the stage  $Y_{m:m:n}$  if  $m$  observed failures are obtained, i.e.  $R_{j+1} = \dots = R_{m-1} = 0$ . The progressive Type-II censoring takes place if  $n \rightarrow \infty$  and the conventional Type-II censoring takes place, if  $n \rightarrow 0$ .

If the practitioner is free to vary the value  $\kappa$ , then this kind of censoring proposal is known as an adaptive progressively Type-II hybrid censoring (APHCT-II) scheme. This variation in  $\kappa$  is completed to regulate the most advantageous of pointed investigation time and a better opportunity of supervising various failures.

### III. Estimation Process

Let  $Y_1, Y_2, \dots, Y_n$  be a life span of  $n$  independently and identically distributed units following the GIRD.  $y_{1:m:n} < y_{2:m:n} < \dots < y_{n_u:m:n} \leq \tau < y_{n_u+1:m:n} \leq \varphi < y_{J+1:m:n} < \dots < y_{m:n:n}$  are completely observed (ordered) lifetimes. Both point and confidence interval estimation is presented in the following subsections:

#### I. Point Estimation

In this section, we used the maximum likelihood estimation method. Under the SSPALT the likelihood function with APHCT-II for GIRD based on  $m$  observed lifetime data takes the following form;

$$L(\lambda, \theta, \beta) \propto \prod_{i=1}^m f_1(y_{i:m:n}) f_2(y_{i:m:n}) \prod_{i=1}^J (S_1(\tau))^{R_i} (S_2(y_{m:m:n}))^{(n-m-\sum_{i=1}^J R_i)} \quad (7)$$

$$\text{where, } S(\tau) = \left[ 1 - e^{-(\theta\tau)^{-2}} \right]^\lambda, S(y_{m:m:n}) = \left[ 1 - e^{-(\theta(\tau + \beta(y_{m:m:n} - \tau)))^{-2}} \right]^\lambda$$

$$\ln L = \ln L(\lambda, \theta, \beta)$$

$n_u$  is the amount of components that are unsuccessful in the normal circumstance and  $n_a$  is the number of components that are unsuccessful in accelerated circumstance.

The log-likelihood function takes the following form;

$$\begin{aligned} \ln L = & -\sum_{i=1}^m (\theta y_i)^{-2} + (\lambda - 1) \sum_{i=1}^m \ln \left[ 1 - e^{-(\theta y_i)^{-2}} \right] - 3 \sum_{i=1}^m \ln [\tau + \beta(y_i - \tau)] m \left[ \ln \left( \frac{4\lambda^2\beta}{\theta^4} \right) \right] \\ & + \sum_{i=1}^J \lambda R_i \ln \left[ 1 - e^{-(\theta\tau)^{-2}} \right] - 3 \sum_{i=1}^m \ln(y_i) - \sum_{i=1}^m (\theta(\tau + \beta(y_i - \tau)))^{-2} \\ & + \sum_{i=1}^J \lambda \left( n - m - \sum_{i=1}^J R_i \right) \ln \left[ 1 - e^{-(\theta(\tau + \beta(y_{m:m:n} - \tau)))^{-2}} \right] + (\lambda - 1) \sum_{i=1}^m \ln \left[ 1 - e^{-(\theta(\tau + \beta(y_i - \tau)))^{-2}} \right] \end{aligned} \quad (8)$$

To obtain maximum likelihood estimates (MLEs) of model parameters and acceleration factor, we differentiate the above equation for parameters  $\phi, \lambda$  and  $\beta$  equating to zero.

where,  $\sigma_i = \tau + \beta(y_i - \tau)$ ,  $\sigma_{m:m:n} = \tau + \beta(y_{m:m:n} - \tau)$ ,  $J = n_u + n_a$

$$\frac{\partial \ln L}{\partial \lambda} = \left(\frac{2m}{\lambda}\right) + \sum_{i=1}^m \ln \left[1 - e^{-(\theta y_i)^{-2}}\right] + \sum_{i=1}^m \ln \left[1 - e^{-(\theta \sigma_i)^{-2}}\right] + \sum_{i=1}^J R_i \ln \left[1 - e^{-(\theta \tau)^{-2}}\right] + \sum_{i=1}^J \left(n - m - \sum_{i=1}^J R_i\right) \ln \left[1 - e^{-(\theta \sigma_m)^{-2}}\right] = 0 \tag{9}$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \theta} = & -\frac{4m}{\theta} + 2 \sum_{i=1}^m \theta^{-3} (y_i)^{-2} - 2(\lambda - 1)\theta^{-3} \sum_{i=1}^m \frac{e^{-(\theta y_i)^{-2}} y_i^{-2}}{\left(1 - e^{-(\theta y_i)^{-2}}\right)} - \sum_{i=1}^m (\sigma_i)^{-2} \theta^{-3} - \\ & 2(\lambda - 1)\theta^{-3} \sum_{i=1}^m \frac{(\sigma_i)^{-2} e^{-(\theta \sigma_i)^{-2}}}{1 - e^{-(\theta \sigma_i)^{-2}}} - 2 \sum_{i=1}^J \lambda R_i \frac{e^{-(\theta \tau)^{-2}} \tau^{-2} \theta^{-3}}{1 - e^{-(\theta \tau)^{-2}}} \\ & - 2 \sum_{i=1}^J \lambda \theta^{-3} \left(n - m - \sum_{i=1}^J R_i\right) \left[\frac{e^{-(\theta \sigma_m)^{-2}} (\sigma_m)^{-2}}{1 - e^{-(\theta \sigma_m)^{-2}}}\right] = 0 \end{aligned} \tag{10}$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \beta} = & 2\theta \sum_{i=1}^m \frac{(\sigma_i - \tau)(\theta \sigma_i)^{-3}}{\beta} - 2(\lambda - 1)\theta \sum_{i=1}^m \frac{(\sigma_i - \tau)(\theta \sigma_i)^{-3} e^{-(\theta \sigma_i)^{-2}}}{\beta(1 - e^{-(\theta \sigma_i)^{-2}})} \\ & - \sum_{i=1}^J \lambda \theta \left(n - m - \sum_{i=1}^J R_i\right) \frac{(\sigma_m - \tau)(\theta \sigma_m)^{-3} e^{-(\theta \sigma_m)^{-2}}}{\beta(1 - e^{-(\theta \sigma_m)^{-2}})} + \frac{m}{\beta} - 3 \sum_{i=1}^m \frac{(\sigma_i - \tau)}{\beta \sigma_i} = 0 \end{aligned} \tag{11}$$

It is an impossible task to solve the above equations manually. Hence, an iterative procedure called the Newton Raphson technique is used to get the MLE of the model parameters and acceleration factor.

## II. Interval Estimation

The interval estimation for the model parameters and acceleration factor based on APHCT-II is obtained. The asymptotic distribution of MLE  $\lambda, \theta$  and  $\beta$  takes the following form presented in the following equations.

$$\left((\hat{\lambda} - \lambda), (\hat{\theta} - \theta)(\hat{\beta} - \beta)\right) \rightarrow N\left(0, I^{-1}(\lambda, \theta, \beta)\right) \tag{12}$$

The above procedure is suggested by Miller and Nelson [34].  $I^{-1}(\lambda, \theta, \beta)$  denotes the variance-covariance matrix of  $\lambda, \theta$  and  $\beta$ . The  $3 \times 3$  matrix  $I^{-1}$  which is approximately equal to  $I$  and the elements  $I_{ij}^{-1}(\lambda, \theta, \beta), i = 1, 2, 3; j = 1, 2, 3$ , closed to  $I_{ij}(\hat{\lambda}, \hat{\theta}, \hat{\beta})$ , under the APHCT-II are given as.

where,  $\frac{\partial^2 \ln L}{\partial \lambda^2} = -\left(\frac{2m}{\lambda^2}\right)$

$$\frac{\partial^2 \ln L}{\partial \lambda \partial \theta} = -2\theta^{-3} \sum_{i=1}^m \frac{y_i^{-2} e^{-(\theta y_i)^{-2}}}{(1 - e^{-(\theta y_i)^{-2}})} - 2\theta^{-3} \sum_{i=1}^m \frac{e^{-(\theta \sigma_i)^{-2}} (\sigma_i)^{-2}}{(1 - e^{-(\theta \sigma_i)^{-2}})} - 2\theta^{-3} \sum_{i=1}^J \left(n - m - \sum_{i=1}^J R_i\right) \frac{(\sigma_m)^{-2} e^{-(\theta \sigma_m)^{-2}}}{(1 - e^{-(\theta \sigma_m)^{-2}})} - 2\theta^{-3} \tau^{-2} \sum_{i=1}^J R_i \frac{e^{-(\theta \tau)^{-2}}}{(1 - e^{-(\theta \tau)^{-2}})}$$

$$\begin{aligned} \frac{\partial^2 \ln L}{\partial \theta^2} &= \frac{4m}{\theta^2} - 6 \sum_{i=1}^m \theta^{-4} (y_i)^{-2} - 2(\lambda - 1)\theta^{-3} \sum_{i=1}^m \frac{e^{-(\theta y_i)^{-2}} y_i^{-2}}{(1 - e^{-(\theta y_i)^{-2}})} \left[ -3\theta^{-1} + 2\theta^{-3} y_i^{-2} - \frac{2y_i^{-2} \theta^{-3} e^{-(\theta y_i)^{-2}}}{1 - e^{-(\theta y_i)^{-2}}} \right] \\ &+ 3 \sum_{i=1}^m \left( (\sigma_i)^{-2} \theta^{-4} - 2(\lambda - 1)\theta^{-3} \sum_{i=1}^m \frac{(\sigma_i)^{-2} e^{-(\theta \sigma_i)^{-2}}}{1 - e^{-(\theta \sigma_i)^{-2}}} - 2 \sum_{i=1}^J \lambda R_i \frac{e^{-(\theta \tau)^{-2}} \tau^{-2} \theta^{-3}}{1 - e^{-(\theta \tau)^{-2}}} \times \right. \\ &\left[ -3\theta^{-1} + 2\theta^{-3} \tau^{-2} - \frac{2\tau^{-2} \theta^{-3} e^{-(\theta \tau)^{-2}}}{1 - e^{-(\theta \tau)^{-2}}} \right] - 2 \sum_{i=1}^J \lambda \theta^{-3} \left( n - m - \sum_{i=1}^J R_i \right) \times \\ &\left. \left[ \frac{e^{-(\theta \sigma_m)^{-2}} (\sigma_m)^{-2}}{1 - e^{-(\theta \sigma_m)^{-2}}} \right] \left[ -3\theta^{-1} + 2\theta^{-3} (\sigma_m)^{-2} - \frac{2\theta^{-3} e^{-(\theta \sigma_m)^{-2}} (\sigma_m)^{-2}}{1 - e^{-(\theta \sigma_m)^{-2}}} \right] \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ln L}{\partial \theta \partial \beta} &= -2 \sum_{i=1}^m (\sigma_i)^{-3} (\sigma_i - \tau) \beta^{-1} \theta^{-3} - 2(\lambda - 1)\theta^{-3} \sum_{i=1}^m \frac{(\sigma_i)^{-2} e^{-(\theta \sigma_i)^{-2}}}{1 - e^{-(\theta \sigma_i)^{-2}}} \times \\ &\left[ -2 \frac{(\sigma_i - \tau)}{(\sigma_i) \beta} + 2\theta^{-2} (\sigma_i - \tau) \beta^{-1} (\theta \sigma_i)^{-3} + \frac{2\theta^{-2} (\sigma_i - \tau) \beta^{-1} (\sigma_i)^{-3} e^{-(\theta \sigma_i)^{-2}}}{1 - e^{-(\theta \sigma_i)^{-2}}} \right] - 2 \sum_{i=1}^J \lambda \theta^{-3} \left( n - m - \sum_{i=1}^J R_i \right) \\ &\left[ -2 \frac{(\sigma_m - \tau)}{(\sigma_m) \beta} + 2\theta^{-2} (\sigma_m - \tau) \beta^{-1} (\theta \sigma_m)^{-3} + \frac{2\theta^{-2} \sigma_m \beta^{-1} (\sigma_m)^{-3} e^{-(\theta \sigma_m)^{-2}}}{1 - e^{-(\theta \sigma_m)^{-2}}} \right] \left[ \frac{e^{-(\theta \sigma_m)^{-2}} (\sigma_m)^{-2}}{1 - e^{-(\theta \sigma_m)^{-2}}} \right] \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ln L}{\partial \beta^2} &= -\frac{m}{\beta^2} + 3 \sum_{i=1}^m \frac{(\sigma_i - \tau)^2}{\beta^2 (\sigma_i)^2} - 6\theta^{-2} \sum_{i=1}^m (\sigma_i)^{-4} \beta^{-2} (\sigma_i - \tau)^2 - 2(\lambda - 1)\theta \times \\ &\sum_{i=1}^m \frac{(\sigma_i - \tau) \beta^{-1} (\theta \sigma_i)^{-3} e^{-(\theta \sigma_i)^{-2}}}{1 - e^{-(\theta \sigma_i)^{-2}}} \left[ 2\theta^{-2} (\sigma_i - \tau) \beta^{-1} (\sigma_i)^{-3} - \frac{3(\sigma_i - \tau)}{\beta \sigma_i} - \frac{2\theta^{-2} (\sigma_i)^{-3} (\sigma_i - \tau) \beta^{-1} e^{-(\theta \sigma_i)^{-2}}}{1 - e^{-(\theta \sigma_i)^{-2}}} \right] \\ &- \sum_{i=1}^J \lambda \theta \left( n - m - \sum_{i=1}^J R_i \right) \frac{(\sigma_m - \tau) \beta^{-1} (\theta \sigma_m)^{-3} e^{-(\theta \sigma_m)^{-2}}}{1 - e^{-(\theta \sigma_m)^{-2}}} \times \\ &\left[ -\frac{3(\sigma_m - \tau) \beta^{-1}}{\sigma_m} + 2\theta^{-2} \beta^{-1} \sigma_m (\sigma_m)^{-3} - \frac{2\theta^{-2} (\sigma_m)^{-3} (\sigma_m - \tau) \beta^{-1} e^{-(\theta \sigma_m)^{-2}}}{1 - e^{-(\theta \sigma_m)^{-2}}} \right] \end{aligned}$$

$$\frac{\partial^2 \ln L}{\partial \beta \partial \lambda} = -2\theta \sum_{i=1}^m \frac{(\sigma_i - \tau) (\theta \sigma_i)^{-3} e^{-(\theta \sigma_i)^{-2}}}{\beta (1 - e^{-(\theta \sigma_i)^{-2}})} - \sum_{i=1}^J \theta \left( n - m - \sum_{i=1}^J R_i \right) \frac{(\sigma_m - \tau) (\theta \sigma_m)^{-3} e^{-(\theta \sigma_m)^{-2}}}{\beta (1 - e^{-(\theta \sigma_m)^{-2}})}$$

The 100(1 - π)% approximated two-sided limits of confidence for parameters λ, θ and β



are given as:

$$\hat{\lambda} \pm Z_{\pi/2} \sqrt{I_{11}^{-1}(\hat{\lambda}, \hat{\theta}, \hat{\beta})}, \hat{\theta} \pm Z_{\pi/2} \sqrt{I_{22}^{-1}(\hat{\lambda}, \hat{\theta}, \hat{\beta})} \text{ and } \hat{\beta} \pm Z_{\pi/2} \sqrt{I_{33}^{-1}(\hat{\lambda}, \hat{\theta}, \hat{\beta})}$$

#### IV. Simulation study

Since it is theoretically not achievable to evaluate the presentation of different censorings for different values of model parameters. For this job, many software and simulation techniques are used. In this segment, the Monte-Carlo simulation procedure is applied to evaluate the efficiency of the MLEs. This efficiency is recorded based on the mean squared error (MSE) and bias of the MLEs. The following three progressive censorings are chosen for this assignment;

- Scheme (I)  $R_1 = R_2 = R_3 = \dots = R_{m-1}, R_m = n - m$
- Scheme (II)  $R_1 = n - m, R_2 = R_3 = R_4 \dots = 0$
- Scheme (III)  $R_1 = R_2 = R_3 = \dots = R_{m-1}, R_m = n - 2m + 1$

For this task, 1000 simulation-based on MSEs and biases are estimated. The steps for this procedure are;

- The values of parameters  $n, m, \tau, \varphi, \lambda, \theta$  and  $\beta$  are specified first.
- After selecting the parameter values, we generate a random sample from GIRD of size  $n$  by the inverse CDF method in both situations (regular and accelerated circumstances).
- Generate the PHC sample for the parameters  $n, m, \tau, \varphi, \lambda, \theta$  and  $\beta$  by using the technique discussed in equation (6).
- The sample data set for the APHCT-II is;  
 $y_{1:m:n} < x_{y:m:n} < \dots y_{n_u:m:n} \leq \tau < y_{n_u+1:m:n} \leq \varphi < y_{J+1:m:n} < \dots < y_{m:n:n}$
- Find the values of the MSEs and the biases associated with MLEs of the parameters, the computing values are presented in Table 2,3,4 and 5 at different values of parameters.

**Table 2:** The average MSEs and biases for  $\lambda, \theta, \beta, \tau$  and  $\varphi$  are set at 0.9, 1.4, 1.76, 2.4 and 6

(n,m)	Schemes	Values of $\lambda$		Values of $\theta$		Values of $\beta$	
		Bias	MSE	Bias	MSE	Bias	MSE
(50,12)	1	0.856	0.929	0.498	0.638	0.574	0.684
	2	0.911	0.998	0.685	0.694	0.633	0.693
	3	0.743	0.873	0.584	0.658	0.593	0.709
(70,12)	1	0.502	0.577	0.476	0.609	0.543	0.644
	2	0.587	0.676	0.632	0.698	0.600	0.676
	3	0.522	0.611	0.564	0.650	0.578	0.687
(90,12)	1	0.411	0.500	0.386	0.521	0.465	0.565
	2	0.599	0.680	0.658	0.705	0.533	0.590
	3	0.431	0.534	0.489	0.580	0.498	0.577
(50,20)	1	0.343	0.344	0.300	0.467	0.398	0.511
	2	0.445	0.587	0.499	0.612	0.466	0.554
	3	0.365	0.398	0.387	0.513	0.440	0.534
(70,20)	1	0.233	0.231	0.190	0.376	0.300	0.432
	2	0.342	0.498	0.409	0.546	0.376	0.467
	3	0.287	0.280	0.298	0.412	0.333	0.442

$(n,m)$	Schemes	Values of $\lambda$		Values of $\theta$		Values of $\beta$	
		Bias	MSE	Bias	MSE	Bias	MSE
(90,20)	1	0.143	0.154	0.122	0.287	0.198	0.365
	2	0.234	0.409	0.265	0.387	0.287	0.398
	3	0.188	0.190	0.198	0.322	0.209	0.370

**Table 3:** Average values of MSEs and biases when  $\lambda, \theta, \beta, \tau$  and  $\varphi$  are set at 0.7, 1.4, 1.76, 2.4 and 9

$(n,m)$	Schemes	Values of $\lambda$		Values of $\theta$		Values of $\beta$	
		Bias	MSE	Bias	MSE	Bias	MSE
(50,12)	1	0.387	0.508	0.609	0.673	0.548	0.698
	2	0.435	0.580	0.655	0.705	0.715	0.775
	3	0.477	0.546	0.640	0.642	0.658	0.739
(70,12)	1	0.323	0.410	0.456	0.521	0.512	0.574
	2	0.408	0.498	0.509	0.578	0.598	0.687
	3	0.397	0.433	0.480	0.547	0.545	0.632
(90,12)	1	0.276	0.324	0.387	0.454	0.431	0.511
	2	0.322	0.413	0.433	0.517	0.508	0.596
	3	0.299	0.356	0.410	0.489	0.474	0.541
(50,20)	1	0.197	0.250	0.311	0.386	0.324	0.434
	2	0.354	0.431	0.465	0.530	0.534	0.608
	3	0.218	0.288	0.327	0.416	0.419	0.487
(70,20)	1	0.113	0.176	0.232	0.318	0.243	0.353
	2	0.265	0.334	0.379	0.464	0.465	0.533
	3	0.175	0.212	0.248	0.354	0.325	0.397
(90,20)	1	0.007	0.119	0.146	0.243	0.154	0.265
	2	0.175	0.254	0.299	0.385	0.386	0.421
	3	0.108	0.175	0.186	0.278	0.256	0.290

**Table 4:** Average values of MSEs and biases when  $\lambda, \theta, \beta, \tau$  and  $\varphi$  are set at 0.7, 1.4, 1.76, 2.8 and 9

$(n,m)$	Schemes	Values of $\lambda$		Values of $\theta$		Values of $\beta$	
		Bias	MSE	Bias	MSE	Bias	MSE
(50,12)	1	0.334	0.398	0.387	0.465	0.480	0.602
	2	0.387	0.446	0.445	0.576	0.587	0.715
	3	0.354	0.431	0.412	0.543	0.535	0.675
(70,12)	1	0.296	0.344	0.297	0.387	0.429	0.519
	2	0.320	0.400	0.365	0.487	0.519	0.630
	3	0.312	0.386	0.345	0.438	0.482	0.567
(90,12)	1	0.230	0.278	0.204	0.316	0.349	0.430
	2	0.266	0.342	0.295	0.416	0.451	0.579
	3	0.240	0.294	0.256	0.398	0.380	0.483
(50,20)	1	0.187	0.238	0.138	0.253	0.227	0.341
	2	0.287	0.360	0.305	0.436	0.465	0.583
	3	0.209	0.267	0.178	0.303	0.283	0.425
(70,20)	1	0.129	0.186	0.008	0.180	0.145	0.265
	2	0.220	0.287	0.221	0.254	0.373	0.454
	3	0.148	0.202	0.120	0.228	0.220	0.374

(n.m)	Schemes	Values of $\lambda$		Values of $\theta$		Values of $\beta$	
		Bias	MSE	Bias	MSE	Bias	MSE
(90,20)	1	0.007	0.129	0.004	0.109	0.100	0.187
	2	0.139	0.198	0.188	0.169	0.270	0.378
	3	0.102	0.149	0.009	0.134	0.139	0.190

**Table 5:** Average values of MSEs and biases when  $\lambda, \theta, \beta, \tau$  and  $\varphi$  are set at 1.5, 1.4, 1.76, 2.4 and 9

(n.m)	Schemes	Values of $\lambda$		Values of $\theta$		Values of $\beta$	
		Bias	MSE	Bias	MSE	Bias	MSE
(50,12)	1	0.560	0.593	0.593	0.656	0.644	0.734
	2	0.677	0.765	0.709	0.788	0.723	0.797
	3	0.600	0.712	0.650	0.693	0.687	0.755
(70,12)	1	0.476	0.499	0.486	0.575	0.563	0.665
	2	0.588	0.691	0.628	0.690	0.659	0.687
	3	0.523	0.633	0.576	0.620	0.581	0.671
(90,12)	1	0.410	0.453	0.399	0.484	0.487	0.556
	2	0.593	0.698	0.645	0.705	0.667	0.710
	3	0.447	0.560	0.480	0.563	0.530	0.600
(50,20)	1	0.334	0.407	0.311	0.422	0.435	0.523
	2	0.450	0.599	0.513	0.567	0.574	0.616
	3	0.389	0.523	0.419	0.497	0.467	0.586
(70,20)	1	0.254	0.306	0.223	0.375	0.370	0.455
	2	0.334	0.492	0.460	0.500	0.479	0.544
	3	0.319	0.345	0.280	0.386	0.417	0.478
(90,20)	1	0.130	0.233	0.155	0.284	0.245	0.374
	2	0.252	0.364	0.359	0.433	0.332	0.407
	3	0.209	0.288	0.197	0.357	0.300	0.431

## V. Application in Real Life Situation

SSPALT is now the most significant procedure of reviewing item trustworthiness rapidly, and the blueprint of capable investigation plans is a serious step to guarantee that SSPALTs can evaluate the item reliability correctly, quickly, and cheaply. With the encouragement of the national approach of civil-military integration, SSPALT will be mostly applied in the research and development of a variety of manufactured goods, and the SSPALT plan design hypothesis will face more challenges. To assist engineers in selecting suitable hypotheses and to stimulate researchers to build up the theories necessary in manufacturing, with the focal point on the demands for theory investigation that happen from the execution of SSPALT, this study reviews and summarizes the expansion of the SSPALT plan. The expansion of the theory and technique for setting up the most favorable SSPALT for shape-scale distribution, which is the most functional and grown-up theory of designing the optimal SSPALT, are explained in detail. Based on the theory of convenience for engineers to choose suitable techniques according to the troubles that originate in practice, this discussed will help to review the progress of optimal ALT plan design theory by taking the engineering problems occurring from the ALT execution as the key thread, provides strategy on choosing suitable theories for engineers, and suggests views about the vital solved theory problems for researchers.

A real life data set is commenced to demonstrate how the ML estimation method works in practice based on real life data set from Nelson [35]. Table 6 is presented the data set and the data

set is correspond to the oil breakdown period of insulating fluid under two stress stages (34 kilovolt (kv) and 36 kv), considering the data set under 34 kv as data under ordinary stress condition. Before further proceeding, we test the strength of GRID to fit the data listed in Table 6 using Kolmogorov-Smirnov (K-S) test statistic and its corresponding p-value for each stress stage. The outcome is presented in Table 7. We can observe that the GRID fits better to the given data in the two stress stages because the p-values are greater than 0.05. The MLEs, p-values and K-S statistic are presented in Table 7.

**Table 6:** Stress values and complete failure data

Stress (in kv)	Complete failure data
34	0.19, 0.78, 0.96, 1.31, 2.78, 3.16, 4.15, 4.67, 4.85, 6.50, 7.35, 8.01, 8.27, 12.06, 31.75, 32.52, 33.91, 36.71, 72.89
36	0.35, 0.59, 0.96, 0.99, 1.69, 1.97, 2.07, 2.58, 2.71, 2.9, 3.67, 3.99, 5.35, 13.77, 25.5

**Table 7:** MLEs of the parameters, p-value and K-S statistic

Parameters	Stress (in kv)	K-S	p-value
$\lambda = 1.2065, \theta = 3.0873, \beta = 1.2189$	34	0.1562	0.5422
	36	0.1752	0.1290

## VI. Results and Conclusion

From Tables 2 to 5, it is concluded that the MLE is consistent and asymptotically normally distributed and one can realize that the biases and MSEs decrease as sample size increase for different values of parameters, which proves the efficiency of MLE.

The study deals with SSPALT by using an adaptive Type-II progressively hybrid censoring scheme for GIRD with a maximum likelihood estimation procedure. The numerical values of MLEs of distribution parameters are attained using the Newton-Raphson technique, and the performances of parameters are recorded in terms of MSEs and biases. Superb efficiency in estimating distribution parameters is examined under APHCT-II due to the huge sample size attained. So, APHCT-II is an excellent option for reliability practitioners to attain a greater efficiency of the distribution parameters. In the future, this work can be extended for different failure distributions under the Bayesian atmosphere.

## References

- [1] Rao, R., (1992). Equivalence of the tampered random variables and tampered failure rate models in ALT for a class of life distribution having the setting the clock back to zero property, *Communication in Statistics – Theory and Methods*, 21, 3: 647-664.
- [2] Lin, C. T., Ng, H. K. T. and Chan, P. S. (2009). Statistical inference of Type-II progressively hybrid censored data with Weibull lifetimes. *Communications in Statistics—Theory and Methods*, 38(10): 1710-1729.
- [3] Balakrishnan, N. and Ng, H.K.T. *Precedence-Type Tests and Applications*, John Wiley & Sons, Hoboken, NJ, 2006.
- [4] Epstein B. (1954). Truncated life-tests in the exponential case. *Annals of Mathematical Statistics*, 25: 555-564.
- [5] Balakrishnan, N. and Kundu, D. (2013). Hybrid censoring: Models, inferential results and applications. *Computational Statistics & Data Analysis*, 57(1): 166-209.
- [6] Childs, A., Chandrasekar, B., Balakrishnan, N. and Kundu, D. (2003). Exact likelihood inference based on Type-I and Type-II hybrid censored samples from the exponential distribution,

*Annals of the Institute of Statistical Mathematics*, 55: 319–330.

[7] Gupta, R.D. and Kundu, D. (1988). Hybrid censoring schemes with exponential failure distribution, *Communications in Statistics - Theory and Methods*, 27: 3065–3083.

[8] Kundu, D. (2007). On hybrid censoring Weibull distribution, *Journal of Statistical Planning and Inference*, 137: 2127–2142.

[9] Dey, S. and Pradhan, B. (2014). Generalized inverted exponential distribution under hybrid censoring. *Statistical methodology*, 18: 101-114.

[10] Salah, M.M., Ahmed, E.A., Alhussain, Z.A., Ahmed, H.H., El-Morshedy, M. and Eliwa, M.S. (2021). Statistical inferences for type-II hybrid censoring data from the alpha power exponential distribution, *PLoS One*, 16(1).

[11] Abdel-Ghaly, El-Khodary, A. E. H. and Ismail, A. A. (2003). Estimation and Optimal Design in Step Partially Accelerated Life Tests for the Pareto Distribution using Type-II Censoring, the Proceedings of the 15th annual conference on Statistics and Computer Modeling in Human and Social Sciences, *Faculty of Economics and Political Science, Cairo University, Egypt*, 16-29.

[12] Alam, I., Intezar, M. A., and Ahmed, A. (2021). Costs of Maintenance Service Policy: a New Approach on Constant Stress Partially Accelerated Life Test for Generalized Inverted Exponential Distribution. *Reliability: Theory & Applications*, 16(2 (62)): 45-57.

[13] Alam, I. and Ahmed, A. (2020). Parametric and interval estimation under step-stress partially accelerated life tests using adaptive type-II progressive hybrid censoring. *Annals of Data Science*, 1-13.

[14] Abd El-Raheem, A.M. (2019a). Optimal plans of constant-stress accelerated life tests for extension of the exponential distribution. *Journal of Testing and Evaluation*, 47(2), (in press).

[15] Abd El-Raheem, A.M. (2019b). Optimal plans and estimation of constant-stress accelerated life tests for the extension of the exponential distribution under type-I censoring. *Journal of Testing and Evaluation*, 47(5), (in press).

[16] Balakrishnan, N., Kundu, D., Ng, H.K.T. and Kannan, N. (2007). Point and interval estimation for a simple step-stress model with type-II censoring. *Journal of Quality Technology*, 39:35-47.

[17] Xiaolin, S. H. I., Pu, L. U., and Yimin, S. H. I. (2018). Inference and optimal design on step-stress partially accelerated life test for hybrid system with masked data. *Journal of Systems Engineering and Electronics*, 29(5): 1089-1100.

[18] Alam, I., and Ahmed, A. (2022). Inference on maintenance service policy under step-stress partially accelerated life tests using progressive censoring. *Journal of Statistical Computation and Simulation*, 92(4): 813-829.

[19] Alam, I. Islam, A.U. and Ahmed, A. (2020). Step Stress Partially Accelerated Life Tests and Estimating Costs of Maintenance Service Policy for the Power Function Distribution under Progressive Type-II Censoring, *Journal of Statistics Applications & Probability*, 9(2): 287-298.

[20] Ismail, A. A. (2014). Inference for a step-stress partially accelerated life test model with an adaptive Type-II progressively hybrid censored data from Weibull distribution. *Journal of Computational and Applied Mathematics*, 260: 533-542.

[21] Ismail, A. A. (2016). Statistical inference for a step-stress partially-accelerated life test model with an adaptive Type-I progressively hybrid censored data from Weibull distribution. *Statistical Papers*, 57(2): 271-301.

[22] Ismail, A. A. (2014). Likelihood inference for a step-stress partially accelerated life test model with Type-I progressively hybrid censored data from Weibull distribution. *Journal of Statistical Computation and Simulation*, 84(11): 2486-2494.

[23] El-Sagheer, R. M., and Ahsanullah, M. (2013). Statistical inference for a step-stress partially accelerated life test model based on progressively type-II-censored data from Lomax distribution. *Journal of Applied Statistical Science*, 21(4): 307-323.

- [24] Zhou, Y., Lu, Z., Shi, Y., and Cheng, K. (2019). The copula-based method for statistical analysis of step-stress accelerated life test with dependent competing failure modes. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 233(3): 401-418.
- [25] Srivastava, P.W., and Neha M. (2010). Optimum step-stress partially accelerated life tests for the truncated logistic distribution with censoring. *Applied Mathematical Modelling*, 34(10): 3166-3178.
- [26] Fatima, K., Naqash, S. and Ahmad, S.P. (2018). Exponentiated Generalized Inverse Rayleigh distribution with Applications in Medical Sciences, *Pakistan Journal of Statistics*, 34(5): 425-439.
- [27] Kumar, K. and Garg, R. (2014). Estimation of the parameters of randomly censored generalized inverted Rayleigh distribution, *International Journal of Agricultural and Statistical Sciences*, 10: 147–155.
- [28] Bakoban, R.A. and Abubaker, M.I. (2015). On the estimation of generalised inverted Rayleigh distribution with real data applications, *International Journal of Electronics Communication and Computer Engineering*, 6: 502–508.
- [29] Bakoban, R.A. and Abubaker, M.I. (2015). Parameters estimation of the generalized inverted Rayleigh distribution based on progressive type II censoring, *Advances and Applications in Statistics*, 47: 19–50.
- [30] DeGroot, M.H. and Goel, P.K. (1979). Bayesian and optimal design in partially accelerated life testing, *Naval Research Logistics*. 16 (2): 223-235.
- [31] Balakrishnan, N. (2007). Progressive censoring methodology: an appraisal, *Test* 16: 211-296.
- [32] Balakrishnan, N. and Aggarwala, R. (2000). *Progressive censoring: Theory, methods and applications*. Birkhauser, Boston.
- [33] Ng, H.K.T., Kundu, D. and Chan, P.S. (2009). Statistical analysis of exponential lifetimes under an adaptive hybrid type-II progressive censoring scheme. *Naval Research Logistics*, 56: 687-698.
- [34] Miller, R. and Nelson, W.B. (1983). Optimum simple step-stress plans for accelerated life testing, *IEEE Transactions on Reliability*, 32: 59-65.
- [35] Nelson W. *Accelerated Testing: Statistical Models, Test Plans and Data Analysis*. Wiley, New York, 1990.

# EFFECT OF CLASSICAL AND ROBUST REGRESSION ESTIMATORS IN THE CONTEXT OF HIGH- DIMENSIONAL DATA WITH MULTICOLLINEARITY AND OUTLIERS

Muthukrishnan R

•

Department of Statistics, Bharathiar University, Coimbatore, 641046, Tamil Nadu, India  
[muthukrishnan1970@gmail.com](mailto:muthukrishnan1970@gmail.com)

Karthika Ramakrishnan

•

Department of Statistics, Bharathiar University, Coimbatore, 641046, Tamil Nadu, India  
[karthikaramakrishnan45@gmail.com](mailto:karthikaramakrishnan45@gmail.com)

## Abstract

*Regression methods are used for the estimation and prediction in various fields of statistical study. It is a statistical method commonly used for determining the degree of relationship between a response and a number of explanatory variables. These explanatory variables may correlate each other and lead to multicollinearity. More than two predictor variables with high correlation show the existence of multicollinearity which results in the estimator having a high variance. Ordinary Least Square estimation fails to give a better regression estimator, when the model's presumptions are not met. This paper explores the various methods which can tolerate the problems of multicollinearity and outliers. This study compares different types of regression estimators such as Ordinary Least Square, Robust, Ridge, and Liu by computing various error values such as Mean Absolute Error, Root Mean Square Error, Mean Absolute Percentage Error and R2 under real environment that has both multicollinearity and outliers. To compare the fit of the aforementioned regression models, the Akaike Information Criterion was also calculated. According to the error measures and AIC this study concludes that the Liu regression estimator performs well when compared with the other estimation methods.*

**Keywords:** Regression, Multicollinearity, Outliers, Ridge, Liu

## I. Introduction

OLS estimator is the commonly used method to predict the parameters of a regression model when all the assumptions of the model are satisfied. The problems that would be affected the results of this method are multicollinearity and outliers. Multicollinearity is the situation where the explanatory variables have highly interdependent. It will increase the error values and thus the estimator may unreliable. Hoerl and Kennard [1] develop a regression procedure to control multicollinearity.

An outlier is a data observation that is unusual. It results the estimator to be not efficient

and changes the sign of the regression coefficients. Mendenhall and Sincich [7] give the definition of outlier as value with absolute standardized error greater than 3. Robust regression methods are usually used to obtain a better result when there are outliers. The main purpose of this paper is to compare different regression methods and identify the good one with better estimator in the presence of both multicollinearity and outlier.

The rest of the paper is structured as follows. In section 2, various regression estimators like OLS, Robust regression, Ridge regression and Liu regression are explained briefly. The performance of these regression procedures is studied under real environments and the results are summarized in section 3 and conclusion of the study is presented in the last section.

## II. Regression Methods

Regression analysis is used to draw inferences from data when there is a connection between the response and the predictor variables, according to Draper and Smith [9]. These approaches in machine learning come in a variety of forms, and their use depends on the type of data being used. It is the primary method to solve the problems in machine learning using data modeling. This paper includes the methods OLS, Robust, Ridge and Liu with the comparison of error measures under different real datasets having the presence of both outliers and multicollinearity. Outliers are identified by the Cook's distance procedure and the analysis has been carried out using R software.

### Ordinary Least Squares (OLS)

Ordinary Least Squares (OLS) is a technique used to predict the dependent variable ( $y$ ) with the help of a number of predictor variables ( $X$ ). It is the popularly used and Best Linear Unbiased Estimator (BLUE) when all the suppositions of the classical regression model are satisfied [Aitken [3]]. The general model of an OLS method with  $k$  independent variables is given by

$$y = X\beta + \varepsilon \quad (1)$$

where  $y$  is the  $(m \times 1)$  vector of response variable,  $X$  is a  $(m \times k)$  matrix,  $\beta$  is a  $(k \times 1)$  vector of an unknown regression parameters and  $\varepsilon$  is a  $(m \times 1)$  vector of residual term that is considered to be independently and identically distributed as normal with mean zero and fixed variance  $\sigma^2$ . The OLS estimator for the unknown parameter is

$$\widehat{\beta}_{OLS} = (X'X)^{-1} (X'y) \quad (2)$$

The performance of  $\widehat{\beta}_{OLS}$  will be statistically insignificant when multicollinearity exists between the explanatory variables.

### Robust Regression

Robust regression is an alternative approach to the classical regression model, when the nature of the data deviates from the key assumptions. The goal of robust regression is to get beyond some of the drawbacks of conventional regression analysis. Under normal distribution with no outliers, this robust method should produce approximately the similar results as OLS. In this section robust regression method like Least Trimmed Square (LTS), Least Median Square (LMS) and M were described.

### Least Trimmed Square (LTS)

Least Trimmed Square (LTS) is a robust regression method developed by Rousseeuw [11].



This method has an objective function of the lowest trimmed of squared residuals as

$$\sum_{i=1}^h r_i^2(\beta) \quad (3)$$

where  $r_1^2(\beta) \leq \dots \leq r_n^2(\beta)$  are the ordered residuals for  $i = 1, 2, \dots, n$ . The number of observations that are not trimmed from the dataset is denoted by  $h$ . LTS minimizes the trimmed sum of these squared residuals. Its estimator is equal to the OLS estimator only when  $h = n$ . The estimator  $\widehat{\beta}_{LTS}$  is estimated by minimizing the sum of residuals over  $\beta$  is given by

$$\widehat{\beta}_{LTS} = \arg \min_{\beta} \sum_{i=1}^h r_i^2(\beta) \quad (4)$$

LTS can be calculated for  $\alpha$ , the trimming proportion tending to 50%. It attains the maximum possible breakdown point at  $h = \binom{n}{2} + [(p+1)/2]$ . The computation of LTS estimator uses an algorithm called FAST-LTS of Rousseeuw and Van Driessen [10].

#### Least Median Square (LMS)

The Least Median Square (LMS) estimator was suggested by Rousseeuw [15]. Being a robust regression technique, the least median of squares method is not sensitive to outliers or other breaches of the normal model's assumption. In this method the sum is replaced by median in the method of least squares. Here the parameters are estimated by reducing the median of the squared residuals. The least median square estimator can be given by

$$\widehat{\beta}_{LMS} = \arg \min_{\beta} r_h^2(\beta) \quad (5)$$

where  $r_h^2(\beta)$  is a median. LMS is robust due to its breakdown value of 50%.

#### M Estimator (M)

M-estimators and their asymptotic properties were introduced by Huber [5]. Here the M stands for "maximum likelihood type". This method laid the foundation for the growth of the other robust methods in the context of regression estimators. M-estimation attempts to reduce the squared residuals  $r_h^2$  in OLS by another function of these  $r_h^2$

$$\min_i \sum_{h=1}^n \rho(r_h) \quad (6)$$

$\rho(r_h)$  is introduced for reducing the effect of outliers, where  $\rho$  is a definite positive, symmetric function with zero as its unique minimum. An algorithm was developed by Susanti et al. [16] for computing the M estimator.

#### Ridge Regression (RR)

Ridge Regression (RR) developed by Hoerl and Kennard [1] to give a reliable regression estimates even in the presence of multicollinearity. It produces an estimator that is biased and will be associated with the constant  $k$  that is used to reduce the bias. Hoerl et al. [2] find out a formula for the calculation of an optimal ridge constant  $k$  such that

$$k = \frac{p\hat{\sigma}^2}{\sum_{i=1}^p \hat{\alpha}_i^2} \quad (7)$$

where  $p$  denotes the number of predictor variables,  $\hat{\sigma}^2$  is the estimated variance and  $\hat{\alpha}_i$  is a conventional OLS regression parameter. Ridge regression depends on this constant  $k$  and will give a biased estimator as given below.

$$\widehat{\beta}_{RR} = (X'X + kI)^{-1} (X'y) \quad (8)$$

Liu Regression (LR)

Liu Regression (LR) is used to deal with datasets having multicollinearity. It was proposed by Liu [6]. It forms a new class of biased estimators called Liu estimators. These estimators are depending upon a biasing parameter  $d$  called the Liu parameter which lies between 0 and 1. The estimator of Liu regression is given by

$$\widehat{\beta}_{LR} = (X'X + I_p)^{-1} (X'y + d \widehat{\beta}_{OLS}) \tag{9}$$

where  $0 \leq d \leq 1$ ,  $I_p$  is the identity matrix of order  $p \times p$  and  $\widehat{\beta}_{OLS}$  is the OLS estimator.  $\widehat{\beta}_{LR}$  is the Liu estimator named by Akdeniz and Kaciranlar [4]. The  $d$  value with the minimum Mean Square Error (MSE) gives an efficient estimator. The R package *liureg* developed by Muhammad Imdadullah et al. [8] provides the tools for the computation of the estimator and the biasing parameter.

### III. Experimental Results

**Table 1:** Computed error measures and AIC under various regression methods (Acetylene Data)

Errors	Regression Methods					
	OLS	LTS	LMS	M	RR	LR
MAPE	0.284	0.195	0.194	0.262	0.210	0.195
MAE	0.008	0.009	0.009	0.007	0.034	0.007
RMSE	0.009	0.016	0.016	0.008	0.044	0.009
R <sup>2</sup>	0.900	0.711	0.728	0.918	0.990	0.992
AIC	93.24	383.09	331.99	183.86	42.10	42.09

**Table 2:** Computed error measures and AIC under various regression methods (Prostate Cancer Data)

Errors	Regression Methods					
	OLS	LTS	LMS	M	RR	LR
MAPE	2.33 (2.17)	2.51 (2.27)	2.59 (2.62)	2.34 (2.32)	2.23 (2.11)	2.21 (2.10)
MAE	0.56 (0.48)	0.64 (0.97)	0.69 (0.93)	0.56 (0.90)	0.56 (0.47)	0.55 (0.47)
RMSE	0.68 (0.58)	0.86 (1.32)	0.97 (1.30)	0.68 (1.25)	0.68 (0.58)	0.67 (0.57)
R <sup>2</sup>	0.66 (0.73)	0.46 (0.70)	0.31 (0.67)	0.66 (0.70)	0.67 (0.73)	0.67 (0.74)
AIC	219.54	91.35	117.08	62.51	62.99	60.29

(.)Without outlier

**Table 3:** Computed error measures and AIC under various regression methods (Hald Data)

Errors	Regression Methods					
	OLS	LTS	LMS	M	RR	LR
MAPE	0.04 (0.01)	0.04 (0.01)	0.12 (0.02)	0.03 (0.01)	0.03 (0.01)	0.03 (0.01)
MAE	0.78 (0.96)	0.91 (0.89)	3.89 (2.53)	0.75 (0.85)	0.81 (0.95)	0.75 (0.90)
RMSE	0.96 (1.29)	1.51 (1.64)	6.43 (3.70)	0.97 (1.28)	0.99 (1.32)	0.95 (1.25)
R <sup>2</sup>	0.996 (0.99)	0.991 (0.99)	0.841 (0.93)	0.996 (0.99)	0.996 (0.99)	0.997 (0.99)
AIC	47.67	39.69	46.95	22.19	10.13	6.72

(.)Without outlier

The experimental studies were carried out under real environments to study and compare the performance of the various regression procedures and thus obtained results were discussed in this section. The numerical studies have been conducted by considering three different case studies under real datasets in which the first data has the presence of multicollinearity and no outliers. The second one has a presence of moderate multicollinearity and outliers. And the third dataset has the presence of high multicollinearity and outliers. The outliers in the real data sets were detected and removed by using cook's distance, Cook [12] and the analysis has been carried out using R software. The presence and absence of multicollinearity has been identified by computing Variance Inflation Factors (VIF). The overall impact of the regressors' dependencies on each term's variance is measured by the VIF for each term in the model. VIF is 1 indicates that there are no correlation between the variables. Moderate correlation is indicated by a VIF between 1 and 5. VIF more than 5 is an indication of high multicollinearity between the variables. The error measures under OLS, LTS, LMS, M, RR and LR estimators were calculated by considering with and without outliers and are summarized in tables.

The Acetylene data set contains the percentage of n-heptane that is converted to acetylene, together with three independent variables. These are typical data from a chemical process, and a full quadratic response surface in each of the three regressors is sometimes regarded as a suitable preliminary model. It has 16 observations and 4 variables in which conversion of n-Heptane to Acetylene ( $y$ ) is considered as the dependent variable and Reactor Temperature ( $X_1$ ), Ratio of H<sub>2</sub> to n- Heptane ( $X_2$ ), Contact Time ( $X_3$ ) are taken as the independent variables. Cook's distance is used to check the presence of outliers in the dataset and there are none to be found. The VIF measures are higher than 10 and hence the dataset has high multicollinearity. The computed error measures and AIC value under various estimators of the dataset are given in Table 1.

The second data come from a study that looked at how males undergoing radial prostatectomy correlated their level of prostate-specific antigen with several clinical measures. The data is available in the R-package "lasso2". It has 97 observations, and 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 lcaivol (log of cancer volume). Seven outliers are found in this dataset and eliminated using Cook's distance. Since the VIFs of the independent variables are in between 1 and 5, there is an indication of moderate multicollinearity. The error measures and AIC calculated under different estimators are shown in Table 2.

Woods et al [17] was introduced the Hald or Portland Cement Data. It has been extensively analysed by Hald (1952), Hamaker (1962) and Kaciranlar et al (1999). This data frame contains 13 observations with four independent variables. They are tricalcium aluminate (X1), tricalcium silicate (X2), tetracalcium aluminoferrite (X3) and  $\beta$ -dicalcium silicate (X4). The response variable Y is the evolved heat after 180 days in a cement mix. Since the Variance Inflation Factors (VIF) of this Hald data set was greater than 10, the explanatory variables are highly correlated. As a result, the dataset has high multicollinearity. Also this data set has three outliers, which are found and eliminated by using Cook's distance. The computed error measures and AIC under various estimators are given in Table 3.

The results from Table 1, Table 2 and Table 3 demonstrate that the error levels for various estimators differ from one another, with LR having the lowest of all these. Also the AIC value of LR is minimum compared to the other estimators. Thus, for a dataset having high multicollinearity and outliers, the Liu (LR) regression estimator is more effective than the other estimators.

#### IV. Conclusion

Statistical learning techniques play a vital role in almost all the field of research study. Regression analysis is one of the statistical learning techniques. In general, the commonly used linear regression procedure 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 dataset. This paper explores various regression procedures such as OLS, Robust, Ridge and Liu. Further, evaluates their performance on different real datasets by considering the problems of multicollinearity and outliers by computing various error measures along with AIC value. On the basis of error and AIC values, the study concluded that the Liu regression procedure gives a better estimator for modelling the data when the dataset having multicollinearity and/or outliers. Further, this regression procedure can be beneficial to researchers, who work on machine learning techniques by considering the factors such as multicollinearity, outliers and high dimensionality.

#### References

- [1] Aitken, A.C. (1936). On least squares and linear combination of observations, *Proceedings of the Royal Statistical Society, Edinburgh*. 55: 42-48.
- [2] Akdeniz, F. and Kaciranlar, 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] Cook, R.D. (2000). Detection of influential observation in linear regression, *Technometrics*. 42: 65-68.
- [4] Draper, N. R. and Smith, H. Applied Regression Analysis, John Wiley & Sons, New York, 1998.
- [5] Hoerl, A. E. and Kennard, R.W. (1970). Ridge regression: Biased estimation for non-orthogonal problems, *Technometrics*.12:55–67.
- [6] Hoerl, A.E., Kennard, R.W. and Baldwin, K.F. (1975). Ridge regression: Some simulations, *Communications in Statistics- Theory and Methods*. 4:105-123.

- [7] Huber, P.J. (1964). Robust estimation of a location parameter, *The Annals of Mathematical Statistics*. 35:492-518.
- [8] Liu, K.J. (1993). A new class of biased estimate in linear regression, *Communications in Statistics*. 22:393-402.
- [9] Maronna, R.A. (2011). Robust ridge regression for high-dimensional data, *Technometrics*.53: 44–53.
- [10] Mendenhall, W. and Sincich, A. Second Course in Statistics: Regression Analysis, 2014.
- [11] Muhammad Imdadullah, Muhammad Aslam and Saima Altaf (2017). liureg: A Comprehensive R Package for the Liu Estimation of Linear Regression Model with Collinear Regressors, *The R Journal*.9:232.
- [12] Rousseeuw, P. J. (1984). Least Median of Squares Regression, *Journal of the American Statistical Association*.79:871–880.
- [13] Rousseeuw, P.J. (1985), Multivariate estimation with high breakdown point, *Mathematical statistics and applications*. 37:283-297.
- [14] Rousseeuw, P. J. and Leroy, A. M. Robust Regression and Outlier Detection, John Wiley & Sons, 1987.
- [15] Rousseeuw, P.J. and Van Driessen, K. (2006). Computing LTS regression for large data sets, *Data mining and knowledge discovery*. 12:29-45.
- [16] 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.
- [17] 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.

# EXPONENTIATED WEIBULL DISTRIBUTION: BAYESIAN ESTIMATION USING PROGRESSIVE TYPE I INTERVAL CENSORING

M. KUMAR, K P ASWATHI



Department of Mathematics, National Institute of Technology Calicut, 673601, Kerala, India  
mahesh@nitc.ac.in, aswathichithra01@gmail.com

## Abstract

*A three-parameter distribution known as the Generalized Weibull (GW) or Exponentiated Weibull distribution is studied in this work. We construct Baye's estimators for the unknown parameters and present reliability function using progressive type I interval censoring data. Two different loss functions, namely, squared error loss and general entropy loss functions are applied to derive Baye's estimators. It is observed that there is no closed-form solution for Baye's estimators as well as for MLE. Hence, Lindley's approximation procedure is applied to obtain Bayesian estimator of unknown parameters, and Newton Rapson method is employed to obtain MLE's numerically. The corresponding reliability function is derived. Monte Carlo simulation is used to obtain MLE. Further, the performance of MLE and Bayes estimators are compared in terms of their respective MSE and Relative errors. It is noted by numerical computation that MLE's performs better than Bayes estimators. In addition to this, Bayes estimators obtained using Squared error loss function and general entropy loss function are compared. It is observed through numerical computation that general entropy loss function is better in terms of MSE.*

**Keywords:** Bayesian inference, Exponentiated Weibull distribution, Lindley's approximation, Maximum likelihood function, Monte Carlo simulation, Relative error.

## 1. INTRODUCTION

When it comes to analyzing data and adapting it to practical situations, statistical distributions are crucial. Weibull or Gamma distributions are typically employed to fit the data in real-world scenarios. In survival analysis, the Gamma distribution has more major applications than all other distributions. But the main drawback of Gamma distribution is that the survival function cannot be obtained in closed form unless the shape parameter is an integer. This makes Weibull distribution more popular than Gamma distribution. Its survival function and failure rate are simple and easy to analyze. And this distribution is easy to handle the censoring data because of that, in recent years Weibull distribution is more popular in analyzing lifetime data. The Exponentiated Weibull distribution (EW) or Generalized Weibull distribution, was first described by [24] as a way to extend the Weibull family of two parameters by one more shape parameter. This distribution yields better fit than classic models such as exponential, gamma, Weibull, and log-normal distribution. Owing to its flexibility in modeling a wide range of industrial data, the EW distribution may be widely and efficiently applied in reliability applications. The fundamental feature of this family is that it supports bathtub-shaped as well as unimodal hazard rates, in addition to numerous monotone hazard rates. The applications of this distribution were first developed by [24]. Using five different classical failure data sets obtained for the Bus-motor

system, [25] demonstrated the potential unfulfillment and flexibility of EW distribution. It is a sub-model of a generic class of exponentiated distributions suggested by [11]. Generalized Weibull distribution was used by [26] to model survival data. The reliability and survival functions of this distribution were studied by [23]. Further statistical features and the importance of this distribution are addressed by [29] and [28]. The moments of the EW distribution were determined by [8]. EW distribution was compared on two-parameter Weibull and Gamma distributions in [32] study with regard to the failure rate. Exponentiated Weibull family distributed lifetime data observed under Type I progressive interval censoring with random removals were analyzed by [6]. Bayesian estimate and prediction for the EW distribution using both informative and non-informative priors was examined by [21]. After fitting a Weibull distribution and an EW distribution to the wind speed data and determining the mean and variance, [9] estimated the parameter using the MLE method. The non-Bayesian estimators methods for parameters of EW distribution studied by [4]. The discrete case of EW distribution studied by [30]. The entropy and stress-strength model of EW distribution studied by [3]. Numerical estimation of parameters of EW distribution based on generalized progressive hybrid censoring scheme studied by [10]. In recent years, estimation of EW distribution under progressive type II censored data studied by [22].

The fundamental feature of this family is that it supports bathtub-shaped as well as unimodal hazard rates, in addition to numerous monotone hazard rates. The EW distribution is defined in the following way.

It has distribution function given by

$$F(x; \alpha, \beta, \lambda) = (1 - e^{-(\lambda x)^\beta})^\alpha, \quad x > 0 \text{ and } \alpha, \beta, \lambda > 0 \quad (1)$$

and therefore its probability density function is of the form

$$f(x; \alpha, \beta, \lambda) = \alpha \beta \lambda^\beta x^{\beta-1} e^{-(\lambda x)^\beta} ((1 - e^{-(\lambda x)^\beta})^{\alpha-1}) \quad (2)$$

The corresponding reliability function is given by

$$R(x; \alpha, \lambda) = 1 - (1 - e^{-(\lambda x)^\beta})^\alpha \quad (3)$$

and the hazard rate is

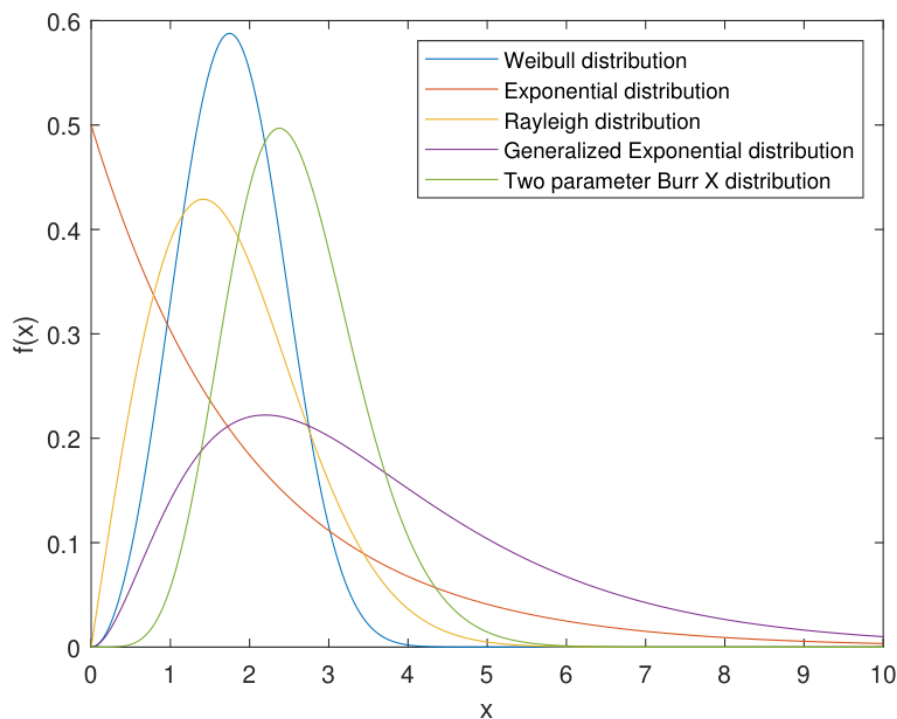
$$h(x) = \frac{f(x)}{1 - F(x)}, \quad x > 0 \quad (4)$$

Note here that, the shape parameters are  $\alpha$  and  $\beta$ , and the scale parameter is  $\lambda$ .

Several well known distributions are particular cases of the EW distribution. For example, the Exponential distribution is the case when  $\alpha = 1$  and  $\beta = 1$ , the Weibull Distribution is defined with  $\alpha = 1$ , Rayleigh Distribution with  $\alpha = 1$  and  $\beta = 2$ ,  $\beta = 1$  Generalized Exponential (GE) Distribution studied by [12], [13], [15], [17] [18], [37] and [39].  $\beta = 2$  Two parameter Burr Type X or Exponentiated Rayleigh(ER) or Generalized Rayleigh(GR) Distribution studied by [2], [36], [16], [14], [43], [38], [5] and [27] among others. Fig.(1) and Fig.(2) represents the many forms of these distributions graphically.

It was discovered that the EW family is a very versatile family that may be utilized to describe many sorts of skewed lifetime data. In reliability analysis, censoring is quite prevalent. It occurs when specific failure times for a subset of test units in an experiment are detected.

In industrial life testing and medical survival analysis, very often the object of interest is lost or withdrawn before failure or the object's lifetime is only known within an interval. Hence, the obtained sample is called a censored sample (or an incomplete sample). The most common censoring schemes are type-I censoring, type-II censoring and progressive censoring. For type-I censoring, life testing ends at a pre-scheduled time and for type-II censoring, life testing ends whenever the number of lifetimes is reached. In type-I and type-II censoring schemes, the tested items are allowed to be withdrawn only at the end-of-life testing. In the progressive censoring



**Figure 1:** Graph of EW distribution for different values of  $\alpha, \beta$  and for fixed  $\lambda = 0.5$

scheme, the tested items are allowed to be withdrawn at some time before the end-of-life testing. See [7] for more information about progressive censoring combined with type-I or type-II and their applications. Using the concepts of progressive censoring, type I censoring, and interval censoring, [1] developed progressive type I interval censoring. Combining progressive censoring and type-II censoring, [18] and [34] investigated Bayesian inference for Weibull distribution and generalized exponential (GE) distribution, respectively. It should be emphasized that in many practical situations, unit lifetime is set on an interval, therefore type I interval censoring is beneficial in these instances (see, [1]). It may be noted that in real-life situations, the lifetime of units may not be recorded precisely due to some reasons, such as technical problems, non-availability of experimental resources or due to some unknown human errors, or some cost-saving measures employed by the industry. Thus such censored data generated can be used effectively in analyzing the reliability characteristics of well-known distribution, such as the more general class of distribution, namely, EW distribution, which gained lots of importance in recent times. The importance of progressive type-I interval censoring in handling practical problem has been studied by authors, namely, [6] and [19]. The concept of progressive type-I interval censoring to the Weibull distribution and compared many different estimation methods for two parameters in the Weibull distribution via simulation introduced by [31]. The recent study about progressive type I interval censoring is On inference and design under progressive type-I interval censoring scheme for inverse Gaussian lifetime model by [40]. A Study on the experimental design for the lifetime performance index of Rayleigh lifetime distribution under progressive type I interval censoring by [44]. Optimal design of accelerated life tests under progressive type I interval censoring with random removals by [46], and experimental design for progressive type I interval censoring on the lifetime performance index of Chen lifetime distribution by [45]. All the works available in the literature aims at obtaining estimators of parameters of EW distribution based upon, either data obtained from complete censoring or from type I censoring, type II censoring, hybrid censoring, etc. No work in the literature addresses the estimation of parameters of EW distribution based upon progressive type I interval-censored data. Therefore we



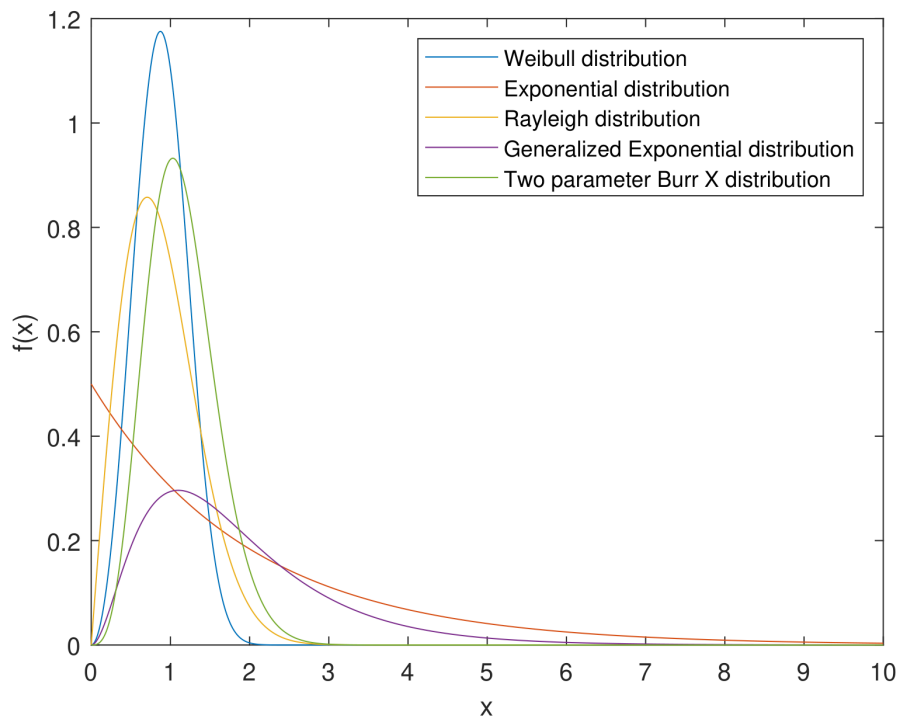


Figure 2: Graph of EW distribution for different values of  $\alpha, \beta$  and for fixed  $\lambda = 1$

consider in the next sections the derivation of MLE and Bayes estimators from data obtained via progressive type I interval censoring for EW distribution. Section 2 provides a brief fundamental required for obtaining estimators based on censored data. Some simulation results and discussion based upon the results obtained are presented in Section 3. The conclusion and future scope of research are given in Section 4.

## 2. BAYESIAN ESTIMATION USING PROGRESSIVE TYPE I INTERVAL CENSORED DATA

In this section, we discuss the brief overview of the terms used in this paper and the procedure of obtaining Bayes's estimators for Parameters and reliability function of EW distribution.

### 2.1. Progressive type I interval censored data and the likelihood function

Statistical inference for exponential distributions using progressive type I interval censored data and pioneered type I interval censoring in a progressive censoring scheme developed by [1]. Under progressive type I interval censoring, observations are only known within two successively pre-scheduled timeframes, and items may be allowed to be deleted at pre-scheduled time points. The progressively type I interval censored sample may be generated in the following manner:

Let  $n$  units be put on a life testing platform simultaneously at time  $t_0 = 0$  and under examination at  $m$  pre-specified time periods  $t_1 < t_2 < \dots < t_m$  where  $t_m$  is the predetermined time to end the experiment. The number of failures  $X_i$  within  $(t_{i-1}, t_i]$  is recorded and  $R_i$  surviving items are randomly removed from the life testing at the  $i^{th}$  inspection time,  $t_i$ , for  $i = 1, 2, \dots, m$ . Because the number of surviving items,  $Y_i$ , is a random variable and the precise number of items removed at time schedule  $t_i$  should not be larger than  $Y_i$ ,  $R_i$  might be calculated by a pre-specified percentage of the remaining surviving units at  $t_i$  for given  $i = 1, 2, \dots, m$ .

For example, given certain pre-specified percentage values say,  $p_1, p_2, \dots, p_{m-1}$  and  $p_m = 1$ ,  $R_i$  can be determined by using  $R_i = \text{floor}[p_i Y_i]$  at each inspection time  $t_i$ , where  $\text{floor}[x]$  yields

$x$ 's biggest integer. Therefore, a progressive type-I interval censored sample with size  $n$ , can be denoted as  $D = (X_i, R_i, t_i)_m, i = 1, 2, \dots, m$ . If  $R_i = 0, i = 1, 2, \dots, m - 1$  and  $R_m = n - \sum_{i=1}^m X_i$ , then the type-I interval-censored sample gradually shrinks to the typical interval-censored sample. Given the progressively type-I censored data,  $D = (X_i, R_i, t_i)_m$  of size  $n$ , from a continuous lifetime distribution with CDF  $F(t; \kappa)$ , then the likelihood function is given as follows

$$L(D | \kappa) \propto \prod_{i=1}^m [F(t_i; \kappa) - F(t_{i-1}; \kappa)]^{X_i} [1 - F(t_i; \kappa)]^{R_i}, \quad (5)$$

where  $t_0 = 0$  and  $\theta$  is the parameter vector. The more details of progressive type I interval censoring can be seen in [33].

For the  $EW(\alpha, \lambda, \beta)$ , the likelihood function (5) can be defined in the following manner:

$$L(D | \alpha, \lambda, \beta) \propto \prod_{i=1}^m [(1 - e^{-(\lambda t_i)^\beta})^\alpha - (1 - e^{-(\lambda t_{i-1})^\beta})^\alpha]^{X_i} [1 - (1 - e^{-(\lambda t_i)^\beta})^\alpha]^{R_i}. \quad (6)$$

The log-likelihood function is thus given by

$$l(\alpha, \lambda, \beta) \propto \sum_{i=1}^m X_i \ln[(1 - e^{-(\lambda t_i)^\beta})^\alpha - (1 - e^{-(\lambda t_{i-1})^\beta})^\alpha] + R_i \ln[1 - (1 - e^{-(\lambda t_i)^\beta})^\alpha]. \quad (7)$$

## 2.2. Maximum likelihood function

In this section, we discuss the Maximum likelihood estimation to estimate unknown parameters  $\alpha, \lambda, \beta$ , and the reliability function  $R(t)$  for EW distribution defined in (1) using the numerical method.

By setting the derivatives of the log likelihood function with respect to  $\alpha, \lambda$  or  $\beta$  to zero, the MLEs of  $\alpha, \lambda$  and  $\beta$  are the solutions to the following likelihood equations

$$\sum_{i=1}^m \left[ X_i \left( \frac{\frac{\partial F_i}{\partial \alpha} - \frac{\partial F_{i-1}}{\partial \alpha}}{F_i - F_{i-1}} \right) \right] = \sum_{i=1}^m \left[ R_i \left( \frac{\frac{\partial F_i}{\partial \alpha}}{1 - F_i} \right) \right]$$

$$\sum_{i=1}^m \left[ X_i \left( \frac{\frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda}}{F_i - F_{i-1}} \right) \right] = \sum_{i=1}^m \left[ R_i \left( \frac{\frac{\partial F_i}{\partial \lambda}}{1 - F_i} \right) \right]$$

and

$$\sum_{i=1}^m \left[ X_i \left( \frac{\frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta}}{F_i - F_{i-1}} \right) \right] = \sum_{i=1}^m \left[ R_i \left( \frac{\frac{\partial F_i}{\partial \beta}}{1 - F_i} \right) \right]$$

There is no closed form of the solution to the above equations and numerical methods can be used to obtain the MLEs from the above likelihood equations. Since there is no closed form of the MLE, Newton-Raphson method is introduced as follows for finding the MLEs of  $\alpha, \lambda$  and  $\beta$ .

One of the most used methods for optimization in statistics is the Newton-Raphson method (or Newton's rule). Assume that  $l$  only involves a one-dimensional parameter and that  $\bar{\theta}$  is our current best guess on the maximum of  $l(\theta)$ .  $l(\theta)$  can be approximated by employing a Taylor series expansion around  $\bar{\theta}$ . Hence we have

$$\bar{l}_{\bar{\theta}}(\theta) = l(\bar{\theta}) + l'(\bar{\theta})(\theta - \bar{\theta}) + \frac{1}{2} l''(\bar{\theta})(\theta - \bar{\theta})^2.$$

When  $\vartheta$  is close to  $\bar{\vartheta}$ , the difference  $l(\vartheta) - \bar{l}_{(\bar{\vartheta})}(\vartheta)$  is small. The maximum value of  $\bar{l}_{(\bar{\vartheta})}(\vartheta)$  is closer to the maximum value of  $l(\vartheta)$  than  $l(\bar{\vartheta})$ .

The gradient of  $\bar{l}_{(\bar{\vartheta})}(\vartheta)$  at  $\vartheta$  is

$$\bar{l}'_{(\bar{\vartheta})}(\vartheta) = l'(\bar{\vartheta}) + l''(\bar{\vartheta})(\vartheta - \bar{\vartheta})$$

and the Hessian or second derivative is

$$\bar{l}''_{(\bar{\vartheta})}(\vartheta) = l''(\bar{\vartheta}).$$

At the point  $\bar{\vartheta}$ ,  $l(\bar{\vartheta})$  and  $\bar{l}_{(\bar{\vartheta})}(\bar{\vartheta})$  have equal first and second derivatives. In the case of log likelihood function Hessian is same as the minus of observed information evaluated at  $\vartheta = \bar{\vartheta}$ ,  $l''(\bar{\vartheta}) = -J(\bar{\vartheta})$ . In the optimum point of the approximation,  $\bar{l}_{(\bar{\vartheta})}(\vartheta)$  has a gradient equal to zero, giving the following equation:

$$l''(\bar{\vartheta})(\vartheta - \bar{\vartheta}) = -l'(\bar{\vartheta}).$$

Solving with respect to  $\vartheta$ , we get

$$\vartheta = \bar{\vartheta} - \frac{l'(\bar{\vartheta})}{l''(\bar{\vartheta})}.$$

This gives a procedure for optimizing  $\bar{l}_{(\bar{\vartheta})}(\vartheta)$ . An iterative procedure for optimizing  $l(\vartheta)$  is given by

$$\vartheta^{(s+1)} = \vartheta^{(s)} - \frac{l'(\vartheta^{(s)})}{l''(\vartheta^{(s)})}$$

which is the Newton-Raphson Method. The procedure is run until there is no significant difference between  $\vartheta^{(s)}$  and  $\vartheta^{(s+1)}$ .

When  $l(\vartheta)$  is a log likelihood function, this algorithm can be written as

$$\vartheta^{(s+1)} = \vartheta^{(s)} - \frac{s(\vartheta^{(s)})}{J(\vartheta^{(s)})}$$

where  $s(\vartheta)$  is the score function while  $J(\vartheta)$  is the observed information matrix.

### 2.3. Bayesian Estimation

In this section, we discuss the Bayesian technique to estimate unknown parameters  $\alpha, \lambda, \beta$ , and the reliability function  $R(t)$  using the Squared error loss and general entropy loss functions. Assume that all parameters, namely,  $\alpha, \lambda$  and  $\beta$  of EW distributions are unknown and independent. We address the problem of constructing Baye's estimators for these parameters. We assume non-informative priors for  $\alpha$  and  $\beta$ , and conjugate prior for  $\lambda$ . The reason for choosing these prior forms is due to their simplicity of in obtaining mathematically treatable posterior distributions. We observe that such priors are successfully applied by many authors, namely, [33] and [35]. The following equations give respective definition of prior densities.

$$\pi_1(\alpha) = \frac{1}{\alpha}, \quad \alpha > 0 \tag{8}$$

$$\pi_2(\lambda) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda}, \quad \lambda > 0, a, b > 0 \tag{9}$$

and

$$\pi_3(\beta) = \frac{1}{\beta}, \quad \beta > 0 \tag{10}$$

respectively where  $\Gamma(\cdot)$  is the gamma function.

We consider two different forms of loss functions in estimating the parameters of EW density. The first one is a symmetric loss function, the squared error loss function (SEL), which is given by

$$L_1(\zeta, \hat{\zeta}) = (\hat{\zeta} - \zeta)^2, \tag{11}$$

where  $\hat{\zeta}$  is the estimate of parameter  $\zeta$ . Then the Bayesian estimate of any function  $q = q(\alpha, \lambda, \beta)$  is obtained by considering the following equation

$$\hat{q} = E(q | D) = \frac{\int_{\alpha} \int_{\lambda} \int_{\beta} q(\alpha, \lambda, \beta) l(\alpha, \lambda, \beta) \pi_1(\alpha) \pi_2(\lambda) \pi_3(\beta) d\alpha d\lambda d\beta}{\int_{\alpha} \int_{\lambda} \int_{\beta} l(\alpha, \lambda, \beta) \pi_1(\alpha) \pi_2(\lambda) \pi_3(\beta) d\alpha d\lambda d\beta} \tag{12}$$

The second loss function, is the generalization of the Entropy loss used by several authors ([41] and [42]). The General Entropy loss (GEL) is defined as:

$$L_2(\zeta, \hat{\zeta}) \propto \left(\frac{\hat{\zeta}}{\zeta}\right)^c - c \log \frac{\hat{\zeta}}{\zeta} - 1, \tag{13}$$

where  $\hat{\zeta}$  is an estimate of parameter  $\zeta$ . It may be noted that when  $c > 0$ , a positive error causes more serious consequences than a negative error. On the other hand, when  $c < 0$ , a negative error causes more serious consequences than a positive error. Then the Bayesian estimator of  $q(\alpha, \lambda, \beta)$  under this general entropy loss function is

$$\hat{q}_{GEL} = [E(q^{-c})]^{-\frac{1}{c}}, \tag{14}$$

provided that  $E(q^{-c})$  exists and is finite. It can be shown that, when  $c = 1$ , the Bayes estimate (12) coincides with the Bayes estimate under the weighted squared error loss function. Similarly, when  $c = -1$  the Bayes estimate (14) coincides with the Bayes estimate under squared error loss function. The equations (12) and (14) cannot be solved for obtaining closed form solutions. Hence, we resort to well known Lindley approximation [20] procedure to evaluate the ratio of integrals involved in (12) and (14). Note that the Lindley approximation procedure is successively employed by authors, such as [18] to obtain Bayesian estimators. Next, the Bayesian posterior expectation function of a parameter vector  $\eta$ , say  $h(\eta)$  is obtained by using the following equation

$$\hat{h}_B = E(h(\eta) | D) = \frac{\int_{\eta} h(\eta) l(\eta) \pi(\eta) d\eta}{\int_{\eta} l(\eta) \pi(\eta) d\eta}, \tag{15}$$

Recall that in the above expression  $l(\eta)$  denotes log likelihood function,  $\pi(\eta)$  denotes prior density and  $D$  denotes the data obtained using progressive type I interval censoring.

By [20], if  $n$ , the sample size is sufficiently large, every ratio of the integral of the form,

$$\begin{aligned} \hat{h} &= E[v(\eta_1, \eta_2, \eta_3)] \\ &= \frac{\int_{\eta_1, \eta_2, \eta_3} v(\eta_1, \eta_2, \eta_3) e^{l(\eta_1, \eta_2, \eta_3) + G(\eta_1, \eta_2, \eta_3)} d(\eta_1, \eta_2, \eta_3)}{\int_{\eta_1, \eta_2, \eta_3} e^{l(\eta_1, \eta_2, \eta_3) + G(\eta_1, \eta_2, \eta_3)} d(\eta_1, \eta_2, \eta_3)} \end{aligned}$$

where

$v(\eta) = v(\eta_1, \eta_2, \eta_3)$  is a function of  $\eta_1, \eta_2$  or  $\eta_3$  only,

$l(\eta_1, \eta_2, \eta_3)$  is log of likelihood function,

and  $G(\eta_1, \eta_2, \eta_3)$  is log joint prior of  $\eta_1, \eta_2$  and  $\eta_3$ ,

can be evaluated as

$$\begin{aligned} \hat{h} &= v(\hat{\eta}_1, \hat{\eta}_2, \hat{\eta}_3) + (v_1 a_1 + v_2 a_2 + v_3 a_3 + a_4 + a_5) + \frac{1}{2} [A(v_1 \sigma_{11} + v_2 \sigma_{12} + v_3 \sigma_{13}) + \\ &\quad B(v_1 \sigma_{21} + v_2 \sigma_{22} + v_3 \sigma_{23}) + C(v_1 \sigma_{31} + v_2 \sigma_{32} + v_3 \sigma_{33})] \end{aligned}$$

where

$\hat{\eta}_1, \hat{\eta}_2$  and  $\hat{\eta}_3$  are the MLE of  $\eta_1, \eta_2$  and  $\eta_3$  respectively.

$$\begin{aligned} a_i &= \rho_1\sigma_{i1} + \rho_2\sigma_{i2} + \rho_3\sigma_{i3}, \quad i = 1, 2, 3, \\ a_4 &= v_{12}\sigma_{12} + v_{13}\sigma_{13} + v_{23}\sigma_{23}, \\ a_5 &= \frac{1}{2}(v_{11}\sigma_{11} + v_{22}\sigma_{22} + v_{33}\sigma_{33}), \\ A &= \sigma_{11}l_{111} + 2\sigma_{12}l_{121} + 2\sigma_{13}l_{131} + 2\sigma_{23}l_{231} + \sigma_{22}l_{221} + \sigma_{33}l_{331}, \\ B &= \sigma_{11}l_{112} + 2\sigma_{12}l_{122} + 2\sigma_{13}l_{132} + 2\sigma_{23}l_{232} + \sigma_{22}l_{222} + \sigma_{33}l_{332}, \\ C &= \sigma_{11}l_{113} + 2\sigma_{12}l_{123} + 2\sigma_{13}l_{133} + 2\sigma_{23}l_{233} + \sigma_{22}l_{223} + \sigma_{33}l_{333} \end{aligned}$$

and subscripts 1,2,3 on the right-hand sides refer to  $\eta_1, \eta_2, \eta_3$  respectively and,

$$\begin{aligned} \rho_i &= \frac{\partial \rho}{\partial \eta_i}, \quad v_i = \frac{\partial v(\eta_1, \eta_2, \eta_3)}{\partial \eta_i}, \quad i = 1, 2, 3, \\ v_{ij} &= \frac{\partial^2 v(\eta_1, \eta_2, \eta_3)}{\partial \eta_i \partial \eta_j}, \quad i, j = 1, 2, 3, \end{aligned}$$

$$l_{ij} = \frac{\partial^2 l(\eta_1, \eta_2, \eta_3)}{\partial \eta_i \partial \eta_j}, \quad i, j = 1, 2, 3, \tag{16}$$

$$l_{ijk} = \frac{\partial^3 l(\eta_1, \eta_2, \eta_3)}{\partial \eta_i \partial \eta_j \partial \eta_k}, \quad i, j, k = 1, 2, 3, \tag{17}$$

and  $\sigma_{ij}$  is the  $(i, j)^{th}$  element of the inverse of the matrix  $\{l_{ij}\}$ , which is given by

$$I(\alpha, \lambda, \beta) = \begin{bmatrix} -\frac{\partial^2 l}{\partial \alpha^2} & -\frac{\partial^2 l}{\partial \alpha \partial \lambda} & -\frac{\partial^2 l}{\partial \alpha \partial \beta} \\ -\frac{\partial^2 l}{\partial \lambda \partial \alpha} & -\frac{\partial^2 l}{\partial \lambda^2} & -\frac{\partial^2 l}{\partial \lambda \partial \beta} \\ -\frac{\partial^2 l}{\partial \alpha \partial \beta} & -\frac{\partial^2 l}{\partial \beta \partial \lambda} & -\frac{\partial^2 l}{\partial \beta^2} \end{bmatrix}$$

Now by equations (8), (9) and (10), by using independence of  $\alpha, \lambda, \beta$ , the joint prior distribution of these three parameters is given by

$$\pi(\alpha, \lambda, \beta) = \frac{b^a \lambda^{a-1} e^{-b\lambda}}{\beta \alpha \Gamma(a)}, \quad \alpha, \lambda, \beta > 0, a, b > 0. \tag{18}$$

Let

$$\begin{aligned} \rho &= \ln \pi(\alpha, \lambda, \beta) \\ &= a \ln b + (a - 1) \ln \lambda - b\lambda - \ln \beta - \ln \alpha - \ln \Gamma(a). \end{aligned} \tag{19}$$

Differentiating (19) with respect to  $\alpha, \lambda, \beta$  respectively, we have

$$\rho_1 = -\frac{1}{\alpha}, \quad \rho_2 = \frac{a-1}{\lambda} - b, \quad \rho_3 = -\frac{1}{\beta}.$$

Observe that while performing progressive type I interval censoring, there are 'm' pre-specified time periods, say,  $t_1 < t_2 < \dots < t_m$ , where  $t_m$  is pre-specified stopping time of experiment. Now let us define the pdf for EW distribution for  $1 \leq i \leq m$  as  $F_i = (1 - e^{-(\lambda x)^\beta})^\alpha$ ,  $i = 1, 2, 3, \dots, m$ . Now from the expression (5) we have

$$l \propto \sum_{i=1}^m \{X_i \ln [F_i - F_{i-1}] + R_i \ln [1 - F_i]\}$$

Then,

$$\begin{aligned}
 l_1 &= \sum_{i=1}^m \left[ X_i \left( \frac{\frac{\partial F_i}{\partial \alpha} - \frac{\partial F_{i-1}}{\partial \alpha}}{F_i - F_{i-1}} \right) - R_i \left( \frac{\frac{\partial F_i}{\partial \alpha}}{1 - F_i} \right) \right] \\
 l_2 &= \sum_{i=1}^m \left[ X_i \left( \frac{\frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda}}{F_i - F_{i-1}} \right) - R_i \left( \frac{\frac{\partial F_i}{\partial \lambda}}{1 - F_i} \right) \right] \\
 l_3 &= \sum_{i=1}^m \left[ X_i \left( \frac{\frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta}}{F_i - F_{i-1}} \right) - R_i \left( \frac{\frac{\partial F_i}{\partial \beta}}{1 - F_i} \right) \right]
 \end{aligned}$$

From equation (16), the values of  $l_{ij}$ , ( $i, j = 1, 2, 3$ ) can be obtained as follows

$$\begin{aligned}
 l_{11} &= \sum_{i=1}^m \left\{ X_i \left[ \frac{(F_i - F_{i-1}) \left( \frac{\partial^2 F_i}{\partial \alpha^2} - \frac{\partial^2 F_{i-1}}{\partial \alpha^2} \right) - \left( \frac{\partial F_i}{\partial \alpha} - \frac{\partial F_{i-1}}{\partial \alpha} \right)^2}{(F_i - F_{i-1})^2} \right] \right. \\
 &\quad \left. - R_i \left[ \frac{(1 - F_i) \frac{\partial^2 F_i}{\partial \alpha^2} + \left( \frac{\partial F_i}{\partial \alpha} \right)^2}{(1 - F_i)^2} \right] \right\}, \\
 l_{12} &= \sum_{i=1}^m \left\{ X_i \left[ \frac{(F_i - F_{i-1}) \left( \frac{\partial^2 F_i}{\partial \alpha \partial \lambda} - \frac{\partial^2 F_{i-1}}{\partial \alpha \partial \lambda} \right) - \left( \frac{\partial F_i}{\partial \alpha} - \frac{\partial F_{i-1}}{\partial \alpha} \right) \left( \frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda} \right)}{(F_i - F_{i-1})^2} \right] \right. \\
 &\quad \left. - R_i \left[ \frac{(1 - F_i) \frac{\partial^2 F_i}{\partial \alpha \lambda} + \left( \frac{\partial F_i}{\partial \alpha} \right) \left( \frac{\partial F_i}{\partial \lambda} \right)}{(1 - F_i)^2} \right] \right\} \\
 &= l_{21}, \\
 l_{13} &= \sum_{i=1}^m \left\{ X_i \left[ \frac{(F_i - F_{i-1}) \left( \frac{\partial^2 F_i}{\partial \alpha \partial \beta} - \frac{\partial^2 F_{i-1}}{\partial \alpha \partial \beta} \right) - \left( \frac{\partial F_i}{\partial \alpha} - \frac{\partial F_{i-1}}{\partial \alpha} \right) \left( \frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta} \right)}{(F_i - F_{i-1})^2} \right] \right. \\
 &\quad \left. - R_i \left[ \frac{(1 - F_i) \frac{\partial^2 F_i}{\partial \alpha \beta} + \left( \frac{\partial F_i}{\partial \alpha} \right) \left( \frac{\partial F_i}{\partial \beta} \right)}{(1 - F_i)^2} \right] \right\} \\
 &= l_{31}, \\
 l_{22} &= \sum_{i=1}^m \left\{ X_i \left[ \frac{(F_i - F_{i-1}) \left( \frac{\partial^2 F_i}{\partial \lambda^2} - \frac{\partial^2 F_{i-1}}{\partial \lambda^2} \right) - \left( \frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda} \right)^2}{(F_i - F_{i-1})^2} \right] \right. \\
 &\quad \left. - R_i \left[ \frac{(1 - F_i) \frac{\partial^2 F_i}{\partial \lambda^2} + \left( \frac{\partial F_i}{\partial \lambda} \right)^2}{(1 - F_i)^2} \right] \right\}, \\
 l_{23} &= \sum_{i=1}^m \left\{ X_i \left[ \frac{(F_i - F_{i-1}) \left( \frac{\partial^2 F_i}{\partial \lambda \partial \beta} - \frac{\partial^2 F_{i-1}}{\partial \lambda \partial \beta} \right) - \left( \frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda} \right) \left( \frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta} \right)}{(F_i - F_{i-1})^2} \right] \right. \\
 &\quad \left. - R_i \left[ \frac{(1 - F_i) \frac{\partial^2 F_i}{\partial \lambda \beta} + \left( \frac{\partial F_i}{\partial \lambda} \right) \left( \frac{\partial F_i}{\partial \beta} \right)}{(1 - F_i)^2} \right] \right\} \\
 &= l_{32},
 \end{aligned}$$

$$l_{33} = \sum_{i=1}^m \left\{ X_i \left[ \frac{(F_i - F_{i-1}) \left( \frac{\partial^2 F_i}{\partial \beta^2} - \frac{\partial^2 F_{i-1}}{\partial \beta^2} \right) - \left( \frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta} \right)^2}{(F_i - F_{i-1})^2} \right] - R_i \left[ \frac{(1 - F_i) \frac{\partial^2 F_i}{\partial \beta^2} + \left( \frac{\partial F_i}{\partial \beta} \right)^2}{(1 - F_i)^2} \right] \right\}.$$

Similarly, from equation (17), the values for  $l_{ijk}(i, j, k = 1, 2, 3)$  can be obtained.

Now we proceed to obtain Bayes estimators of the parameters  $\alpha, \lambda, \beta$  of EW distribution function, and the reliability function  $R(t)$  under squared error loss function. Recall that  $v(\hat{\alpha}_s, \hat{\lambda}_s, \hat{\beta}_s)$  denotes a function MLE's for  $\alpha, \lambda, \beta$ . Hence we present here the Bayes estimators of  $\alpha, \lambda, \beta$  and  $R(t)$  via following equations:

- $v(\hat{\alpha}, \hat{\lambda}, \hat{\beta}) = \hat{\alpha}$  then

$$\hat{\alpha}_s = \hat{\alpha} - \frac{1}{\hat{\alpha}} \sigma_{11} + \frac{a-1-b\hat{\lambda}}{\hat{\lambda}} \sigma_{12} - \frac{1}{\hat{\beta}} \sigma_{13} + \frac{1}{2} [A\sigma_{11} + B\sigma_{21} + C\sigma_{31}], \quad (20)$$

- $v(\hat{\alpha}, \hat{\lambda}, \hat{\beta}) = \hat{\lambda}$  then

$$\hat{\lambda}_s = \hat{\lambda} - \frac{1}{\hat{\alpha}} \sigma_{21} + \frac{a-1-b\hat{\lambda}}{\hat{\lambda}} \sigma_{22} - \frac{1}{\hat{\beta}} \sigma_{23} + \frac{1}{2} [A\sigma_{12} + B\sigma_{22} + C\sigma_{32}], \quad (21)$$

- $v(\hat{\alpha}, \hat{\lambda}, \hat{\beta}) = \hat{\beta}$  then

$$\hat{\beta}_s = \hat{\beta} - \frac{1}{\hat{\alpha}} \sigma_{31} + \frac{a-1-b\hat{\lambda}}{\hat{\lambda}} \sigma_{32} - \frac{1}{\hat{\beta}} \sigma_{33} + \frac{1}{2} [A\sigma_{13} + B\sigma_{23} + C\sigma_{33}], \quad (22)$$

- $v(\hat{\alpha}, \hat{\lambda}, \hat{\beta}) = R(\hat{x})$  then

$$\hat{R}_s = \hat{R} + (\hat{R}_1 a_1 + \hat{R}_2 a_2 + \hat{R}_3 a_3 + a_4 + a_5) + \frac{1}{2} [A(\hat{R}_1 \sigma_{11} + \hat{R}_2 \sigma_{12} + \hat{R}_3 \sigma_{13}) + B(\hat{R}_1 \sigma_{21} + \hat{R}_2 \sigma_{22} + \hat{R}_3 \sigma_{23}) + C(\hat{R}_1 \sigma_{31} + \hat{R}_2 \sigma_{32} + \hat{R}_3 \sigma_{33})], \quad (23)$$

where,

$$\begin{aligned} \hat{R}_1 &= \frac{\partial \hat{R}}{\partial \hat{\alpha}} \\ &= -\left(1 - e^{-(\hat{\lambda}x)^{\hat{\beta}}}\right)^{\hat{\alpha}} \log\left(1 - e^{-(\hat{\lambda}x)^{\hat{\beta}}}\right), \\ \hat{R}_2 &= \frac{\partial \hat{R}}{\partial \hat{\lambda}} \\ &= \hat{\alpha} \hat{\beta} x \left(-e^{-(\hat{\lambda}x)^{\hat{\beta}}}\right) (\hat{\lambda}x)^{\hat{\beta}-1} \left(1 - e^{-(\hat{\lambda}x)^{\hat{\beta}}}\right)^{\hat{\alpha}-1}, \\ \hat{R}_3 &= \frac{\partial \hat{R}}{\partial \hat{\beta}} \\ &= \hat{\alpha} \left(-e^{-(\hat{\lambda}x)^{\hat{\beta}}}\right) (\hat{\lambda}x)^{\hat{\beta}} \log(\hat{\lambda}x) \left(1 - e^{-(\hat{\lambda}x)^{\hat{\beta}}}\right)^{\hat{\alpha}-1}. \end{aligned}$$

Next, we present Baye's estimators using GEL function. Let  $\hat{\alpha}_g, \hat{\lambda}_g, \hat{\beta}_g$  and  $\hat{R}_g$  denote Baye's estimators of  $\alpha, \lambda, \beta$  and  $R(t)$  respectively. The following steps, for various choice of  $v(\hat{\alpha}, \hat{\lambda}, \hat{\beta})$  Bayes estimators for  $\alpha, \lambda, \beta$  and  $R(t)$  respectively,

- $v(\hat{\alpha}, \hat{\lambda}, \hat{\beta}) = \hat{\alpha}^{-c}$  then

$$\hat{\alpha}_g = \hat{\alpha}^{-c} - c\hat{\alpha}^{-(c+1)} \left( -\frac{1}{\hat{\alpha}}\sigma_{11} + \left( \frac{a-1}{\hat{\lambda}} - b \right) \sigma_{12} - \frac{1}{\hat{\beta}}\sigma_{13} \right) + \frac{1}{2} \left( c(c+1)\hat{\alpha}^{-(c+2)}\sigma_{11} \right) - \frac{c\hat{\alpha}^{-(c+1)}}{2} [A\sigma_{11} + B\sigma_{21} + C\sigma_{31}] \quad (24)$$

- $v(\hat{\alpha}, \hat{\lambda}, \hat{\beta}) = \hat{\lambda}^{-c}$  then

$$\hat{\lambda}_g = \hat{\lambda}^{-c} - c\hat{\lambda}^{-(c+1)} \left( -\frac{1}{\hat{\alpha}}\sigma_{21} + \left( \frac{a-1}{\hat{\lambda}} - b \right) \sigma_{22} - \frac{1}{\hat{\beta}}\sigma_{23} \right) + \frac{1}{2} \left( c(c+1)\hat{\lambda}^{-(c+2)}\sigma_{22} \right) - \frac{c\hat{\lambda}^{-(c+1)}}{2} [A\sigma_{12} + B\sigma_{22} + C\sigma_{32}] \quad (25)$$

- $v(\hat{\alpha}, \hat{\lambda}, \hat{\beta}) = \hat{\beta}^{-c}$  then

$$\hat{\beta}_g = \hat{\beta}^{-c} - c\hat{\beta}^{-(c+1)} \left( -\frac{1}{\hat{\alpha}}\sigma_{31} + \left( \frac{a-1}{\hat{\lambda}} - b \right) \sigma_{32} - \frac{1}{\hat{\beta}}\sigma_{33} \right) + \frac{1}{2} \left( c(c+1)\hat{\beta}^{-(c+2)}\sigma_{33} \right) - \frac{c\hat{\beta}^{-(c+1)}}{2} [A\sigma_{13} + B\sigma_{23} + C\sigma_{33}] \quad (26)$$

- $v(\hat{\alpha}, \hat{\lambda}, \hat{\beta}) = \hat{R}^{-c}$  then

$$\hat{R}_g = \hat{R}^{-c} + (\hat{R}_1 a_1 + \hat{R}_2 a_2 + \hat{R}_3 a_3 + a_4 + a_5) + \frac{1}{2} [A(\hat{R}_1 \sigma_{11} + \hat{R}_2 \sigma_{12} + \hat{R}_3 \sigma_{13}) + B(\hat{R}_1 \sigma_{21} + \hat{R}_2 \sigma_{22} + \hat{R}_3 \sigma_{23}) + C(\hat{R}_1 \sigma_{31} + \hat{R}_2 \sigma_{32} + \hat{R}_3 \sigma_{33})], \quad (27)$$

where

$$\hat{R}_i = \frac{\partial \hat{R}}{\partial \hat{\eta}_i}, i = 1, 2, 3 \text{ and } (\hat{\eta}_1, \hat{\eta}_2, \hat{\eta}_3) = (\hat{\alpha}, \hat{\lambda}, \hat{\beta}).$$

Observe that all equations define above depends upon MLEs of  $\alpha, \lambda$  and  $\beta$ . The detailed procedure for obtaining MLE is discussed in Section 2.2. Moreover, these MLEs don't have closed form studies. Note that we resorted to using Newton Raphson method for solving equations for obtaining MLEs numerically. Then next Section present the simulation study to obtain Bayes estimators for various parameters of EW distribution and the reliability function  $R(t)$ .

### 3. SIMULATION

In this Section, The results obtained in previous section, are illustrated by means of simulation. The data simulated by using R programming language are used to obtain Bayes estimators of parameters of EW distribution, namely,  $\alpha, \lambda, \beta$  and  $R(t)$ . Further, the performance of these estimators are studied by computing their respective mean square error and standard deviation. The following subsection will describe the details of simulation procedure.

#### 3.1. Simulation Algorithm

Let us assume that prior distribution for  $\alpha \sim U(0, 1), \lambda \sim \text{Gamma}(a, b)$  and  $\beta \sim U(0, 1)$  are chosen at random.

If the random variable  $U$  follows a uniform distribution in  $(0, 1)$ , then  $X = \left[ -\frac{1}{\lambda} \log \left( 1 - U^{\frac{1}{\alpha}} \right) \right]^{\frac{1}{\beta}}$  follows the  $GW(\alpha, \lambda, \beta)$ . Next, progressive type-I interval censored sampling data,  $D = (X_i, R_i, t_i)_m$ , of the  $GW(\alpha, \lambda, \beta)$ , are generated as follows. First, the random variables,  $U_1, U_2, \dots, U_n, n \leq m$ , are generated from  $U(0, 1)$ , and then  $GW(\alpha, \lambda, \beta)$  data  $t'_1, t'_2, \dots, t'_k, \dots, t'_n$  are calculated by inverting



$t'_k = \left[ -\frac{1}{\lambda} \log \left( 1 - U_k^{\frac{1}{\alpha}} \right) \right]^{\frac{1}{\beta}}$ . Now, the number,  $X_i$ , of failures within  $(t_{(i-1)}, t_i]$  are generated and  $R_i$  surviving items are randomly removed from the testing based on the pre-specified inspection times  $t_1 < \dots < t_m$  and the pre-specified percentage  $p = (p_1, p_2, \dots, p_{m-1}, 1)$ , respectively. The specific steps are as given below. (see, Aggarwala [?])

- Set  $X_0 = 0$  and  $R_0 = 0$  and for  $i = 1, 2, \dots, m$
- $X_i \mid X_{i-1}, \dots, X_0, R_{(i-1)}, \dots, R_0 \sim rbinom \left( n - \sum_{j=1}^{i-1} (X_j + R_j), \frac{F_i - F_{(i-1)}}{1 - F_{(i-1)}} \right)$
- $R_i \mid X_i, \dots, X_0, R_{(i-1)}, \dots, R_0 = floor \left[ p_i * \left( n - \sum_{j=1}^i X_j - \sum_{j=1}^{i-1} R_j \right) \right]$

where  $rbinom(n, p)$  generates a random variable from the binomial distribution with parameters  $n$  and  $p$ .

### 3.2. Example

Let the priors  $\alpha \sim U(0, 1)$ ,  $\lambda \sim Gamma(1, 2)$  and  $\beta \sim U(0, 1)$  and a set of parameters  $\alpha, \lambda$  and  $\beta$  are generated from these distributions. Let us assume that values for  $\alpha = 0.4650936$ ,  $\lambda = 0.09790184$ ,  $\beta = 0.2090737$  and  $R(t; \alpha, \lambda, \beta)_{t=1} = 0.1592157$  are selected from this set as true values. Let us assume that  $m=8$ . Then, the randomly generated data are chosen from the Uniform distribution  $U(0,1)$  as follows:

$$U = (0.8716594, 0.6916711, 0.3129649, 0.3065460, 0.7183383, 0.3928726, 0.4819814, 0.6090094)$$

To generate the inspection time set of the gradually type-I interval censored sample by applying

$$t'_k = \left[ -\frac{1}{\lambda} \log \left( 1 - U_k^{\frac{1}{\alpha}} \right) \right]^{\frac{1}{\beta}}$$
 is given by,

$$T = (0.4273016, 0.5336827, 6.341113, 10.02617, 63.84012, 108.4094, 223.2485, 595.9245)$$

To create distinct progressive type-I interval censored samples, four group sample sizes  $n=10, 15, 20, 25, 30, 35, 40, 45$  and five pre-specified percentages  $p: p_{(1)}$  and  $p_{(2)}$  are considered, where

$$p_{(1)} = (0, 0, 0, 0, 0, 0, 0, 1), \quad p_{(2)} = (0.1, 0, 0, 0, 0, 0, 0, 1)$$

In Tables 1 and 2, for specific  $p_{(1)}$  and  $p_{(2)}$  in progressive type I interval censoring, relative error (Re) and mean square error (MSE) of Bayesian estimators under SEL function ( $B_S$ ) and Linex Loss function ( $B_L$ ) with  $c = 0.5$ , are permitted. Note that Re is given by

$$Re = \frac{|\hat{g} - g|}{g}$$

and MSE is given by

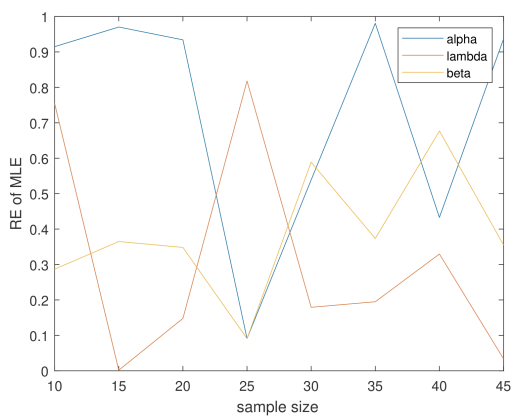
$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{g}_i - g_i)^2,$$

where  $\hat{g}$  denote the MLEs or Bayesian estimates of  $g$ .

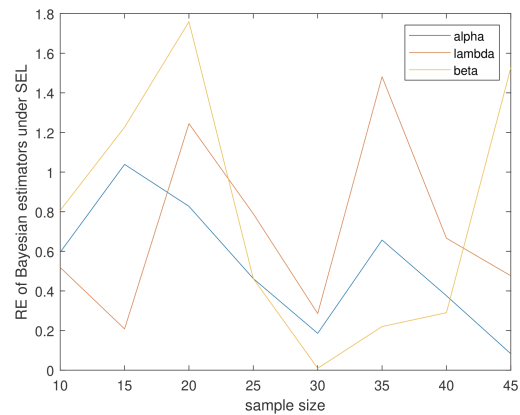
After an extensive study of the results thus obtained, conclusions are drawn regarding the behavior of the errors of estimators, which are summarized below graphically (see Figure 3- Figure 14).

**Table 1:** RE and MSE of the Example for fixed  $p = p_{(1)}$

Item	n	RE				MSE			
		$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{R}$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{R}$
MLE	10	0.9149	0.7544	0.2869	0.7365	0.0902	0.0098	0.0023	0.0154
	15	0.9703	0.0017	0.3649	0.9462	0.0231	0.0000	0.0027	0.0076
	20	0.9341	0.1479	0.3482	0.8148	0.0536	0.0002	0.0029	0.0097
	25	0.0909	0.8181	0.909	0.1052	0.0000	0.1202	0.0005	0.0007
	30	0.5389	0.1793	0.5890	0.6659	0.0004	0.0022	0.0098	0.0161
	35	0.9808	0.1949	0.3735	0.9434	0.1066	0.0000	0.0102	0.0114
	40	0.4328	0.3296	0.6769	0.6602	0.0024	0.0061	0.0113	0.0176
$B_s$	45	0.9366	0.0389	0.3552	0.9200	0.0833	0.0000	0.0085	0.0047
	10	0.5954	0.5185	0.8065	0.5529	0.0382	0.2128	0.0179	0.0087
	15	1.0388	0.2083	1.2265	0.2629	0.0265	0.0336	0.3157	0.0589
	20	0.8271	1.2446	1.7592	0.5632	0.0420	0.0123	0.3435	0.0047
	25	0.4623	0.7895	0.4622	0.8032	0.0162	0.1044	1.2557	0.0381
	30	0.1860	0.2859	0.0094	0.0636	0.0000	0.5612	0.0000	0.0002
	35	0.6569	1.4812	0.2202	0.0294	0.0478	0.0007	0.0262	0.0000
$B_g$	40	0.3758	0.6661	0.2903	0.0621	0.0000	0.0244	0.0021	0.0002
	45	0.0828	0.4764	1.5272	0.1365	0.0007	0.0000	0.1572	0.0001
	10	0.1253	1.7028	1.1261	0.5510	0.0017	0.0498	0.0349	0.0086
	15	0.3227	0.3916	1.4387	0.2632	0.0026	0.1188	0.0434	0.0591
	20	0.0519	1.9679	1.2674	0.5606	0.0002	0.0699	0.0390	0.0046
	25	1.1834	0.2332	0.4580	0.2034	0.1062	0.0091	0.0123	0.2439
	30	0.0138	0.3522	0.7898	0.7545	0.0000	0.0084	0.0343	0.0207
35	0.7894	1.8737	0.3078	0.2895	0.0690	0.0005	0.0069	0.0000	
40	1.1188	0.4942	1.2519	1.3603	0.0161	0.0137	0.0386	0.0745	
45	0.0184	0.4489	0.3615	0.1628	0.0000	0.0000	0.0088	0.0001	



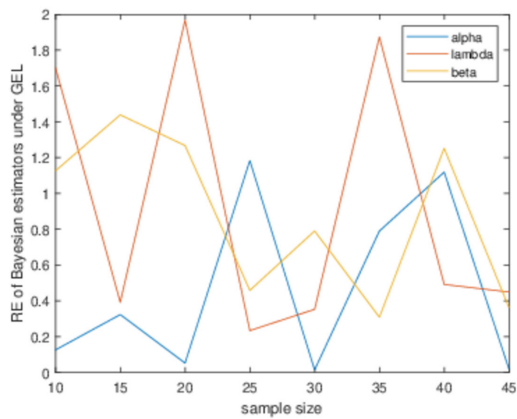
**Figure 3:** Relative Error of MLE for  $p(1)$



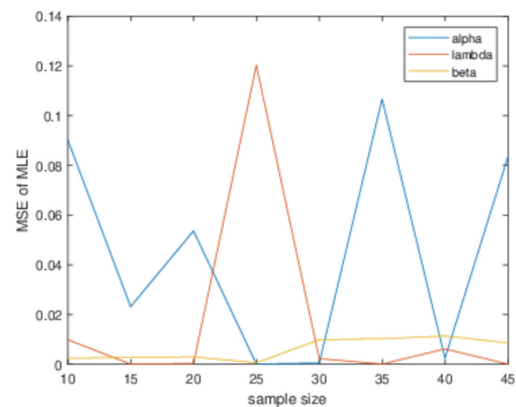
**Figure 4:** Relative Error of  $B_s$  for  $p(1)$

**Table 2:** RE and MSE of the Example for fixed  $p = p_{(2)}$

Item	n	RE				MSE			
		$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{R}$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{R}$
MLE	10	0.9009	0.3422	0.3621	0.7676	0.0837	0.0003	0.0066	0.0098
	15	0.9481	0.3285	0.3743	0.8974	0.0134	0.0004	0.0003	0.0070
	20	0.9152	0.4706	0.2872	0.7947	0.0629	0.0172	0.0014	0.0081
	25	0.9561	0.0317	0.3609	0.8741	0.0663	0.0000	0.0016	0.0085
	30	0.6986	0.4347	0.8745	0.7925	0.0252	0.0160	0.0822	0.0247
	35	0.9852	0.9342	0.6004	0.8572	0.0715	0.0013	0.0333	0.0127
	40	0.9431	0.3594	0.2985	0.8779	0.0284	0.0033	0.0007	0.0073
	45	0.9467	0.6701	0.2837	0.8926	0.0244	0.0007	0.0014	0.0059
	$B_s$	10	0.0605	0.4409	0.3096	1.2529	0.0004	0.0005	0.0048
15		1.9421	0.1121	1.5398	1.9250	0.0563	0.0000	0.0249	0.0744
20		0.1961	0.0578	0.4542	0.7478	0.0029	0.0003	0.0035	0.0072
25		0.5377	0.2024	0.4913	0.2352	0.0210	0.0338	0.0030	0.0006
30		0.3550	0.0855	0.6875	0.1301	0.0065	0.0999	0.0508	0.0006
35		1.5559	1.5346	0.5596	1.3814	0.1783	0.3580	0.0289	0.0330
40		0.512	0.7987	0.4375	1.0897	0.0084	0.0165	0.1403	0.0413
45		0.4568	0.8569	0.1748	1.8877	0.0057	0.0314	0.0005	0.0265
$B_g$		10	0.0173	0.4638	0.5745	1.2778	0.0000	0.0005	0.0166
	15	0.7190	1.6573	1.0796	1.1535	0.0077	0.0840	0.0086	0.0403
	20	0.1677	0.2676	1.7243	0.7769	0.0021	0.0056	0.0501	0.0078
	25	0.1207	0.3234	1.1491	0.2385	0.0011	0.0729	0.0582	0.0006
	30	0.5552	0.2137	0.0639	0.2737	0.0159	0.0038	0.0004	0.0000
	35	1.2341	1.1229	0.1248	1.3814	0.1122	0.1917	0.0014	0.0329
	40	0.4512	1.2013	0.3081	1.3114	0.0065	0.0372	0.0712	0.0466
	45	0.5546	2.6507	1.6666	1.9275	0.0083	0.0978	0.0488	0.0276



**Figure 5:** Relative Error of  $B_g$  for  $p(1)$



**Figure 6:** Mean Squared Error of MLE for  $p(1)$

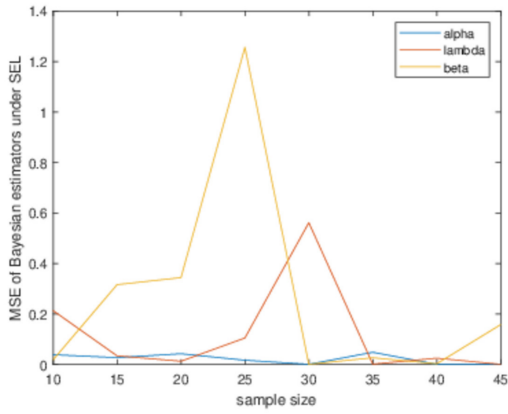


Figure 7: Mean Squared Error of  $B_s$  for  $p(1)$

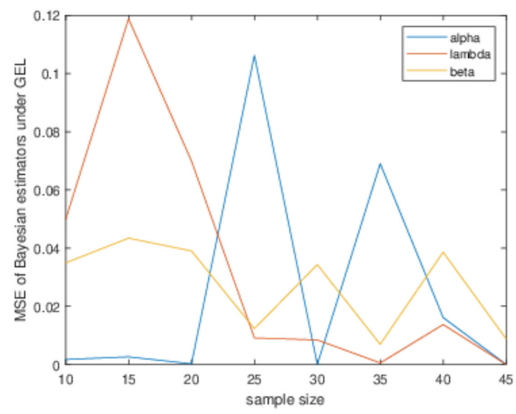


Figure 8: Mean Squared Error of  $B_g$  for  $p(1)$

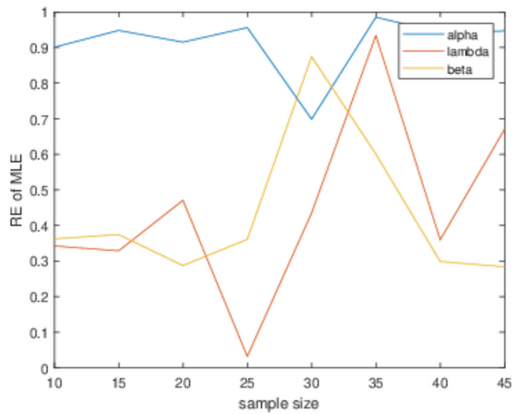


Figure 9: Relative Error of MLE for  $p(2)$

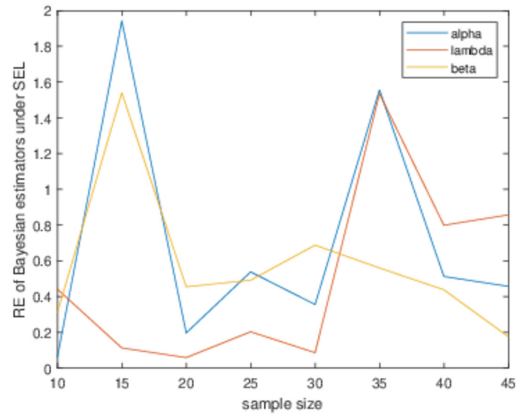


Figure 10: Relative Error of  $B_s$  for  $p(2)$

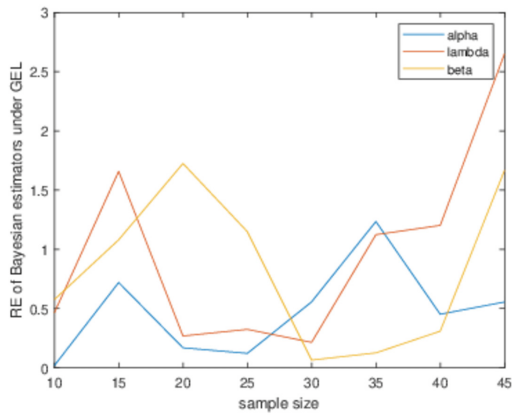


Figure 11: Relative Error of  $B_g$  for  $p(2)$

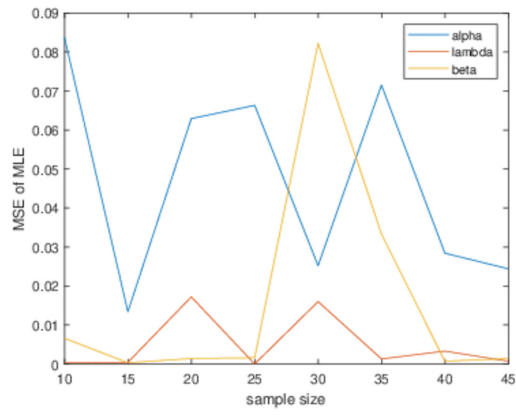


Figure 12: Mean Squared Error of MLE for  $p(2)$

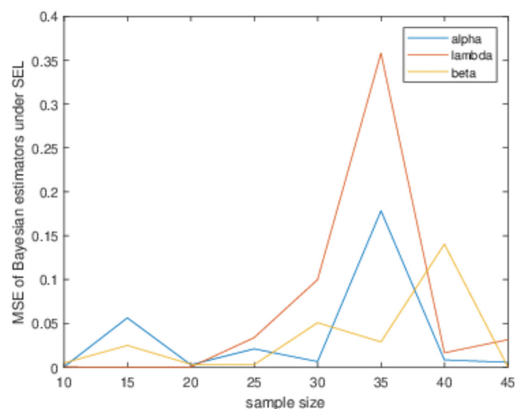


Figure 13: Mean Squared Error of  $B_s$  for  $p(2)$

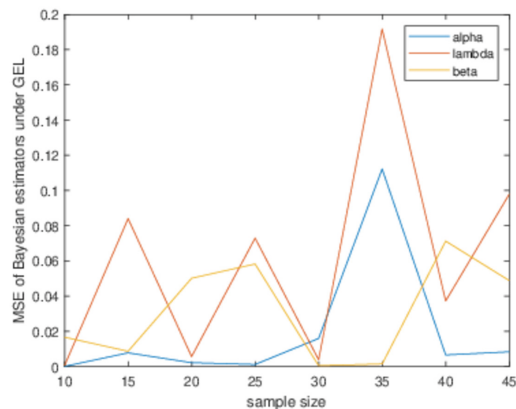


Figure 14: Mean Squared Error of  $B_g$  for  $p(2)$

#### 4. CONCLUSION

In this article, the performance of the proposed Bayes estimators has been compared to the maximum likelihood estimator of the  $EWD(\alpha, \lambda, \beta)$  under the progressive type-I interval censoring based on the squared error loss function and general entropy loss function using Lindley's approximation. The simulation result indicates that this approach is better suited for small sample sizes. MLE is the best choice when compared to Bayesian estimators. From Table 1, it is observed that the general entropy loss function in Bayesian estimation is better as compared to the squared error loss function in terms of MSE. From Table 2, it is noted that the squared error loss function in Bayesian estimation is better as compared to the general entropy loss function in terms of MSE. It can be seen from Figures 4, 5, 10 and 11 that the RE of Bayes estimators show fluctuating trend, and one can not see continuously decreasing or increasing trend for RE.

It is observed in practice, especially while modeling lifetime of electronic products, this three-parameter EW distribution describes the lifetime in the best possible way as compared to commonly used lifetime distributions such as Exponential distribution or Weibull distribution. Moreover, practically progressive type I interval censoring is the most convenient way of obtaining data of lifetimes as compared to traditional censoring schemes such as type I or type II or hybrid censoring. Further, the results obtained in this paper can be used for applications in the field of economics or analysis of clinical data in the medical field.

The results obtained in this paper use the approximation process such as Lindley approximation to obtain Bayes estimators of parameters of EW distribution. As future scope of research an analytical solution for deriving Bayes estimators can be considered by using suitable choice of prior distributions.

#### REFERENCES

- [1] Aggarwala, R. (2001). Progressive interval censoring: Some mathematical results with applications to inference. *Communications in Statistics-Theory and Methods*, 30(8-9):1921-1935.
- [2] Ahmad Sartawi, H. and Abu-Salih, M. S.(1991). Bayesian prediction bounds for the burr type x model. *Communications in Statistics-Theory and Methods*, 20(7):2307-2330.
- [3] Al-Noor, N. H., Abid, S. H. and Boshi, M. A. A. (2019). On the exponentiated weibull distribution. *In Aip conference proceedings*, volume 2183, page 110003. AIP Publishing LLC.
- [4] ALkanani, I. H. and Abbas, M. S. (2014). The non-bayesian estimators methods for parameters of exponentiated weibull (ew) distribution. *International Journal of Mathematics and Statistical Studies*, Vol.2, No.5, pp. 81-94.

- [5] Aludaat, K. M., Alodat, M. T. and Alodat, T. T. (2008). Parameter estimation of Burr type X distribution for grouped data. *Applied Mathematical Sciences*, 2(9), 415–423.
- [6] Ashour, S. K. and Afif, W. M. (2007). Statistical analysis of exponentiated Weibull family under type I progressive interval censoring with random removals. *Journal of Applied Sciences Research*, 3(12), 1851–1863.
- [7] Balakrishnan, N. and Aggarwala, R. Progressive censoring: theory, methods, and applications. Springer Science & Business Media. 2000.
- [8] Choudhury, A. (2005). A simple derivation of moments of the exponentiated Weibull distribution. *Metrika*, 62, 17–22.
- [9] Datta, D. and Datta, D. (2013). Comparison of Weibull distribution and exponentiated Weibull distribution based estimation of mean and variance of wind data. *International Journal of Energy, Information and Communications*, 4(4), 1–12.
- [10] Elshahhat, A. (2017). Parameters estimation for the exponentiated Weibull distribution based on generalized progressive hybrid censoring schemes. *American Journal of Applied Mathematics and Statistics*, 5(2), 33–48.
- [11] 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.
- [12] Gupta, R. D. and Kundu, D. (1999). Theory & methods: Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, 41(2), 173–188.
- [13] Gupta, R. D. and Kundu, D. (2001). Generalized exponential distribution: different method of estimations. *Journal of Statistical Computation and Simulation*, 69(4), 315–337.
- [14] Gupta\*, R. D. and Kundu, D. (2004). Discriminating between gamma and generalized exponential distributions. *Journal of Statistical Computation & Simulation*, 74(2), 107–121.
- [15] Gupta, R. D. and Kundu, D. (2007). Generalized exponential distribution: Existing results and some recent developments. *Journal of Statistical planning and inference*, 137(11), 3537–3547.
- [16] Jaheen, Z. F. and Al-Matrafy B. N. (2002). Bayesian prediction bounds from the scaled Burr type X model. *Computers & Mathematics with Applications*, 44(5-6), 587–594.
- [17] Kundu, D. and Gupta, R. D. (2005). Estimation of  $P[Y < X]$  for generalized exponential distribution. *Metrika*, 61, 291–308.
- [18] Kundu, D. and Gupta, R. D. (2008). Generalized exponential distribution: Bayesian estimations. *Computational Statistics & Data Analysis*, 52(4), 1873–1883.
- [19] Lin, C. T., Wu, S. J. and Balakrishnan, N. (2009). Planning life tests with progressively Type-I interval censored data from the log-normal distribution. *Journal of Statistical Planning and Inference*, 139(1), 54–61.
- [20] Lindley, D. V. (1980). Approximate bayesian methods. *Trabajos de estadística y de investigación operativa*, 31, 223–245.
- [21] Madi, M. T. and Raqab, M. Z. (2009). Bayesian inference for the generalized exponential distribution based on progressively censored data. *Communications in Statistics-Theory and Methods*, 38(12), 2016–2029.
- [22] Moradi, N., Panahi, H. and Habibirad, A. (2022). Estimation for the Three-Parameter Exponentiated Weibull Distribution under Progressive Censored Data. *Journal of the Iranian Statistical Society*, 21(1), 153–177.
- [23] Mudholkar, G. S. and Hutson, A. D. (1996). The exponentiated Weibull family: some properties and a flood data application. *Communications in Statistics-Theory and Methods*, 25(12), 3059–3083.
- [24] Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE transactions on reliability*, 42(2), 299–302.
- [25] Mudholkar, G. S., Srivastava, D. K. and Freimer, M. (1995). The exponentiated Weibull family: A reanalysis of the bus-motor-failure data. *Technometrics*, 37(4), 436–445.
- [26] Mudholkar, G. S., Srivastava, D. K. and Kollia, G. D. (1996). A generalization of the Weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association*, 91(436), 1575–1583.

- [27] Nadarajah, S. (2011). exponentiated exponential distribution: a survey. *AStA Advances in Statistical Analysis*, 95, 219–251.
- [28] Nadarajah, S. and Gupta, A. K. (2005). On the moments of the exponentiated Weibull distribution. *Communications in Statistics-Theory and Methods*, 34(2), 253–256.
- [29] Nassar, M. M. and Eissa, F. H. (2003). On the exponentiated Weibull distribution. *Communications in Statistics-Theory and Methods*, 32(7), 1317–1336.
- [30] Nekoukhou, V. and Bidram, H. (2015). The exponentiated discrete Weibull distribution. *Sort*, 39, 127–146.
- [31] Ng, H. K. T. and Wang, Z. (2009). Statistical estimation for the parameters of Weibull distribution based on progressively type-I interval censored sample. *Journal of Statistical Computation and Simulation*, 79(2), 145–159.
- [32] Pal, M., Ali, M. M. and Woo, J. (2006). Exponentiated weibull distribution. *Statistica*, 66(2), 139–147.
- [33] Peng, X. Y. and Yan, Z. Z. (2013). Bayesian estimation for generalized exponential distribution based on progressive type-I interval censoring. *Acta Mathematicae Applicatae Sinica*, English Series, 29(2), 391–402.
- [34] Pradhan, B. and Kundu, D. (2009). On progressively censored generalized exponential distribution. *Test*, 18, 497–515.
- [35] Preda, V., Panaitescu, E. and Constantinescu, A. (2010). Bayes estimators of modified-Weibull distribution parameters using Lindley's approximation. *Wseas transactions on mathematics*, 9(7), 539–549.
- [36] Raqab, M. Z. (1998). Order statistics from the Burr type X model. *Computers & Mathematics with Applications*, 36(4), 111–120.
- [37] Raqab, M. Z. (2002). Inferences for generalized exponential distribution based on record statistics. *Journal of statistical planning and inference*, 104(2), 339–350.
- [38] Raqab, M. Z. and Kundu, D. (2006). Burr type X distribution: revisited. *Journal of probability and statistical sciences*, 4(2), 179–193.
- [39] Raqab, M. Z. and Madi, M. T. (2005). Bayesian inference for the generalized exponential distribution. *Journal of Statistical computation and Simulation*, 75(10), 841–852.
- [40] Roy, S., Pradhan, B. and Purakayastha, A. (2022). On inference and design under progressive type-I interval censoring scheme for inverse Gaussian lifetime model. *International Journal of Quality & Reliability Management*, 39(8), 1937–1962.
- [41] Singh, P. K., Singh, S. K. and Singh, U. (2008). Bayes estimator of inverse Gaussian parameters under general entropy loss function using Lindley's approximation. *Communications in Statistics-Simulation and Computation*, 37(9), 1750–1762.
- [42] Singh, S. K. (2011). Estimation of parameters and reliability function of exponentiated exponential distribution: Bayesian approach under general entropy loss function. *Pakistan journal of statistics and operation research*, 217–232.
- [43] Surles, J. G. and Padgett, W. J. (2005). Some properties of a scaled Burr type X distribution. *Journal of statistical planning and inference*, 128(1), 271–280.
- [44] Wu, S. F., Liu, T. H., Lai, Y. H. and Chang, W. T. (2022). A study on the experimental design for the lifetime performance index of Rayleigh lifetime distribution under progressive type I interval censoring. *Mathematics*, 10(3), 517.
- [45] Wu, S. F. and Song, M. Z. (2023). Experimental Design for Progressive Type I Interval Censoring on the Lifetime Performance Index of Chen Lifetime Distribution. *Mathematics*, 11(6), 1554.
- [46] Yang, C. and Tse, S. K. (2005). Planning accelerated life tests under progressive type I interval censoring with random removals. *Communications in Statistics-Simulation and Computation*, 34(4), 1001–1025.

# METHODS FOR ENSURING AND PROVING FUNCTIONAL SAFETY OF AUTOMATIC TRAIN OPERATION SYSTEMS

I.B. Shubinsky<sup>1</sup>, E.N. Rozenberg<sup>2</sup>, H. Schäbe<sup>3</sup>

<sup>1</sup>DsC, prof., NIIAS, Moscow, Russia, Igor-shubinsky@yandex.ru

<sup>2</sup>DsC, prof., NIIAS, Moscow, Russia, s,lavruhina@vniias.ru

<sup>3</sup>Doctor of Natural Sciences, TÜV Rheinland InterTraffic, Cologne, Germany,  
dr.hendrik.schaebe@gmail.com

## Abstract

*The paper examines the specificity of artificial intelligence-based automatic train operation systems. Justifying the functional safety (FS) of such systems is quite difficult. The paper proposes a process for proving the functional safety of intelligent systems. A hybrid control system for a shunting locomotive was developed and analysed. It combines machine vision (MV), train protection devices and manual control by a driver. A model is presented that allows examining the functional safety of a locomotive control system layer by layer, i.e., evaluating the time to safety degradation depending on the component failure and subsequent requirement of bringing the locomotive to a complete stop. This allows to improve the FS of the shunting locomotive control system with machine vision from SIL 2 to SIL 3 and maintaining it during sufficiently long periods of time (over a quarter of the mean time to system failure). The mean time of faultless operation of a locomotive control system until it has to be brought to a complete stop for safety reasons can be increased three times. A general approach is proposed to design the functional safety of automatic train operation systems. It is based on the division of the information processing process into two subprocesses, i.e., internal intelligent information processing onboard the locomotive for the purpose of decision-making regarding track vacancy and communication of initial visual information to the operating driver for decision-making. The division of this process must be combined with redundant machine vision facilities, regular comparison of the outputs of the onboard and fixed machine vision facilities, redundant comparison outputs, smoothing of the outputs in the process of locomotive movement.*

**Keywords:** Functional safety, artificial intelligence, automatic train operation system, machine vision, dependability, safety justification, safety case, statistical and experimental methods, expert methods, simulation methods, heuristic semi-Markov graph methods, process methods of compliance confirmation, safety device, control system, Markov model, standards.

## 1. Introduction

The problem of ensuring functional safety of any technical system consists of two integral components. The first one consists in the development of proposals, techniques, procedures, methods for improving FS. The second component is intended for verifying the efficiency of the chosen method of improving safety. Essentially, the second component of the problem consists in



proving the acceptability of the achieved level of FS. Substantiating FS for railway control systems with the grade of automation GoA 2/3 (from manual driving with the function of automatic train operation to automatic train driving with no human driver) is quite difficult. These systems use artificial intelligence-based methods for the purpose of training information processing algorithms. One of the first works on artificial intelligence aimed at recognising patterns by training recognition algorithms was the monograph by Vapnik and Chervonenkis titled "Pattern recognition theory (statistical problems of training)" [1]. In [2 – 4], it was shown that an automatic train operation (ATO) system has a number of distinctive features. Those include the following:

1. Distributed system architecture.
2. Availability of machine vision and effect of weather conditions.
3. Close information interaction between the system and the environment via information communication channels.
4. Presence of a large and not always definite number of vulnerabilities within a system closely connected to the environment.
5. A high probability of evolving environmental effects and resulting changed system behaviour.
6. The changed control algorithm parameters as the result of neural network training using the incoming information flows and accumulated databases.
7. Branching software of both the generic part of the system, and, especially, rolling stock detection and control facilities.

Braband and Schäbe [2] note that due to the specificity of the ATO it requires special methods for proving the FS. It should be noted that one of the key features of the system is that, along with its distributed architecture, the connections within the system change significantly. The latter noticeably reduces the options to prove the safety of such a system.

Given the great uncertainty associated with the operation of ATO it is quite difficult to prove its FS using conventional methods, i.e., those set forth in STO RZD 1.19.009-2009 [5] that were largely applied to devices and simple systems with a known and limited number of vulnerabilities. The recommendations of IEC 61508-1-2012 (sections 6, 8) [6], IEC 61508-2-2012 [7], IEC 61508-3-2018 [8], and IEC 62279-2016 [9] may prove to be very helpful in this situation. Along the conventional methods of safety case preparation, the above standards suggest taking into account the design and manufacture process, quality and functional safety assurance organisation of complex hardware and software systems and their components for the purpose of evaluating the functional safety level of such systems. Such measures and procedures jointly solve the problem of *safety justification*. One of the components of a safety justification involves confirming the compliance with the specified requirements, which is largely ensured using the safety case. A development of this approach combined with the guidance material accumulated by the railway industry is reflected in GOST 33432-2015 [10].

As regards intelligent systems with the above distinctive features, the standard recommends the following scope of safety justification:

1. Development of an FS policy;
2. Development of an FS program;
3. Development of a safety case.

An *FS policy* is to be in place at the ATO system manufacturer and is to be generally applied to all the products developed by such an organisation. It is to make provisions for solving the following main problems:

- tasks and objectives of FS assurance;
- principles and approaches to ensuring FS;
- principles of FS-related risk management;
- organisation of FS assurance.

In [11], Braband and Schäbe suggest using the outputs of the ATO-RISK project ordered by the Deutsche Zentrum für Schienenverkehrsforschung for the purpose of managing risks. The project aims to define the criteria of risk acceptability as regards automatic train operation. As described in [11], the risk level is evaluated through a function-specific explicit risk analysis or using reference systems. Explicit risk analysis is performed by evaluating various scenarios using the semi-quantitative approach and a risk score matrix. The matrix qualitatively differentiates the expected severity of harm depending on the category of the accidents. That approach can be recommended for the purpose of system safety policy definition.

*FS assurance and FS case program* are developed for each product autonomously and are intended to be supplied to the customer as proof of the product being of high quality in accordance with the requirements of the FS standards and corresponds to the declared safety integrity level (SIL). The ultimate goal of the ATO FS measures consists in the preparation of a safety case.

## 2. Characteristic features of the functional safety case of automatic train operation systems

The scope of FS case preparation includes reports on not only the FS status, but on the measures taken by the ATO manufacturer for managing quality and ensuring FS. Those reports allow the customer to evaluate the engineering level and manufacturing quality of the system, including the supply of components, organisation and process quality of the FS assurance activities, risk evaluation results, depth and quality of the FS requirements verification and validation activities.

A conclusion of an ATO's compliance with the FS requirements is built upon the FS status report taking into consideration the above reports on the quality and FS management measures. That is a very important consideration. The matter is that the distributed system architecture, changing parameters of the control algorithms as the result of neural network training and other functional features of intelligent systems do not contribute to a guaranteed evaluation of their FS status. The use of reports of the adopted quality management and FS measures significantly enhance the informational description of the system and corroborates the confidence in the assessment of its FS state.

Confirmed compliance with the specified FS requirements plays a crucial role in the system FS case document. To that end, the following methods are used: statistical, experimental, expert, simulation, analytical, process.

*The statistical and experimental methods* enable the most objective, quantitative evaluation of a system's FS as long as their feasibility and reliability are ascertained. The matter of feasibility directly depends on the FS requirements. The required safety integrity level of an ATO system with continuous performance requests is typically SIL 2 [6, 12], which corresponds to the required range of a system's dangerous failure rates  $\lambda_{WS} = (10^{-7}10^{-6})/h$ . The probability of the system's dangerous failures within an hour of operation should be within the range  $Q_{WS}(1) = 10^{-7}10^{-6}$  [6]. Under the above requirements, an experimental identification of a single dangerous failure would require at least  $N \geq \frac{1}{Q(1)} = (10^710^6)$  tests, taking into account a statistical confidence level of 90% this will be even 3 106...3 107 hours.. As the duration of each test should be at least one hour, identifying a single

dangerous failure would take over 100 years. Even if the testing is carried out on many systems in parallel, it is complicated to accumulate a significant testing time is needed,

In principle, *experimental methods* allow indirectly confirming or disproving an ATO's compliance with the specified FS requirements. Naturally, an indirect estimate can only be used as an addition to other estimates rather than individually. The process of expert evaluation of complex system parameters is mature, as is the algebra of processing of experts' opinions. However, applying such methods to the ATO FS estimation has a number of difficulties. To begin with, the experience of ATO operation is still insignificant. The accumulated knowledge is clearly insufficient. Subsequently, it is difficult to presume an acceptable level of subject-area competence in the experts. Additionally, in various industries, including railway transportation, the number of FS experts is limited. Therefore, it is very difficult to involve a sufficient number of experts and evaluate the coherence of their opinions. However, we must strive for a situation, whereas expert methods can, to a certain extent, be used for confirming ATO compliance.

*Simulation methods* are widely used in the course of development and testing. They are based on the Monte Carlo method. The Monte Carlo simulation method allows using pseudorandom number generators to simulate practically the entire known spectrum of input, intermediate, and disturbance effects on a system. They are processed using software simulation of the system to generate outputs depending on the simulated data. However, that method has a serious drawback, i.e., the results contain a spread between the outputs of different simulations. Reducing the spread, i.e., reducing the dispersion, requires a large number of executions of the model, which, in turn, causes a sharp increase in the duration of the simulation. A number of methods of reduction of dispersion has been developed for the purpose of cutting the simulation time. Those include the following: Monte Carlo simulation (e.g., data and output value simulation), method of augmented variables, stratified sampling method, etc. Weighted sampling provides better results in terms of dispersion reduction. Drawing from that method, we have developed a simulation method based on semi-field testing [13] by means of artificial introduction of malfunctions (faults, perturbations, program errors) into the system. Despite the obvious advancements in simulation, those methods have a number of significant drawbacks that restrict their applicability in ATO research.

The main factors that restrict the application of simulation in ATO research are as follows:

1. A detailed description of the system and its features is required, which, for a system as complex as an ATO, requires significant efforts and associated large scope of work. Additionally, due to the complex system architecture, a clear description of such system is extremely complicated.
2. The high cost of developing a simulation model for the system.
3. Evidence of adequacy of the model to the actual system is required.
4. Each update of the system's structure and improvement of its algorithms require to repeat the activities specified above in Items 1 and 3. Practically, that comes down to the development of new simulation models.

*Analytical methods* are the main tool for safety case preparation. However, their applicability to ATO safety justification raises certain doubts. That is primarily due to the distributed architecture of such systems and, subsequently, the difficulty (or sometimes impossibility) to formalise the task of safety justification. In order to solve that problem, we propose the following. *Heuristic semi-Markov (Markov) graph methods*. The matter is that non-formalised problems of safety justification of systems with complex architectures are solved using heuristics, i.e., *a person's own ideas, rules that allow reducing the scope of potential solutions*. The essence of the developed methods [4, 12] consists in a combination of heuristics in the data representations and mathematical models of the system's safety and dependability with strict mathematical methods of analysis.

Under uncertainty or absence of certain data, an analytical estimation of system safety indicators is achieved through multi-stage calculations that consist in the implementation of the following sequence of actions:

1. Construction of the Markov graph of the ATO.
2. Definition of the mathematical models of the graph's edges and nodes.
3. Definition of equations for FS.
4. Expert evaluation of initial data.
5. Calculation, analysis of the results.
6. Determination of the most significant factors.
7. Simplification of the obtained calculation formulas maintaining an acceptable error.
8. Analytical evaluation of compliance with the required SIL.
9. Finalisation of the procedure in case of confidence in the results of evaluation or improvement of the examined ATO functional model (FS graph). If necessary, the model will be refined and steps 1-9 of the analytical estimation of safety parameters for above for the updated model will be repeated.

If reliable information and data are available, individual actions will suffice, e.g., graph construction, definition of formulas, calculation and analysis of the results. Other actions, e.g., expert evaluation of the initial data, identification of the most significant factors, simplification of calculation formulas, improvement of graph construction conditions, repeated construction(s) of a FS graph arise as needed depending on the availability or non-availability of information to the system's safety analyst.

Due to the above distinctive features of an ATO and in order to improve the confidence in the FS examination results along the recommendations of EN 50129 [25] *the process methods of FS compliance assurance* should be widely used.

Evaluating the achieved SIL for each safety function of an ATO's hardware components is possible based on the recommendations of EN 50129 [25] chapter 7..

The applied methods and means of failure management are evaluated based on the recommendations of EN 50129 [25] annex B..

The applied methods and means for preventing systematic errors can be evaluated based on annexe E of EN 50129 [25].

Regarding the software of an ATO, EN 50128 [26] recommends a number of procedures (annexe A) whose application significantly improves the confidence in the FS state estimate.

### 3. Methods of ensuring safe and uninterrupted operation of a shunting locomotive control system with machine vision

#### 3.1. Introduction

Railway signalling systems are undergoing a new stage of their development in order to solve one of the key problems in railway transportation that consist in the creation of unattended train operation. Along with the conventional means of functional safety, they include sufficiently complex ATO systems [16]. Now, using control algorithms based on logical and certain arithmetical operations is insufficient to ensure safe train operation. The technological development of control systems is associated with solving complex mathematical problems and eventually with the use of neural networks for information processing.

The simplicity of process-related tasks for the first safety integrity level allowed using mature methods to ensure compliance with functional safety requirements by means of hardware and software redundancy [6]. A clear advantage of using simpler technology in the form of hardwired logic and microcontrollers was the simplicity of built-in online testing and, subsequently, to achieve the required rate of dangerous failures [17].

In the process of automatic train operation system development, it became clear that their rate of dangerous failures will not be below the SIL2 threshold. The application of prototypes of such systems in railway transportation has shown that, in principle, they are man-machine systems, in which automatic train operation facilities cannot be fully trusted with ensuring train protection without an operator's involvement.

Systems used to ensure operational safety in stations as part of shunting operations require a SIL 2. That is due to the fact that the speed of train movement in the course of shunting operations is significantly lower than in the course of mainline operations [18]. Meanwhile, the demanding work performed by a driver in the course of shunting operations should be taken into consideration. When and where possible such operations should be automated. Thus, even if a driver is present onboard, the future requirements must be close to SIL3 or a new, more detailed SIL 2+ classification of safety is to be introduced. For information processing facilities, this level can be achieved with the help of a real-time operating system and high-performance microprocessors. In this context, the matters of validity of information processing and completeness of online tests arise. The tendency for using complex computer-based systems, on the one hand, and the expectation of their high redundancy, on the other hand, complicate such control. Indeed, within the information processing circuit, a small amount of memory and limited number of commands are used. In this context, a high test coverage cannot be guaranteed, as many elements of the information processing structure are not utilised. That, in turn, causes limitations in the assurance of an acceptable level of correct detection of failures of the automatic shunting cab signalling system (ASCSS).

#### 3.2. Problem definition

Currently, the ASCSS shunting locomotive control system is single-channel, system, which does not allow to raise its safety integrity level above SIL 2. By using information redundancy, a virtual second channel can be created to ensure additional monitoring of this computer-based system [19]. That will enable a high probability of correct detection of ASCSS failures. The monitoring process is to be designed in a way as to not affect the operation of the control algorithm of a shunting locomotive. The safety device (SD) software generates an ordered sequence of computer instructions that, within the ASCSS system, are implemented as a series of reference signatures, which allows additionally monitoring of the operation of complex ASCSS devices, thus enhancing its SIL.

By building upon that principle, such SIL2 and SIL3 devices can be used for monitoring even more complex devices with video cameras and neural networks. It must be noted that such a complex device itself, e.g., ASCSS, implements functions of the type “*prevention passing shunting signal at danger*” that align with the typical purpose of ASCSS and additionally detect obstacles using machine vision. After identifying the basic technical function of information processing, it can be used as a functional test for the equipment of a complex unattended system. That can be seen as a segment within the space of possible solutions of unattended systems.

Thus, the shunting signals themselves become a functional test of an even more complex system [20]. Additionally, within the examined system, the hardware and software machine vision facilities are to be additionally monitored by comparing the readings of onboard and trackside machine vision sensors [21]. Such a hierarchy may prove to be useful in reducing the cost of hardware and simplifying the safety case preparation as compared with the situation when all functions are implemented within a single processor [22]. It must be noted that, if no innovative solutions are used, ensuring system dependability becomes an issue, as machine vision facilities significantly increase the scope of system hardware, which causes a reduction of its dependability.

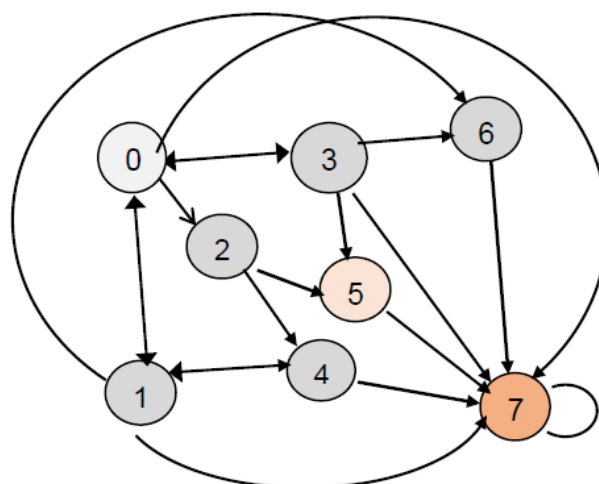
### 3.3. Research model and findings

Any information that can be depicted as objects and connections can be conveniently represented in graph form. Graphs are commonly used for visualising information, involving the transformation of large amounts of complex types of abstract information into a user-friendly visual form.

The authors built the model based on the following criteria:

*a safe failure* involving the failure of ASCSS and MV facilities, control of the locomotive is assigned to the driver; *a dangerous failure* involving the failure of ASCSS, MV facilities and SD, the shunting locomotive is brought to a complete stop. The question of the criticality of a dangerous failure is not discussed in this paper. It should be examined individually.

Figure 1 shows the state graph of functional safety of interaction between ASCSS and the SD and MV facilities.



**Figure 1.** FS state graph of the interaction between ASCSS and the safety device and machine vision facilities

*States of the system model:*

1. All objects of the control system are up;
2. SD has failed and is recovering; all the other system facilities are up;
3. ASCSS has failed, all the other system facilities are up;
4. MV facilities have failed and are recovering; all the other system facilities are up;
5. MALS and SD have failed; ASCSS is recovering;
6. MV and ASCSS have failed, locomotive control is assigned to the driver (safe failure);
7. MV and SD have failed;
8. All three systems have failed.

### **Dangerous failure**

System safe states are marked with the following colours:

	SIL 3,		safe failure (SIL1)
	SIL 2,		dangerous failure

In the system model, the following transitions are provided for:

0-1, SD failure;

0-2, ASCSS failure detected using built-in tests and/or signature analysis;

0-3, MV failure detected using built-in tests and/or by comparing the readings with the ASCSS program;

0-5, undetected MV failure;

0-7 and 1-7, undetected ASCSS failure;

1-4, ASCSS failure subject to SD failure;

1-6, MV failure subject to SD failure;

1-0, SD repair;

2-4, SD failure subject to ASCSS failure;

2-5, MV failure subject to ASCSS failure;

3-5, ASCSS failure subject to MV failure;

3-6, SD failure subject to ASCSS failure;

3-0, MV repair;

4-7, MV failure subject to MALS and SD failure;

4-1, ASCSS repair;

5-7, SD failure subject to MALS and MV failure;

6-7, MALS failure subject to MV and SD failure;

7-0, transition into the original state as the result of possible modification of ASCSS, if the risk of dangerous failures is acceptable.

The adopted premises and assumptions, defined mathematical models of the graph's edges and

nodes, the FS formulae, as well as the expert evaluation of the initial data are set forth by us in [23].

That model allows examining the FS of an ATO layer by layer, i.e., evaluating the time to safety integrity level degradation depending on failures of components and onset of a complete dangerous failure (dangerous failure of the second type in Fig.1). Thus, in particular, in [23] it was established that

- the mean time of the system being in SIL3 (state 0 in the graph in Fig.1) is described with the formula

$$T_0 = \frac{1}{\lambda_{SD} + \lambda_M + \lambda_{MV}} \quad (1)$$

where  $\lambda_{SD}$ ,  $\lambda_M$ ,  $\lambda_{MV}$  are the failure rates of the safety device, locomotive control system and machine vision, respectively;

- the mean time of faultless system operation at a level at least as high as SIL2, the mean time to a safe failure of type 1

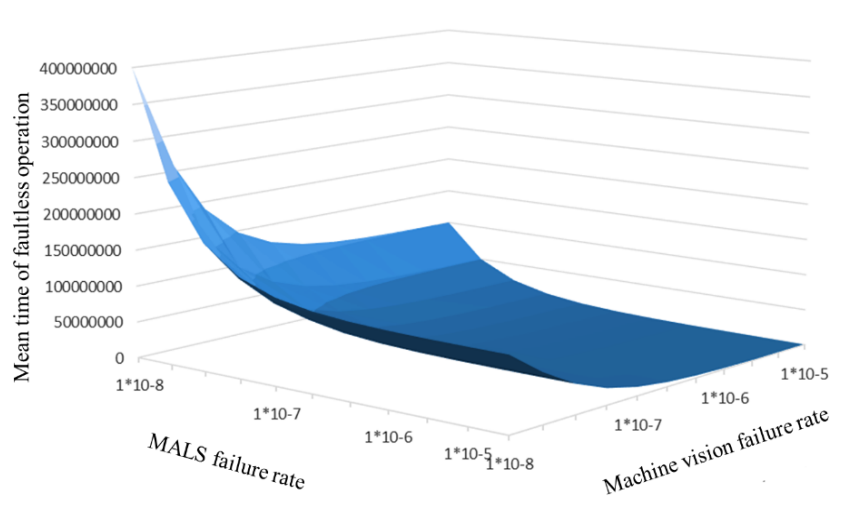
$$T_{\geq SIL2}^{system} \approx \frac{\lambda_{SD}(2\lambda_{SD} + 3\lambda_{MV}) + \lambda_{MV}^2}{(\lambda + \lambda_M)\lambda} \quad (2)$$

- the mean time of faultless system operation to a dangerous failure (complete stop of the shunting locomotive)

$$T_{wrong-side} \approx \frac{\lambda_{SD}(2\lambda_{SD} + 3\lambda_{MV}) + \lambda_{MV}^2 + 2\lambda_M\lambda_{MV}}{(\lambda + \lambda_M)\lambda} \quad (3)$$

Formulae (2) and (3) were obtained with an error not exceeding the first order of magnitude assuming that the failure detection parameters of ASCSS and machine vision facilities are close to one. That assumption is based on the fact that the monitoring of ASCSS operation using additional signature analysis procedures, as well as regular comparison of the ASCSS outputs with the machine vision outputs ensure complete and reliable performance monitoring of both the ASCSS control system, and the machine vision facilities.

Figure 2 shows the time of faultless operation of the shunting locomotive to a complete stop-vs-the failure rate of machine vision and ASCSS equipment curve. The failure rate of safety devices is taken equal to  $\lambda_{SD}=1*10^{-8}$ , which corresponds to the safety integrity rate of SIL3.



**Figure 2.** Time of faultless shunting locomotive operation to a complete stop-vs-the failure rate of machine vision and ASCSS equipment curve (“MALS failure rate”)



The key task of the research consists in evaluating the level of functional safety of the automatic train operation system of a shunting locomotive. Such a comprehensive assessment can be enabled by a research of the system's safety coefficient. The probability of an opposite event, i.e. a dangerous failure, is the system's danger coefficient. That coefficient, under the same assumptions that were used for deducing formulas (2) and (3), was obtained in [23] with an error not exceeding the first order of magnitude. It was established that the hazard coefficient significantly depends on the repair rates of facilities  $\mu$  and ASCSS repair rate upon a hazardous failure  $\mu_1$ .

The three-dimensional graphs of a system's hazard coefficient against parameters  $\mu$  and  $\mu_1$  subject to  $\lambda_M = 10^{-5} 1/h$  and  $\lambda_{MV} = 10^{-5} 1/h$  are shown in Fig. 3.

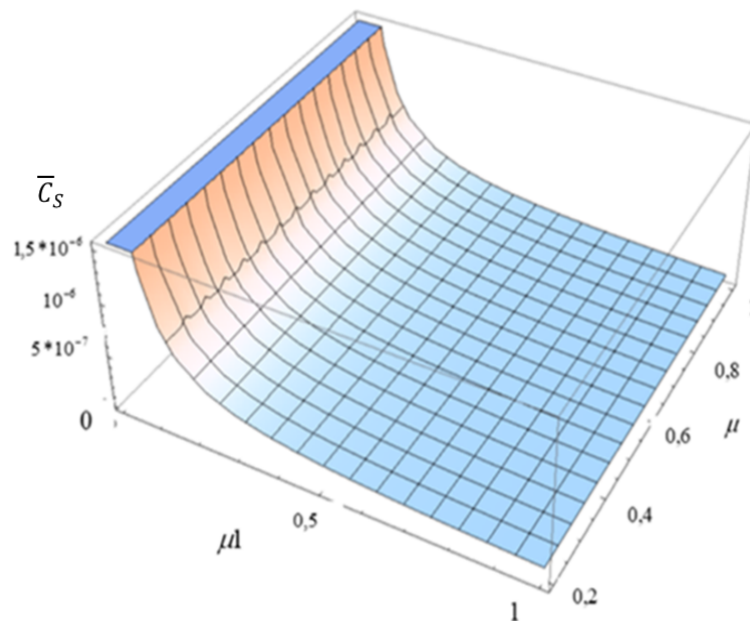


Figure 3. Graphs of hazard coefficient  $\bar{C}_S$  against repair rates  $\mu_1$  and  $\mu$ .

From these graphs it follows that as the system's repair rate  $\mu_1$  increases from 0.0059 /h to 1 /h the hazard coefficient decreases almost 30 times. The chosen limit values of the repair rate correspond to the system's repair times from an hour to a week. That range was chosen based on the nature of the malfunction. Thus, if a set of spare parts is available, hardware failures can be rectified within an hour, while rectifying software errors may take up to 7 days. Therefore, timely and prompt rectification of malfunctions may significantly improve a system's safety indicators.

#### 4. A general approach to designing the functional safety of automatic train operation systems

##### 4.1. Methods for designing the functional safety of automatic train operation systems

The main problem in the development of that approach consists in the fact that such a system has many distinctive features associated with the complex architecture and information processing algorithms, the incompleteness and fuzziness of initial data. Therefore, it is difficult to apply classical methods of probabilistic evaluation in the form of two or more independent hardware and software

information processors [20]. That is the exact reason why the redundancy of such information processors in the form of onboard machine vision cameras for the purpose of safe detection of obstacles on the track is unlikely to achieve the required safety level due to the unknown testing time of such a learning, i.e., constantly changing vital information processing system.

Braband and Schäbe [2] assumed the mandatory presence, , of an additional device within the processing system, whose safety could be proven using conventional methods owing to its constant structure.

Rozenberg and Shubinsky [12, 15] suggested using the so-called multi-level structures in order to ensure FS. This approach has shown good results in the development of advanced onboard and trackside safety systems. Additionally, an extremely important property of a system's safety evaluation was used that consists in obtaining reliable information on a facility's state history in terms of safety.

As regards the safety cases of neural network-enabled automatic train operation systems the principles of multi-level safety should be used. The difference consists in the fact that a complete set of technical equipment within a locomotive's operating environment is to be examined rather than an individual smart device, e.g., a machine vision camera on such a locomotive.

Indeed, a camera with a predesigned program for processing information on obstacles on the track does not depend only on the previously taken neural network training measures, but on the specific factors that affect the operability of the camera's hardware, software faults, etc. Additionally, it should be noted that the effect of the environment, i.e., snow, fog, and rain influences the obstacle detection zone, which directly affects the safety, as it is associated with the braking distance.

Under such conditions, the situation ahead of the train is additionally monitored from a special control centre, where an operating driver monitors several locomotives [21].

The complexity of this method consists in the fact that the reaction of the operating driver becomes a critical component, while he/she depends on the stable onboard camera image and dependability of the broadband communication at a particular location.

On the other hand, the division of the information processing process into two subprocesses, i.e., internal intelligent information processing onboard the locomotive for the purpose of decision-making regarding track vacancy and communication of initial visual information to the operating driver for decision-making allows improving safety. The criterion in this case is that the onboard system should have a high probability of false alarm, while the operating driver can rectify this situation using a special command transmitted to the locomotive by radio. In practice, if this principle was not used, an ATO system would stop, for instance, because of a plastic bag on the track.

It should be noted that the system includes trackside devices that monitor track vacancy in places with poor visibility [20]. Information on such fixed systems is communicated to the locomotive in real time, which significantly improves train traffic safety. Thus, the used model is simplified, but it enables an analytical study of the problem. That constitutes the advantage of this approach to developing the research model over more complex models. An interesting feature of the interaction between the fixed and onboard machine vision facilities is that, under identical environmental conditions, they can see the same objects, within the line of sight or under various, interesting angles.

The availability of objects detected by two independent systems allows using this property for cross-comparison of intelligent technical facilities, especially for the purpose of making correct decisions

by intelligent onboard systems that operate in more difficult operating conditions (traffic speed, limited visibility zone, etc.). An object comparison can be in the form of images processed by fixed and onboard cameras or it can contain the expected inversion of an image of the same object if two machine vision cameras point at it from opposite directions. Such a predefined property for a system for safe comparison of results enables better independence of information processing. Each technical facility, including video cameras, contains self-testing features that necessarily contribute to the calculation of their safe operation. Given that, as regards an intelligent system that employs neural networks, it is difficult to talk about complete testing, self-diagnostics using observed objects known beforehand should be used. For instance, next to the railway track, within the scanning zone of machine vision cameras or lidars, there are signals, control cabinets, catenary masts, and communications posts that are strictly referenced to the track coordinates, which is even more relevant if a 3D map of the infrastructure facilities is used onboard.

Thus, capturing such objects allows testing onboard cameras and sensors taking into account the detection distance and identification of the type of objects. If the frequency of object acquisition is high enough within the distance between such locations, the probability of no failure or no error of the information processing algorithm can be calculated for a moving object. The advantage of this method consists in the complete processing of information, when, along internal testing of hardware components, the required level of system safety can be achieved. In that case, the system itself appears to be a “black box”, but with perfectly known outputs at an absolutely known spatial coordinate.

#### 4.2. Conceptual safety model of an automatic train operation system

An ATO system includes the following key facilities:

- onboard train control and protection equipment;
- monitoring centre equipment;
- trackside machine vision facilities;
- onboard machine vision facilities.

The conceptual safety model of an automatic train operation system contains a description of the dependability and safety states of the system’s component facilities, their interrelations, as well as the effects of disturbing weather effects. This model is presented in the form of a system safety state graph (Fig. 4).

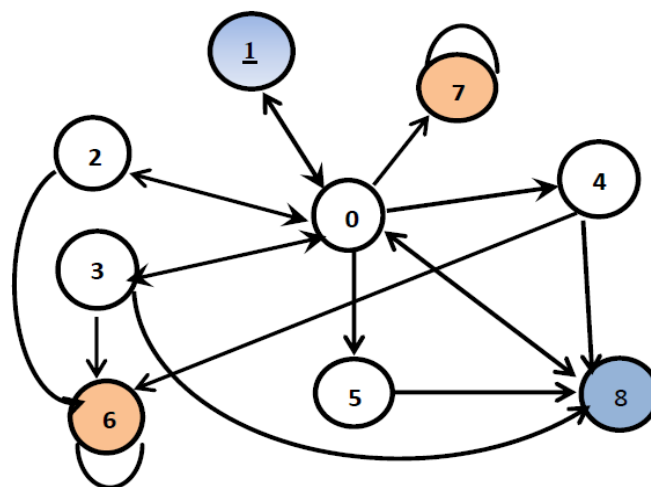


Figure 4. Safety state graph of an automatic train operation system

While building the system safety model, the following *criterion of dangerous failure* was adopted: failure of all machine vision facilities and monitoring centre or an undetected failure of a locomotive control and protection system. *Criterion of safe failure*: failure of fixed machine vision facilities, monitoring centre and effect of disturbing weather conditions or detected failure of the locomotive control and protection system.

Graph states:

0, up state, no disturbing weather effects;

1, detected locomotive control and protection system failure, *a safe failure*;

2, failure of monitoring centre facilities;

3, failure of trackside machine vision facilities;

4, failure of onboard machine vision facilities;

5, disturbing weather effects;

6, failure of all machine vision facilities and monitoring centre, *a dangerous failure* of the automatic train operation system;

7, non-detected locomotive control and protection system failure, *a dangerous failure*;

8, failure of trackside machine vision facilities, monitoring centre and disturbing weather effects, *a safe failure*.

The entire set of system states according to the state graph in Fig. 4 is divided into the following subsets: - the subset of up states  $S_U = \{0,2,3,4,5\}$ ; - the subset of safety states  $S_S = \{1,8\}$ ; - the subset of hazardous states  $S_H = \{6,7\}$ .

The up and safe states form the set of good states.

Given below are the model's good state transitions that need clarification: 1-0, 2-0, 3-0, 8-0, repair of facilities after failures; 3-8, monitoring centre failure subject to trackside machine vision facilities failure; 4-8, monitoring centre failure subject to onboard machine vision facilities failure; 7-8, failure of trackside machine vision facilities subject to disturbing weather conditions.

The mathematical formulation of the model takes into account the following considerations. The system is new and unique, no statistical information about it is available. Therefore, the distribution functions of system parameters are not established. Based on the existing experience in railway control and management systems, it can be safely assumed that failures of electronic devices, such as devices of a train control and protection system, monitoring centre facilities, and machine vision facilities are exponentially distributed. This assumption does not apply to random values of time to device repair restoration after failures, much less to random adverse weather effects. The problem of disturbing effects was theoretically examined by Schäbe and Viertl in [23]. Those models are also applicable to disturbing weather effects. In order to ensure adequate results, the authors were forced to use a complex mathematical description of the random process of adverse effects on the locomotive's control system. The above circumstances complicate their practical application in mathematical simulation of the safety of the automatic train operation system.

In the absence of practical information, it is very difficult to predict the quantitative safety indicators of the automatic train operation system. In this paper, in the context of great uncertainty, we aim to identify the most significant factors affecting the system's safety. The assumption of the Poisson process of random events in the automatic train operation system fits this purpose. The Poisson processes are ordinary, stationary and have no aftereffect. On the one hand, due to the significant uncertainty in the initial conditions, their application does not contribute to accurate prediction of the safety of a system's behaviour characteristics. On the other hand, the obtained outputs can be regarded as conservative bounds for constructing a safe ATO system by neutralising the identified

most significant adverse factors. Thus, the used model is simplified, but it enables an analytical study of the problem. That constitutes the advantage of this approach over more complex models.

The adopted assumptions defined mathematical models of the graph's edges and nodes, the FS formulas, as well as the expert evaluation of the initial data are provided by us in [25].

The limit value of an automatic train operation system's time to dangerous failure takes place subject to the absence of destructive disturbing weather conditions ( $\gamma \rightarrow 0$ ) and compliance with the requirements of IEC 61508-2 [6] ( $\alpha \rightarrow 0$ ).

Under those conditions, the probability of dangerous failure with an error not exceeding the first order of smallness tends to the following form:

$$G_{WS}(t) \cong \lambda_{WS} \cdot t \rightarrow \frac{\lambda}{2} t,$$

where  $\lambda$  is the failure rate of the machine vision facilities (it is assumed that the onboard and fixed facilities have about the same dependability).

## 5. Conclusion

Ensuring the FS of an automatic train operation requires not only developing or applying known methods of designing a safe system, but, most importantly, proving the acceptability of the achieved level of FS. In respect to automatic train operation systems, the conventional methods of proving the FS (statistical, experimental, expert, simulation) are of limited use. That is due to the distributed architecture of the systems, changing information processing algorithms in the course of training, large number of vulnerabilities, etc. For the purpose of improving the confidence in the FS evaluation results, it is proposed to focus on the technological methods and use the widely applied analytical expert semi-Markov method, proposed here.

The proposed process of monitoring the operation of ASCSS and machine vision facilities, creation of a second, virtual channel allow improving the FS of the shunting locomotive control system with machine vision from SIL 2 to SIL 3 and maintaining it over a sufficiently long period of time (over a quarter of the mean time to failure of the ASCSS). The mean time of faultless operation of the shunting locomotive control system may grow almost three times as long as the achieved level of the system's FS remains unchanged. Additionally, the time of faultless operation of the locomotive until it has to be brought to a complete stop for safety reasons can also increase over three times. This important result can be practically achieved despite the increased amount of the system's equipment due to the introduction of machine vision facilities.

A general approach to ensuring the FS of an ATO is proposed. It is based on the division of the information processing process into two subprocesses, i.e., internal intelligent information processing onboard the locomotive for the purpose of decision-making regarding track vacancy and communication of initial visual information to the locomotive driver for decision-making. The division of this process must be combined with redundant machine vision facilities, regular comparison of the outputs of the onboard and fixed machine vision facilities, redundant comparison outputs, smoothing of the outputs in the process of locomotive movement. The EN 50129 functional safety requirements for the locomotive control and protection system and SIL4 requirements for the machine vision facilities are to be fulfilled as well. Additionally, adverse weather effects are to be countered by improving the efficiency of machine learning of the machine vision software.

## References

1. Vapnik V.N., Chervonenkis A.Ya. Pattern recognition theory (statistical problems of training). Moscow: Nauka; 1974. (in Russ.)
2. Braband J., Schäbe H. On safety assessment of artificial intelligence. *Dependability* 2020;4:25-34.
3. I.B. Shubinsky, E.N. Rozenberg, H. Schabe INNOVATIVE METHODS OF ENSURING THE FUNCTIONAL SAFETY OF TRAIN CONTROL SYSTEMS. *Reliability: Theory & Applications*. 2023, December 4(76): 909-920, DOI: <https://doi.org/10.24412/1932-2321-2023-476-909-920>
4. Shubinsky I.B., Rozenberg E.N. General provisions of the substantiation of functional safety of intelligent systems in railway transportation. *Dependability* 2023;23(3):38-45.
5. STO RZD 1-19.009-2009. Railway signalling devices and systems. Safety case.
6. IEC 61508-1:2012. Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems. Part 1. General requirements. Moscow: Standartinform; 2014.
7. IEC 61508-2:2012. Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems. Part 2. System requirements. Moscow: Standartinform; 2014.
8. IEC 61508-3:2018. Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems. Part 3. Software requirements. Moscow: Standartinform; 2018.
9. IEC 62279-2016. Railway applications. Communication, signalling and processing systems. Software for railway control and protection systems. Moscow: Standartinform; 2017.
10. GOST 33432-2015. Functional safety. Policy and programme of safety provision. Safety proof of the railway objects. Moscow: Standartinform; 2019. (in Russ.)
11. Braband J., Schäbe H. Risk analysis for automated driving – validation and findings. *Signal+ Draht* 2023;115(4).
12. Shubinsky I.B., Rozenberg E.N. Functional safety of railway control systems. Vologda: Infra-Inzheneria; 2023. (in Russ.)
13. Shubinsky I.B. Dependable, fault-tolerant information systems. Methods of synthesis. Moscow: Dependability; 2016. (in Russ.)
14. Shubinsky I.B. Structural dependability of information systems. Methods of analysis. Moscow: Dependability; 2012. (in Russ.)
15. Rozenberg E.N. Multi-level train control and protection system: Doctoral Thesis (Doctor of Technical Sciences): 05.13.06, 05.22.08. Moscow; 2004. (in Russ.)
16. Okhotnikov A.L., Popov P.A. Self-driving: yesterday, today and tomorrow. *Automation, Communications, Informatics* 2019;8. (in Russ.)
17. Shvir V. Dependability of electronic circuits in railway signalling devices. *Rail International* 1986;1:59-67. (in Russ.)
18. Kalinin A.V. Driverless shunting locomotives control. Main principles and prospects of technology development. *Intellektualnye IT upravleniya ITNOU* 2017;1:12-14. (in Russ.)
19. Shubinsky I.B., Rozenberg E.N., Korovin A.S., Penkova N.G. On a method for ensuring functional safety of a system with single-channel information processing. *Dependability* 2022;22(3):44-52.
20. Sapozhnikov V.V., Sapozhnikov V.I., Khristov Kh.A., Gavzov D.V. Methods for constructing safe computer-based railway signalling systems. Moscow: Transport; 1995. (in Russ.)
21. Mylnikov P.D., Okhotnikov A.P., Popov P.A. Onboard information system. Pat. no. 2742960 dated 12.02.2021. Bul. no. 5 N. (in Russ.)
22. Shubinsky I.B. Functional dependability of information systems. Methods of analysis. Moscow: Dependability; 2012. (in Russ.)

23. Shubinsky I.B., Rozenberg E.N., Panfiorov I.A., Boyarinova N.A., Kuzmin A.I. Estimating the safety and reliability of the control system of a locomotive with machine vision. *Dependability* 2023;23(1):30-37. (in Russ.)
24. Schäbe H., Viertl R. An axiomatic approach to models of accelerated life testing. *Eng. Fract. Mechanics* 1995;50(2):203-217.
25. Shubinsky I.B., Schäbe H., Rozenberg E.N. On the safety assessment of an automatic train operation system. *Dependability* 2021;21(4):31-37.

# SENSITIVITY AND PROFITABILITY ANALYSIS OF TWO-UNITS AMMONIA/UREA PLANT

SARA SALIM AL ORAIMI

Department of Mechanical and Industrial Engineering, National University of Science and Technology, Sultanate of Oman

SYED MOHD RIZWAN

Department of Applied Mathematics and Science, National University of Science and Technology, Sultanate of Oman

KAJAL SACHDEVA

Department of Mathematics, Maharshi Dayanand University, Rohtak, 124001, Haryana, India

## Abstract

*This paper presents a reliability modelling of a two-unit ammonia/urea plant. Real maintenance data of the production plant have been used for this purpose. Four types of failure were noted: process, electrical, mechanical and instrumental failures. Both ammonia/urea formation units work in parallel and do not fail simultaneously. Various reliability indices of the plant, such as availability, busy period for repair, and expected number of repairs for each type of failure, have been obtained. Markov processes and regenerative point techniques are used for analysis. Profit analysis for the plant is also done, along with a graphical representation of various parameters. Finally, sensitivity analysis is carried out to see the impact of varied parameters on the profit function of the plant.*

**Keywords:**Ammonia plant; Markov process; regenerative point techniques; repairs; failures

## 1. INTRODUCTION

Many researchers have studied complex industrial systems under various operating conditions and presumptions, contributing to the discipline of reliability modelling and analysis. Rizwan et al. [1,2] presented a reliability modelling strategy and its application to industries, specifically focusing on a biscuit manufacturing factory controlled by a single-unit and two-unit hot standby PLC system. Mathew et al. [3,4] discussed the reliability modelling of single-unit and two-unit systems in a continuous casting plant. They utilized real maintenance data from a steel production plant to analyze the system's performance and identify different types of failures. Also, Mathew et al. [5] did a comparative analysis of the two models of the CC plant. Mathew et al. [6,7] presented reliability modelling in an actual CC plant with different installed and full installed capacities, where two EOT cranes operate in parallel. Padmavati et al. [8–13] evaluated the impact of prioritizing repair over maintenance on the overall reliability and availability of the desalination plant. Also, they discussed the implications of the research for the design and operation of desalination plants in terms of cost-effectiveness and efficiency by taking different assumptions for failures and repairs. Rizwan et al. [14–16] analyzed the reliability of a wastewater treatment plant. They estimated various reliability indices associated with the plant. Also, they



highlighted the importance of regular monitoring, repairs, and replacements in maintaining the reliability of systems.

Al Rahbi et al. [17–23] analyzed the reliability of single unit/ multiple units of a rodding anode plant in the aluminium industry with single/numerous repairers and optimized maintenance strategies, improved system reliability, and reduced downtime, leading to increased productivity and cost savings. Taj et al. [24–30, 33] evaluated reliability analysis conducted on the cable plant's single- or three-unit machine subsystem with repair priority over maintenance. Rizwan et al. [31] examined the three pumps' performance in distributing desalinated water. The study included maintenance data collected over five years, encompassing various failure reasons, restoration times, and waiting times. Rizwan et al. [32] explored the reliability and sensitivity analysis of Membrane Biofilm Fuel Cells. Thus, the literature has widely discussed the reliability modelling and analysis of complex industrial systems in various failure/maintenance circumstances. But the concept of reliability analysis for ammonia/urea manufacturing plants has yet to be discussed.

The demand for fertilizers is growing day by day around the world to meet agricultural requirements. The most used or consumed fertilizer is UREA, which is manufactured from ammonia, from different industrial chemical reactions. The UREA fertilizer manufacturing facilities consist ammonia manufacturing plant along with a urea plant. To meet the growing demand for urea in the current market, these facilities 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 cause 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 [34, 35], etc., may be carried out to ensure continuous and smooth plant operations. This paper provides sensitivity and profitability analysis, along with reliability analysis of parallel ammonia and urea plants worldwide that have operated for more than 15 years. The research is based on the actual plant data, with some assumptions, failure rate, or probability.

## 2. NOTATIONS

The following are the notations used in the analysis:

$\lambda_u$  = failure rate of ammonia plant

$p_1/ p_2/ p_3/ p_4$  = probability of process failure/ electrical failure/ mechanical failure/ instrumental failure in unit 1.

$p_5/ p_6/ p_7/ p_8$  = probability of process/ electrical/ mechanical/ instrumental failure in unit 2.

$\alpha_1/ \alpha_2/ \alpha_3/ \alpha_4$  = repair rate of process/ electrical/ mechanical/ instrumental failure in unit 1.

$\alpha_5/ \alpha_6/ \alpha_7/ \alpha_8$  = repair rate of process/ electrical/ mechanical/ instrumental failure in unit 2.

$f^u(t)$  = 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 process/ electrical/ mechanical/ instrumental failure in unit 1.

$g_5(t)/ g_6(t)/ g_7(t)/ g_8(t)$  = p.d.f. of repair time due to process/ electrical/ mechanical/ instrumental failure in unit 2.

## 3. DATA SUMMARY

The real data from a urea manufacturing company is summarized as follows:

Probability of process failure in unit 1,  $p_1 = 0.2088$

Probability of electrical failure in unit 1,  $p_2 = 0.0220$

Probability of mechanical failure in unit 1,  $p_3 = 0.1758$

Probability of instrumental failure in unit 1,  $p_4 = 0.1978$

Probability of process failure in unit 2,  $p_5 = 0.1648$

- Probability of electrical failure in unit 2,  $p_6 = 0.0220$
- Probability of mechanical failure in unit 2,  $p_7 = 0.1868$
- Probability of instrumental failure in unit 2,  $p_8 = 0.0220$
- Failure rate of urea plant,  $\lambda_u = 0.00031$  per hour
- Repair rate of process failure in unit 1,  $\alpha_1 = 0.0249$  per hour
- Repair rate of electrical failure in unit 1,  $\alpha_2 = 0.2$  per hour
- Repair rate of mechanical failure in unit 1,  $\alpha_3 = 0.0081$  per hour
- Repair rate of instrumental failure in unit 1,  $\alpha_4 = 0.01833$  per hour
- Repair rate of process failure in unit 2,  $\alpha_5 = 0.0150$  per hour
- Repair rate of electrical failure in unit 2,  $\alpha_6 = 0.0175$  per hour
- Repair rate of mechanical failure in unit 2,  $\alpha_7 = 0.0057$  per hour
- Repair rate of instrumental failure in unit 2,  $\alpha_8 = 0.0392$  per hour

#### 4. MODEL DESCRIPTION AND ASSUMPTIONS

- Initially, we have an operative ammonia manufacturing plant composed of two parallel units: Unit 1 and Unit 2.
- The four types of failures are observed in both units, i.e., process, electrical, mechanical, and instrumental.
- Both units cannot fail simultaneously.
- Repair is carried out upon failures.
- Failure rate and repair rates all are taken as general.

#### 5. STOCHASTIC MODEL

Table 1 shows the rates of transition from state  $i (S_i)$  to state  $j (S_j)$ . The set of states  $\{0,1,2,3,\dots,8\}$  all are operative and regenerative.

**Table 1**  
 State Transition Table

$S_i / S_j$	$S_0$	$S_1$	$S_2$	$S_3$	$S_4$	$S_4$	$S_6$	$S_7$	$S_8$
$S_0$	0	$p_1 f^u(t)$	$p_2 f^u(t)$	$p_3 f^u(t)$	$p_4 f^u(t)$	$p_5 f^u(t)$	$p_6 f^u(t)$	$p_7 f^u(t)$	$p_8 f^u(t)$
$S_1$	$g_1(t)$	0	0	0	0	0	0	0	0
$S_2$	$g_2(t)$	0	0	0	0	0	0	0	0
$S_3$	$g_3(t)$	0	0	0	0	0	0	0	0
$S_4$	$g_4(t)$	0	0	0	0	0	0	0	0
$S_5$	$g_5(t)$	0	0	0	0	0	0	0	0
$S_6$	$g_6(t)$	0	0	0	0	0	0	0	0
$S_7$	$g_7(t)$	0	0	0	0	0	0	0	0
$S_8$	$g_8(t)$	0	0	0	0	0	0	0	0

where,

- State 0 ( $S_0$ ) - Both urea processing machines unit 1 and 2 operative.
- State 1 ( $S_1$ ) - Unit 1 failed due to process failure, and Unit 2 is still operative.
- State 2 ( $S_2$ ) - Unit 1 failed due to electrical failure, and Unit 2 is still operative.
- State 3 ( $S_3$ ) - Unit 1 failed due to mechanical failure, and Unit 2 is still operative.
- State 4 ( $S_4$ ) - Unit 1 failed due to instrumental failure, and Unit 2 is still operative.
- State 5 ( $S_5$ ) - Unit 1 is operative, and Unit 2 failed due to process failure.
- State 6 ( $S_6$ ) - Unit 1 is operative, and Unit 2 failed due to electrical failure.
- State 7 ( $S_7$ ) - Unit 1 is operative, and Unit 2 failed due to mechanical failure.
- State 8 ( $S_8$ ) - Unit 1 is operative, and Unit 2 failed due to instrumental failure.

The transition probability from state  $i (S_i)$  to state  $j (S_j)$ ,  $q_{ij}(t)$  is given by

$$\begin{aligned} q_{01}(t) &= p_1 f^u(t), & q_{10}(t) &= g_1(t), \\ q_{02}(t) &= p_2 f^u(t), & q_{20}(t) &= g_2(t), \\ q_{03}(t) &= p_3 f^u(t), & q_{30}(t) &= g_3(t), \\ q_{04}(t) &= p_4 f^u(t), & q_{40}(t) &= g_4(t), \\ q_{05}(t) &= p_5 f^u(t), & q_{50}(t) &= g_5(t), \\ q_{06}(t) &= p_6 f^u(t), & q_{60}(t) &= g_6(t), \\ q_{07}(t) &= p_7 f^u(t), & q_{70}(t) &= g_7(t), \\ q_{08}(t) &= p_8 f^u(t), & q_{80}(t) &= g_8(t) \end{aligned}$$

The steady-state probability,  $p_{ij}$  as

$$\begin{aligned} p_{01} = p_1, \quad p_{02} = p_2, \quad p_{03} = p_3, \quad p_{04} = p_4, \quad p_{05} = p_5, \quad p_{06} = p_6, \quad p_{07} = p_7, \quad p_{08} = p_8 \\ p_{10} = p_{20} = p_{30} = p_{40} = p_{50} = p_{60} = p_{70} = p_{80} = 1 \end{aligned} \quad (1)$$

Sojourn time ( $\mu_i$ ), i.e., mean stay time in particular state  $i$ , is given as

$$\begin{aligned} \mu_0 &= \int_0^\infty t \cdot f^u(t) dt, & \mu_5 &= \int_0^\infty t \cdot g_5(t) dt, \\ \mu_1 &= \int_0^\infty t \cdot g_1(t) dt, & \mu_6 &= \int_0^\infty t \cdot g_6(t) dt, \\ \mu_2 &= \int_0^\infty t \cdot g_2(t) dt, & \mu_7 &= \int_0^\infty t \cdot g_7(t) dt, \\ \mu_3 &= \int_0^\infty t \cdot g_3(t) dt, & \mu_8 &= \int_0^\infty t \cdot g_8(t) dt, \\ \mu_4 &= \int_0^\infty t \cdot g_4(t) dt, \end{aligned}$$

The contribution to mean sojourn time,  $m_{ij}$ , is given by

$$m_{ij} = \int_0^\infty t \cdot q_{ij}(t) dt. \text{ It can be verified that}$$

$$m_{01} + m_{02} + m_{03} + m_{04} + m_{05} + m_{06} + m_{07} + m_{08} = \mu_0,$$

$$\begin{aligned} m_{10} = \mu_1, \quad m_{20} = \mu_2, \quad m_{30} = \mu_3, \quad m_{40} = \mu_4, \\ m_{50} = \mu_5, \quad m_{60} = \mu_6, \quad m_{70} = \mu_7, \quad m_{80} = \mu_8. \end{aligned}$$

## 6. SYSTEM PERFORMANCE MEASURES

### 6.1. Availability of the System

Define

$A_i^u(t)$  = probability that it is operative at time  $t$ , given that the system is in state  $i$  at time  $t = 0$ .

Using the state transitions, we get the following equations:

$$\begin{aligned}
 A_0^u(t) &= M_0(t) + q_{01}(t) \otimes A_1^u(t) + q_{02}(t) \otimes A_2^u(t) + q_{03}(t) \otimes A_3^u(t) + q_{04}(t) \otimes A_4^u(t) + q_{05}(t) \otimes A_5^u(t) \\
 &\quad + q_{06}(t) \otimes A_6^u(t) + q_{07}(t) \otimes A_7^u(t) + q_{08}(t) \otimes A_8^u(t) \\
 A_1^u(t) &= M_1(t) + q_{10}(t) \otimes A_0^u(t) \\
 A_2^u(t) &= M_2(t) + q_{20}(t) \otimes A_0^u(t) \\
 A_3^u(t) &= M_3(t) + q_{30}(t) \otimes A_0^u(t) \\
 A_4^u(t) &= M_4(t) + q_{40}(t) \otimes A_0^u(t) \\
 A_5^u(t) &= M_5(t) + q_{50}(t) \otimes A_0^u(t) \\
 A_6^u(t) &= M_6(t) + q_{60}(t) \otimes A_0^u(t) \\
 A_7^u(t) &= M_7(t) + q_{70}(t) \otimes A_0^u(t) \\
 A_8^u(t) &= M_8(t) + q_{80}(t) \otimes A_0^u(t)
 \end{aligned} \tag{2}$$

where

$M_i(t)$  = probability that the system stays in state  $i$  while operating rather than transferring to any other state.

Taking Laplace transform of equations (28)-(36) and solving for  $A_0^{u*}(s)$ , we get

$$A_0^{u*}(s) = \frac{N_1^u(s)}{D_1^u(s)}$$

where

$$\begin{aligned}
 N_1^u(s) &= M_0^*(s) + q_{01}^*(s)M_1^*(s) + q_{02}^*(s)M_2^*(s) + q_{03}^*(s)M_3^*(s) + q_{04}^*(s)M_4^*(s) \\
 &\quad + q_{05}^*(s)M_5^*(s) + q_{06}^*(s)M_6^*(s) + q_{07}^*(s)M_7^*(s) + q_{08}^*(s)M_8^*(s)
 \end{aligned} \tag{3}$$

$$\begin{aligned}
 D_1^u(s) &= 1 - q_{01}^*(s)q_{10}^*(s) - q_{02}^*(s)q_{20}^*(s) - q_{03}^*(s)q_{30}^*(s) - q_{04}^*(s)q_{40}^*(s) - q_{05}^*(s) \\
 &\quad - q_{06}^*(s)q_{60}^*(s) - q_{07}^*(s)q_{70}^*(s) - q_{08}^*(s)q_{80}^*(s)
 \end{aligned} \tag{4}$$

The steady-state availability of the system is given by :

$$A_0^u = \lim_{s \rightarrow 0} s.A_0^{u*}(s) = \lim_{s \rightarrow 0} s \cdot \frac{N_1^u(s)}{D_1^u(s)} = \frac{N_1^{u'}(0)}{D_1^{u'}(0)} = \frac{N_1^u}{D_1^u} \text{ (say)} \tag{5}$$

where

$$N_1^u = \mu_0 + p_1\mu_1 + p_2\mu_2 + p_3\mu_3 + p_4\mu_4 + p_5\mu_5 + p_6\mu_6 + p_7\mu_7 + p_8\mu_8 \tag{6}$$

$$D_1^u = p_1\mu_1 + p_2\mu_2 + p_3\mu_3 + p_4\mu_4 + p_5\mu_5 + p_6\mu_6 + p_7\mu_7 + p_8\mu_8 + \mu_0$$

## 6.2. Busy Period for Repair

The expected time for which the repair man is busy for the repair of unit 1 and unit 2 due to process failure in steady state is given by:

$$PB^{u1} = PN_2^{u1} / D_1^u \quad \text{and} \quad PB^{u2} = PN_2^{u2} / D_1^u$$

where

$$PN_2^{u1} = p_{01}u_1 = p_1u_1$$

$$PN_2^{u2} = p_{05}u_5 = p_5u_5$$

The expected time for which the repair man is busy for the repair of unit 1 and unit 2 due to electrical failure in steady state is given by:

$$EB^{u1} = EN_2^{u1} / D_1^u \quad \text{and} \quad EB^{u2} = EN_2^{u2} / D_1^u$$

where

$$EN^{u1} = p_{02}u_2 = p_2u_2$$

$$EN^{u2} = p_{06}u_6 = p_6u_6$$

The expected time for which the repair man is busy for the repair of unit 1 and unit 2 due to mechanical failure in steady state is given by:

$$MB^{u1} = MN_2^{u1} / D_1^u \quad \text{and} \quad MB^{u2} = MN_2^{u2} / D_1^u$$

where

$$MN_2^{u1} = p_{03}u_3 = p_3u_3$$

$$MN_2^{u2} = p_{07}u_7 = p_7u_7$$

The expected time for which the repair man is busy for the repair of unit 1 and unit 2 due to instrumental failure in steady state is given by:

$$IB^{u1} = IN_2^{u1} / D_1^u \quad \text{and} \quad IB^{u2} = IN_2^{u2} / D_1^u \quad \text{where}$$

$$IN_2^{u1} = p_{04}u_4 = p_4u_4$$

$$IN_2^{u2} = p_{08}u_8 = p_8u_8$$

### 6.3. Expected Number of Repairs

The expected number of repairs in unit 1 and unit 2 due to process failure in steady state is given by:

$$PR^{u1} = PN_3^{u1} / D_1^u \quad \text{and} \quad PR^{u2} = PN_3^{u2} / D_1^u$$

where

$$PN_3^{u1} = p_{01}p_{10} = p_1$$

$$PN_3^{u2} = p_{05}p_{50} = p_5$$

The expected number of repairs in unit 1 and unit 2 due to electrical failure in steady state is given by:

$$ER^{u1} = EN_3^{u1} / D_1^u \quad \text{and} \quad ER^{u2} = EN_3^{u2} / D_1^u$$

where

$$EN_3^{u1} = p_{02}p_{20} = p_2$$

$$EN_3^{u2} = p_{06}p_{60} = p_6$$

The expected number of repairs in unit 1 and unit 2 due to mechanical failure in steady state is given by:

$$MR^{u1} = MN_3^{u1} / D_1^u \quad \text{and} \quad MR^{u2} = MN_3^{u2} / D_1^u$$

where

$$MN_3^{u1} = p_{03}p_{30} = p_3$$

$$MN_3^{u2} = p_{07}p_{70} = p_7$$

The expected number of repairs in unit 1 and unit 2 due to instrumental failure in steady state is given by:

$$IR^{u1} = IN_3^{u1} / D_1^u \quad \text{and} \quad IR^{u2} = IN_3^{u2} / D_1^u$$

where

$$IN_3^{u1} = p_{04}p_{40} = p_4$$

$$IN_3^{u2} = p_{08}p_{80} = p_8$$

## 7. PROFIT ANALYSIS OF THE SYSTEM

The profit equation of the system is as follows:

$$P^u = C_0A_0^u - C_1(PB^{u1} + PR^{u1}) - C_2(PB^{u2} + PR^{u2}) - C_3(EB^{u1} + ER^{u1}) - C_4(EB^{u2} + ER^{u2}) \\ - C_5(MB^{u1} + MR^{u1}) - C_6(MB^{u2} + MR^{u2}) - C_7(IB^{u1} + IR^{u1}) - C_8(IB^{u2} + IR^{u2})$$

where

$C_0$  = Revenue generated by the system

$C_1(C_2)/C_3(C_4)/C_5(C_6)/C_7(C_8)$ : - Cost per unit time for engaging the repair man and cost for repair due to process/electrical/mechanical/instrumental failure in unit 1 (unit 2).

## 8. NUMERICAL ANALYSIS

In this section, interpretation from graphs and tables has been made for the above-obtained system measures in Section 4 and Section 5. Let us assume all the failures and repair times follow exponential distribution along with their p.d.f. as:

$$f^u(t) = \lambda_u e^{-\lambda_u t},$$

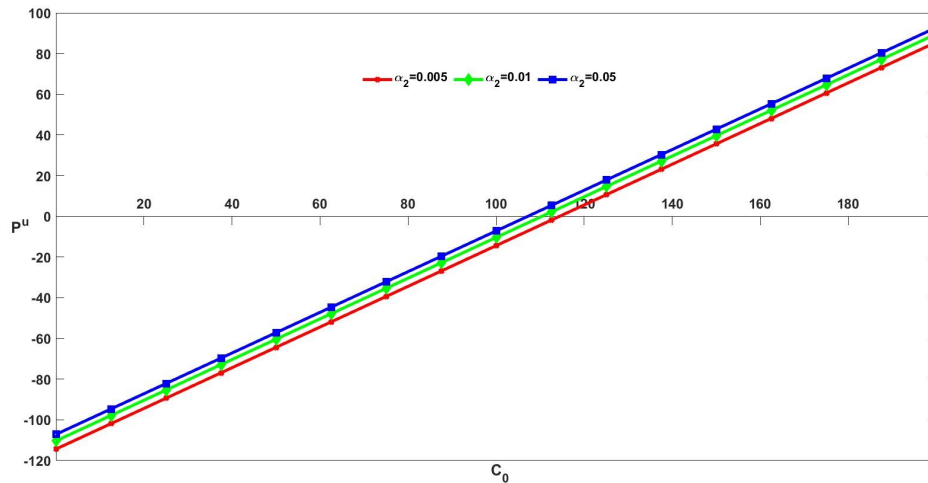
$$g_i(t) = \alpha_i e^{-\alpha_i t}, i = 1, 2, 3, \dots, 8.$$

Using the values as written in Section 3 that is calculated from real data from a manufacturing company, we get system effectiveness measures as:

- Availability of Ammonia Plant,  $A_0^u = 1$
- Busy Period for Repair of Unit 1 due to Process Failure,  $PB^{u1} = 0.0025$
- Busy Period for Repair of Unit 2 due to Process Failure,  $PB^{u2} = 0.0033$
- Busy Period for Repair of Unit 1 due to Electrical Failure,  $EB^{u1} = 3.3209 * 10^{-5}$
- Busy Period for Repair of Unit 2 due to Electrical Failure,  $EB^{u2} = 3.7953 * 10^{-4}$
- Busy Period for Repair of Unit 1 due to Mechanical Failure,  $MB^{u1} = 0.0066$
- Busy Period for Repair of Unit 2 due to Mechanical Failure,  $MB^{u2} = 0.0099$
- Busy Period for Repair of Unit 1 due to Instrumental Failure,  $IB^{u1} = 0.0033$
- Busy Period for Repair of Unit 2 due to Instrumental Failure,  $IB^{u2} = 1.6943 * 10^{-4}$
- Expected no. of Repair of Unit 1 due to Process Failure,  $PR^{u1} = 6.3036 * 10^{-5}$
- Expected no. of Repair of Unit 2 due to Process Failure,  $PR^{u2} = 4.9753 * 10^{-5}$
- Expected no. of Repair of Unit 1 due to Electrical Failure,  $ER^{u1} = 6.6418 * 10^{-6}$
- Expected no. of Repair of Unit 2 due to Electrical Failure,  $ER^{u2} = 6.6418 * 10^{-6}$

- Expected no. of Repair of Unit 1 due to Mechanical Failure,  $MR^{u1} = 5.3074 * 10^{-5}$
- Expected no. of Repair of Unit 2 due to Mechanical Failure,  $MR^{u2} = 5.6395 * 10^{-5}$
- Expected no. of Repair of Unit 1 due to Instrumental Failure,  $IR^{u1} = 5.9715 * 10^{-5}$
- Expected no. of Repair of Unit 2 due to Instrumental Failure,  $IR^{u2} = 5.8712 * 10^{-5}$

The graph of profit function ( $P^u$ ) w.r.t. revenue ( $C_0$ ) for different values of repair rate ( $\alpha_1$ ) has been shown in Figure 1.



**Figure 1:** Change in Profit w.r.t. Revenue and Repair Rate

It shows that the increase in revenue and repair rate increases profit. Also, the cut-off points for the system to be profitable can be observed in Fig. 1.:

- 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.

**Table 2.** Sensitivity and Relative Sensitivity Analysis of Profit Function

Parameter	Sensitivity Analysis	Relative Sensitivity Analysis
$\lambda_u$	$-3.3476 * 10^5$	-0.0212
$\alpha_1$	59.2593	$3.0154 * 10^{-4}$
$\alpha_2$	1.0047	$4.1063 * 10^{-5}$
$\alpha_3$	$3.8409 * 10^3$	0.0064
$\alpha_4$	$3.5789 * 10^3$	0.0134
$\alpha_5$	17.3446	$5.3167 * 10^{-5}$
$\alpha_6$	5.0857	$1.8188 * 10^{-5}$
$\alpha_7$	578.7666	$6.7416 * 10^{-4}$
$\alpha_8$	3.7234	$2.9827 * 10^{-5}$
$C_0$	1	1.0218
$C_1$	-0.0026	$-3.6631 * 10^{-4}$
$C_2$	-0.0034	$-1.2854 * 10^{-4}$
$C_3$	$-3.9851 * 10^{-5}$	$-5.0143 * 10^{-5}$
$C_4$	$-3.8617 * 10^{-4}$	$-2.6915 * 10^{-5}$
$C_5$	-0.0066	-0.0065
$C_6$	-0.01	$-8.9916 * 10^{-4}$
$C_7$	-0.0033	-0.0137
$C_8$	$-1.7607 * 10^{-4}$	$-3.4829 * 10^{-5}$

Table 2 shows the sensitivity and relative sensitivity analysis [34] of the profit function concerning different parameters that affect the system's profit. It shows that the profit function decreases rapidly with the change in  $C_3$  and increases with the shift in  $\alpha_3$ .

Also, the decreasing order in which parameters affect the profit function from Table 2 as:

$C_0 > \lambda_u > C_7 > \alpha_4 > C_5 > \alpha_3 > C_6 > \alpha_7 > C_1 > \alpha_1 > C_2 > \alpha_5 > C_3 > \alpha_2 > C_8 > \alpha_8 > C_4 > \alpha_6$ .

## 9. CONCLUSION

In this paper, the parallel functioning of two units of an ammonia/urea plant reliability modelling has been examined. The availability of the plant, the busy period for repairs, and the anticipated number of repairs for each type of failure have all been obtained as reliability indices. Profit analysis for the plant is also carried out along with the graphical representation with respect to various parameters. Profit increases when revenue and repair rates both increase. The cut-off point is also drawn to determine when a system is profitable. Finally, sensitivity analysis is performed to assess the effect of various parameters on the plant's profit function. It demonstrates that, in comparison to other factors, revenue 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.

### Acknowledgement

This research is funded under a student's research grant from the College of Engineering, National University of Science and Technology, Sultanate of Oman.

### REFERENCES

- [1] Rizwan, S. M., Khurana, V., and Taneja, G. (2007). Modelling and Optimization of a Single-Unit plc System. *International Journal of Modelling and Simulation*, 27(4), 361-368.
- [2] Rizwan, S. M., Khurana, V., and Taneja, G. (2010). Reliability analysis of a hot standby industrial system. *International Journal of Modelling and Simulation*, 30(3), 315-322.



- [3] Mathew, A. G., Rizwan, S. M., Majumder, M. C., and Ramachandran, K. P. (2011). Reliability modelling and analysis of a two unit continuous casting plant. *Journal of the Franklin Institute*, 348(7), 1488-1505.
- [4] Mathew, A. G., Rizwan, S. M., Majumder, M. C., Ramachandran, K. P., Taneja, G. (2009). Profit evaluation of a single unit CC plant with scheduled maintenance. *Caledonian Journal of Engineering*, 5(1), 25-33.
- [5] Mathew, A. G., Rizwan, S. M., Majumder, M. C., Ramachandran, K. P., and Taneja, G. (2010, October). Comparative analysis between profit of the two models of a CC plant. *American Institute of Physics*. In AIP Conference Proceedings (Vol. 1298, No. 1, pp. 226-231).
- [6] Mathew, A. G., Rizwan, S. M., Majumder, M. C., Ramachandran, K. P., and Taneja, G. (2010). Reliability modeling and analysis of a two-unit parallel CC plant with different installed capacities. *Journal of Manufacturing Engineering*, 5(3), 197-204.
- [7] Mathew, A. G., Rizwan, S. M., Majumder, M. C., Ramachandran, K. P., and Taneja, G. (2011). Reliability analysis of identical two-unit parallel CC plant system operative with full installed capacity. *International Journal of Performability Engineering* 7(2): 179.
- [8] Padmavathi, N., Rizwan, S. M., Anita Pal, and Gulshan Taneja. (2012) Reliability analysis of an evaporator of a desalination plant with online repair and emergency shutdowns. *Aryabhata Journal of Mathematics and Informatics*, 4(1): 1-12.
- [9] Padmavathi, N., Rizwan, S. M., Pal, A., and Taneja, G. (2013, October). Comparative analysis of the two models of an evaporator of a desalination plant. In *Proc. of International Conference on Information and Mathematical Science, Punjab, India* (pp. 418-422).
- [10] Rizwan, S. M., Pal, A., and Taneja, G. (2013). Probabilistic analysis of an evaporator of a desalination plant with priority for repair over maintenance. *International Journal of Scientific and Statistical Computing*, 4(1), 1-9.
- [11] Rizwan, S. M., Pal, A., and Taneja, G. Probabilistic Analysis of a Desalination Plant with Major and Minor Failures and Shutdown During Winter Season. *i-manager's Journal on Mathematics*, 3(2): 21-26.
- [12] RIZWAN, V. M., Padmavathi, N., Pal, A., and Taneja, G. (2013). Reliability analysis of a seven unit desalination plant with shutdown during winter season and repair/maintenance on FCFS basis. *International Journal of Performability Engineering*, 9(5), 523.
- [13] Padmavathi, N., Rizwan, S. M., and Senguttuvan, A. (2015). Comparative analysis between the reliability models portraying two operating conditions of a desalination plant. *International Journal of Core Engineering and Management*, 1(12), 1-10.
- [14] Rizwan, S. M., and Thanikal, J. V. (2014). Reliability analysis of a waste water treatment plant with inspection. *i-manager's Journal on Mathematics* 3(2), 21-26.
- [15] Rizwan, S. M., Thanikal, J. V., and Torrijos, M. (2014). A general model for reliability analysis of a domestic waste water treatment plant. *International Journal of Condition Monitoring and Diagnostic Engineering Management*, 17(3), 3-6.
- [16] Rizwan, S. M., Thanikal, J. V., Padmavathi, N., and Yazidi, H. (2015). Reliability and availability analysis of an anaerobic batch reactor treating fruit and vegetable waste. *International Journal of Applied Engineering Research*, 10(24), 44075-44079.
- [17] Al Rahbi, Y., Rizwan, S. M., Alkali, B. M., Cowell, A., and Taneja, G. (2017, September). Reliability analysis of a subsystem in aluminium industry plant. In *2017 6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)* (pp. 199-203). IEEE.
- [18] Al Rahbi, Y., Rizwan, S. M., Alkali, B. M., Cowell, A., and Taneja, G. (2017). Reliability analysis of rodding anode plant in aluminium industry. *International Journal of Applied Engineering Research*, 12(16), 5616-5623.
- [19] Al Rahbi, Y., Rizwan, S., Alkali, B., Cowell, A., and Gulshan, T. (2018). Maintenance analysis of a butt thimble removal system in aluminium plant. *International Journal of Mechanical Engineering and Technology*, 9(4), 695-703.

- [20] Yaqoob Al Rahbi, R., SM, A., BM, A. C., and Taneja, G. (2018). Reliability analysis of rodding anode plant in an aluminum industry with multiple repair men. *Advances and Applications in Statistics*, 53(5), 569-597.
- [21] Al Rahbi, Y., Rizwan, S. M., Alkali, B. M., Cowell, A., and Taneja, G. (2019). Reliability analysis of a rodding anode plant in aluminum industry with multiple units failure and single repair man. *International Journal of System Assurance Engineering and Management*, 10, 97-109.
- [22] Al Rahbi, Y., Rizwan, S. M., Alkali, B., Cowell, A., and Taneja, G. (2019). Reliability analysis of multiple units with multiple repair men of rodding anode plant in aluminum industry. *Advances and Applications in Statistics*, 54(1), 151-178.
- [23] Al Rahbi, Y., and Rizwan, S. M. (2020, July). A Comparative Analysis between the Models of a Single Component with Single Repair man and Multiple Repair men of an Aluminium Industry. In *2020 International Conference on Computational Performance Evaluation (ComPE)* (pp. 132-135). IEEE.
- [24] Taj, S. Z., Rizwan, S. M., Alkali, B. M., Harrison, D. K., and Taneja, G. (2017, April). Reliability modelling and analysis of a single machine subsystem of a cable plant. In *2017 7th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO)* (pp. 1-4). IEEE.
- [25] Taj, S. M., Rizwan, S. M., Alkali, B. M., Harrison, D. K., and Taneja, G. L. (2017). Reliability analysis of a single machine subsystem of a cable plant with six maintenance categories. *International Journal of Applied Engineering Research*, 12(8), 1752-1757.
- [26] Taj, S. Z., Rizwan, S. M., Alkali, B. M., Harrison, D. K., and Taneja, G. (2017). Probabilistic modeling and analysis of a cable plant subsystem with priority to repair over preventive maintenance. *i-manager's Journal on Mathematics (JMAT)*, 6(3), 12-21.
- [27] Taj, S. Z., Rizwan, S., Alkali, B., Harrison, D., and Taneja, G. (2018). Reliability analysis of a 3-unit subsystem of a cable plant. *Advances and Applications in Statistics*, 52(6), 413-429.
- [28] Taj, S. Z., Rizwan, S. M., Alkali, B. M., Harrison, D. K., and Taneja, G. (2018). Performance and cost benefit analysis of a cable plant with storage of surplus produce. *International Journal of Mechanical Engineering and Technology*, 9(8), 814-826.
- [29] Taj, S. Z., Rizwan, S. M., Alkali, B. M., Harrison, D. K., and Taneja, G. (2018). Profit analysis of a cable manufacturing plant portraying the winter operating strategy. *International Journal of Mechanical Engineering and Technology*, 9(11), 370-381.
- [30] Taj, S. Z., Rizwan, S. M., Alkali, B. M., Harrison, D. K., and Taneja, G. (2020). Three reliability models of a building cable manufacturing plant: a comparative analysis. *International Journal of System Assurance Engineering and Management*, 11, 239-246.
- [31] Rizwan, S. M., Sachdeva, K., Alagirisamy, S., and Al Rahbi, Y. (2023). Performance and Sensitivity Analysis of the Three Pumps of a Desalination Water Pumping Station. *International Journal of Engineering Trends and Technology*, 71(1), 283-292.
- [32] Rizwan, S. M., Sachdeva, K., Al Balushi, N., Al Rashdi, S., and Taj, S. Z. Reliability and Sensitivity Analysis of Membrane Biofilm Fuel Cell. *International Journal of Engineering Trends and Technology*, 71(3), 73-80.
- [33] Taj, S. Z., and Rizwan, S. M. (2023). Comparative Analysis Between Three Reliability Models of a Two-Unit Complex Industrial System. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 30(2), 243-254.
- [34] Sachdeva, K., Taneja, G., and Manocha, A. (2022). Sensitivity and Economic Analysis of an Insured System with Extended Conditional Warranty. *Reliability: Theory and Applications*, 17(3 (69)), 315-327.
- [35] Sachdeva, K., Taneja, G., and Manocha, A. (2023). Reliability and Sensitivity Analysis of a System with Conditional and Extended Warranty. *Reliability: Theory and Applications*, 18(3 (74)), 689-707.

# STUDIES ON A NEW MANPOWER MODEL WITH NON- HOMOGENEOUS POISSON RECRUITMENT, PROMOTION AND LEAVING PROCESSES

K. Suryanarayana Rao<sup>1</sup>, K. Srinivasa Rao<sup>2</sup>

•

<sup>1</sup>Department of Basic Science & Humanities, Vignan's Institute of Engineering for women,  
Visakhapatnam, Andhra Pradesh, India. Email: suryanarayanarao1@gmail.com

<sup>2</sup>Department of Statistics, Andhra University, Visakhapatnam, Andhra Pradesh, India.  
Email: ksraoau@yahoo.co.in

## Abstract

*For proper utilization of manpower in any organization manpower modeling is needed. This paper addresses the two graded manpower model with non-stationary recruitment, promotion and leaving processes. Here it is assumed that the recruitment process in the first grade follows a NHP process which is further assumed that the promotion and leaving processes are also NHP processes. Using the difference-differential equations, the joint p.g.f of the number of employees in the organization at any time 't' is derived. The characteristics of the model such as the average number of employees in each grade, the average waiting time of an employee in each grade, the variance of the number of employees in each grade and the C.V of an employee in each grade are derived explicitly. The sensitivity analysis of the model with respect to the changes in parameter is also studied through numerical illustration. The comparative study between homogeneous Poisson recruitment and NHP recruitment is also discussed. This model also improves some of the earlier models as particular cases.*

**Keywords:** NHP process, two-graded manpower model, duration of stays any grade, performance of the model.

## 1. Introduction

An optimal utilization of Human Resources planning of manpower structure is a prerequisite for any organization. Hence, several works have been reported in literature regarding manpower models with various assumptions on the constituent processes. Graded manpower systems and its analysis are more important in order to develop policies of the organization with respect to manpower. Starting with the pioneering work by Seal [1] with manpower modeling of human resources much work has been reported in literature regarding graded manpower systems (Srinivasa Rao et al. [2]). The different approaches in manpower modeling are explained by Ugwuowo [3] and Wang [4]. Parthasarathy et al. [5] have analyzed the two grade system and tried to use to represent the threshold as a specific case of the exponentiated exponential distribution (EE distribution). Jeeva and Geetha [6], Gulzarul Hasan [7, 8] studied the manpower models governed by a fuzzy environment. Kannan Nilakantan [9] analyzed the manpower models with staffing policies. Maijamma [10] is approach has the benefit of being the first to use linear programming and determined the ideal number of hires and promotions to make in order to

reduce the overall cost of the manpower planning system, particularly the cost of hiring and promoting people. This study specifically examined how applying the linear programming model can result in lower recruitment and promotion costs. In terms of dependability and attainability, the actionable model has been found to be effective and reliable. Sathiyamoorthi and Elangovan [11], Lalithadevi and Srinivasan [12] have utilized geometrical process and shock models for analyzing single graded manpower models. Parameswari [13] studied the estimation of the variability of the time to recruitment for a two-graded personnel system. Ravichandran [14], Sendhamilzselvi et al. [15, 16] studied on calculating the mean and variance of the time to recruitment in a two graded manpower system with two continuous thresholds for depletion.

The Poisson process is extensively utilized in manpower models for analyzing the manpower system with respect to various organization by Srinivasa Rao et al. [17], Kondababu and Srinivasa Rao [18], Srinivasa Rao and Kondababu [19], Govinda Rao et al. [20, 21]. Srinivasa Rao and Mallikharjuna Rao [22] have studied two graded manpower models with NHP recruitments. NHP processes can be used to incorporate time-varying complexity. In order to reflect potential recruitment patterns over time, one can use this method. The time spent on trial recruitment modelling has many advantages. Saral et al. [23] has studied manpower models with two graded systems with respect to recruitment policy and thresholds. Jayanthi [24] studied and analyzed the single graded system by considering time to recruitment with breakdown thresholds. Thilaka et al. [25] studied a method by deriving the characteristics of a two-grade human resource system under the conditions that (a) personnel can move from one grade to the next for training and skill improvement, and (b) people who previously left the system can be hired in both grades. The steady state and transient behaviors are discussed. Srinivasa Rao and Ganapathi Swamy [26, 27] studied the manpower models with Duane recruitment processes. They considered that the leaving or promotion processes are stationary and independent of time. But in many practical situations it is observed that the employee leaving and promotion is dependent on time for example in corporate and public sector offices having the graded system employee promotions or leaving is done based on the time and duration of their stay in the organization. Hence, in analyzing the manpower models ignoring the non-stationary influence off promotion or leaving process may lead to falsification in the model and may not estimate the characteristic of the model accurately if the system is governed by non-stationary.

To have an accurate analysis one has to consider the non homogeneity of the recruitment/promotion/leaving processes of the models. Very little work has been reported in literature regarding manpower models with non-homogeneous recruitment/promotion/leaving processes in graded systems. Therefore in this paper, the model with NHP recruitment, promotion and leaving processes is developed and analyzed. The rest of the paper is arranged as follows: Section 2 deals with the development of the two graded manpower model using the difference differential equations. Section 3 deals with the derivation of the characteristics of the model such as probability of extinction, probability of at least one employee in grade 1 and grade 2, average number of employees in each grade, the variance of the number of employees in the organization and the variance of the number of employees in the organization. Section 4 deals with numerical illustration and discussion on the characteristics of the model. Section 5 deals with sensitivity analysis of the model. Section 6 is to compare the proposed model with that of the manpower model with homogeneous poisson recruitment and promotion/leaving processes. Section 7 deals with conclusions.

## 2. Two graded manpower model

Consider a two graded manpower model in which the organization is having two grades namely, grade-1 and grade-2. The recruitment process of grade-1 is assumed to follows a NHP process with mean recruitment rate is  $\lambda(t) = \lambda_1 + \lambda_2 t$ . The promotion process from grade-1 to grade-2 follows a NHP process with mean promotion rate  $\alpha(t) = a_1 + a_2 t$ . The leaving processes in grade-2 follow a NHP process with mean leaving rate  $\beta(t) = b_1 + b_2 t$ .

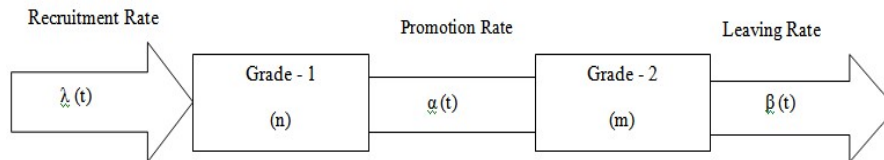


Figure 1: Manpower model

With these suppositions, the model postulates are:

- The probability that an employee will be recruited in grade-1 at random intervals of time  $h$  is  $[\lambda(t) h + o(h)]$ .
- When there are 'n' employees in grade 1, the probability of a promotion from grade-1 to grade-2 during an random interval of time 'h' is  $[n \alpha(t) h + o(h)]$ .
- When there are 'm' employees in grade 2, the probability of an employee quitting the company from grade-2 during an random interval of time 'h' is  $[m \beta(t) h + o(h)]$ .
- When there are 'n' employees in grade 1 and 'm' employees in grade-2, the probability that no employee will join or leave the company during a tiny interval of time 'h' is  $[1 - \lambda(t)h - n \mu(t)h - m \beta(t) h + o(h)]$ .
- The probability that an event other than those listed above took place within a tiny period of time 'h' is  $o(h)$ .

Let  $P_{n,m}(t)$  represent the probability that the organization will have 'n' employees in grade-1 and 'm' employees in grade-2 at time  $t$ . The difference-differential equations of the model with this structure are:

$$\frac{\partial P_{n,m}(t)}{\partial t} = -[\lambda(t) + n \alpha(t) + m \beta(t)]P_{n,m}(t) + \lambda(t)P_{n-1,m}(t) + (n+1)\alpha(t)P_{n+1,m-1}(t) + (m+1)\beta(t)P_{n,m+1}(t) \forall n, m \geq 0 \quad (1)$$

$$\frac{\partial P_{n,0}(t)}{\partial t} = -[\lambda(t) + n \alpha(t)]P_{n,0}(t) + \lambda(t)P_{n-1,0}(t) + \beta(t)P_{n,1}(t) \forall n > 0, m = 0 \quad (2)$$

$$\frac{\partial P_{0,m}(t)}{\partial t} = -[\lambda(t) + m \beta(t)]P_{0,m}(t) + \alpha(t)P_{1,m-1}(t) + (m+1)\beta(t)P_{0,m+1}(t) \forall n = 0, m > 0 \quad (3)$$

$$\frac{\partial P_{0,0}(t)}{\partial t} = -[\lambda(t)]P_{0,0}(t) + \beta(t)P_{0,1}(t) \forall n = 0, m = 0 \quad (4)$$

$P(z_1, z_2; t)$  be the joint p.g.f of  $P_{n,m}(t)$ . Then

$$P(z_1, z_2; t) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} P_{n,m}(t) z_1^n z_2^m \quad (5)$$

This implies

$$\begin{aligned} \frac{\partial P_{n,m}(t)}{\partial t} = & - \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} [\lambda(t) + n \alpha(t) + m \beta(t)] p_{n,m}(t) z_1^n z_2^m \\ & + \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \lambda(t) p_{n-1,m}(t) z_1^{n-1} z_2^m + \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} (n+1) \alpha(t) p_{n+1,m-1}(t) z_1^{n+1} z_2^{m-1} \\ & + \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} (m+1) \beta(t) p_{n,m+1}(t) z_1^n z_2^{m+1} \end{aligned} \quad (6)$$

This implies

$$\frac{\partial p(z_1, z_2; t)}{\partial t} = [\alpha(t)(z_2 - z_1)] \frac{\partial p}{\partial z_1} + [\beta(t)(1 - z_2)] \frac{\partial p}{\partial z_2} + \lambda(t)(z - 1) P(z_1, z_2; t) \quad (7)$$

Solving the equation (7) by Lagrangian's method, the auxiliary equation is

$$\frac{dt}{1} = \frac{dz_1}{-\alpha(t)(z_2 - z_1)} = \frac{dz_2}{-\beta(t)(1 - z_2)} = \frac{dP}{-\lambda(t)(1 - z_1)P(z_1, z_2, t)} \quad (8)$$

Consider the recruitment rate, promotion rate and leaving rates are linear and time dependent and is of the form.

$$\lambda(t) = \lambda_1 + \lambda_2 t$$

$$\alpha(t) = a_1 + a_2 t, \text{ Where } a_1 > 0, a_2 > 0$$

$$\beta(t) = b_1 + b_2 t, \text{ Where } b_1 > 0, b_2 > 0$$

First and third terms in equation (8), will give

$$A = (z_2 - 1) e^{-\int \beta(t) dt} \quad (9)$$

$$B = z_1 e^{-\int \alpha(t) dt} + (z_2 - 1) e^{-\int \beta(t) dt} \left( \int \alpha(t) e^{\int [\beta(t) - \alpha(t)] dt} dt \right) + \int \alpha(t) e^{-\int \alpha(t) dt} dt \quad (10)$$

First and fourth terms in equation (8), will give

$$\begin{aligned} C = & P(z_1, z_2; t) \exp\left(-\left[z_1 e^{-\int \alpha(t) dt} + (z_2 - 1) e^{-\int \beta(t) dt} \left( \int \alpha(t) e^{\int [\beta(t) - \alpha(t)] dt} dt \right)\right]\right) \\ & + \int \alpha(t) \cdot e^{-\int \alpha(t) dt} dt \left[ \int \lambda(t) \cdot e^{\int \alpha(t) dt} dt \right] \\ & + \left[ (z_2 - 1) e^{-\int \beta(t) dt} \int \lambda(t) \cdot e^{\int \alpha(t) dt} \left( \int \alpha(t) e^{\int [\beta(t) - \alpha(t)] dt} dt \right) dt \right] \\ & + \left[ \int \lambda(t) \cdot e^{\int \alpha(t) dt} \left( \int \alpha(t) e^{-\int \alpha(t) dt} dt \right) dt \right] + \int \lambda(t) dt \end{aligned} \quad (11)$$

Where A, B & C are arbitrary constants. With the initial conditions  $P_{00}(0) = 1, P_{00}(t) = 0, \forall t > 0$ . We have the joint p.g.f for the number of employees in the grade-1 and the number of employees in the grade-2 at time 't' is

$$\begin{aligned}
 P(z_1, z_2; t) = & \exp[\lambda_1 [(z_1 - 1)e^{-\left(a_1 t + a_2 \frac{t^2}{2}\right) \left(\frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv}{\lambda_1} - \frac{1}{a_1}\right)} \\
 & + (z_2 - 1)e^{-\left(b_1 t + b_2 \frac{t^2}{2}\right) \left(\frac{1}{b_1 - a_1} - \frac{\int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{a_1}\right)} \\
 & - (z_2 - 1)e^{-\left(b_1 t + b_2 \frac{t^2}{2}\right) \left(\frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{\lambda_1}\right)} \\
 & - \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv \left(\int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv\right)}{\lambda_1} - \frac{1}{b_1}]] \quad (12)
 \end{aligned}$$

### 3. Characteristics of the model

Expanding  $P(z_1, z_2, t)$ , we obtain the probability that there are no employee in the organization as.

$$\begin{aligned}
 P_{0,0}(t) = & \exp[-\lambda_1 [e^{-\left(a_1 t + a_2 \frac{t^2}{2}\right) \left(\frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv}{\lambda_1} - \frac{1}{a_1}\right)} \\
 & + e^{-\left(b_1 t + b_2 \frac{t^2}{2}\right) \left(\frac{1}{b_1 - a_1} - \frac{\int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{a_1}\right)} \\
 & + e^{-\left(b_1 t + b_2 \frac{t^2}{2}\right) \left(\frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{\lambda_1}\right)} \\
 & - \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv \left(\int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv\right)}{\lambda_1} - \frac{1}{b_1}]] \quad (13)
 \end{aligned}$$

Taking  $z_2 = 1$  in  $P(z_1, z_2; t)$ , we obtain the p.g.f of employees in the grade-1 in the organization as

$$P(z_1, t) = \exp \left[ \lambda_1 (z_1 - 1) e^{-\left(a_1 t + a_2 \frac{t^2}{2}\right) \left(\frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv}{\lambda_1} - \frac{1}{a_1}\right)} \right] \quad (14)$$

Expanding  $P(z_1, t)$  and collecting the constant terms, we obtain the probability that there is no employee in grade -1 of the organization as

$$P_0.(t) = \exp \left[ -\lambda_1 e^{-\left(a_1 t + a_2 \frac{t^2}{2}\right) \left(\frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv}{\lambda_1} - \frac{1}{a_1}\right)} \right] \quad (15)$$

In grade-1 the average number of employees in organization is

$$L_1(t) = \lambda_1 e^{-\left(a_1 t + a_2 \frac{t^2}{2}\right)} \left( \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv}{\lambda_1} - \frac{1}{a_1} \right) \quad (16)$$

The probability that there the existence of employees in grade-1 of the organization is

$$U_1(t) = 1 - \exp \left[ -\lambda_1 e^{-\left(a_1 t + a_2 \frac{t^2}{2}\right)} \left( \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv}{\lambda_1} - \frac{1}{a_1} \right) \right] \quad (17)$$

The average waiting time of an employee in grade-1 of the organization is

$$W_1(t) = \frac{L_1(t)}{\alpha(t)[1-P_0(t)]}$$

$$W_1(t) = \frac{\lambda_1 e^{-\left(a_1 t + a_2 \frac{t^2}{2}\right)} \left( \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv}{\lambda_1} - \frac{1}{a_1} \right)}{(a_1 + a_2 t) \left[ 1 - \exp \left[ -\lambda_1 e^{-\left(a_1 t + a_2 \frac{t^2}{2}\right)} \left( \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv}{\lambda_1} - \frac{1}{a_1} \right) \right] \right]} \quad (18)$$

The variance of the number of employees in grade-1 of the organization is

$$V_1(t) = \lambda_1 e^{-\left(a_1 t + a_2 \frac{t^2}{2}\right)} \left( \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv}{\lambda_1} - \frac{1}{a_1} \right) \quad (19)$$

The C.V of the number of employees in grade-1 of the organization is

$$CV_1(t) = \left[ \lambda_1 e^{-\left(a_1 t + a_2 \frac{t^2}{2}\right)} \left( \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv}{\lambda_1} - \frac{1}{a_1} \right) \right]^{-\frac{1}{2}} \quad (20)$$

Similarly, taking  $z_1 = 1$  in  $P(z_1, z_2; t)$ , we obtain the p.g.f of the number of employees in grade-2 of the organization as

$$P(z_2, t) = \exp[\lambda_1 [(z_2 - 1) e^{-\left(b_1 t + b_2 \frac{t^2}{2}\right)} \left( \frac{1}{b_1 - a_1} - \frac{\int_0^t (a_1 + a_2 v) \cdot e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{a_1} \right) + (z_2 - 1) e^{-\left(b_1 t + b_2 \frac{t^2}{2}\right)} \left( \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv}{\lambda_1} - \frac{\int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{\lambda_1} \right) - \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv}{\lambda_1} \left( \frac{\int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{\lambda_1} - \frac{1}{b_1} \right) \right] \quad (21)$$



Expanding  $P(z, t)$  and collecting the constant terms, we obtain the probability that there is no employee in grade-2 of the organization as

$$\begin{aligned}
 P_0(t) = & \exp[-\lambda_1] e^{-\left(b_1 t + b_2 \frac{t^2}{2}\right)} \left( \frac{1}{b_1 - a_1} - \frac{\int_0^t (a_1 + a_2 v) \cdot e^{(b_1 - a_1)v + (b_2 - a_2) \frac{v^2}{2}} dv}{a_1} \right) \\
 & + e^{-\left(b_1 t + b_2 \frac{t^2}{2}\right)} \left( \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2) \frac{v^2}{2}} dv}{\lambda_1} \right. \\
 & \left. - \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv \left( \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2) \frac{v^2}{2}} dv \right)}{\lambda_1} - \frac{1}{b_1} \right) \quad (22)
 \end{aligned}$$

In grade-2 the average number of employees in organization is

$$\begin{aligned}
 L_2(t) = & \lambda_1 e^{-\left(b_1 t + b_2 \frac{t^2}{2}\right)} \left( \frac{1}{b_1 - a_1} - \frac{\int_0^t (a_1 + a_2 v) \cdot e^{(b_1 - a_1)v + (b_2 - a_2) \frac{v^2}{2}} dv}{a_1} \right) \\
 & + \lambda_1 e^{-\left(b_1 t + b_2 \frac{t^2}{2}\right)} \left[ \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2) \frac{v^2}{2}} dv}{\lambda_1} \right. \\
 & \left. - \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv \left( \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2) \frac{v^2}{2}} dv \right)}{\lambda_1} - \frac{1}{b_1} \right] \quad (23)
 \end{aligned}$$

The probability that there the existence of employees in grade-2 of the organization is

$$\begin{aligned}
 U_2(t) = & 1 - \exp[-\lambda_1] e^{-\left(b_1 t + b_2 \frac{t^2}{2}\right)} \left( \frac{1}{b_1 - a_1} - \frac{\int_0^t (a_1 + a_2 v) \cdot e^{(b_1 - a_1)v + (b_2 - a_2) \frac{v^2}{2}} dv}{a_1} \right) \\
 & + e^{-\left(b_1 t + b_2 \frac{t^2}{2}\right)} \left[ \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2) \frac{v^2}{2}} dv}{\lambda_1} \right. \\
 & \left. - \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv \left( \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2) \frac{v^2}{2}} dv \right)}{\lambda_1} - \frac{1}{b_1} \right] \quad (24)
 \end{aligned}$$

The average waiting time of an employee in grade-2 of the organization is

$$W_2(t) = \frac{L_2(t)}{(b_1 + b_2 t) [U_2(t)]}$$

Where  $L_2(t)$  and  $U_2(t)$  are given in equation (23) and (24) respectively.

The variance of the number of employees in grade-2 of the organization is

$$\begin{aligned}
 V_2(t) = & \lambda_1 e^{-(b_1 t + b_2 \frac{t^2}{2})} \left( \frac{1}{b_1 - a_1} - \frac{\int_0^t (a_1 + a_2 v) \cdot e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{a_1} \right) \\
 & + \lambda_1 e^{-(b_1 t + b_2 \frac{t^2}{2})} \left[ \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(\frac{a_1 v + a_2 v^2}{2}\right)} dv \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{\lambda_1} \right. \\
 & \left. - \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(\frac{a_1 v + a_2 v^2}{2}\right)} dv \left( \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv \right) dv}{\lambda_1} - \frac{1}{b_1} \right] \quad (25)
 \end{aligned}$$

The C.V of the number of employees in grade-2 of the organization is

$$\begin{aligned}
 CV_2(t) = & \left[ \lambda_1 e^{-(b_1 t + b_2 \frac{t^2}{2})} \left( \frac{1}{b_1 - a_1} - \frac{\int_0^t (a_1 + a_2 v) \cdot e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{a_1} \right) \right. \\
 & + \lambda_1 e^{-(b_1 t + b_2 \frac{t^2}{2})} \left[ \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(\frac{a_1 v + a_2 v^2}{2}\right)} dv \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{\lambda_1} \right. \\
 & \left. \left. - \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(\frac{a_1 v + a_2 v^2}{2}\right)} dv \left( \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv \right) dv}{\lambda_1} - \frac{1}{b_1} \right] \right]^{\frac{-1}{2}} \quad (26)
 \end{aligned}$$

The average number of employees in the organization is

$$\begin{aligned}
 L(t) = & \lambda_1 e^{-(a_1 t + a_2 \frac{t^2}{2})} \left( \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(\frac{a_1 v + a_2 v^2}{2}\right)} dv}{\lambda_1} - \frac{1}{a_1} \right) \\
 & + \lambda_1 e^{-(b_1 t + b_2 \frac{t^2}{2})} \left( \frac{1}{b_1 - a_1} - \frac{\int_0^t (a_1 + a_2 v) \cdot e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{a_1} \right) \\
 & + \lambda_1 e^{-(b_1 t + b_2 \frac{t^2}{2})} \left[ \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(\frac{a_1 v + a_2 v^2}{2}\right)} dv \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{\lambda_1} \right. \\
 & \left. - \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(\frac{a_1 v + a_2 v^2}{2}\right)} dv \left( \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv \right) dv}{\lambda_1} - \frac{1}{b_1} \right] \quad (27)
 \end{aligned}$$

The variance of the number of employees in the organization is

$$\begin{aligned}
 V(t) = & \lambda_1 e^{-(a_1 t + a_2 \frac{t^2}{2})} \left( \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv}{\lambda_1} - \frac{1}{a_1} \right) \\
 & + \lambda_1 e^{-(b_1 t + b_2 \frac{t^2}{2})} \left( \frac{1}{b_1 - a_1} - \frac{\int_0^t (a_1 + a_2 v) \cdot e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{a_1} \right) \\
 & + \lambda_1 e^{-(b_1 t + b_2 \frac{t^2}{2})} \left[ \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv}{\lambda_1} \right. \\
 & \left. - \frac{\int_0^t (\lambda_1 + \lambda_2 v) e^{\left(a_1 v + a_2 \frac{v^2}{2}\right)} dv \left( \int_0^t (a_1 + a_2 v) e^{(b_1 - a_1)v + (b_2 - a_2)\frac{v^2}{2}} dv \right)}{\lambda_1} - \frac{1}{b_1} \right] \quad (28)
 \end{aligned}$$

#### 4. Numerical illustration and results

The behavior of the proposed manpower model is discussed through a numerical illustration. Since the performance characteristics of the manpower model are highly sensitive with respect to time; the transient behavior of the model is studied through computing the performance measures with the following set of values for the model parameters:

$t = 0.13, 0.14, 0.15, 0.16$ ;  $\lambda_1 = 2, 3, 4, 5, 6$ ;  $\lambda_2 = 3, 4, 5, 6, 7$ ;  $a_1 = 7, 7.4, 7.8, 8.2, 8.6$   
 $a_2 = 5, 7, 9, 11, 13$ ;  $b_1 = 9, 9.4, 9.8, 10.2, 10.6$ ;  $b_2 = 9, 12, 15, 17, 20$

For different values of parameters  $t, \lambda_1, \lambda_2, a_1, a_2, b_1, b_2$  and using the equations, the performance measures such as the average number of employees in grade-1 and in grade-2, the average waiting time of an employee in grade-1 and in grade-2, the variance of the number of employees in grade-1 and in grade-2 and the C.V of the number of employees in both grade-1 and in grade-2 are computed and presented in Table 1 and Table 2. The relationship between the parameters and performance measures are represented in the Figure 1 and Figure 2.

From the Table 1, As time ( $t$ ) varies from 0.13 to 0.16, the average number of employees in grade-1 increases from 0.07505 to 0.12475 and in grade 2 decreases from 0.13306 to 0.07818, the average waiting time of an employee in grade-1 increases from 0.13569 to 0.13637 and grade-2 decreases from 0.10502 to 0.09958, when all the other parameters are fixed.

As the recruitment rate ( $\lambda_1$ ) varies from 3 to 6, the average number of employees in grade1 and in grade-2 raises from 0.17384 to 0.32109 and 0.11727 to 0.23454 respectively, the average waiting time of an employee in grade-1 and in grade-2 raises from 0.13967 to 0.14989 and 0.10151 to 0.10746 respectively, when all the other parameters are fixed.

As the recruitment rate ( $\lambda_2$ ) varies from 4 to 7, the average number of employees in grade-1 increases from 0.32995 to 0.35654 and in grade-2 it remains constant , the average waiting time of an employee in grade-1 increases from 0.15052 to 0.15242 and in grade-2 it remains constant, when all the other parameters are fixed.

As the promotion rate parameter ( $a_1$ ) varies from 7.4 to 8.6, the average number of employees in grade-1 and in grade-2 increases from 0.37094 to 0.39615 and 0.40239 to 2.80333 respectively, the average waiting time of an employee in grade-1 decreases from 0.14596 to 0.12884 and in grade-2 increases from 0.11635 to 0.28584, when all the other parameters are fixed.

**Table 1 :** Value of  $L_1(t), L_2(t), W_1(t)$  and  $W_2(t)$  for different value of parameters

t	$\lambda_1$	$\lambda_2$	$a_1$	$a_2$	$b_1$	$b_2$	$L_1(t)$	$L_2(t)$	$W_1(t)$	$W_2(t)$
<b>0.13</b>	2	3	7	5	9	9	0.07505	0.13306	0.13569	0.10502
<b>0.14</b>	2	3	7	5	9	9	0.09266	0.11241	0.13598	0.10305
<b>0.15</b>	2	3	7	5	9	9	0.10921	0.09419	0.13621	0.10124
<b>0.16</b>	2	3	7	5	9	9	0.12475	0.07818	0.13637	0.09958
0.16	<b>3</b>	3	7	5	9	9	0.17384	0.11727	0.13967	0.10151
0.16	<b>4</b>	3	7	5	9	9	0.22292	0.15636	0.14303	0.10347
0.16	<b>5</b>	3	7	5	9	9	0.27200	0.19545	0.14643	0.10545
0.16	<b>6</b>	3	7	5	9	9	0.32109	0.23454	0.14989	0.10746
0.16	6	<b>4</b>	7	5	9	9	0.32995	0.23454	0.15052	0.10746
0.16	6	<b>5</b>	7	5	9	9	0.33881	0.23454	0.15115	0.10746
0.16	6	<b>6</b>	7	5	9	9	0.34767	0.23454	0.15178	0.10746
0.16	6	<b>7</b>	7	5	9	9	0.35654	0.23454	0.15242	0.10746
0.16	6	7	<b>7.4</b>	5	9	9	0.37094	0.40239	0.14596	0.11635
0.16	6	7	<b>7.8</b>	5	9	9	0.38196	0.67530	0.13990	0.13174
0.16	6	7	<b>8.2</b>	5	9	9	0.39020	1.21167	0.13419	0.16526
0.16	6	7	<b>8.6</b>	5	9	9	0.39615	2.80333	0.12884	0.28584
0.16	6	7	8.6	<b>7</b>	9	9	0.39277	2.80133	0.12440	0.28568
0.16	6	7	8.6	<b>9</b>	9	9	0.38942	2.79940	0.12025	0.28551
0.16	6	7	8.6	<b>11</b>	9	9	0.38611	2.79754	0.11636	0.28536
0.16	6	7	8.6	<b>13</b>	9	9	0.38282	2.79575	0.11270	0.28521
0.16	6	7	8.6	13	<b>9.4</b>	9	0.38282	1.13528	0.11270	0.15432
0.16	6	7	8.6	13	<b>9.8</b>	9	0.38282	0.59809	0.1127	0.11821
0.16	6	7	8.6	13	<b>10.2</b>	9	0.38282	0.34032	0.11270	0.10136
0.16	6	7	8.6	13	<b>10.6</b>	9	0.38282	0.19334	0.11270	0.09134
0.16	6	7	8.6	13	10.6	<b>12</b>	0.38282	0.18320	0.11270	0.08741
0.16	6	7	8.6	13	10.6	<b>15</b>	0.38282	0.17348	0.11270	0.08379
0.16	6	7	8.6	13	10.6	<b>18</b>	0.38282	0.16417	0.11270	0.08044
0.16	6	7	8.6	13	10.6	<b>21</b>	0.38282	0.15524	0.11270	0.07734

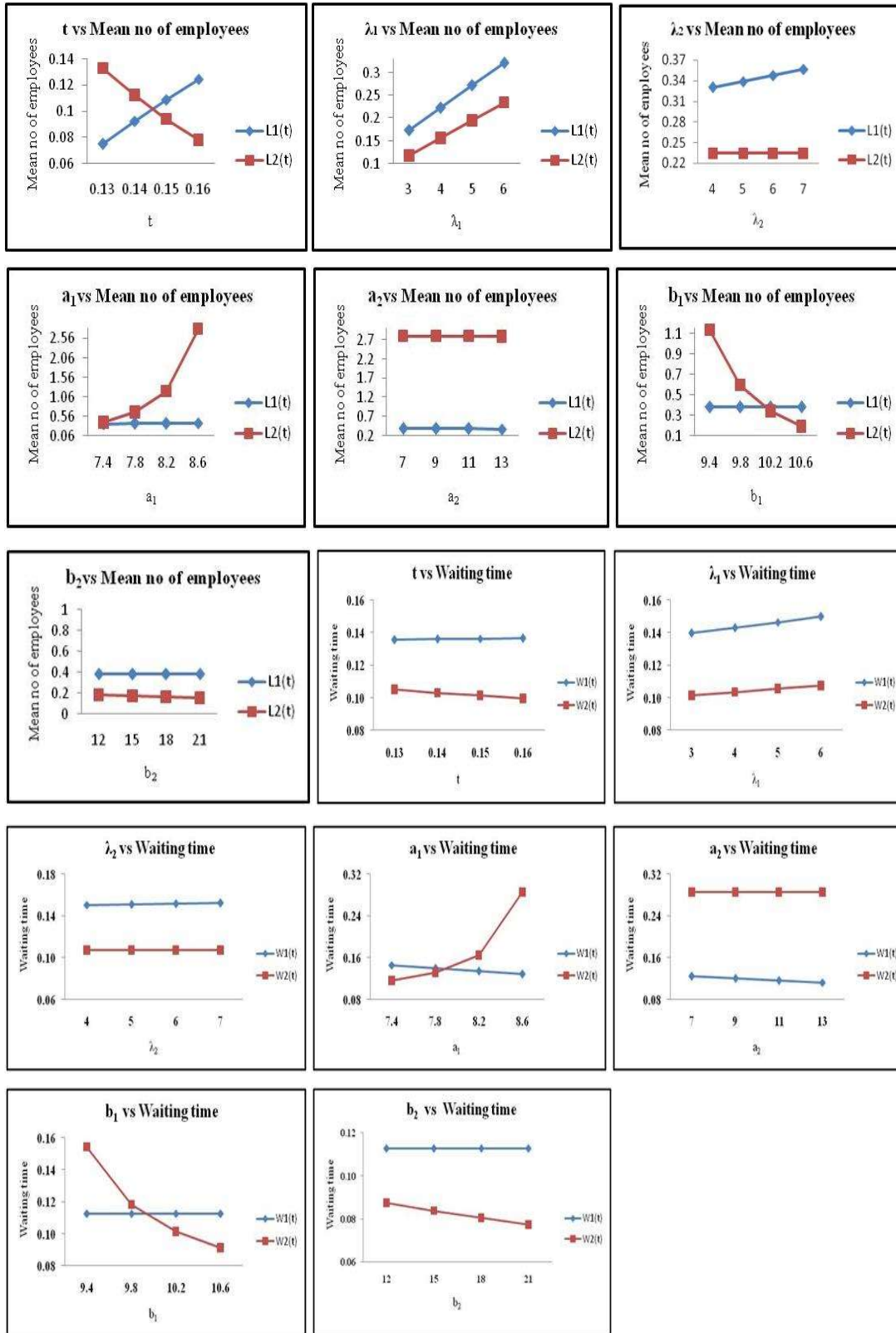


Figure 2: Relation between the parameters and performance measures

As the promotion rate parameter ( $a_2$ ) varies from 7 to 13, the average number of employees in grade-1 and in grade-2 reduces from 0.39277 to 0.38282 and 2.80133 to 2.79575 respectively, the average waiting time of an employee in grade-1 and in grade-2 reduces from 0.12440 to 0.11270 and 0.28568 to 0.28521 respectively, when all the other parameters are fixed.

As the leaving rate parameter ( $b_1$ ) varies from 9.4 to 10.6, the average number of employees in grade-1 remains constant and in grade-2 decreases from 1.13528 to 0.19334, the average waiting time of an employee in grade-1 is not affected and in grade-2 it is decreasing from 0.15432 to 0.09134, when all the other parameters are fixed.

As the promotion rate parameter ( $b_2$ ) varies from 12 to 21, the average number of employees in grade-1 remains constant and in grade-2 decreases from 0.18320 to 0.15524, the average waiting time of an employee in grade-1 is not affected and in grade-2 it is decreasing from 0.08741 to 0.07734, when all the other parameters are fixed.

From Table 2, As time (t) varies from 0.13 to 0.16, the variance of the number of employees in grade-1 increases from 0.07505 to 0.12475 and in grade-2 decreases from 0.13306 to 0.07818, C.V of the number employees in grade-1 decreases from 4.21372 to 2.83122 and in grade-2 increases from 2.74138 to 3.57642, When all the other parameters are fixed.

As the recruitment rate parameter ( $\lambda_1$ ) varies from 3 to 6, the variance of the number of employees in grade-1 and in grade-2 raises from 0.17384 to 0.32109 and 0.11727 to 0.23454 respectively, the C.V of the number of employees in grade-1 and in grade-2 reduces from 2.39844 to 1.76477 and 2.92013 to 2.06485 respectively, when all the other parameters are fixed.

As the recruitment rate parameter ( $\lambda_2$ ) varies from 4 to 7, the variance of the number of employees in grade-1 increases from 0.32995 to 0.35654 and in grade-2 remains constant, the C.V of the number employees in grade-1 decreases from 1.74091 to 1.67474 and in grade-2 remains constant, when all the other parameters are fixed.

As the promotion rate parameter ( $a_1$ ) varies from 7.4 to 8.6, the variance of the number of employees in grade-1 and in grade-2 raises from 0.37094 to 0.39615 and 0.40239 to 2.80333 respectively, the C.V of the number of employees in grade-1 and in grade-2 reduces from 1.64191 to 1.58881 and 1.57643 to 0.59726 respectively, when all the other parameters are fixed.

As the promotion rate parameter ( $a_2$ ) variation from 7 to 13, the variance of the number of employees in grade-1 and in grade-2 raises from 0.39277 to 0.38282 and 2.80133 to 2.79575 respectively, the C.V of the number of employees in grade-1 and in grade-2 raises from 1.59563 to 1.61623 and 1.59747 to 0.59807 respectively, when all the other parameters are fixed.

As the promotion rate parameter ( $b_1$ ) varies from 9.4 to 10.6, the variance of the number of employees in grade-1 remains constant and in grade-2 decreases from 1.13528 to 0.19334, the C.V of the number of employees in grade-1 remains constant and in grade-2 increases from 0.93853 to 2.27427, when all the other parameters are fixed.

As the promotion rate parameter ( $b_2$ ) varies from 12 to 21, the variance of the number of employees in grade-1 remains constant and in grade-2 decreases from 0.18320 to 0.15524, the C.V of the number of employees in grade-1 remains constant and in grade-2 increases from 2.33637 to 2.53801, when all the other parameters are fixed.

**Table 2:** Values of  $V_1(t), V_2(t), CV_1(t)$  and  $CV_2(t)$  for different values of parameters

t	$\lambda_1$	$\lambda_2$	$a_1$	$a_2$	$b_1$	$b_2$	$V_1(t)$	$V_2(t)$	$CV_1(t)$	$CV_2(t)$
<b>0.13</b>	2	3	7	5	9	9	0.07505	0.13306	3.65017	2.74138
<b>0.14</b>	2	3	7	5	9	9	0.09266	0.11241	3.28514	2.98268
<b>0.15</b>	2	3	7	5	9	9	0.10921	0.09419	3.02606	3.25833
<b>0.16</b>	2	3	7	5	9	9	0.12475	0.07818	2.83122	3.57642
0.16	<b>3</b>	3	7	5	9	9	0.17384	0.11727	2.39844	2.92013
0.16	<b>4</b>	3	7	5	9	9	0.22292	0.15636	2.11799	2.52891
0.16	<b>5</b>	3	7	5	9	9	0.27200	0.19545	1.91740	2.26193
0.16	<b>6</b>	3	7	5	9	9	0.32109	0.23454	1.76477	2.06485
0.16	6	<b>4</b>	7	5	9	9	0.32995	0.23454	1.74091	2.06485
0.16	6	<b>5</b>	7	5	9	9	0.33881	0.23454	1.71799	2.06485
0.16	6	<b>6</b>	7	5	9	9	0.34767	0.23454	1.69595	2.06485
0.16	6	<b>7</b>	7	5	9	9	0.35654	0.23454	1.67474	2.06485
0.16	6	7	<b>7.4</b>	5	9	9	0.37094	0.40239	1.64191	1.57643
0.16	6	7	<b>7.8</b>	5	9	9	0.38196	0.67530	1.61806	1.21689
0.16	6	7	<b>8.2</b>	5	9	9	0.39020	1.21167	1.60088	0.90847
0.16	6	7	<b>8.6</b>	5	9	9	0.39615	2.80333	1.58881	0.59726
0.16	6	7	8.6	<b>7</b>	9	9	0.39277	2.80133	1.59563	0.59747
0.16	6	7	8.6	<b>9</b>	9	9	0.38942	2.79940	1.60247	0.59768
0.16	6	7	8.6	<b>11</b>	9	9	0.38611	2.79754	1.60934	0.59788
0.16	6	7	8.6	<b>13</b>	9	9	0.38282	2.79575	1.61623	0.59807
0.16	6	7	8.6	13	<b>9.4</b>	9	0.38282	1.13528	1.61623	0.93853
0.16	6	7	8.6	13	<b>9.8</b>	9	0.38282	0.59809	1.61623	1.29305
0.16	6	7	8.6	13	<b>10.2</b>	9	0.38282	0.34032	1.61623	1.71419
0.16	6	7	8.6	13	<b>10.6</b>	9	0.38282	0.19334	1.61623	2.27427
0.16	6	7	8.6	13	10.6	<b>12</b>	0.38282	0.18320	1.61623	2.33637
0.16	6	7	8.6	13	10.6	<b>15</b>	0.38282	0.17348	1.61623	2.40092
0.16	6	7	8.6	13	10.6	<b>18</b>	0.38282	0.16417	1.61623	2.46808
0.16	6	7	8.6	13	10.6	<b>21</b>	0.38282	0.15524	1.61623	2.53801

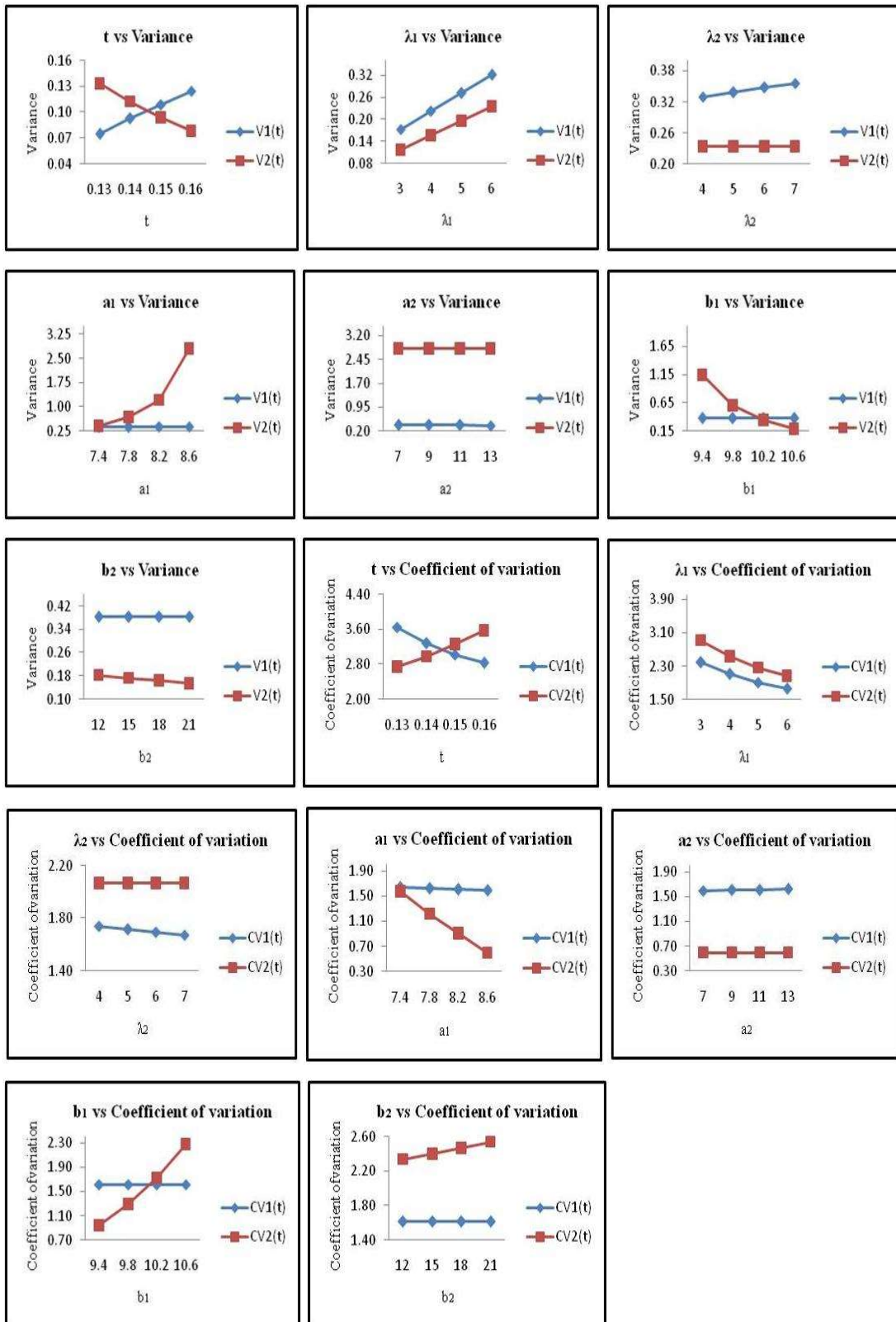


Figure 3: Relation between the parameters and performance measures



### 5. Sensitivity analysis of the model

**Table3:** The values of  $L_1(t), L_2(t), W_1(t), W_2(t), V_1(t)$  and  $V_2(t)$  for different values of  $t, \lambda_1, \lambda_2, a_1, a_2, b_1$  and  $b_2$

Parameter	Performance Measures	-15%	-10%	-5%	0%	5%	10%	15%
t=0.2	$L_1(t)$	0.20768	0.22948	0.25000	0.26931	0.28747	0.30456	0.32064
	$L_2(t)$	0.33263	0.29217	0.25585	0.22330	0.19420	0.16825	0.14516
	$W_1(t)$	0.14345	0.14385	0.14416	0.14439	0.14455	0.14464	0.14466
	$W_2(t)$	0.12032	0.11672	0.11345	0.11048	0.10776	0.10528	0.10300
	$V_1(t)$	0.20768	0.22948	0.25000	0.26931	0.28747	0.30456	0.32064
	$V_2(t)$	0.33263	0.29217	0.25585	0.22330	0.19420	0.16825	0.14516
$\lambda_1=3$	$L_1(t)$	0.23848	0.24876	0.25903	0.26931	0.27958	0.28986	0.30013
	$L_2(t)$	0.18980	0.20097	0.21213	0.22330	0.23446	0.24563	0.25679
	$W_1(t)$	0.14228	0.14298	0.14368	0.14439	0.14510	0.14581	0.14653
	$W_2(t)$	0.10870	0.10929	0.10988	0.11048	0.11107	0.11167	0.11227
	$V_1(t)$	0.23848	0.24876	0.25903	0.26931	0.27958	0.28986	0.30013
	$V_2(t)$	0.18980	0.20097	0.21213	0.22330	0.23446	0.24563	0.25679
$\lambda_2=5$	$L_1(t)$	0.25974	0.26293	0.26612	0.26931	0.27250	0.27569	0.27888
	$L_2(t)$	0.22330	0.22330	0.22330	0.22330	0.22330	0.22330	0.22330
	$W_1(t)$	0.14373	0.14395	0.14417	0.14439	0.14461	0.14483	0.14505
	$W_2(t)$	0.11048	0.11048	0.11048	0.11048	0.11048	0.11048	0.11048
	$V_1(t)$	0.25974	0.26293	0.26612	0.26931	0.27250	0.27569	0.27888
	$V_2(t)$	0.22330	0.22330	0.22330	0.22330	0.22330	0.22330	0.22330
$a_1=6.7$	$L_1(t)$	0.25225	0.25993	0.26537	0.26931	0.27180	0.27323	0.27373
	$L_2(t)$	0.01724	0.06421	0.12665	0.22330	0.38499	0.75544	2.30366
	$W_1(t)$	0.16421	0.15707	0.15060	0.14439	0.13877	0.13335	0.12844
	$W_2(t)$	0.09987	0.10222	0.10541	0.11048	0.11929	0.14107	0.25340
	$V_1(t)$	0.25225	0.25993	0.26537	0.26931	0.27180	0.27323	0.27373
	$V_2(t)$	0.01724	0.06421	0.12665	0.22330	0.38499	0.75544	2.30366
$a_2=6$	$L_1(t)$	0.27093	0.27039	0.26985	0.26931	0.26877	0.26823	0.26770
	$L_2(t)$	0.22410	0.22383	0.22356	0.22330	0.22304	0.22278	0.22252
	$W_1(t)$	0.14787	0.14669	0.14553	0.14439	0.14327	0.14216	0.14107
	$W_2(t)$	0.11052	0.11050	0.11049	0.11048	0.11046	0.11045	0.11043
	$V_1(t)$	0.27093	0.27039	0.26985	0.26931	0.26877	0.26823	0.26770
	$V_2(t)$	0.22410	0.22383	0.22356	0.22330	0.22304	0.22278	0.22252
$b_1=7.9$	$L_1(t)$	0.26931	0.26931	0.26931	0.26931	0.26931	0.26931	0.26931
	$L_2(t)$	62.67895	1.19799	0.46792	0.2233	0.11041	0.04563	0.00701
	$W_1(t)$	0.14439	0.14439	0.14439	0.14439	0.14439	0.14439	0.14439
	$W_2(t)$	7.03467	0.1843	0.12909	0.11048	0.10069	0.09394	0.08896
	$V_1(t)$	0.26931	0.26931	0.26931	0.26931	0.26931	0.26931	0.26931
	$V_2(t)$	62.67895	1.19799	0.46792	0.2233	0.11041	0.04563	0.00701
$b_2=11$	$L_1(t)$	0.26931	0.26931	0.26931	0.26931	0.26931	0.26931	0.26931
	$L_2(t)$	0.23241	0.22934	0.22630	0.22330	0.22033	0.21740	0.21451
	$W_1(t)$	0.14439	0.14439	0.14439	0.14439	0.14439	0.14439	0.14439
	$W_2(t)$	0.11471	0.11326	0.11185	0.11048	0.10913	0.10781	0.10653
	$V_1(t)$	0.26931	0.26931	0.26931	0.26931	0.26931	0.26931	0.26931
	$V_2(t)$	0.23241	0.22934	0.22630	0.22330	0.22033	0.21740	0.21451

The sensitivity of the model is performed with respect to the value of time, recruitment rate, promotion rate and leaving rate of the both grade-1 and grade-2.

For different values of  $t$ ,  $\lambda_1$ ,  $\lambda_2$ ,  $a_1$ ,  $a_2$ ,  $b_1$  and  $b_2$  the average number of employees in grade-1 and in grade-2, average waiting time of an employee in grade-1 and in grade-2, the variance of the number of employees in grade-1 and in grade-2 are computed and presented in Table-3 with variation of -15%, -10%, -5%, 0%, 5%, 10%, 15% of the model parameters.

The performance measures are highly influenced by time ( $t$ ). As  $t$  increases from -15% to +15%, the average number of employees along with the average waiting time of employees, the variance of the number of employees increases in grade-1. The average number of employees along with the average waiting time of employees, the variance of the number of employees decreases in grade-2.

As the recruitment rate parameter  $\lambda_1$  increases from -15% to +15%, the average number of employees, average waiting time of employees and the variance of the number of employees increasing in grade-1 and in grade-2.

As the recruitment rate parameter  $\lambda_2$  increases from -15% to +15%, the average number of employees along with the average waiting time of employees, the variance of the number of employees are increases in grade-1 and there is no change with respect to the performance measures in grade-2.

When the promotion rate parameter  $a_1$  increases from -15% to +15%, the average number of employees along with the variance of the number of employees increases, the average waiting time of employees decreases in grade-1 and the average number of employees along with average waiting time of employees, the variance of the number of employees increases in grade-2.

When the promotion rate parameter  $a_2$  increases from -15% to +15%, the average number of employees, average waiting time of employees and the variance of the number of employees decreasing in grade-1 and in grade-2.

When the leaving rate parameter  $b_1$  increases from -15% to +15%, the average number of employees, average waiting time of employees and the variance of the number of employees in grade-1 remain constant and in grade-2 are decreasing.

When the leaving rate parameter  $b_2$  increases from -15% to +15%, the average number of employee, average waiting time of employee and the variance of the number of employees in grade-1 are not influenced and in grade-2 are decreasing.

## 6. Comparative study of the models

The comparative study of the developed model with that of homogeneous Poisson recruitment is presented in this section. The performance measures of both the models are presented in Table 4 for different values of  $t=0.18, 0.19, 0.20, 0.21, \text{ and } 0.22$ .

From the Table 4, As time ( $t$ ) increases the percentage variation of the performance measures between two models also increasing. The model with NHP recruitment can predict the performance measure more accurately than the model with homogeneous Poisson recruitment. It is also observe that the assumption of NHP recruitment has a significant influence on all the performance measure of the model. Time also has a significant effect on the system performance measures.

**Table-4:** Comparative study of models with non-homogeneous and homogeneous recruitment

<b>t</b>	<b>Parameter Measure</b>	<b>Non-Homogeneous recruitment</b>	<b>Homogeneous recruitment</b>	<b>Difference</b>	<b>Percentage of Variation</b>
<b>t=0.18</b>	L <sub>1</sub> (t)	0.62162	0.62475	0.00313	0.50100
	L <sub>2</sub> (t)	0.09084	0.20145	0.11061	54.90692
	W <sub>1</sub> (t)	0.10049	0.12807	0.02758	21.53510
	W <sub>2</sub> (t)	0.04519	0.08833	0.04314	48.83958
	V <sub>1</sub> (t)	0.62162	0.62475	0.00313	0.50100
	V <sub>2</sub> (t)	0.09084	0.20145	0.11061	54.90692
<b>t=0.19</b>	L <sub>1</sub> (t)	0.64298	0.65171	0.00873	1.33955
	L <sub>2</sub> (t)	0.06385	0.16513	0.10128	61.33349
	W <sub>1</sub> (t)	0.10027	0.12962	0.02935	22.64311
	W <sub>2</sub> (t)	0.04188	0.08679	0.04491	51.74559
	V <sub>1</sub> (t)	0.64298	0.65171	0.00873	1.33955
	V <sub>2</sub> (t)	0.06385	0.16513	0.10128	61.33349
<b>t=0.20</b>	L <sub>1</sub> (t)	0.66132	0.67598	0.01466	2.16870
	L <sub>2</sub> (t)	0.04256	0.13433	0.09177	68.31683
	W <sub>1</sub> (t)	0.09992	0.13103	0.03111	23.74265
	W <sub>2</sub> (t)	0.03812	0.08549	0.04737	55.40999
	V <sub>1</sub> (t)	0.66132	0.67598	0.01466	2.16870
	V <sub>2</sub> (t)	0.04256	0.13433	0.09177	68.31683
<b>t=0.21</b>	L <sub>1</sub> (t)	0.67695	0.69784	0.02089	2.99352
	L <sub>2</sub> (t)	0.02599	0.10828	0.08229	75.99741
	W <sub>1</sub> (t)	0.09946	0.13230	0.03284	24.82237
	W <sub>2</sub> (t)	0.03335	0.08441	0.05106	60.49046
	V <sub>1</sub> (t)	0.67695	0.69784	0.02089	2.99352
	V <sub>2</sub> (t)	0.02599	0.10828	0.08229	75.99741
<b>t=0.22</b>	L <sub>1</sub> (t)	0.69019	0.71751	0.02732	3.80761
	L <sub>2</sub> (t)	0.01330	0.08632	0.07302	84.59222
	W <sub>1</sub> (t)	0.09891	0.13346	0.03455	25.88791
	W <sub>2</sub> (t)	0.02629	0.08350	0.05721	68.51497
	V <sub>1</sub> (t)	0.69019	0.71751	0.02732	3.80761
	V <sub>2</sub> (t)	0.01330	0.08632	0.07302	84.59222

## 7. Conclusion

In this paper, a novel model with two grades of manpower is developed and examined. This procedure has the ability to describe time-dependent recruiting. The model's characteristics, such as the average number of employees in each grade, the average waiting time for an employee in each grade, the number of employees in each grade's variance and the number of employees in each grade's C.V in the organization are explicitly derived. The sensitivity analysis of the model revealed that the system performance metrics are significantly influenced by non-homogeneous recruitment rate.

When recruiting is done in a time-dependent manner, the performance measures can be predicted more correctly and realistically by employing the developing model. This model also incorporates few of the prior models as special instances for particular values of parameters. This model can also be improved by taking cost factors into account and determining the ideal values for the model's parameters, which will be considered later. This model can be utilized to predict the human resource characteristics of the organization at defense and IT sectors as the recruitment, promotion and leaving processes in these organizations are time dependent.

## References

- [1] Seal, H. (1945). The mathematics of a population composed of  $k$  stationary strata each recruited from the stratum below the supported at the lowest level by a uniform number. *Biometrika*, 33: 226-230.
- [2] Srinivasa Rao, K., Srinivasa Rao, V. and Vivekananda Murthy, M. (2006). On two graded manpower planning model. *OPSEARCH*, 43(3):117-130.
- [3] Ugwuowo, F.I. and Mc Clean, S.I. (2000). Modeling heterogeneity in a manpower system: A review. *Applied Stochastic Models in Business and Industry*, 16(2):99-110.
- [4] Wang, J. (2005). A review of operations research applications in workforce planning and potential modeling of military training. *DSTO Systems Sciences Laboratory, Edinburgh South Australia 5111*,1-37.
- [5] Parthasarathy, S., Ravichandran, M. K., and Vinoth, R. (2010). An Application of Stochastic Models – Grading System in Manpower Planning. *International Business Research*, 3(2): 79-86.
- [6] Jeeva, M., Geetha, N. (2013). Recruitment Model in Manpower Planning Under Fuzzy Environment. *British Journal of Applied Science & Technology*, 3(4):1380–1390.
- [7] Gulzarul Hasan, Md., Suhaib Hasan, S. (2016). Manpower planning with annualized hours flexibility: a fuzzy mathematical programming approach. *Decision Making in Manufacturing and Services*, 10(1-2):5–29.
- [8] Gulzarul Hasan, Md., Qayyum, Z., Suhaib Hasan, S. (2019). Multi-objective annualized hours manpower planning model: a modified fuzzy goal programming approach. *Ind. Eng. Manag. Syst.* 18(1): 52–66.
- [9] Kannan Nilakantan. (2015). Evaluation of staffing policies in Markov manpower systems and their extension to organizations with outsource personnel. *Journal of the Operational Research Society*, 66:1324–1340.
- [10] Maijamma, B., Otinya, G. (2022). Decision Making for Recruitment and promotion policies using linear programming. *Jurnal Aplikasi Manajemen, Ekonomi dan Bisnis*, 6(2): 48-60.
- [11] Sathiyamoorthi, R. and Elangovan, R. (1998). Shock model approach to determine the expected time for recruitment. *Journal of Decision and Matematika Sciences*, 3(1-3): 67-78.

- [12] Lalitha, A. D., Srinivasan, A. (2014). A stochastic model on the time to recruitment for a single grade manpower system with attrition generated by a geometric process of inter decision times using univariate policy of recruitment. *Blue Ocean Research Journals*, 3(7): 12-15.
- [13] Parameswari, K. and Srinivasan, A. (2016). Estimation of variance of time to recruitment for a two-grade manpower system with two types of decisions when the wastages form a geometric process. *International Journal of Mathematics Trends and Technology*, 33(3):161-165.
- [14] Ravichandran, G., Srinivasan, A. (2021). Variance Of Time To Recruitment In A Single Grade Marketing Organization With Non-Instantaneous Exits And Dependent Wastage Having Two Thresholds, *International Journal of Aquatic Science*, 12(02):139-148.
- [15] SendhamizhSelvi, S. and Jenita, S. (2016). Estimation of Mean Time to Recruitment for a Two Graded Manpower System Involving Independent and Non-identically Distributed Random Variables with Thresholds Having SCBZ Property. *International Journal of Science and Research*, 6(3): 1001-1005.
- [16] SendhamizhSelvi, S. and Jenita, S. (2018). Variance of Time to Recruitment for a Two Graded Manpower System with Different Distribution Thresholds having Non-identically Distributed Wastages and Correlated Inter-Decision Times. *Journal of Global Research in Mathematical Archivists*, 5(7): 63-68.
- [17] Srinivasa Rao, K., Vivekananda Murthy, M. and Srinivasa Rao, V. (2003). Three graded manpower planning model. *Proceeding of APORS*, 410- 418.
- [18] Konda Babu, P. and Srinivasa Rao, K. (2013). Studies on two graded manpower model with bulk recruitment in both the grades. *International Journal of Human Resource Management Research*, 1(2):10-15.
- [19] Srinivasa Rao, K. and Konda Babu, P. (2014). Grade Manpower Model with Bulk Recruitment in First Grade. *International Journal of Human Resource Management Research and Development*, 4(1):1-37.
- [20] Govinda Rao, S. and Srinivasa Rao, K. (2013). Bivariate manpower model for permanent and temporary grades under equilibrium. *Indian Journal of Applied Research*, 3(8):63 – 66.
- [21] Govinda Rao, S. and Srinivasa Rao, K. (2014). Manpower model with three level recruitment in the initial grade. *International Research Journal of Management Science and Technology*, 5(9):79-102.
- [22] Srinivasa Rao, K., Mallikharjuna Rao, K. (2015). On two graded manpower model with non-homogeneous Poisson recruitments. *International Journal of Advanced Computer and Mathematical Sciences*, 6(3): 40-66.
- [23] Saral, L., SendhamizhSelvi, S. and Srinivasan, A. (2017). Estimation of mean time to recruitment for a two graded manpower system with two thresholds, different epoch for exits and correlated inter-decisions under correlated wastage. *International Educational Scientific Research Journal*, 3(3):1-6.
- [24] Jayanthi, L. S. and Uma, K. P. (2018). Determination of manpower model for a single grade system with two sources of depletion and two components for threshold. *EAI Endorsed Transactions on Energy Web and information Technologies*, 5(20):1-10.
- [25] Thilaka, B., Janaki, S., Udayabhaskaran, S. (2023) Grade sizes in two-grade models. *Advances and Applications in Mathematical Sciences*, 22: 837-859.
- [26] Srinivasa rao, K. and GanapathiSwamy, Ch. (2019). On Two Grade Manpower Model With Duane Recruitment Process. *Journal of Applied Science and Computations*, VI(I): 220-235.
- [27] GanapathiSwamy, Ch., Srinivasa Rao, K. (2022). A novel application of duane process for modeling two graded manpower system with direct recruitment in both the grades. *Reliability: Theory & Applications*, 17(2): 38-55.

# THE ROLE OF RECORD VALUES IN STATISTICAL INFERENCE: A REVIEW ARTICLE

Mahmoud A. Selim Alsanea

•

Applied College, King Khalid University, Saudi Arabia &  
Department of Statistics, Faculty of Commerce, Al-Azhar University, Egypt  
Selim.one@gmail.com

## Abstract

*The record values data have received the attention of researchers in statistics for over seven decades. Through these decades the records have played a significant and widely utilized role for statistical inference in parameter estimation, predicting future values, hypothesis tests, as well as stress-strength tests, and characterizing distributions. In this paper, the types of record values, some distributional properties, and statistical inferences of record values and their applications are reviewed. The purpose of this paper is to shed light on the role of record values in statistical inference. Therefore, we will examine this issue from two perspectives, the first perspective being estimation and the second perspective being prediction. These are through some of the most important lifetime distributions are Exponential, Weibull, Gumbel, Geometric, Pareto, Generalized exponential, Rayleigh, Lomax, and Nadarajah-Haghighi distributions. I hope that the findings of this paper will be useful for researchers in various fields and lead to further enhancement of research in record values theory and its applications.*

**Keywords:** maximum likelihood estimation; maximum likelihood predictor; record values; point prediction; probability distribution; Bayesian estimation; Bayesian prediction

## 1. Introduction

In statistics, a record value or record statistic is the largest or smallest value obtained from a sequence of random variables. Record values arise naturally in both theoretical and practical areas of probability and statistics. On the practical side, Record values are of interest and importance in several branches of studies such as, hydrology, seismology, psychology, medicine, engineering. All of us constantly hear of new records being created in events such as stock market prices, rainfall, temperature, flood level, athletic events, oil, and mining surveys etc. In any field, whenever a new high or a new low value is observed, in connection with the phenomena under study, it becomes a part of history and will be called as a record. On the theoretical side, various statistical inference procedures such as point or interval estimation and prediction as well as hypothesis testing can be developed based on observed record sequences.

Records become extremely important and necessary in some cases, including when we only want to study the value of the events that exceed the previous ones, or when observations are destroyed by experimental tests, or it is impossible to obtain a complete sample. Overall, records can be useful in any situation where there is a need to track and analyze data over time.

The aim of this study to shed light on the role of record values in statistical inference. The rest of this paper reviews the records from the following aspects: Section 2 introduces the definition of record values and their distributional properties. Section 3 introduces a literature review on uses of record values in statistical inference. Section 4 reviews the record values data in practical applications. Section 5 reviews the computer software for records. Section 6 discusses the future works. Section 7 introduces some conclusions.

## 2. Definition of Record Values and their Distributional Properties

The formal study of record value theory probably started with the pioneering paper by Chandler [1]. In this section, the definition of upper and lower and k-th record values and their distributional properties are introduced.

### 2.1 Upper Record Values

Let  $X_1, X_2, \dots, X_n, \dots$  be a sequence of independent and identically distributed (iid) random variables that have cumulative distribution function (cdf)  $F(x)$  and probability density function (pdf)  $f(x)$ . Let  $Y_n = \max\{X_1, X_2, \dots, X_n\}$  for  $n \geq 1$ . We say  $X_j$  is an upper record value of this sequence if  $Y_j > Y_{j-1}, j > 1$ . Thus  $X_j$  will be called an upper record value if its value exceeds that of all previous observations. The first record  $Y_1 = X_1$  is called the trivial record.

The times at which records appear are of interest and are called record times. The random variables  $U(0) = 1$ , and  $U(m) = \min\{j: j > U(m-1), X_j > X_{U(m-1)}\}$  are called the upper record times, and the sequence  $\{U(m), m \geq 0\}$  is called the sequence of upper record times.

The sequence of inter-record times, denoted by  $\{T_n, n \geq 1\}$ , is defined as  $T_n = U(n+1) - U(n), n = 1, 2, \dots$

Many distributional properties of upper record values in the sequence of iid continuous random variables  $X_1, X_2, \dots, X_{U(m)}$  with cdf  $F(x)$  and pdf  $f(x)$  have been expressed in terms of the function  $R(x) = -\ln[1 - F(x)]$ . The pdf of the upper record value  $X_{U(m)}$  (see Arnold, et al. [2]) is

$$f_m(x) = \frac{(R(x))^{m-1}}{(m-1)!} f(x), \quad -\infty < x < \infty \quad (2.1)$$

and the joint pdf of the first (m) upper record values  $X_{U(1)} = x_1, X_{U(2)} = x_2, \dots, X_{U(m)} = x_m$  is given by

$$f_{1,2,\dots,m}(x_1, x_2, \dots, x_m) = f(x_m) \prod_{i=1}^{m-1} \frac{f(x_i)}{1 - F(x_i)}, \quad (2.2)$$

and the joint pdf of the upper record values  $X_{U(n)}$  and  $X_{U(m)}$  ( $n < m$ ) is

$$f_{n,m}(x, y) = \frac{[R(x)]^{n-1}}{(n-1)!(m-n-1)!} \cdot \frac{f(x)}{1 - F(x)} \cdot [R(y) - R(x)]^{m-n-1} f(y), \quad (2.3)$$

$$-\infty < x < y < \infty, \quad n = 0, 1, \dots, n < m,$$

the conditional probability density function of the upper record values  $X_{U(j)}$  given  $X_{U(i)}$  can be expressed as follows

$$f(x_j | x_i) = \frac{(R(x_j) - R(x_i))^{j-i-1}}{(j-i-1)!} \frac{f(x_j; \theta)}{1 - F(x_i; \theta)}, \quad (2.4)$$

$$-\infty < x_i < x_j < \infty$$

where  $R(\cdot) = -\ln(1 - F(\cdot))$

## 2.2 Lower Record Values

Let  $X_1, X_2, \dots, X_n, \dots$  be a sequence of iid random variables from a continuous distribution with cdf  $F(x)$  and pdf  $f(x)$ . Let  $Y_n = \min\{X_1, X_2, \dots, X_n\}$  for  $n \geq 1$ . We say  $X_j$  is a lower record value of this sequence if  $Y_j < Y_{j-1}, j > 1$ . Thus  $X_j$  will be called a lower record value if its value is lower than of all previous observations. By definition  $X_1$  is a lower record value. The times at which record appear are of interest which called a record times.

The random variables  $L(0) = 1$ , and  $L(m) = \min\{j: j > L(m-1), X_j < X_{L(m-1)}\}$  are called the lower record times, and the sequence  $\{L(m), m \geq 0\}$  is called the sequence of lower record times.

Many distributional properties of lower record values in the sequence of iid continuous random variables  $X_1, X_2, \dots, X_{L(m)}$  with cdf  $F(x)$  and pdf  $f(x)$  have been expressed in terms of the function  $G(x) = -\ln F(x)$ . The pdf of the lower record value  $X_{L(m)}$ , is

$$f_m(x) = \frac{(G(x))^{m-1}}{(m-1)!} f(x), \quad -\infty < x < \infty \quad (2.5)$$

and the joint pdf of the first (m) lower record values  $X_{L(1)} = x_1, X_{L(2)} = x_2, \dots, X_{L(m)} = x_m$  is given by

$$f_{1,2,\dots,m}(x_1, x_2, \dots, x_m) = f(x_m) \prod_{i=1}^{m-1} \frac{f(x_{(i)})}{F(x_{(i)})}, \quad (2.6)$$

$$-\infty < x_m < x_{m-1} < \dots < x_1 < \infty$$

and the joint pdf of the lower record values  $X_{L(s)}$  and  $X_{L(r)}$  ( $r < s$ ) is

$$f_{r,s}(x, y) = \frac{[G(x)]^{r-1}}{(r-1)!(s-r-1)!} \cdot \frac{f(x)}{F(x)} [G(y) - G(x)]^{s-r-1} f(y) \quad (2.7)$$

$$-\infty < y < x < \infty$$

where  $x = X_{L(r)}$  and  $y = X_{L(s)}$ .

and the conditional pdf of the lower record values  $X_{L(j)}$  given  $X_{L(i)}$  can be expressed as follow

$$f(x_j|x_i) = \frac{(G(x_j) - G(x_i))^{j-i-1}}{(j-i-1)!} \frac{f(x_j; \theta)}{F(x_i; \theta)}, \quad -\infty < x_j < x_i < \infty \quad (2.8)$$

where  $G(.) = -\ln F(.)$  for more details, see for example, Ahsanullah and Nevzorov [3].

## 2.3 The K-th Upper Record Values

Let  $\{X_n, n \geq 1\}$  be a sequence of iid random variables with a cdf  $F(x)$  and pdf  $f(x)$ . The j-th order statistic of the sample  $X_1, X_2, \dots, X_n$  is denoted by  $X_{j:n}$ . For a fixed positive integer k, Dziubdziela and Kopociński [4] defined the sequence  $\{U_n^{(k)}, n \geq 1\}$  of k-th upper record times for the sequence  $\{X_n, n \geq 1\}$  as follows:

$$U_1^{(k)} = 1$$

$$U_{n+1}^{(k)} = \min\{j > U_n^{(k)}: X_{j:j+k-1} > X_{U_n^{(k)}:U_n^{(k)}+k-1}\},$$

Then the sequence  $\{Y_n^{(k)}, n \geq 1\}$ , where  $Y_n^{(k)} = X_{U_n^{(k)}:U_n^{(k)}+k-1}$  is called a sequence of k-th upper record values of  $\{X_n, n \geq 1\}$ . For convenience, we also take  $Y_0^{(k)} = 0$ . Note that for  $k = 1$ , we get the usual upper record values as defined in Chandler [1].

The pdf of  $Y_n^{(k)}$  ( $n \geq 1$ ) as given by Grudzien [5] is

$$f_{Y_n^{(k)}}(x) = \frac{k^n}{(n-1)!} [-\ln \bar{F}(x)]^{n-1} [\bar{F}(x)]^{k-1} f(x), \quad -\infty < x < \infty \quad (2.9)$$

and the joint pdf of  $Y_m^{(k)}$  and  $Y_n^{(k)}, 1 \leq m < n, n \geq 2$ , is



$$f_{Y_m^{(k)}, Y_n^{(k)}}(x, y) = \frac{k^n}{(m-1)!(n-m-1)!} [-\ln \bar{F}(y) + \ln \bar{F}(x)]^{n-m-1} \times [-\ln \bar{F}(x)]^{m-1} \frac{f(x)}{\bar{F}(x)} [\bar{F}(y)]^{k-1} f(y), \quad x < y \quad (2.10)$$

## 2.4 The K-th Lower Record Values

Let  $\{X_n, n \geq 1\}$  be a sequence of iid random variables with a cdf  $F(x)$  and pdf  $f(x)$ . The  $j$ -th order statistic of a sample  $X_1, X_2, \dots, X_n$  is denoted by  $X_{j:n}$ . For a fixed positive integer  $k$ , we defined the sequence  $\{L_k(n), n \geq 1\}$  as  $k$ -th lower record times of  $\{X_n, n \geq 1\}$  as follows:

$$L_k(n) = 1 \\
L_k(n+1) = \min\{j > L_k(n) : X_{k:L_k(n)+k-1} > X_{k:j+k-1}\},$$

The sequence  $\{Y_n^{(k)}, n \geq 1\}$ , where  $Y_n^{(k)} = X_{k:L_k(n)+k-1}$  is called a sequence of  $k$ -th lower record values of  $\{X_n, n \geq 1\}$ . For convenience, we also take  $Y_0^{(k)} = 0$ . Note that for  $k = 1$ , we get the usual lower record values as defined in Chandler [1].

The pdf of  $Y(k)n$  ( $n \geq 1$ ) as given by Grudzien [5] is

$$f_{Y_n^{(k)}}(x) = \frac{k^n}{(n-1)!} [-\ln F(x)]^{n-1} [F(x)]^{k-1} f(x), \quad n \geq 1 \quad (2.11)$$

and the joint pdf of  $Y_m^{(k)}$  and  $Y_n^{(k)}, 1 \leq m < n, n \geq 2$ , is

$$f_{Y_m^{(k)}, Y_n^{(k)}}(x, y) = \frac{k^n}{(m-1)!(n-m-1)!} [-\ln F(y) + \ln F(x)]^{n-m-1} \times [-\ln F(x)]^{m-1} \frac{f(x)}{F(x)} [F(y)]^{k-1} f(y), \quad y < x \quad (2.12)$$

## 3. Literature Review of Statistical Inference of Record Values

This section presents a review of the statistical literature to emphasize the role of record values in statistical inference. Therefore, we will examine this issue from two perspectives, the first perspective being estimation and the second perspective being prediction. These are through some of the certain distributions are Exponential, Weibull, Gumbel, Geometric, Pareto, Generalized exponential, Rayleigh, Lomax, and Nadarajah-Haghighi distributions.

### 3.1 Review of Previous Studies on Estimation based on Record Values

In this subsection, previous studies on estimation problems based on record values for some of the certain distributions are reviewed.

#### 3.1.1 Exponential distribution

Jaheen [6] obtained the ML and empirical Bayes estimate for the parameter of the exponential model based on record statistics. The estimate is obtained using the squared error loss and Varian's linear-exponential (LINEX) loss functions. Ahmadi and Doostparast [7] presented Bayes estimation when the data consist of  $k$  record values from a two-parameter exponential distribution under linear exponential loss function. Balakrishnan and Stepanov [8] discussed the Fisher information contained in records. In the case when the initial distribution belongs to the exponential family. Doostparast [9] derived the Bayesian and non-Bayesian estimates for the two parameters of the exponential distribution based on lower record values, with respect to the squared error (SE) and LINEX loss functions, and then compared with together. Arnold, et al. [2] obtained the ML estimates and best linear unbiased estimator (BLUE) for the exponential distribution. Wu [10] presented the interval estimation for the scale parameter of two-parameter exponential

distribution using upper record values. In addition, two methods for the joint confidence region of two parameters are proposed. Asgharzadeh, et al. [11] proposed two families of optimal confidence regions for the location and scale parameters of the two-parameter exponential distribution based on upper records. Constrained optimization problems are used to find the smallest-area confidence regions for the exponential parameters with a specified confidence level. Baklizi [12] considered the stress-strength reliability when the available data is in the form of record values from the one parameter and two parameters exponential distribution. The ML estimators and the associated confidence intervals are derived. Ahsanullah and Aliev [13] considered several distributional properties of the upper records from the exponential distribution and presented some characterizations of the exponential distribution.

### 3.1.2 Weibull distribution

Abd-El-Hakim and Sultan [14] obtained the maximum likelihood estimators for the location and scale parameters of Weibull distribution based on upper record values. Soliman, et al. [15] discussed a Bayesian analysis in the context of record statistics values from the two-parameter Weibull distribution. The ML and the Bayes estimates based on record values are derived for the two unknown parameters and some survival time parameters e.g., reliability and hazard functions. The Bayes estimates are obtained based on a conjugate prior for the scale parameter and a discrete prior for the shape parameter of this model. This is done with respect to both symmetric loss function (squared error loss), and asymmetric loss function (linear-exponential (LINEX)) loss function. Jafari and Zakerzadeh [16] proposed a simple and exact test and a confidence interval for the shape parameter. In addition to a generalized confidence interval, a generalized test variable is derived for the scale parameter when the shape parameter is unknown. The paper presents a simple and exact joint confidence region as well. Wang and Ye [17] investigated point estimation and confidence intervals estimation for the Weibull distribution based on record data. The uniformly minimum variance unbiased estimator for the Weibull shape is derived. Based on this estimator, a bias-corrected estimator for the Weibull scale is obtained and it is shown to have much smaller bias and mean squared error compared with the maximum likelihood estimator. Confidence intervals for parameters and reliability characteristics of interest are constructed using pivotal or generalized pivotal quantities. Then the results are extended to the stress-strength model involving two Weibull populations with different parameter values. Construction of confidence intervals for the stress-strength reliability is discussed. Raqab, et al. [18] considered the problem of the estimation for the 3-parameter Weibull distribution based on record data. The maximum likelihood method is used for the estimation of all parameters involved in the model. Hassan, et al. [19] investigated the estimation of multicomponent stress-strength reliability following Weibull distribution based on upper record values. Al-Duais [20] developed a LINEX loss function to estimate the parameters and reliability function of the Weibull distribution based on upper record values when both shape and scale parameters are unknown. They performed this by merging a weight into LINEX to produce a new loss function called the weighted linear exponential (WLINEX) loss function. Then, they utilized WLINEX to derive the parameters and reliability function of the Weibull distribution. The results revealed that the proposed method is the best for estimating parameters and has good performance for estimating reliability.

### 3.1.3 Gumbel distribution

Ahsanullah [21] obtained ML, best linear invariant and minimum variance unbiased (MVU) estimators of the Gumbel location and scale parameters. Mousa, et al. [22] obtained the Bayesian estimators for the two parameters of the Gumbel distribution based on lower record values. Malinowska and Szynal [23] obtained Bayesian estimation for the two parameters of a Gumbel distribution based on k-th lower record values. Seo and Kim [24] addressed inference problems for

Gumbel distribution when the available data are lower record values. they first derive unbiased estimators of unknown parameters, and then, they construct an exact confidence interval for the scale parameter by deriving certain properties and pivotal quantities. For Bayesian inference, they derive noninformative priors such as the Jeffreys and reference priors for unknown parameters and examine whether they satisfy the probability matching criteria; then, they apply them to develop objective Bayesian analysis. Asgharzadeh, et al. [25] presented exact confidence intervals and joint confidence regions for the parameters of Gumbel distribution based on record data. Exact confidence intervals and joint confidence regions for the parameters of inverse Weibull distribution are also discussed. Three numerical examples with climate data are presented to illustrate the proposed methods.

### 3.1.4 Geometric distribution

Ahsanullah and Holland [26] discussed some distributional properties of the record values of non-identically distributed random variables having geometric distributions. Three theorems dealing with the characterization of the geometric distribution based on these distributional properties are presented. The unique minimum variance unbiased estimators of some functions of the parameters of the distribution are studied. Ahmadi and Doostparast [7] obtained Bayesian and non-Bayesian estimators of the parameter of geometric distribution based on upper record values. Okasha and Wang [27] E-Bayesian and Bayesian methods have been used for estimating the parameter, reliability, and hazard functions of the geometric distribution based on upper record value samples. Francis, et al. [28] obtained the shrinkage estimate of  $R = P(X \leq Y)$  when  $X$  the stress and  $Y$  the strength are independent geometric variable and the sample on  $Y$  the strength is upper records.

### 3.1.5 Pareto distribution

Arnold and Press [29] discussed the Bayesian estimation for Pareto data based on record values. El-Qasem [30] used the upper record values to obtain the ML estimator for the uniform, the exponential and the Pareto distribution with one parameter. Sultan and Moshref [31] obtained the best linear unbiased estimates for the location and scale parameters of record values from the generalized Pareto distribution. Raqab, et al. [32], Raqab [33] obtained the ML and Bayes estimators from the two-parameter Pareto distribution for the two unknown parameters based on record values. Doostparast, et al. [34] on the basis of record values from the two-parameter Pareto distribution, ML and Bayes estimators as well as credible regions are developed for the two parameters of the Pareto distribution. Ahsanullah and Shakil [35] established some new results on the characterizations of the Pareto distribution by upper record values. Azhad, et al. [36] discussed inferences about the multicomponent stress strength reliability are drawn under the assumption that strength and stress follow independent Pareto distribution under the setup of upper record values. The ML estimator, Bayes estimator under squared error and LINEX loss functions, of multicomponent stress-strength reliability are constructed.

### 3.1.6 Generalized exponential distribution

Jaheen [6] derived Bayes and empirical Bayes estimators for the one-parameter of the generalized exponential distribution based on lower record values. These estimates are obtained based on squared error and LINEX loss functions. Madi and Raqab [37] used the importance sampling to estimate the model parameters. Baklizi [38] considered the ML and Bayesian estimation of the stress-strength reliability based on lower record values from the generalized exponential distribution. Confidence intervals, exact and approximate, as well as the Bayesian credible sets for the stress-strength reliability are obtained. Dey, et al. [39] derived the ML estimates and the Bayes estimates based on lower records for the unknown parameters of the

generalized exponential distribution. The Bayesian estimation of the parameters of the generalized exponential distribution has been studied with respect to both symmetric and asymmetric loss functions. They have also derived the Bayes interval. Sana and Faizan [40] obtained ML estimators for the two unknown parameters of the generalized exponential distribution based on lower record values. They also obtained the Bayes estimators of the unknown parameters using Lindley's approximation under symmetric and asymmetric loss functions.

### 3.1.7 Rayleigh distribution

Balakrishnan and Chan [41] derived explicit expressions for the means, variances and covariances from a Rayleigh distribution. They also established some recurrence relationships for the single and product moments. These results are then used to derive explicitly the best linear unbiased estimators for the scale-parameter as well as the location-scale parameter cases. Hendi, et al. [42] obtained the Bayes estimators for the parameter, reliability function, and failure rate function based on upper record values of Rayleigh distribution. These estimators are obtained on the basis of square error and LINEX loss functions. Soliman and Al-Aboud [43] obtained the estimators of the parameter of Rayleigh distribution based on upper record values, Bayesian and non-Bayesian approaches have been used to obtain the estimators of the parameter, and some lifetime parameters such as the reliability and hazard functions. Ahsanullah and Shakil [44] established some results on characterizations of Rayleigh distribution based on order statistics and record values. Seo, et al. [45] provided the exact confidence intervals for unknown by providing some pivotal quantities in the two-parameter Rayleigh distribution based on the upper record values. Finally, the validity of the proposed inference methods was examined from Monte Carlo simulations and real data. Seo and Kim [46] provided an objective Bayesian analysis method based on the objective priors (the Jeffreys and reference priors, and the second-order PMP) for unknown parameters of the two-parameter Rayleigh distribution when the upper record values are observed. Abdi and Asgharzadeh [47] presented exact joint confidence regions for the parameters of the Rayleigh distribution based on record data. By providing some appropriate pivotal quantities, they construct several joint confidence regions for the Rayleigh parameters. These joint confidence regions are useful for constructing confidence regions for functions of the unknown parameters.

### 3.1.8 Lomax distribution

Lee and Lim [48] characterized the Lomax distribution by conditional expectations of record values. Nasiri and Hosseini [49] obtained ML estimation based on records and a proper prior distribution to attain a Bayes estimation (both informative and non-informative) based on records for quadratic loss and squared error loss functions. The study considers the shortest confidence interval and highest posterior distribution confidence interval based on records. Mahmoud, et al. [50] considered the Bayes estimators of the unknown parameters of the Lomax distribution under the assumptions of gamma priors on both the shape and scale parameters. The Bayes estimators cannot be obtained in explicit forms. So, they propose Markov Chain Monte Carlo (MCMC) techniques to generate samples from the posterior distributions and in turn computing the Bayes estimators. Point estimation and confidence intervals based on ML and bootstrap methods are also used. Mahmoud, et al. [51] addressed the problem of estimating  $R = P[Y < X]$  for the Lomax distributions, and classical and MCMC Bayesian analysis for R were developed when both samples on X and Y are in the form of upper record values, observed from the Lomax distribution. Hassan and Zaky [52] considered estimation of entropy for Lomax distribution based on upper record values. Bayesian estimator of Shannon entropy is discussed under informative and non-informative priors. The entropy Bayesian estimator and the corresponding credible interval based on a LINEX, squared error loss functions are derived.

### 3.1.9 Nadarajah-Haghighi distribution

Selim [53] discussed maximum likelihood and Bayes estimation of the two unknown parameters of Nadarajah and Haghighi distribution based on record values. It assumed that in Bayes case, the unknown parameters of Nadarajah and Haghighi distribution have gamma prior densities. Lindley approximation is exploited to obtain point estimators for the unknown parameters. Sana and Faizan [54] discussed maximum likelihood and Bayes estimation of the two unknown parameters of Nadarajah and Haghighi distribution based on record values. Different Bayes estimates are derived under squared error, balanced squared error and general entropy loss functions by using Jeffreys' prior information and extension of Jeffreys' prior information. Tierney and Kadane approximation method used to compute these estimates. MirMostafae, et al. [55] obtained exact explicit expressions as well as several recurrence relations for the single and product moments of record values and then these results are used to compute the means, variances and the covariances of the upper record values. Also, these calculated moments are used to find the best linear unbiased estimators of the location and scale parameters of NH distribution. Confidence intervals for the unknown parameters are also discussed.

### 3.1.10 General classes of distributions

Abu-Youssef [56] characterized general classes of continuous distribution by considering the conditional expectation of function of record values. The specific distribution considered as a particular case of the general class of distribution are Weibull, Pareto, power function, Burr, beta of the first kind, Cauchy, rectangular, Rayleigh, Lomax, and inverse Weibull distributions. Ahmadi and Doostparast [7] obtained Bayesian estimation for the two parameters of some life distributions, including Exponential, Weibull, Pareto and Burr type XII, based on upper record values. Ahmadi, et al. [57] showed how to develop Bayes estimation in the context of upper k-record data from a semi-parametric class of distributions that includes several well-known lifetime distributions such as exponential, Weibull (one parameter), Pareto and Burr type XII under some balanced type of loss functions. Malinowska and Szydal [58] characterized general classes of continuous distributions by the conditional expectation of the kth lower record values. Specific distributions inverse exponential, inverse Weibull, inverse Pareto, negative exponential, negative Weibull, negative Pareto, negative power, Gumbel, exponentiated-Weibull, loglogistic, Burr X, inverse Burr XII and inverse paralogistic distributions.

## 3.2 Reviewing Previous Studies on Prediction based on Recorded Values

In this section, to accentuate the role of recorded values in statistical prediction, we will review the literature on prediction problems based on record values for some of the certain distributions.

### 3.2.1 Exponential distribution

Ahsanullah [59] obtained best linear unbiased predictor and best linear invariance predictor of future records  $X_s$  based on  $X_i, 1 \leq i \leq m$ , for  $m < s$ , using the standard least squares theory. Dunsmore [60] studied the problem of predicting future records from the Bayesian viewpoint and derived classical results for the exponential and the gamma models. Awad and Raqab [61] considered the prediction problem of the future nth record value based on the first m ( $m < n$ ) observed record values from one parameter exponential distribution. Jaheen [6] obtained empirical Bayes prediction bounds for future record values. Ahmadi and Doostparast [7] presented Bayes prediction procedures when the data consist of k record values from a two-parameter exponential distribution under linear exponential loss function. Ahmadi and MirMostafae [62] studied the

problem of predicting future records based on observed order statistics from two-parameter exponential distribution. The prediction intervals for the future order statistics as well as for the total lifetime in a future sample of size  $m$  from two parameter exponential distribution are obtained on the basis of the first  $n$  records coming from the same distribution. Asgharzadeh, et al. [11] proposed two families of optimal confidence regions for the location and scale parameters of the two-parameter exponential distribution based on upper records.

### 3.2.2 Weibull distribution

Soliman, et al. [15] derived Bayesian predictive density function for Weibull distribution, which is necessary to obtain bounds for predictive interval of future record. Paul and Thomas [63] studied prediction of a future record for Weibull distribution using best linear unbiased predictor. Raqab, et al. [18] considered the problem of prediction for the 3-parameter Weibull distribution based on record data. The ML method is used for the joint prediction of future records along with the estimation of all parameters involved in the model. The existence and uniqueness of the MLPs of future records as well as the PMLEs of all unknown quantities were discussed in detail. Volovskiy and Kamps [64] considered point prediction of future record values from a sequence of independent and identically distributed two-parameter Weibull random variables using the maximum likelihood method. Two likelihood functions for prediction, the predictive and the observed predictive likelihood functions, are considered and the associated predictors are derived. Mean squared error and Pitman closeness criterion are used for comparing the prediction procedures.

### 3.2.3 Gumbel distribution

Ahsanullah [65] gave two types of predictors of the  $n$ -th record value based on the first  $m$  ( $m < n$ ) record values. Mousa, et al. [22] obtained the Bayesian predictions, either point or interval, for future lower record values. Malinowska and Szynal [23] obtained Bayesian prediction, either point or interval, for future  $n$ -th lower record values. Seo and Kim [24] addressed inference problems for Gumbel distribution when the available data are lower record values. They first derive unbiased estimators of unknown parameters, and then, they construct a predictive interval for the next lower value by deriving certain properties and pivotal quantities.

### 3.2.4 Geometric distribution

Ahsanullah and Holland [26] discussed some distributional properties of the record values of non-identically distributed random variables having geometric distributions. Three theorems dealing with the characterization of the geometric distribution based on these distributional properties are presented. Also various predictors of the  $n$ th record valued utilizing the first  $m$  ( $m < n$ ) record values are studied. Ahmadi and Doostparast [7] considered Bayesian and non-Bayesian prediction, either point or interval, of geometric distribution based on the past record values observed.

### 3.2.5 Pareto distribution

Arnold and Press [29] discussed the Bayesian prediction for Pareto data based on record values. Madi and Raqab [66] used the Bayesian approach to establish future predictions for the Pareto records. Raqab, et al. [32], Raqab [33] used the Bayesian approach to predicting future record values, either point or interval, from the Pareto distribution based on the past record values observed. Also, the ML prediction of the future records and other classical methods are used for obtaining prediction intervals for the future records. Paul and Thomas [67] studied prediction of future records of Pareto distribution by using best linear unbiased predictors. Shafay, et al. [68] discussed the problem of prediction of the two-parameter Pareto distribution from a future

sample. The Bayesian approach is applied to construct predictors based on observed  $k$ -record values for the cases when the future sample size is fixed and when it is random. Several Bayesian prediction intervals are derived.

### 3.2.6 Generalized exponential distribution

Jaheen [6] obtained the prediction bounds for future lower record values from the generalized exponential distribution by using Bayes and empirical Bayes techniques. Madi and Raqab [37] described and used a Bayesian parametric approach to predict the behavior of further Los Angeles rainfall records from generalized exponential distribution. Importance sampling is used to estimate the model parameters, and the Gibbs and Metropolis samplers are used to implement the prediction procedure. Dey, et al. [39] derived the Bayes interval and discussed the Bayesian prediction intervals of the future record values based on the observed record values. Vidović [69] investigated Bayesian point predictors of order statistics from a future sample based on the  $k$ -th lower record values from generalized exponential distribution. Sana and Faizan [40] derived the Bayesian prediction for the future record values from generalized exponential distribution.

### 3.2.7 Rayleigh distribution

Balakrishnan and Chan [41] developed the prediction of a future record value and the test for superiority of the current record values from a Rayleigh distribution. Soliman and Al-Aboud [43] obtained Bayesian prediction intervals of the future record values from Rayleigh distribution. Seo, et al. [45] provided the exact predictive intervals for the future upper record values by providing some pivotal quantities in the two-parameter Rayleigh distribution based on the upper record values. Finally, the validity of the proposed inference methods was examined from Monte Carlo simulations and real data. Seo and Kim [46] provided an objective Bayesian analysis method based on the objective priors (the Jeffreys and reference priors, and the second-order PMP) for unknown parameters of the two-parameter Rayleigh distribution when the upper record values are observed. Abdi and Asgharzadeh [47] presented exact joint confidence regions for the parameters of the Rayleigh distribution based on record data. By providing some appropriate pivotal quantities, they construct several joint confidence regions for the Rayleigh parameters. These joint confidence regions are useful for constructing confidence regions for functions of the unknown parameters.

### 3.2.8 Lomax distribution

Volovskiy and Kamps [70] Point prediction of future record values from a sequence of independent and identically distributed Pareto and Lomax random variables is addressed. The focus is on likelihood-based prediction techniques; in particular, the maximum likelihood as well as the maximum observed likelihood prediction principles are invoked to derive predictors. Moreover, one-sided prediction intervals are also addressed.

### 3.2.9 Nadarajah-Haghighi distribution

MirMostafaei, et al. [55] investigated based on the observed records, how to obtain best linear unbiased predictor for the future record values. prediction intervals for future records are also discussed. Selim [53] discussed the Bayesian and non-Bayesian predictions of both point and interval predictions of the future record values.

### 3.2.10 General classes of distributions

AL-Hussaini and Ahmad [71] obtained Bayesian prediction bounds for the  $n$ th future record value based on the one-sample scheme, all of the informative and future observations are assumed

to be obtained from a general class of distributions which includes the Weibull, compound Weibull, Pareto, beta, Gompertz, compound Gompertz among other distributions. Ahmadi and Doostparast [7] obtained prediction, either point or interval, for future upper record values from a Bayesian view point of some life distributions, including Exponential, Weibull, Pareto and Burr type XII. Ahmadi, et al. [72] discussed the problem of predicting future k-records based on k-record data for a large class of distributions, which includes several well-known distributions such as: exponential, Weibull (one parameter), Pareto, Burr type XII, among others.

## 4. The Record Values Data in Applications

This section reviews the applications of records in various disciplines. The purpose of this review is to show the widespread use of real records data in statistical inference.

### 4.1 Applications on weather, rainfall, and floods

Raqab and Balakrishnan [73] considered the record values of daily temperatures (in degrees Fahrenheit) recorded at the National Center of Atmospheric Research (NCAR) during the year 2005. Nadar, et al. [74] considered the data set represents the monthly water capacity data from the Shasta reservoir in California, USA and were taken for the month of February from 1991 to 2010. Chacko and Mary [75] used the data which represent the records of the total annual rainfall (in inches) at Oxford, England, for the years 1858-1903. Seo and Song [76] analyzed two real data sets: one is the average annual temperatures (in degrees centigrade) recorded at Daejeon in Korea from 1969 to 2016). The other is carbon dioxide (CO<sub>2</sub>) emissions in Trinidad and Tobago from 1971 to 2016. Volovski and Kamps [77] considered data collected by the German Federal Office of Hydrology in its role as a scientific advisor to the Federal Waterways and Shipping Administration. The data set contains hourly measurements (in cm) of water level for the time period from January 1918 to February 2019 collected at the measurement site Cuxhaven-Steubenhöft located at the river Elbe. Selim [53] considered the real data set which represent the total annual rainfall (in inches) during the month of January from 1880 to 1916 recorded at Los Angeles Civic Center. Asgharzadeh, et al. [78] analyzed the total annual rainfall (in inches) during March recorded at Los Angeles Civic Center from 1973 to 2006 (see the website of Los Angeles Almanac: [www.laalman-ac.com/weather/we08aa.htm](http://www.laalman-ac.com/weather/we08aa.htm)). Raqab, et al. [79] discussed the analysis of real life data representing the water level exceedances over the level 65m by the River Nidd at Hunsingore Weir which is located in North Yorkshire, England from 1934 to 1970. Tripathi, et al. [80], Awwad, et al. [81] considered a real data set regarding the March precipitation measured in inches, over a period of 30 years which was reported by Hinkley [82].

### 4.2 Applications in industry and life-testing

Salehi and Doostparast [83] considered a data set on life testing of an given electrical equipment, planned for quality control purposes. Singh, et al. [84] considered the data set represents the failure times (in h) of 59 conductors from an accelerated life test from Lawless [85]. Vidović [86] considered the case where failure and running times (1000 of cycles) of a sample of 30 units of a larger electrical system are under study. Wang, et al. [87] considered the real-life data set from Lawless [25, p. 3] which represents the times to breakdown of an electrical insulating fluid subjected to 30 kilovolts. Wu [10] considered the data for times between successive failures of air conditioning equipment in a Boeing 70 airplane.

### 4.3 Applications in health and medicine

Salehi and Doostparast [83] considered a data set representing the Hemoglobin of the Australian 102 men athletes data. Seo and Kim [88], Awwad, et al. [81] discussed the analysis of



the data of the survival times in days of a group of lung cancer patients provided in Lawless. Kumar, et al. [89] reanalyzing Efron's data pertaining to a head-and-neck cancer clinical trial. EL-Sagheer, et al. [90], Fayyazishishavan and Kılıç Depren [91] used the data represents a COVID-19 data belonging to the Netherlands of 30 days, which recorded from 31 March to 30 April 2020.

#### 4.4 Miscellaneous applications

Carlin and Gelfand [92] considered the record-breaking Olympic high jumps since 1896, as presented in the World Almanac and Book of Facts 1989 by Hoffman [93]. Tanış [94] considered the data includes of the monthly actual taxes revenue (in million Egyptian pounds) in Egypt from January 2006 to November 2010. Volovskiy and Kamps [70] applied the proposed prediction procedures to the well-known Danish reinsurance claims dataset to predict record fire losses. The data were collected at Copenhagen Reinsurance and consist of 2167 fire losses in millions of Danish Krone between 1980 and 1990.

### 5. Computer Software for Records

Software specialized in calculating record values is quite rare. In program R, for example, the built-in routines for computing the records are available in two packages:

**5.1 Package "Records":** This package includes Functions for producing lower k-record times; lower k-record values; upper k-record times; upper k-record values for given samples (See Appendix). <https://CRAN.R-project.org/package=Records>.

**5.2 Package "RecordTest":** This package includes statistical tools based on the probabilistic properties of the record occurrence in a sequence of independent and identically distributed continuous random variables. That is tools to prepare a time series as well as distribution-free trend and change-point tests and graphical tools to study the record occurrence. Details about the implemented tools can be found in Castillo-Mateo, et al. [95]. <https://CRAN.R-project.org/package=RecordTest>.

### 6. Future Work

Although much has been done with respect to record values theory, there is still scope for more work. Here, we discuss some open problems that the researchers may like to work on.

- i. There is little work with respect to the theory of records for bivariate or multivariate random sequences. Therefore, we recommend more studies in this direction.
- ii. Does the type of records affect the estimates of the parameters?. Tripathi, et al. [96] noted that the performance of the estimator depends on the type of records. However, the suitability of the type of record varies from one distribution to distribution. Therefore, this topic needs further study.
- iii. Develop a methodology for conducting inference based on record values and record times. Where the record times and record values jointly contain considerably more information about distribution than do the record values alone, see Feuerverger and Hall [97].
- iv. Instead of just using record values in inference, we suggest using the record values with their corresponding inter-record times, see Kızılaslan and Nadar [98], Arashi and Emadi [99].
- v. Investigating the concept of records with respect to using the kernel density approach to characterize the behavior of records is an interesting extension of the theory of records. This approach will be very useful in cases where a classical distribution

cannot be identified to statistically fit the underlying data from which the record observations are obtained.

- vi. Serious difficulties arise for statistical inference based on records due to the fact that the occurrences of record data are very rare in practical situations [since the mean of the number of records in a random sample of size  $n$  is equal to  $1 + 2^{-1} + \dots + n^{-1}$  (see Arnold, et al. [100])] and the expected waiting time is infinite for every record after the first. Although, these problems can be avoided if we consider the model of  $k$ -record statistics introduced by Dziubdziela and Kopociński [4]. However, research is still open to investigating this problem.

## 7. Conclusion

By reviewing the literature (we only mentioned some of them) on the record values in statistical inference, we conclude that the records have played a significant and widely utilized role for statistical inference in parameter estimation, predicting future values, hypothesis tests, as well as stress-strength tests and characterizing of distributions. For this purpose, various known statistical inference methods have been used, including Bayesian and non-Bayesian methods. We also conclude from the applications in previous studies that the records are not limited to a specific field, but rather comprehend all aspects of life including sports, health, medicine, insurance, economy, industry, climate, environment, floods, and rainfall. Overall, records play a critical role in statistical inference by providing the data needed to make informed decisions and draw accurate conclusions.

## References

- [1] K. Chandler, "The distribution and frequency of record values," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 14, no. 2, pp. 220-228, 1952.
- [2] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja, *Records*. John Wiley & Sons, 2011.
- [3] M. Ahsanullah and V. B. Nevzorov, "Records via probability theory," 2015.
- [4] W. Dziubdziela and B. Kopociński, "Limiting properties of the  $k$ -th record values," *Applicationes Mathematicae*, vol. 2, no. 15, pp. 187-190, 1976.
- [5] Z. Grudzien, "Characterization of distribution of time limits in record statistics as well as distributions and moments of linear record statistics from the samples of random numbers," *Praca Doktorska, UMCS, Lublin*, 1982.
- [6] Z. F. Jaheen, "Empirical Bayes inference for generalized exponential distribution based on records," *Communications in Statistics-Theory and Methods*, vol. 33, no. 8, pp. 1851-1861, 2004.
- [7] J. Ahmadi and M. Doostparast, "Bayesian estimation and prediction for some life distributions based on record values," *Statistical Papers*, vol. 47, pp. 373-392, 2006.
- [8] N. Balakrishnan and A. Stepanov, "On the Fisher information in record data," *Statistics & probability letters*, vol. 76, no. 5, pp. 537-545, 2006.
- [9] M. Doostparast, "A note on estimation based on record data," *Metrika*, vol. 69, no. 1, pp. 69-80, 2009.
- [10] S.-F. Wu, "Interval Estimation for the Two-Parameter Exponential Distribution Based on the Upper Record Values," *Symmetry*, vol. 14, no. 9, p. 1906, 2022.
- [11] A. Asgharzadeh, S. Bagheri, N. Ibrahim, and M. Abubakar, "Optimal confidence regions for the two-parameter exponential distribution based on records," *Computational Statistics*, vol. 35, pp. 309-326, 2020.
- [12] A. Baklizi, "Estimation of  $\Pr(X < Y)$  using record values in the one and two parameter exponential distributions," *Communications in Statistics—Theory and Methods*, vol. 37, no. 5, pp. 692-698, 2008.

- [13] M. Ahsanullah and F. Aliev, "Some characterizations of exponential distribution by record values," *Journal of Statistical Research*, vol. 42, no. 2, pp. 41-46, 2008.
- [14] N. Abd-El-Hakim and K. Sultan, "Maximum likelihood estimates of Weibull parameters based on record values," *J. Egypt. Math. Soc.*, vol. 9, no. 1, pp. 79-89, 2001.
- [15] A. A. Soliman, A. H. Abd Ellah, and K. S. Sultan, "Comparison of estimates using record statistics from Weibull model: Bayesian and non-Bayesian approaches," *Computational Statistics & Data Analysis*, vol. 51, no. 3, pp. 2065-2077, 2006.
- [16] A. A. Jafari and H. Zakerzadeh, "Inference on the parameters of the Weibull distribution using records," *arXiv preprint arXiv:1501.02201*, 2015.
- [17] B. X. Wang and Z.-S. Ye, "Inference on the Weibull distribution based on record values," *Computational Statistics & Data Analysis*, vol. 83, pp. 26-36, 2015.
- [18] M. Z. Raqab, L. A. Alkhalfan, O. M. Bdair, and N. Balakrishnan, "Maximum likelihood prediction of records from 3-parameter Weibull distribution and some approximations," *Journal of computational and applied mathematics*, vol. 356, pp. 118-132, 2019.
- [19] A. S. Hassan, H. F. Nagy, H. Z. Muhammed, and M. S. Saad, "Estimation of multicomponent stress-strength reliability following Weibull distribution based on upper record values," *Journal of Taibah University for Science*, vol. 14, no. 1, pp. 244-253, 2020.
- [20] F. S. Al-Duais, "Bayesian estimations under the weighted LINEX loss function based on upper record values," *Complexity*, vol. 2021, pp. 1-7, 2021.
- [21] M. Ahsanullah, "Estimation of the parameters of the Gumbel distribution based on the m record values," *Comput. Statist. Quart.*, vol. 6, pp. 231-239, 1990.
- [22] M. A. Mousa, Z. Jaheen, and A. Ahmad, "Bayesian estimation, prediction and characterization for the Gumbel model based on records," *Statistics: A Journal of Theoretical and Applied Statistics*, vol. 36, no. 1, pp. 65-74, 2002.
- [23] I. Malinowska and D. Szydal, "On a family of Bayesian estimators and predictors for a Gumbel model based on the kth lower records," *Applicationes Mathematicae*, vol. 1, no. 31, pp. 107-115, 2004.
- [24] J. I. Seo and Y. Kim, "Statistical inference on Gumbel distribution using record values," *Journal of the Korean Statistical Society*, vol. 45, no. 3, pp. 342-357, 2016.
- [25] A. Asgharzadeh, M. Abdi, and S. Nadarajah, "Interval estimation for Gumbel distribution using climate records," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 39, pp. 257-270, 2016.
- [26] M. Ahsanullah and B. Holland, "Distributional properties of record values from the geometric distribution," *Statistica neerlandica*, vol. 41, no. 2, pp. 129-137, 1987.
- [27] H. M. Okasha and J. Wang, "E-Bayesian estimation for the geometric model based on record statistics," *Applied Mathematical Modelling*, vol. 40, no. 1, pp. 658-670, 2016.
- [28] G. Francis, E. Anjana, and E. Jeevanand, "Shrinkage Estimation of Strength Reliability for Geometric Distribution Using Record Values," *Acta Scientific COMPUTER SCIENCES Volume*, vol. 4, no. 4, 2022.
- [29] B. C. Arnold and S. J. Press, "Bayesian estimation and prediction for Pareto data," *Journal of the American Statistical Association*, vol. 84, no. 408, pp. 1079-1084, 1989.
- [30] A. El-Qasem, "Estimation via record values," *Journal of Information and Optimization Sciences*, vol. 17, no. 3, pp. 541-548, 1996.
- [31] K. S. Sultan and M. E. Moshref, "Record values from generalized Pareto distribution and associated inference," *Metrika*, vol. 51, no. 2, pp. 105-116, 2000.
- [32] M. Z. Raqab, J. Ahmadi, and M. Doostparast, "Statistical inference based on record data from Pareto model," *Statistics*, vol. 41, no. 2, pp. 105-118, 2007.
- [33] M. Z. Raqab, "Distribution-free prediction intervals for the future current record statistics," *Statistical Papers*, vol. 50, pp. 429-439, 2009.
- [34] M. Doostparast, M. G. Akbari, and N. Balakrishna, "Bayesian analysis for the two-parameter Pareto distribution based on record values and times," *Journal of Statistical Computation and Simulation*, vol. 81, no. 11, pp. 1393-1403, 2011.
- [35] M. Ahsanullah and M. Shakil, "A note on the characterizations of Pareto distribution by upper record values," *Communications of the Korean Mathematical Society*, vol. 27, no. 4, pp. 835-842, 2012.

- [36] Q. J. Azhad, M. Arshad, and N. Khandelwal, "Statistical inference of reliability in multicomponent stress strength model for pareto distribution based on upper record values," *International Journal of Modelling and Simulation*, vol. 42, no. 2, pp. 319-334, 2022.
- [37] M. T. Madi and M. Z. Raqab, "Bayesian prediction of rainfall records using the generalized exponential distribution," *Environmetrics: The official journal of the International Environmetrics Society*, vol. 18, no. 5, pp. 541-549, 2007.
- [38] A. Baklizi, "Likelihood and Bayesian estimation of  $\Pr(X < Y)$  using lower record values from the generalized exponential distribution," *Computational Statistics & Data Analysis*, vol. 52, no. 7, pp. 3468-3473, 2008.
- [39] S. Dey, T. Dey, M. Salehi, and J. Ahmadi, "Bayesian inference of generalized exponential distribution based on lower record values," *American Journal of Mathematical and Management Sciences*, vol. 32, no. 1, pp. 1-18, 2013.
- [40] S. Sana and M. Faizan, "Bayesian estimation using lindley's approximation and prediction of generalized exponential distribution based on lower record values," *Journal of Statistics Applications & Probability*, vol. 10, no. 1, pp. 61-75, 2021.
- [41] N. Balakrishnan and P. Chan, "Record values from Rayleigh and Weibull distributions and associated inference," *NIST special publication SP*, pp. 41-41, 1994.
- [42] M. Hendi, S. Abu-Youssef, and A. Alraddadi, "A Bayesian analysis of record statistics from the Rayleigh model," in *International Mathematical Forum*, 2007, vol. 2, no. 13, pp. 619-631.
- [43] A. A. Soliman and F. M. Al-Aboud, "Bayesian inference using record values from Rayleigh model with application," *European Journal of Operational Research*, vol. 185, no. 2, pp. 659-672, 2008.
- [44] M. Ahsanullah and M. Shakil, "Characterizations of Rayleigh distribution based on order statistics and record values," *Bull. Malays. Math. Sci. Soc.*, vol. 36, no. 3, pp. 625-635, 2013.
- [45] J.-I. Seo, J.-W. Jeon, and S.-B. Kang, "Exact interval inference for the two-parameter Rayleigh distribution based on the upper record values," *Journal of Probability and Statistics*, vol. 2016, 2016.
- [46] J. I. Seo and Y. Kim, "Objective Bayesian inference based on upper record values from Rayleigh distribution," *Communications for Statistical Applications and Methods*, vol. 25, no. 4, pp. 411-430, 2018.
- [47] M. Abdi and A. Asgharzadeh, "Rayleigh confidence regions based on record data," *Journal of Statistical Research of Iran JSRI*, vol. 14, no. 2, pp. 171-188, 2018.
- [48] M.-Y. Lee and E.-H. Lim, "Characterizations of the Lomax, exponential and Pareto distributions by conditional expectations of record values," *Journal of the Chungcheong Mathematical Society*, vol. 22, no. 2, pp. 149-149, 2009.
- [49] P. Nasiri and S. Hosseini, "Statistical inferences for Lomax distribution based on record values (Bayesian and classical)," *Journal of Modern Applied Statistical Methods*, vol. 11, no. 1, p. 15, 2012.
- [50] M. A. Mahmoud, A. A. Soliman, A. H. Abd Ellah, and R. M. El-sagheer, "MCMC technique to study the Bayesian estimation using record values from the Lomax distribution," *International Journal of Computer Applications*, vol. 73, no. 5, 2013.
- [51] M. A. Mahmoud, R. M. El-Sagheer, A. A. Soliman, and A. H. Abd Ellah, "Bayesian estimation of  $P[Y < X]$  based on record values from the Lomax distribution and MCMC technique," *Journal of Modern Applied Statistical Methods*, vol. 15, no. 1, p. 25, 2016.
- [52] A. S. Hassan and A. N. Zaky, "Entropy Bayesian estimation for Lomax distribution based on record," *Thailand Statistician*, vol. 19, no. 1, pp. 95-114, 2021.
- [53] M. A. Selim, "Estimation and prediction for Nadarajah-Haghighi distribution based on record values," *Pak. J. Statist.*, vol. 34, no. 1, pp. 77-90, 2018.
- [54] M. Sana and M. Faizan, "Bayesian estimation for Nadarajah-Haghighi distribution based on upper record values," *Pakistan Journal of Statistics and Operation Research*, pp. 217-230, 2019.

- [55] S. T. MirMostafae, A. Asgharzadeh, and A. Fallah, "Record values from NH distribution and associated inference," *Metron*, vol. 74, pp. 37-59, 2016.
- [56] S. E. Abu-Youssef, "On characterization of certain distributions of record values," *Applied mathematics and computation*, vol. 145, no. 2-3, pp. 443-450, 2003.
- [57] J. Ahmadi, M. J. Jozani, É. Marchand, and A. Parsian, "Bayes estimation based on k-record data from a general class of distributions under balanced type loss functions," *Journal of Statistical Planning and Inference*, vol. 139, no. 3, pp. 1180-1189, 2009.
- [58] I. Malinowska and D. Szynal, "On characterization of certain distributions of kth lower (upper) record values," *Applied Mathematics and Computation*, vol. 202, no. 1, pp. 338-347, 2008.
- [59] M. Ahsanullah, "Linear prediction of record values for the two parameter exponential distribution," *Annals of the Institute of Statistical Mathematics*, vol. 32, pp. 363-368, 1980.
- [60] I. R. Dunsmore, "The future occurrence of records," *Annals of the Institute of Statistical Mathematics*, vol. 35, pp. 267-277, 1983.
- [61] A. M. Awad and M. Z. Raqab, "Prediction intervals for the future record values from exponential distribution: comparative study," *Journal of Statistical Computation and Simulation*, vol. 65, no. 1-4, pp. 325-340, 2000.
- [62] J. Ahmadi and S. MirMostafae, "Prediction intervals for future records and order statistics coming from two parameter exponential distribution," *Statistics & Probability Letters*, vol. 79, no. 7, pp. 977-983, 2009.
- [63] J. Paul and P. Y. Thomas, "On generalized upper (k) record values from Weibull distribution," *Statistica*, vol. 75, no. 3, pp. 313-330, 2015.
- [64] G. Volovskiy and U. Kamps, "Likelihood-Based Prediction of Future Weibull Record Values," *REVSTAT-Statistical Journal*, vol. 21, no. 3, pp. 425-445, 2023.
- [65] M. Ahsanullah, "Inference and prediction of the Gumbel distribution based on record values," *Pakistan Journal of Statistics*, vol. 7, no. 3, pp. 53-62, 1991.
- [66] M. T. Madi and M. Z. Raqab, "Bayesian prediction of temperature records using the Pareto model," *Environmetrics*, vol. 15, no. 7, pp. 701-710, 2004.
- [67] J. Paul and P. Y. Thomas, "On generalized (k) record values from Pareto distribution," *Aligarh J Statist*, vol. 36, no. 1, pp. 63-78, 2016.
- [68] A. R. Shafay, N. Balakrishnan, and J. Ahmadi, "Bayesian prediction of order statistics with fixed and random sample sizes based on k-record values from Pareto distribution," *Communications in Statistics-Theory and Methods*, vol. 46, no. 2, pp. 721-735, 2017.
- [69] Z. Vidović, "Bayesian Prediction of Order Statistics Based on k-Record Values from a Generalized Exponential Distribution," *Stats*, vol. 2, no. 4, pp. 447-456, 2019.
- [70] G. Volovskiy and U. Kamps, "Comparison of likelihood-based predictors of future Pareto and Lomax record values in terms of Pitman closeness," *Communications in Statistics-Theory and Methods*, vol. 52, no. 6, pp. 1905-1922, 2023.
- [71] E. K. AL-Hussaini and A. E.-B. A. Ahmad, "On Bayesian predictive distributions of generalized order statistics," *Metrika*, vol. 57, pp. 165-176, 2003.
- [72] J. Ahmadi, M. Jafari Jozani, É. Marchand, and A. Parsian, "Prediction of k-records from a general class of distributions under balanced type loss functions," *Metrika*, vol. 70, no. 1, pp. 19-33, 2009.
- [73] M. Z. Raqab and N. Balakrishnan, "Prediction intervals for future records," *Statistics & Probability Letters*, vol. 78, no. 13, pp. 1955-1963, 2008.
- [74] M. Nadar, A. Papadopoulos, and F. Kızılaslan, "Statistical analysis for Kumaraswamy's distribution based on record data," *Statistical Papers*, vol. 54, pp. 355-369, 2013.
- [75] M. Chacko and M. S. Mary, "Estimation and prediction based on k-record values from normal distribution," *Statistica*, vol. 73, no. 4, pp. 505-516, 2013.
- [76] J.-I. Seo and J. J. Song, "A bayesian nonparametric model for upper record data," *Applied Mathematical Modelling*, vol. 71, pp. 363-374, 2019.
- [77] G. Volovskiy and U. Kamps, "Maximum product of spacings prediction of future record values," *Metrika*, vol. 83, no. 7, pp. 853-868, 2020.
- [78] A. Asgharzadeh, A. Fallah, M. Raqab, and R. Valiollahi, "Statistical inference based on Lindley record data," *Statistical Papers*, vol. 59, pp. 759-779, 2018.

- [79] M. Z. Raqab, O. M. Bdair, and F. M. Al-Aboud, "Inference for the two-parameter bathtub-shaped distribution based on record data," *Metrika*, vol. 81, pp. 229-253, 2018.
- [80] A. Tripathi, U. Singh, and S. K. Singh, "Inferences for the DUS-exponential distribution based on upper record values," *Annals of Data Science*, vol. 8, pp. 387-403, 2021.
- [81] R. R. A. Awwad, O. M. Bdair, and G. K. Abufoudeh, "Bayesian estimation and prediction based on Rayleigh record data with applications," *Statistics in Transition new series*, vol. 22, no. 3, pp. 59-79, 2021.
- [82] D. Hinkley, "On quick choice of power transformation," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 26, no. 1, pp. 67-69, 1977.
- [83] M. Salehi and M. Doostparast, "Expressions for the mean of the order statistics from the skew-normal distribution and their application," *The Proceeding of Refereed and Invited Papers*, p. 477, 2015.
- [84] S. Singh, Y. Mani Tripathi, and S.-J. Wu, "Bayesian estimation and prediction based on lognormal record values," *Journal of Applied Statistics*, vol. 44, no. 5, pp. 916-940, 2017.
- [85] J. F. Lawless, *Statistical models and methods for lifetime data*. John Wiley & Sons, 2011.
- [86] Z. Vidović, "On MLEs of the parameters of a modified Weibull distribution based on record values," *Journal of Applied Statistics*, vol. 46, no. 4, pp. 715-724, 2019.
- [87] L. Wang, Y. M. Tripathi, S.-J. Wu, and M. Zhang, "Inference for confidence sets of the generalized inverted exponential distribution under k-record values," *Journal of Computational and Applied Mathematics*, vol. 380, p. 112969, 2020.
- [88] J. I. Seo and Y. Kim, "Objective Bayesian analysis based on upper record values from two-parameter Rayleigh distribution with partial information," *Journal of Applied Statistics*, vol. 44, no. 12, pp. 2222-2237, 2017.
- [89] D. Kumar, M. Kumar, and J. Saran, "Power Generalized Weibull Distribution Based on Record Values and Associated Inferences with Bladder Cancer Data Example," *Communications in Mathematics and Statistics*, pp. 1-26, 2022.
- [90] R. M. EL-Sagheer, M. S. Eliwa, K. M. Alqahtani, and M. El-Morshedy, "Bayesian and non-Bayesian inferential approaches under lower-recorded data with application to model COVID-19 data," *AIMS Mathematics*, vol. 7, no. 9, pp. 15965-15981, 2022.
- [91] E. Fayyazishishavan and S. Kılıç Depren, "Inference of stress-strength reliability for two-parameter of exponentiated Gumbel distribution based on lower record values," *Plos one*, vol. 16, no. 4, p. e0249028, 2021.
- [92] B. P. Carlin and A. E. Gelfand, "Parametric likelihood inference for record breaking problems," *Biometrika*, vol. 80, no. 3, pp. 507-515, 1993.
- [93] M. S. Hoffman, "The world almanac and book of facts 1989. New York: Newspaper Enterprise Association," ed: Inc, 1988.
- [94] C. Tanış, "Transmuted lower record type inverse rayleigh distribution: estimation, characterizations and applications," *Ricerche di Matematica*, vol. 71, no. 2, pp. 777-802, 2022.
- [95] J. Castillo-Mateo, A. C. Cebrián, and J. Asín, "RecordTest: An R Package to Analyze Non-Stationarity in the Extremes Based on Record-Breaking Events," *Journal of Statistical Software*, vol. 106, pp. 1-28, 2023.
- [96] A. Tripathi, U. Singh, and S. K. Singh, "Does the Type of Records Affect the Estimates of the Parameters?," *Journal of Modern Applied Statistical Methods*, vol. 19, no. 1, p. 27, 2022.
- [97] A. Feuerverger and P. Hall, "On statistical inference based on record values," *Extremes*, vol. 1, pp. 169-190, 1998.
- [98] F. Kızılaslan and M. Nadar, "Estimation and prediction of the Kumaraswamy distribution based on record values and inter-record times," *Journal of Statistical Computation and Simulation*, vol. 86, no. 12, pp. 2471-2493, 2016.
- [99] M. Arashi and M. Emadi, "Evidential inference based on record data and inter-record times," *Statistical Papers*, vol. 49, pp. 291-301, 2008.
- [100] B. Arnold, N. Balakrishnan, and H. Nagaraja, "Records. John Wiley&Sons," *New York*, 1998.

### Appendix: R program

```
rinfal<-c(11.30, 20.34, 13.13, 10.4, 12.11, 38.18, 9.21, 22.31, 14.05, 13.87, 19.28, 34.84, 13.36,
11.85, 26.28, 6.73, 16.11, 8.51, 16.86, 7.06, 5.59, 7.91, 16.29, 10.6, 19.32, 8.72, 19.52, 18.65, 19.3, 11.72,
19.18, 12.63, 16.18, 11.6, 13.42, 23.65, 17.05, 19.92, 15.26, 13.86, 8.58, 12.52, 13.71, 19.66, 9.59, 6.67,
7.38, 17.56, 17.44, 9.77, 12.66, 12.5, 12.53, 16.95, 11.84, 14.55, 21.66, 12.07, 22.41, 23.43, 13.06, 18.96,
32.76, 11.18, 19.17, 19.21, 11.58, 12.13, 12.61, 7.22, 7.99, 10.6, 8.21, 26.21, 9.46, 11.99, 11.94, 16, 9.54,
21.13, 5.58, 8.18, 4.85, 18.79, 8.38, 7.93, 13.69, 20.44, 22, 16.58, 27.47, 7.77, 12.32, 7.17, 21.26, 14.92,
14.35, 7.22, 12.31, 33.44, 19.67, 26.98, 8.98, 10.71, 31.25, 10.43, 12.82, 17.86, 7.66, 12.48, 8.08, 7.35,
11.47, 21, 27.36, 8.11, 24.35, 12.46, 12.4, 31.01, 9.09, 11.57, 17.94, 4.42, 16.49, 9.24, 37.25, 13.19, 3.21,
13.53, 9.08, 16.36, 20.2, 8.69, 5.85, 6.08, 8.52, 9.65, 19, 4.79, 18.82, 14.86, 5.82, 12.4, 27.85)
> library(Records)
> lower.record.times(rinfal, 1)
[1] 1 4 7 16 21 81 83 124 129
lower.record.values(rinfal, 1)
[1] 11.30 10.40 9.21 6.73 5.59 5.58 4.85 4.42 3.21
> upper.record.times(rinfal, 1)
[1] 1 2 6
> upper.record.values(rinfal, 1)
[1] 11.30 20.34 38.18
> lower.record.times(rinfal, 2)
[1] 2 3 4 7 16 18 20 21 46 81 83 124 129
> lower.record.values(rinfal, 2)
[1] 20.34 13.13 11.30 10.40 9.21 8.51 7.06 6.73 6.67 5.59 5.58 4.85 4.42
> upper.record.times(rinfal, 2)
[1] 2 3 6 8 12 127
> upper.record.values(rinfal, 2)
[1] 11.30 13.13 20.34 22.31 34.84 37.25
```

# APPLICATION OF THE FUZZY-SET THEORY TO ASSESS THE KNOWLEDGE OF ELECTRIC POWER INDUSTRY SPECIALISTS

V.Kh. Nasibov R.R. Alizade I.Y. Mastaliyev A.M. Ramazanli

Azerbaijan Scientific-Research and Design-Prospecting Power Engineering Institute  
nvaleh@mail.ru , rena\_alizade@mail.ru , ismayil\_inf@mail.ru , ahmet.ramazanli@gmail.com

## Abstract

*In most cases, the assessment of the knowledge of electric power industry workers is carried out according to a test scheme, where the correct answer is selected from the list of answers. All questions have the same difficulty and only the single correct answer gives a certain score [1]. The article developed a universal model for assessing the knowledge of electric power industry workers, where using the theory of fuzzy logic and fuzzy inference, both the complexity of questions and the possibility of a partial correct answer are taken into account.*

**Keywords:** Knowledge Assessment, Training, Fuzzy Knowledge Base

## I. Introduction

Articles Personnel training, advanced training of electric power industry workers is a necessary task to improve the efficiency and safety of the operation of electric power facilities. Refresher courses for employees, which should be held at least once every 3-5 years, are necessary for employees and the head of energy enterprises in the electric power industry. Terms should be determined by the internal regulations of the enterprise, as well as the requirements of standards. Refresher courses are held; electricians, technologists, power engineers and heads of departments. Upon completion of advanced training courses, knowledge is tested by conducting an appropriate exam, where a test scheme of answers to the questions posed is mainly implemented. Only one correct answer is selected from the submitted answers, all other answers are considered incorrect. With this approach to testing knowledge, the complexity of the questions is not taken into account, and the possibility of a partial correct answer is also excluded.

The need to take into account the complexity of questions and a partial correct answer makes it possible to use the theory of fuzzy sets and fuzzy inference to assess the level of preparedness of electric power industry workers [2].

## II. Knowledge assessment

To account for the complexity of the questions, the questions are divided into four groups: relatively easy questions, normal questions, questions of medium difficulty and difficult questions. The weight coefficients of correct answers are ranked according to the level of difficulty of the questions. The partial correct answer for groups of normal and questions of average difficulty has



a smaller total in the resulting assessment of knowledge than the partial correct answer for complex questions. To obtain a quantitative value of knowledge assessment based on linguistic information, one can use the provisions of the theory of fuzzy sets and fuzzy logic [3].

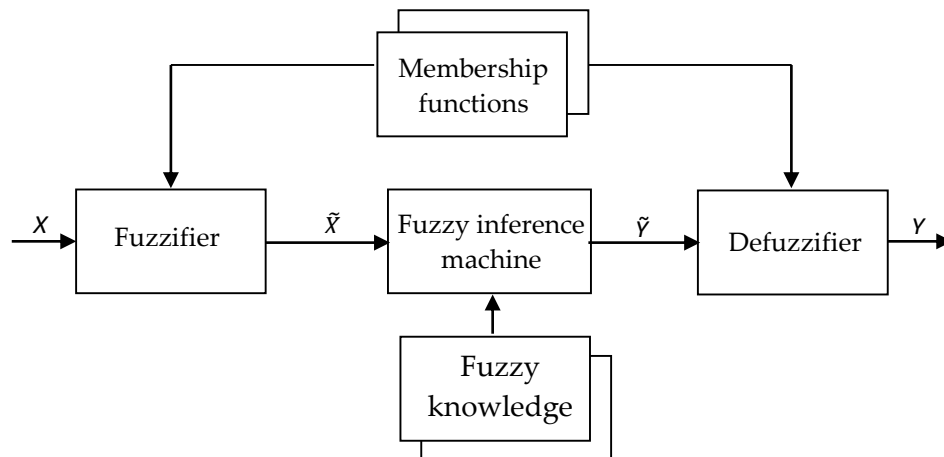


Figure 1: Fuzzy inference circuit

The fuzzy model contains the following blocks: a fuzzifier that converts a fixed vector of influencing factors  $X$  into a vector of fuzzy sets  $\tilde{X}$  required to perform fuzzy inference;

$$Y = f(X)$$

fuzzy knowledge base containing information about dependence in the form of linguistic rules of the "IF-THEN" type;

a fuzzy inference machine that, based on the rules of the knowledge base, determines the value of the output variable in the form of a fuzzy set  $\tilde{Y}$  corresponding to the fuzzy values of the input variables  $\tilde{X}$ ;

a defuzzifier that converts the output fuzzy set  $\tilde{Y}$  into a clear number  $Y$ . The Matlab program contains the Fuzzy Logic Toolbox package, which implements two types of fuzzy models, the Mamdani and Sugeno types. For our case, a Mamdani-type fuzzy model is preferable.

In the Mamdani-type model, the relationship between inputs  $X=(x_1, x_2, \dots, x_n)$  and output  $Y$  is determined by a fuzzy knowledge base of the following format:

$$\text{if } (x_1 = a_{1,j1}) \text{ and } (x_2 = a_{2,j1}) \text{ and...and } (x_n = a_{n,j1})$$

$$\text{or } (x_1 = a_{1,j2}) \text{ and } (x_2 = a_{2,j2}) \text{ and...and } (x_n = a_{n,j2})$$

$$\text{or } (x_1 = a_{1,jk_j}) \text{ and } (x_2 = a_{2,jk_j}) \text{ and...and } (x_n = a_{n,jk_j})$$

That

$$y = d_j, \quad i = 1, m,$$

Where

$a_{i,jp}$  – linguistic term, which evaluates the variable  $X_i$  in the line with the number

$$jp \quad (p = \overline{1, k_j})$$

$k_j$  – number of rows – conjunctions in which the output  $y$  evaluated by linguistic term  $d_j$ ;

$m$  – the number of terms used for the linguistic evaluation of the output variable  $y$ .

All linguistic terms in the knowledge base are represented as fuzzy sets defined by the corresponding membership functions:

$\mu_{jp}(x_i)$  – input membership function  $x_i$  fuzzy term  $a_{i,jp}$ ,  $i = \overline{1, n}$ ,  $j = \overline{1, m}$ ,  $p = \overline{1, k_j}$ , those.

$$a_{i,jp} = \int_{x_i}^{\overline{x_i}} \mu_{jp}(x_i) / x_i, \quad x_i \in [x_i, \overline{x_i}]$$

$\mu_{d_j}(y)$  – output membership function  $y$  fuzzy term  $d_j$ ,  $j = \overline{1, m}$ , those.

$$d_j = \int_y^{\overline{y}} \mu_{d_j}(y) / y, \quad y \in [y, \overline{y}]$$

Degree input vector accessories  $X^* = (x_1^*, x_2^*, \dots, x_n^*)$  fuzzy terms  $d_j$  from the fuzzy knowledge base is determined by the following system of fuzzy logical equations:

$$\mu_{d_j}(X^*) = \underset{p=1, k_j}{\mathbf{V}} \underset{i=1, n}{\mathbf{\Delta}} [\mu_{jp}(x_i^*)], \quad j = \overline{1, m},$$

Where  $\mathbf{V}$  ( $\mathbf{\Delta}$ ) –operation from the s-norm (t-norm), i.e. from a set of implementations of logical operations OR (AND). The following implementations are most often used: for the OR operation - finding the maximum, for the AND operation - finding the minimum.

The fuzzy set  $\tilde{y}$  corresponding input vector  $X^*$ , is defined as follows:

$$\tilde{y} = \underset{j=1, m}{\mathbf{agg}} \left( \int_y^{\overline{y}} \mathbf{imp}(\mu_{d_j}(X^*), \mu_{d_j}(y)) / y \right),$$

Where

$\mathbf{imp}$  – implication, usually implemented as a minimum finding operation;

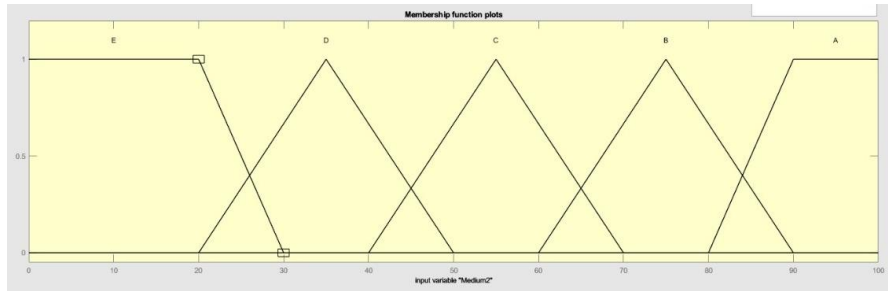
$\mathbf{agg}$  – aggregation of fuzzy sets, which is most often implemented by the operation of finding the maximum.

### III. Fuzzy inference models

Clear output value  $y$ , corresponding to the input vector  $X^*$ , is determined as a result of defuzzification of the fuzzy set  $\tilde{y}$ . The most commonly used defuzzification is the center of gravity method:

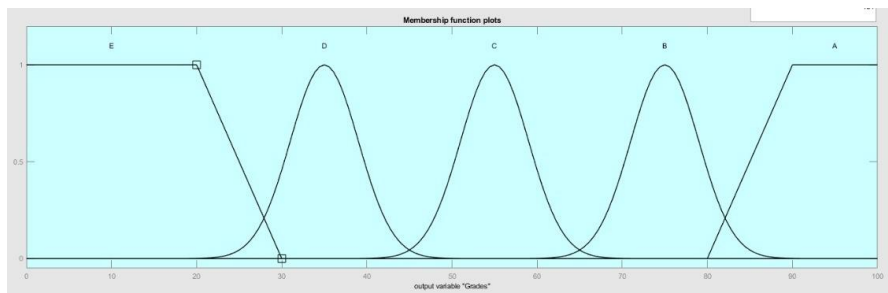
The choice of the membership function affects the accuracy of the fuzzy inference model. Figures 2-9 show various membership functions for input and output variables [4-6].

Figure 2 shows the function of input variables (answers to all four groups of questions by complexity) in the form of a triangle. The optimal output membership function is shown in Figure 3, in which the adequacy of the result to the rules corresponds to 88.3%.



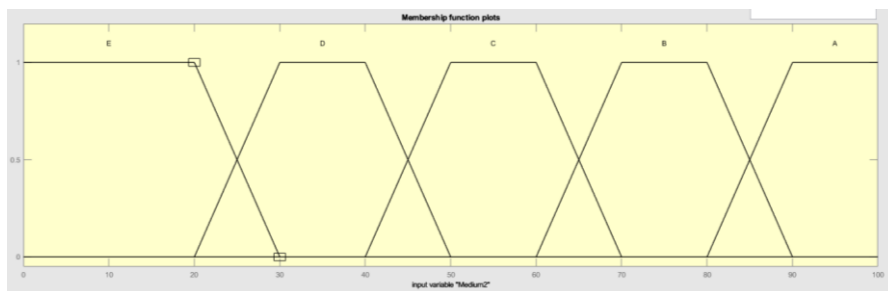
**Figure 2:** Linear membership function for the input

According to the rules (Rules), the program compared the data with the forms of relations and received the results. The triangular membership function results were compared with the rules and it was observed that the result was about 88.3% correct.

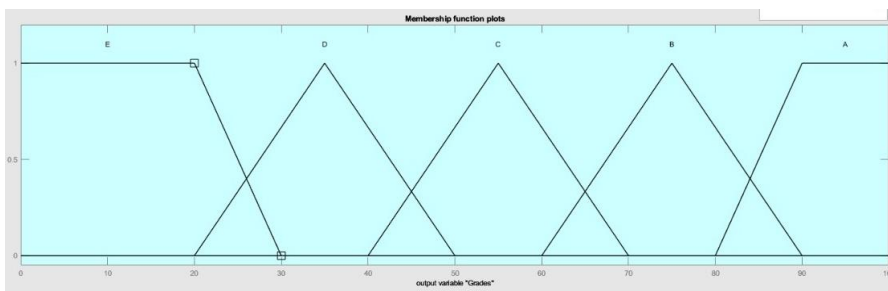


**Figure 3:** Linear membership function for the output

Figure 4 shows the membership function of the inputs in the form of a trapezoid, and Figure 5 corresponds to the membership function of the output in the form of a triangle. Such a choice of the output membership function leads to a high indicator of the adequacy of the output to the rules - 91%.

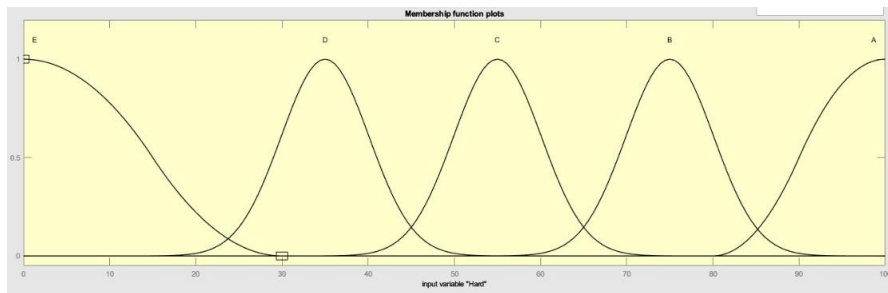


**Figure 4:** Trapezoidal membership function for the input

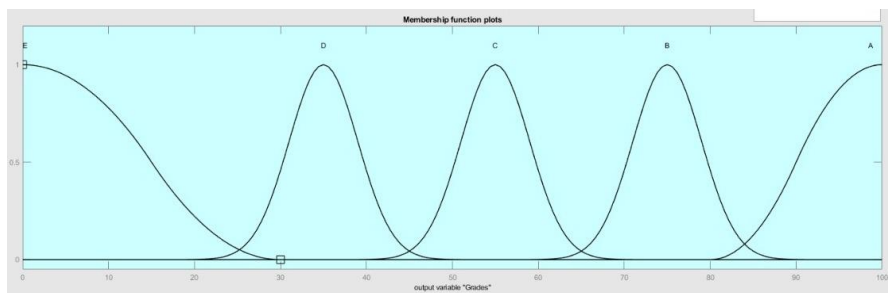


**Figure 5:** Trapezoidal membership function for the output

One of the most commonly used membership functions is the Gbell function (a Gaussian type function). The graphs of the membership functions for the inputs and for the output are shown in Figures 6 and 7, respectively, the result is 88.1% adequate.

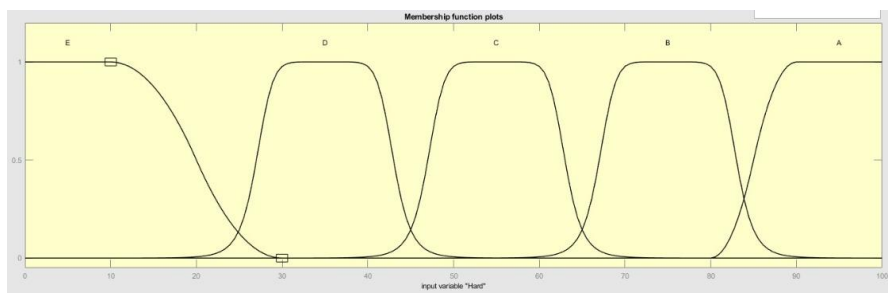


**Figure 6:** Gaussian membership function for the input

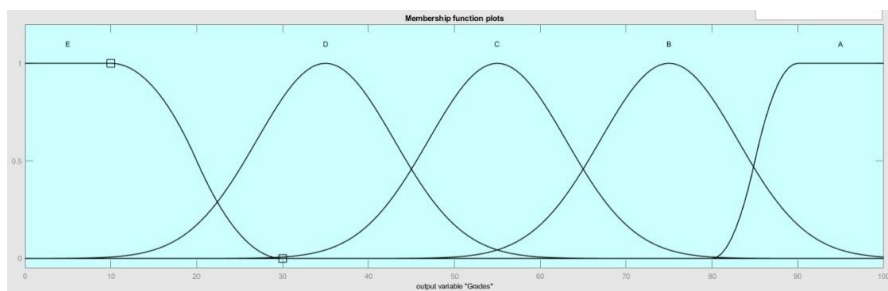


**Figure 7:** Gaussian membership function for the output

Numerous studies on the choice of membership functions for inputs and outputs have shown that the maximum adequacy is achieved when using Gauss-Linear functions, which is formed by combining the Gauss and limf functions (Gauss and limf), which are shown in Figure 7 and 8. For these membership functions, the adequacy of the output to the rules is 97.7%.



**Figure 8:** Gauss-Linear membership function for the input



**Figure 9:** Gauss-Linear membership function for the output

Table 1 shows a comparison of the adequacy of the output to the rules for various membership functions. As can be seen from Table 1, the maximum adequacy of the model is achieved when using membership functions of the Gauss-Linear type [7-9].

**Table 1:** Percentages, based on the rules of the membership functions

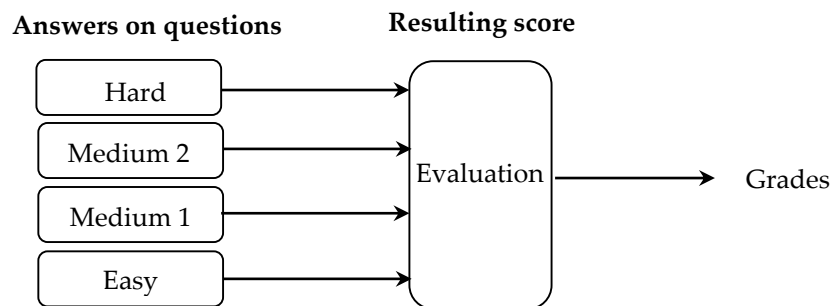
Membership Function	Accuracy percentages
Triangular	88.3
Trapezoid	91.0
Gaussian	88.1
Gauss Linear	97.7

Table 2 presents a fragment of the learning rules (knowledge base) of the fuzzy inference model for assessing the level of preparedness of electric power industry workers.

**Table 2:** Fragment of model training rules

Examples	Hard	Medium 2	Medium 1	Easy	Result (with rules)	Result (Gbell-mf)
Examples 1	85	85	75	65	A	A(84.84)
Examples 2	85	75	85	55	A	A(84.84)
Examples 3	75	85	35	45	B	B(64.94)
Examples 4	65	55	65	45	B	B(64.94)
Examples 5	65	35	75	55	C	C(55.06)
Examples 6	15	25	85	65	C	D(37.87)
Examples 7	45	35	25	35	D	D(35.10)
Examples 8	15	25	45	95	D	D(27.86)
Examples 9	15	25	15	55	E	E(13.23)
Examples 10	5	15	65	55	E	E(13.23)

With the selected answers to questions from four blocks, you can get the corresponding score, as shown in Figure 10. Here, the first column shows the score for difficult questions, the fourth for easy questions, and the last column the resulting score.



**Figure 10:** a) Evaluation by points



Figure 10: b) Evaluation by points

#### IV. Discussion

1. A universal method for assessing the level of preparedness of electric power industry workers has been developed, where, using the theory of fuzzy logic and fuzzy inference, one can take into account the complexity of questions, as well as the possibility of a partial correct answer.

2. By choosing membership functions for the inputs and outputs of functions of the Gauss-Linear type, you can achieve the maximum adequacy of the fuzzy inference model - 97.7%.

#### References

- [1] Jianchang Lu & Yinchun Lvy, Research of Electric Power Enterprise Knowledge Workers' Incentive Based on Fuzzy Model. *Procedia Engineering*, Volume 16, 2011
- [2] Ling Li Talking about the Electric Power Enterprise Knowledge Staff Management Advantages and Challenges. *J. China power education*, 2006
- [3] Fuzzy Logic Toolbox™ User's Guide Matlab.
- [4] Nadia Rousseau, Carl Beaudoin, Corina Borri-Anadon & oth. Enquête nationale sur les services éducatifs complémentaires à la formation professionnelle. Rapport de recherche <https://www.researchgate.net/publication/344671920> Enquete nationale sur les services éducatifs complémentaires a la formation professionnelle Rapport de recherche
- [5] Zhiyong Xu, Xudong Zhang, Ming Zeng, Fan Yan Application of ANP-based Multilevel Fuzzy Comprehensive Evaluation Methods to Post-evaluation for Grid Construction Projects. *J. East China Electric Power*, 2009
- [6] Mei-Hsiang Wang & Tarng-Yao Yang Investigating the success of knowledge management: An empirical study of small- and medium-sized enterprises. *Asia Pacific Management Review* Volume 21, Issue 2, June 2016
- [7] I. Chan & C.K. Chao Knowledge management in small and medium-sized enterprises. *Communications of the ACM*, 51 (4), 2008
- [8] J.C. Anderson & D.W. Gerbing Structural equation modeling in practice: a review and recommended two-step approach. *Psychological Bulletin*, 103 (3), 1988
- [9] R.E. Bohn Measuring and managing technological knowledge. *Sloan Management Review*, 36 (1), 1994

# A CRITICAL LITERATURE REVIEW AND FUTURE PERSPECTIVE OF RAM APPROACHES FOR COMPLEX SYSTEMS IN VARIOUS PROCESS INDUSTRIES

Mausoof Sheikh<sup>1</sup>

•

<sup>1</sup>Ph.D. Scholar, Department of Production and Industrial Engineering, National Institute of Technology, Kurukshetra-136119, INDIA  
[shaikhmsf22@gmail.com](mailto:shaikhmsf22@gmail.com)

Dr. P.C. Tewari

•

Professor, Department of Mechanical Engineering, National Institute of Technology, Kurukshetra-136119, INDIA  
[pctewari1@gmail.com](mailto:pctewari1@gmail.com)

## Abstract

*In the industrial systems there is a requirement that systems should work efficiently for long time. System performance is an important aspect for failure free operation but in real practice complete failure free operation of any production system is seldom possible. Detailed critical literature review for the past thirty three years of Reliability, Maintainability and Availability (RAM) approaches has been carried out which can help to improve performance of Complex systems. Review of some papers provided the detailed information about past and current scenario of RAM practices in research field and industries. Different RAM tools and techniques extracted from the review may be helpful in qualitative and quantitative analysis of the complex systems. In this paper, author tried to focuss on some major aspects of RAM approaches.*

**Keywords:** Reliability, Availability, Maintainability, Safety, Markov, Petri Nets, Dependability.

## I. Introduction

All assets would be developed in a perfect world with low failure rates, low maintenance costs and simplicity of use in mind. There should be adequate balance in the productivity of an asset with the cost of its purchase and maintenance. A team of design, systems, and reliability professionals will usually do RAM analysis during the design phase. Over the course of the asset's life, the study can be repeated by maintenance and service reliability engineers, who have vital information on the performance and health of the asset.

These three factors (Reliability, Availability and Maintainability) are all equally crucial and frequently complement one another. Machine is operational if it is available, reliable if it is likely to function correctly, and maintainable if it can be quickly rectified even if something goes wrong.

**Reliability** can be defined as the probability that a system will complete the task and run flawlessly in a specific environment for a set period of time. In reliability engineering, high levels of "lifetime" engineering uncertainty and failure hazards are addressed through prediction, prevention, and risk management. Although stochastic parameters influence and determine dependability, mathematics and statistics are not the only ways to achieve it. Reliability engineering is closely related to quality engineering, safety engineering and system safety.

**Availability** is a special parameter that combines serviceability and reliability criteria. It indicates the probability that an asset will be operational (neither maintained nor repaired) at any given time. Availability is measured at steady state, taking into account potential downtime during which the service may (and will) become unavailable during its expected useful life. In reliability engineering calculations, the failure rate is taken as the predicted strength of failure, assuming the component is fully functional in its original state.

**Maintainability** discusses the ease of maintenance on an asset and the resource requirements. The probability that an asset will resume its intended state after a maintenance duty can also be calculated using this method. Mean Time to Repair (MTTR) is a common metric used to evaluate it.

System engineers, logisticians and users are particularly interested in the three RAM (Reliability, Maintainability and Availability) characteristics of a system. Together these characteristics have an impact on a product or system's usefulness as well as its life-cycle costs. A decision-making tool known as Research on Reliability, Availability and Maintainability (RAM) is utilized to increase system availability, which in turn increase overall profitability and reduces cycle costs life. In engineering, the term "Reliability, Availability, Maintainability and Safety" is frequently used to describe a property of a product or system.

## II. Critical Literature Review

**Ciaro et. al. (1990)** performed analysis of processing systems using semi-markov reward processes. The semi-markov reward process is an extension of an algorithm proposed by Beaudry, it was presented for the computation of accumulated reward in a semi-markov process [1]. **Viswanadham et. al. (1991)** formulated the performability of the fault-tolerant manufacturing system. Through examples, the authors try to show the importance of performability in automated manufacturing system design. Performability measures considered deal with throughput and manufacturing lead time, which essentially determine the competitiveness of a plant [2]. **Kumar et. al. (1992)** studied the analytic behavior of reliability and availability of the crystallizer system in sugar plants. The model was based on Chapman-Kolmogorov equations. The Laplace transform was used to derive steady-state availability and various state probabilities. The effect of failure and repair rates on availability has been studied [3]. **Sharma and Bazovsky (1993)** performed analyses of large and complex systems using Markov method. Laplace transform method was used to solve the differential equation. After the modeling, design engineers were able to evaluate their own design to increase the reliability of the system [4]. **Behera et. al. (1994)** used deterministic and stochastic based petri net to modelled the flexible manufacturing system. Performance evaluation of system has also been done. Generalized stochastic petri net was also used to modelled the flexible manufacturing systems. Performance measure obtained was almost equal for both deterministic petri net and generalized stochastic petri net [5].



**Murty and Naikan (1995)** investigated the optimization of a manufacturing plant's availability and maintenance costs. It has been concluded that before making any major decisions regarding the formation of the maintenance budget, it is prudent to carefully examine the economic viability of overspending on maintenance in order to increase plant availability [6]. **Kumar et. al. (1996)** assessed the shell gasification and carbon recovery processes in a urea fertilizer plant from a behavioral perspective. The author of this paper used straight forward probability considerations to formulate the issue. The equations for steady-state availability were derived, and they provided the equipment's behavior based on the analyses of the results. Based on the failure results, advise the maintenance manager with guidelines for carrying out the workplan for repairs, among other things [7]. **Arora and Mehta (1997)** evaluated the steam and power generation capacity of thermal power plants. The authors devised the expressions for steady-state availability and mean time between failures. Graphs illustrate how failure and repair rates affect system availability. The Chapman-Kolmogorov birth-death process and a probabilistic method were used in the modeling procedure. The critical system and subsystem made a decision to limit failures based on the results, and the plant staff was informed of the results so they could make plans for the system's failure-free operation [8]. **Pellegrini et. al. (1998)** used a statistical approach based on the semi-Markov technique to assess the availability and performance of electronic complex systems. The electronic system's resolution model of a semi-Markov process was identified from the findings, and the mean Laplace transform was used to calculate the asymptotic availability value [9]. **Singh and Mahajan (1999)** evaluated a production facility for utensils for availability and dependability. Differential equations resolved using the Laplace transformation. The Markovian method was used to investigate the effect of different parameter availability. The findings demonstrated that availability impacted when repair and failure rates are disrupted [10].

**Borgnovo et. al. (2000)** proposed modeling through Monte Carlo. The plant's tool management and operation were advised by Monte Carlo modeling. The paper's analysis looks at the operation and maintenance plan [11]. **Zhang and Horigome (2001)** looked at how the availability and dependability of the system's failure and repair rates evolved over time. The solution shows the system's availability and dependability with varying failure and repair rates [12]. **Wang and Loman (2002)** critically examined the K-out-of-N system's availability and dependability using M cold steady units. The design process for such power systems has been investigated and it has been found that this kind of design is capable of eliminating Single Point Failure (SPF), Common Mode Failure (CMF) and the greatest likelihood of human error [13]. **Dai et. al. (2003)** presented a model of a centralized heterogeneous distributed system in order to learn more about distributed system service availability and dependability. The model parameter's sensitivity was investigated. Conclusion has been made that the service reliability function can assist in appropriately allocating testing resources [14]. **Rauzy [2004]** described six approaches for calculating the time-dependent probability of Markov models. After a thorough investigation of techniques such full matrix exponentiation, Euler approach, Runge-Kutta method, and Adams-Bashford multi-steps methods of order 2 and 4, it was shown that computers nowadays could potentially manage Markov networks with millions of transitions [15].

**Marseguerra et. al. (2004)** studied multi-objective optimization, which takes into account parameter uncertainty and is primarily based on genetic algorithms. This method gives the decision-maker a tool to use in order to find a solution that is also optimal in terms of expected safety behavior and allows for a high degree of assurance in the actual system performance after applying the procedure to more complex systems [16]. **Gupta et. al. (2005)** utilized the mathematical formulation of the model to propose a numerical analysis of the process's availability and reliability in the bute-oil processing plant. After discussion, it may indicate that the proposed

technique is applicable to complicated systems that are also governed by substantial differential equations [17]. **Majeed and Sadiq (2006)** utilized the Markovian method to create a model for the Dokan hydro power station. The discussion and modeling of the issue led to the conclusion that power station reliability decreased annually. Conclusion has been made that a poor maintenance program and the inexperience of engineers and technicians affected availability adversely [18]. **Chuan ke and Kuangkhu (2007)** conducted a comparison of the availability of repairable redundant systems. Four bootstrap approaches were used to compute a comparison of confidence intervals for steady-state availability [19]. **Sharma and Kumar et. al. (2008)** presented the Markovian method of obtaining system behavior through the use of RAM analysis in crucial engineering systems. The transition diagram was used to create the differential equations. Based on the results, it has been advised to the managerial staff that characteristics like MTBF and MTTR are important for the system's planning and maintenance [20].

**Goyal et. al. (2009)** carried out an availability analysis of a part of Rubber tube production system under pre-emptive resume priority repair. The methodology used was Markov modelling. The purpose of the paper was to improve operational availability. Based on the results, the effect of failure and repair rates on availability was found. This information helps maintenance management improve the overall reliability and availability of the system [21]. **Adhikary et. al. (2010)** analyzed a coal-fired power plant's RAM. Before the data are fitted best with a probability distribution, a trend test and a serial correlation test are used to verify the distribution of failure and repair data. The significant subsystem has been identified through the use of Pareto analysis. The findings led to the conclusion that a rise in MTBF and a decrease in MTTR increase the power plant's availability [22]. **Vora et. al. (2011)** evaluated performance of turbo generator system of thermal plant using probabilistic approach. Markov approach has been used for problem formulation through transition diagram. Based on result availability graphs of failure and repair for maximum availability has been analyzed [23]. **Garg and Sharma (2012)** analyzed the performance of the synthesis unit in a fertilizer plant. The system's behavioral sensitivity has also been investigated. The Lamda Tau-Technique was used to investigate the behavior of a complex system that could be improved. Eight significant dependability parameters were also registered as fuzzy membership function [24]. **Wolde et. al. (2013)** discussed the issue of railway carrier inspections and maintenance. Using mathematical modelling, this study ties failure and repair rates to system performance. This modelling was used to evaluate inspection plans for any system, further optimizing its cost [25].

**Suleiman K et. al. (2013)** dealt with applying a probabilistic strategy to analyze stochastic data and evaluate thermal power plant performance. According to the findings of the analysis, availability decreases as the failure rate rises, while availability rises as the repair rate rises and vice versa. The plant management can use the result-based system for system availability analysis [26]. **Dewangan et. al. (2014)** investigated the reliability of thermal power plant's steam turbines. Investigation has been done based on failure database of five year. Failure modes and effect analysis (FMEA) used to categorize critical components. Based on investigation it has been concluded that well planning and regular scheduled maintenance can improve the reliability of plant [27]. **Aggarwal et. al. (2015)** proposed a performance model based on the Markov birth-death process for calculating RAM, dependability, MTBF and MTTR. Modeling has been done mathematically using Chapman-Kolmogorov differential equations and probabilistic considerations. Most critical subsystem pointed and suggested management to take utmost care [28] **Talebborouane et. al. (2016)** applied sophisticated fault tree and stochastic Petri Net formalisms to examine the availability of safety-critical systems. Generalized stochastic Petri Nets and fault tree driven Markov processes were utilized for analysis to get over the drawbacks of the Markov process and Petri Nets. It concluded that Petri Net is better for modeling as compared to

fault tree driven Markov process [29]. **Kumar and Tewari (2017)** utilized Particle Swarm Optimization (PSO) to optimize and analyse the performance of a beverage plant system. Exponential distribution is considered for repair and failure rate, and Markov approach is used for mathematical modelling. Results have been discussed with plant management for the improvement of system performance [30].

**Malik and Tewari (2018)** modeled and prioritized maintenance for a coal-fired thermal power plant's water flow system. Chapman-Kolmogorov equations were derived to obtain performance modelling using the Markov approach. The authors demonstrated the proposed approach to assisting in this kind of decision-making process through the case study [31]. **Singhal and Sharma (2018)** used the Markov process and generalized fuzzy numbers to analyse the availability of industrial systems. The uncertainty of data has been dealt with generalized fuzzy numbers. Availability analysis has been analysed through different arithmetic operations. The system analyst observed the impact of failure and repair rates on the system [32]. **Velmurugan et. al. (2019)** used Markov process to analyze the reliability, availability and maintainability of the forming industry. MATLAB software was used to solve mathematical functions. Based on the results, most critical subsystem was established. The best maintenance policy has also been provided to the maintenance manager for optimal maintenance [33]. **Elusakin and Shafiee (2020)** estimated the reliability of subsea blowout preventers using advanced analysis method stochastic Petri Nets with different failure modes. MTBF, availability and reliability terms obtained and analyzed. Sensitivity analysis was carried out to assess the impact that the redundancy design and fault coverage factor have on system performance. Based on the results, it defined that system availability and MTBF were significantly influenced by fault coverage and redundancy [34]. **Jagtap et. al. (2020)** optimized the availability of the boiler furnace system in coal-fired thermal power plant using Particle Swarm Optimization (PSO). The Markov method was used for the analysis of the system. Based on the results maintenance priority has been handed over to plant management [35].

**Maihulla et. al. (2021)** utilized RAMD (Reliability, Availability, Maintainability and Dependability) analysis to evaluate the efficiency of the complex system of reverse osmosis water purification equipment. The primary objective was to optimize the economy. The components were determined through sensitivity analysis. The RAM of a subsystem (the high-pressure pump) has a substantial effect on the system's overall availability, it was found after the discovery [36]. **Kumar et. al. (2021)** utilized the Petri Nets modeling method to examine the performance of a complex manufacturing system in order to influence the actual behavioral patterns of the many subsystems deployed in the plant. Subsystem that has been severely impacted by availability has been determined by the results [37]. **Parkash and Tewari (2022)** conducted modeling using the Markovian method and employed a probabilistic approach to design the Decision Support System (DSS) for assembly line maintenance. Probabilities for the steady state were determined using a transition diagram and by solving differential equations. The most important subsystem was found and subsystem maintenance priorities were finalized [38]. **Kumar and Tewari (2022)** evaluated performability features of ash handling system of a coal based thermal power plant using Petri Nets based techniques. Failure and Repair rate impact has been determined. Stochastic Petri Nets (SPN) applied for modeling. Based on the results, vital part of system has been identified. Petri Nets were found to reduce the time-consuming computational efforts required by Markov and other modeling methods while also ensuring better results [39]. **Behnamfar et. al. (2023)** presented a continuous Markov process-based reliability analysis of wireless power transfer for electric vehicle charging. To determine overall system reliability, five subsystems were individually analyzed on individual reliability. Based on the results, it was discovered that the system was highly reliable over a twenty-year lifespan, with 66.31% availability [40]. **Malik et. al.**

(2023) evaluated the performability of the veneer cutting system of the Plywood Plant using a stochastic approach, and Particle Swarm Optimization (PSO) was used for the optimization of the results. Based on the analysis, maintenance engineers get help optimizing overall maintenance costs and overall production costs [41].

### III. Research Gaps

In this section, the RAM approach has been used to discuss the brief findings of a literature survey conducted over the last three decades.

1. In this survey, it has been carried out that Researchers primarily focused on RAM approaches however very limited work is reported regarding RAMD (Reliability, Availability, Maintainability and Dependability) and RAMS (Reliability, Availability, Maintainability and Safety). Researchers missed the importance of effect of Dependability and Safety on Reliability. Safety is an important aspect of working in a safe environment; it also increases the motivation of team members to work in any hazardous environment, whereas dependability is also an important aspect or parameter which effected reliability in positive way by accomplishing its assigned mission or services.
2. Several researchers discussed their efforts to increase plant availability through the use of suitable maintenance procedures, policies, and different operational schedules. But very few researchers reported the relation between cost and maintenance policies with operations schedule. Factors which affected cost also need to be focused.
3. It is observed from the literature review that many techniques, including fault tree analysis, Markov models, and Lambda tau technology, have been applied. Each of these techniques has a variety of benefits and drawbacks. But there is a tool Markovian Petri Nets which can make good balance between modeling and decision making power. Application of this kind of tool is very limited in the literature survey.

### IV. Concluding Remarks

Detailed overview of the literature illustrates various RAM issues, tools and techniques applied in various plants and process industries. In literature survey authors majorly focused on maintenance plan, lowering maintenance cost, production costs and increasing performability and productivity etc. In order to further improve the plant's performability, various RAM tools and techniques can be utilized in both the design and the operational stages.

In order to ensure that the systems remain operational for an extended period of time, each plant is divided into a number of systems or subsystems for effective maintenance planning. Markov Analysis, Failure Mode and Effects Analysis (FMEA), Fault Tree Analysis, Reliability Growth Analysis, Fuzzy Model, Monte Carlo technique, Chapman Kolmogorov birth- death process, Stochastic Petri Nets, Particle Swarm Optimization (PSO) and other techniques were utilized for the analysis and modeling. The paper also discusses the advantages and disadvantages of each of these techniques.

**Acknowledgements:** The National Institute of Technology, Kurukshetra Director's outstanding assistance made it feasible to complete this work. The referees' suggestions and comments allowed the authors to improve the quality of the work, therefore they would like to express their sincere gratitude for that. I wish to thank my wife and parents for their love and assistance with this work during that time.

### Disclosure Statement

The authors declare that they have no conflict of interest.

### References

- [1] Ciardo, G., Marie, R. A., Sericola, B., and Trivedi, K.S., (1990). Performability Analysis Using Semi-Markov Reward Processes. *IEEE Transactions on Computers*, 1251–1264.
- [2] Viswanadham, N., Narahari, Y., and Ram, R., (1991). Performability of Automated Manufacturing Systems. *Control and Dynamic Systems*, 77–120.
- [3] Kumar, D., Singh, J., and Pandey, P.C., (1992). Availability of the Crystallization System in the Sugar Industry under Common-Cause Failure. *IEEE Transactions on Reliability*, 85–91.
- [4] Sharma, T.C., and Bazovsky, I., (1993). Reliability analysis of large system by Markov techniques. *Annual Reliability and Maintainability Symposium*, 260–267.
- [5] Behera, T.K., Mishra, B.S., Patnaik, L.M., and Girault, C., (1994). Modelling and performance evaluation of flexible manufacturing systems using deterministic and stochastic timed Petri nets. *IEE Conf. Publ.*, 362–368.
- [6] Murty, A.S.R., and Naikan, V.N.A., (1995). Availability and maintenance cost optimization of a production plant. *International Journal of Quality and Reliability Management*, 28–35.
- [7] Kumar, Sunand., Kumar, Dinesh., and Mehta, N.P., (1996). Behavioral analysis of shell gasification and carbon recovery process in a urea fertilizer plant. *Microelectronics Reliability*, 671-673.
- [8] Arora, Navneet., and Kumar, Dinesh., (1997). Availability analysis of steam and power generation systems in the thermal power plant. *Microelectronics Reliability*, 795-799.
- [9] Pellegrini, G. G., Catelani, M., and Iuculano, G., (1998). Measurement of the availability performance for electronic complex systems. *IEEE Instrumentation and Measurement Technology Conference, IEEE*, 51–54.
- [10] Singh, J., Mahajan, P. (1999). Reliability of Utensils Manufacturing Plant. *A Case Study, Operational Research Society of India*, 260-269.
- [11] Borgonovo, E., Marseguerra, M., and E. Zio, E., (2000). A Monte Carlo methodological approach to plant availability modeling with maintenance. *Reliability Engineering and System Safety*, 61-73.
- [12] Zhang, T., and Horigome, M., (2001). Availability and reliability of system with dependent components and time-varying failure and repair rates. *IEEE Transactions on Reliability*, 151-158.
- [13] Wang, Wendai., and Loman, James., (2002). Reliability/availability of K-out-of-N system with M cold standby units. *Annual Reliability and Maintainability Symposium*, 450-455.
- [14] Dai, Y.S., Xie, M., Poh, K.L., and Liu, G.Q., (2003). A study of service reliability and availability for distributed systems. *Reliability Engineering and System Safety*, 103-112
- [15] A. Rauzy., (2004) "An experimental study on iterative methods to compute transient solutions of large Markov models, 1-25.
- [16] Marseguerra, M., Zio, E., and Podofillini, L., (2004). Optimal reliability/availability of

uncertain systems via multi-objective genetic algorithms. *IEEE Transactions on Reliability*, 424–434.

[17] Gupta, P., Singh, J., and Singh, I.P., (2005). Mission Reliability and Availability Prediction of Flexible Polymer Powder Production System, *Operational Research Society of India*, 152-167.

[18] Majeed, A. R., and Sadiq, N. M., (2006). Availability & Reliability evaluation of Dokan hydro power station. *IEEE/PES Transmission & Distribution Conference and Exposition*, 1-6.

[19] Ke, J.C., and Chu, Y.K., (2007). Comparative analysis of availability for a redundant repairable system. *Applied Mathematics and Computation*, 332–338.

[20] Sharma, R. K., & Kumar, S., (2008). Performance modeling in critical engineering systems using RAM analysis. *Reliability Engineering & System Safety*, 913-919.

[21] Goyal, A., Sharma, S. K., and Gupta, P., (2009). Availability analysis of a part of rubber tube production system under preemptive resume priority repair. *International Journal of Industrial Engineering*, 260-269.

[22] Adhikary, D.D., Bose, G., Mitra, S., and Bose, D., (2010). Reliability, Maintainability & Availability analysis of a coal fired power plant in eastern region of India. *2nd International Conference on Production and Industrial Engineering*, 1505–1513.

[23] Vora, Y., Patel, M.B., and Tewari, P., (2011). Simulation Model for Stochastic Analysis and Performance Evaluation of Steam Generator System of a Thermal Power Plant. *International Journal of Engineering Science and Technology*, 5141-5149.

[24] H. Garg, H., and S. P. Sharma, S.P., (2012) Behavior analysis of synthesis unit in fertilizer plant. *International Journal of Quality and Reliability Management*, 217–232.

[25] Wolde, M. Ten., and Ghobbar, A.A., (2013). Optimizing inspection intervals - Reliability and availability in terms of a cost model: A case study on railway carriers. *Reliability Engineering and System Safety*, 137–147.

[26] Suleiman, K., Ali, U.A., and Yusuf, I., (2013) Stochastic Analysis and Performance Evaluation of a Complex Thermal Power Plant. *Innovative Systems Design and Engineering*, 21–32.

[27] Dewangan, D.N., Jha, M.K., and Banjare, Y.P., (2014) Reliability Investigation of Steam Turbine Used in Thermal Power Plant. *International Journal of Innovative Research in Science, Engineering and Technology*, 14915–14923.

[28] Aggarwal, A., Kumar, S., and Singh, V., (2015). Performance modeling of the skim milk powder production system of a dairy plant using RAMD analysis, *International Journal of Quality and Reliability Management*, 167–181.

[29] Talebberouane, M., Khan, F., and Lounis, Z., (2016). Availability analysis of safety critical systems using advanced fault tree and stochastic Petri net formalisms. *Journal of Loss Prevention in the Process Industries*, 193–203.

[30] Kumar, P., and Tewari, P.C., (2017) Performance analysis and optimization for CSDGB filling system of a beverage plant using particle swarm optimization. *International Journal of Industrial Engineering Computations*, 303–314.

[31] Malik, S., and Tewari, P.C., (2018) Performance modeling and maintenance priorities decision for the water flow system of a coal-based thermal power plant. *International Journal of Quality and Reliability Management*, 996–1010.

[32] Singhal, N., and Sharma, S.P., (2019). Availability Analysis of Industrial Systems Using Markov Process and Generalized Fuzzy Numbers. *Journal of Metrology Society of India*, 79–91.

[33] Velmurugan, K., Venkumar, P., & Sudhakarapandian, R. (2019). Reliability availability maintainability analysis in forming industry. *International Journal of Engineering and Advanced Technology*, 822-828.

[34] Elusakin, T., and Shafiee, M., (2020). Reliability analysis of subsea blowout preventers with condition-based maintenance using stochastic Petri nets. *Journal of Loss Prevention in the Process Industries*, 1-16.

- [35] Jagtap, H., Bewoor, A., Kumar, R., Ahmadi, M.H., and Lorenzini, G., (2020). Markov-based performance evaluation and availability optimization of the boiler–furnace system in coal-fired thermal power plant using PSO, *Energy Reports*, 1124–1134.
- [36] Maihulla, A.S., Yusuf, I., and Bala, S.I., (2021). Performance evaluation of a complex reverse osmosis machine system in water purification using reliability, availability, maintainability and dependability analysis, *Reliability: Theory & Application*, 115-131.
- [37] Kumar, A., Kumar, V., Modgil, V., Kumar, A., and Sharma, A., (2021). Performance Analysis of Complex Manufacturing System using Petri Nets Modeling Method. In *Journal of Physics: Conference Series*, IOP Publishing Ltd, 1-9.
- [38] Parkash, S., and Tewari, P.C., (2022). Performance modeling and dss for assembly line system of leaf spring manufacturing plant. *Reliability: Theory & Application*, 403–412.
- [39] Kumar, S., and Tewari, P.C., (2022) Performability Analysis of Multistate Ash Handling System of Thermal Power Plant With Hot Redundancy Using Stochastic Petrinets. *Reliability: Theory & Application*, 190–201.
- [40] Behnamfar, M., Taher, M.A., Polowsky, A., Roy, S., Tariq, M., and Sarwat, A., (2023) Reliability Analysis of Wireless Power Transfer for Electric Vehicle Charging based on Continuous Markov Process. *Fourth International Symposium on 3D Power Electronics Integration and Manufacturing (3D-PEIM)*, 1–5.
- [41] Malik, S., Kumar, N., and Kumar, S., (2023) Process Modeling and Numerical Investigation of Veneer Cutting System of a Plywood Plant With Stochastic Approach, *Reliability: Theory & Application*, 141–153.

# EXACT AND CONDITIONAL BOUNDS FOR GENERALIZED CUMULATIVE ENTROPY

ALEXEY V. LEBEDEV



Lomonosov Moscow State University, Russia  
avlebed@yandex.ru

## Abstract

*The differential entropy is a natural analog of the Shannon entropy for discrete distributions in respect to absolutely continuous distributions (with density). In modern studies, many other kinds of entropy have been introduced and analyzed, including various cumulative entropies, which are based not on the density but on the (cumulative) distribution function of random variable. Such characteristics can be used, for example, in computer vision, reliability theory, risk analysis, etc. We consider some generalizations of cumulative entropy, for a wide class of entropy generators. We use the methods of probability theory, calculus of variations and Cauchy–Bunyakovsky–Schwarz inequality. In the class of centered and normalized random variables, exact and conditional bounds are found as well as the distributions on which they are attained. By conditional bounds we understand bounds for one generalized cumulative entropy given the value of another entropy (in the class of random variables with zero mean and unit variance). This problem is analogous to the previously posed and partly solved problem on conditional bounds for expectations of sample maxima when we know the expected maximum of a sample of another size or expected maxima of two smaller samples.*

**Keywords:** cumulative entropy, exact bounds, conditional bounds, calculus of variations

## 1. INTRODUCTION

The *differential entropy* is a natural analog of the Shannon entropy for discrete distributions in respect to absolutely continuous distributions [14, 6]. For a random variable  $X$  with probability density function  $p(x)$ , it is given by

$$H(X) = - \int_{-\infty}^{+\infty} p(x) \ln p(x) dx.$$

For a given variance  $\sigma^2$ , the differential entropy attains its maximum on Gaussian distributions  $\mathcal{N}(\mu, \sigma^2)$  [14, §20]; then

$$H(X) = \frac{1 + \ln(2\pi\sigma^2)}{2}.$$

In modern studies, many other kinds of entropy have been introduced and analyzed, including various *cumulative entropies*, which are based not on the density but on the (cumulative) distribution function. Such characteristics can be used, for example, in computer vision [13], reliability theory and risk analysis [4, 5], etc. Even medical applications have been noted [1].

In [13], for nonnegative random variables there was introduced the *cumulative residual entropy* (CRE)

$$\mathcal{E}(X) = - \int_0^{+\infty} \bar{F}(x) \ln \bar{F}(x) dx,$$



where  $\bar{F}(x) = 1 - F(x)$ ,  $F$  being the (cumulative) distribution function (CDF) of a random variable  $X$ , and in [4] there was introduced the *cumulative entropy* (CE)

$$\mathcal{CE}(X) = - \int_0^{+\infty} F(x) \ln F(x) dx,$$

which was afterwards also called the *direct cumulative entropy* (in contrast to the residual one). In such expressions it is assumed that  $0 \ln 0 = 0$ .

It is clear that these functionals can be extended from nonnegative to arbitrary random variables by taking integrals over the entire axis:

$$\mathcal{E}(X) = - \int_{-\infty}^{+\infty} \bar{F}(x) \ln \bar{F}(x) dx, \quad \mathcal{CE}(X) = - \int_{-\infty}^{+\infty} F(x) \ln F(x) dx. \quad (1)$$

In the general case, the integrals may both converge or diverge. For these cumulative entropies, there is symmetry

$$\mathcal{E}(X) = \mathcal{CE}(-X). \quad (2)$$

Note that cumulative entropies (as well as the differential entropy) are traditionally written as numerical characteristics of a random variable  $X$ , though they actually depend on its distribution function  $F$  only.

In [3], representations for  $\mathcal{E}(X)$  and  $\mathcal{CE}(X)$  through moments of order statistics (using the power series expansion of the logarithm) have been obtained and upper bounds on these entropies were constructed assuming that  $X$  has mean  $\mu$  and variance  $\sigma^2$  (taking into account classical estimates for order statistics [7, 8]).

Namely, there were obtained the inequality [3, Theorem 1]

$$\mathcal{E}(X) \leq \sum_{n=1}^{+\infty} \frac{\sigma}{(n+1)\sqrt{2n+1}} \approx 1.21\sigma, \quad (3)$$

which is also valid for  $\mathcal{CE}(X)$  due to symmetry (2), and the inequality [3, Theorem 3]

$$\mathcal{E}(X) + \mathcal{CE}(X) \leq \sum_{n=1}^{+\infty} \frac{\sigma\sqrt{2}}{n\sqrt{n+1}} \approx 3.09\sigma. \quad (4)$$

Also, various classes of generalized cumulative entropies have been considered [9, 10].

In particular, in [9] there were introduced the *cumulative residual STM (Sharma–Taneja–Mittal) entropy*

$$SR_{\alpha,\beta}(X) = \frac{1}{\beta - \alpha} \int_0^{\infty} (\bar{F}^{\alpha}(x) - \bar{F}^{\beta}(x)) dx, \quad \alpha, \beta > 0, \quad \alpha \neq \beta,$$

and the *cumulative STM entropy*

$$SP_{\alpha,\beta}(X) = \frac{1}{\beta - \alpha} \int_0^{\infty} (F^{\alpha}(x) - F^{\beta}(x)) dx, \quad \alpha, \beta > 0, \quad \alpha \neq \beta.$$

Clearly, they can also be extended from nonnegative to arbitrary random variables:

$$\begin{aligned} SR_{\alpha,\beta}(X) &= \frac{1}{\beta - \alpha} \int_{-\infty}^{+\infty} (\bar{F}^{\alpha}(x) - \bar{F}^{\beta}(x)) dx, \\ SP_{\alpha,\beta}(X) &= \frac{1}{\beta - \alpha} \int_{-\infty}^{+\infty} (F^{\alpha}(x) - F^{\beta}(x)) dx, \\ &\alpha, \beta > 0, \quad \alpha \neq \beta. \end{aligned}$$

In [10], for a broad class of generalized cumulative entropies, optimal distributions (with given means and variances) that maximize these entropies (i.e., give their exact upper limits) have been obtained by methods of calculus of variations; however, the corresponding maximum values

of the entropies have not been derived. If they are derived, for example, for  $\mathcal{E}(X)$ ,  $\mathcal{CE}(X)$ , and  $\mathcal{E}(X) + \mathcal{CE}(X)$ , it turns out that these bounds are stronger than (3) and (4).

We will consider for simplicity the class of distributions with zero mean and unit variance. Following [10], one can easily deduce that the maximum value of  $\mathcal{E}(X)$  is 1 and the maximum is attained at the shifted exponential distribution with the CDF

$$F(x) = 1 - e^{-(x+1)}, \quad x \geq -1; \tag{5}$$

the maximum value of  $\mathcal{CE}(X)$  is the same, and it is attained at the distribution with the CDF

$$F(x) = e^{x-1}, \quad x \leq 1; \tag{6}$$

and the maximum value of  $\mathcal{E}(X) + \mathcal{CE}(X)$  is  $\pi/\sqrt{3} \approx 1.81$ , the maximum being attained at the logistic distribution with the CDF

$$F(x) = \frac{e^{\pi x/\sqrt{3}}}{e^{\pi x/\sqrt{3}} + 1}. \tag{7}$$

Next we formulate a simple statement that allows us to obtain an upper bound on the generalized cumulative entropy without deriving the corresponding optimal distribution; we will demonstrate it by an example of the cumulative residual STM-entropy.

Then we solve a new problem about the range in which one generalized cumulative entropy of a random variable can lie provided that another entropy of this random variable is known (for random variables with zero mean and unit variance). Besides the general theorem, we in detail analyze the case of the relationship of the entropies  $\mathcal{E}(X)$  and  $\mathcal{CE}(X)$ .

This problem is analogous to the previously posed and partly solved problem on conditional bounds for expectation of sample maxima when we know the expected maximum of a sample of another size [11] or the expected maxima of two smaller samples [12]. In this case, the corresponding characteristics are also expressed as integral functionals of the distribution function.

From the point of view of calculus of variations, the arising problems belong to the class of isoperimetric problems and are solved by the method of Lagrange multipliers (Euler–Lagrange equations).

## 2. MAIN RESULTS

Consider the class  $CN$  of centered and normalized random variables, i.e.,

$$CN = \{X : \mathbf{E}X = 0, \mathbf{Var}X = 1\}.$$

It is clear that for all the above-mentioned entropies, in order to establish bounds, it suffices to consider random variables in this class. Indeed, let a random variable  $X$  have mean  $\mu$  and variance  $\sigma^2$ ; then it admits a representation  $X = \mu + \sigma X_0$  with  $X_0 \in CN$ , and it follows from definition (1) that  $\mathcal{E}(X) = \sigma \mathcal{E}(X_0)$ , and so on.

Introduce a notation for the generalized inverse distribution function (also called the quantile function)

$$x(u) = \inf\{x : F(x) \geq u\}, \quad u \in [0, 1],$$

where  $F$  is the CDF of the random variable  $X$ . Then

$$X \stackrel{d}{=} x(U),$$

where  $U$  is uniformly distributed on  $[0, 1]$ , and the condition  $X \in CN$  is equivalent to the following constraints on  $x(u)$ :

$$\mathbf{E}X = \int_0^1 x(u) du = 0, \quad \mathbf{Var}X = \int_0^1 x^2(u) du = 1,$$

where the function  $x(u)$ ,  $u \in [0, 1]$ , is nondecreasing and right continuous.

We will consider functions  $g$  (entropy generators) satisfying the following conditions:

(\*)  $g(u)$  is a nonnegative continuous concave function on  $[0, 1]$  which is piecewise smooth on  $(0, 1)$ , with  $g(0) = g(1) = 0$ , and such that

$$G = \int_0^1 (g'(u))^2 du < \infty.$$

Introduce generalized cumulative entropies represented by the integral (if it converges)

$$\mathcal{E}_g(X) = \int_{-\infty}^{+\infty} g(\bar{F}(x)) dx, \tag{8}$$

where  $F$  is the CDF of the random variable  $X$ .

Using integration by parts and the change of variables  $u = F(x)$ , we can obtain the following representations:

$$\mathcal{E}_g(X) = - \int_{-\infty}^{+\infty} x dg(\bar{F}(x)) = \int_0^1 x(u)g'(1-u) du = \int_0^1 g(1-u) dx(u),$$

which gives a particular case of the generalized cumulative  $\Phi$ -entropy

$$CE_{\Phi}(F) = \int_0^1 \Phi(u) dx(u)$$

introduced in [10], with the only difference that in [10] it was not required that  $\Phi(0) = \Phi(1) = 0$  (though it was actually the case in all examples considered there).

Definition (8) also implies  $\mathcal{E}_g(X) = \sigma \mathcal{E}_g(X_0)$ ,  $X_0 = (X - \mu) / \sigma$ ,  $\sigma > 0$ .

**Proposition 1.** *Let  $g$  satisfy condition (\*); then*

$$\max_{X \in CN} \mathcal{E}_g(X) = \sqrt{G}.$$

The proposition follows from the fact that according to [10, Theorem 1] this maximum is attained at the distribution with the inverse CDF

$$x(u) = \frac{g'(1-u)}{\sqrt{G}}, \quad u \in [0, 1].$$

**Corollary 1.** Let  $1/2 < \min\{\alpha, \beta\} \leq 1$ ,  $\alpha \neq \beta$ ; then

$$\max_{X \in CN} SR_{\alpha, \beta} = \sqrt{\frac{2\alpha\beta - \alpha - \beta + 1}{(2\alpha - 1)(2\beta - 1)(\alpha + \beta - 1)}}, \tag{9}$$

and the maximum is attained at the distribution with the inverse CDF

$$x(u) = \frac{\alpha(1-u)^{\alpha-1} - \beta(1-u)^{\beta-1}}{(\beta - \alpha)\sqrt{G}}, \quad u \in [0, 1]. \tag{10}$$

In this case an optimal distribution  $F$  is not found explicitly, but it can be obtained, for example, for  $\alpha = 1$  or  $\beta = 1$ , when all expressions become simpler (this was made in [10]).

Note that for  $\min\{\alpha, \beta\} > 1$  the concavity condition for  $g$  is violated, and for  $0 < \min\{\alpha, \beta\} \leq 1/2$  the entropy  $SP_{\alpha, \beta}(X)$  may take infinitely large values on  $X \in CN$  (when the corresponding integrals diverge).

Clearly, analogous statements hold as well for  $SP_{\alpha, \beta}$ , since  $SP_{\alpha, \beta}(X) = SR_{\alpha, \beta}(-X)$ .

**Theorem 1.** Assume that  $g_1$  and  $g_2$  satisfy conditions (\*), the integrals

$$G_{ij} = \int_0^1 g'_i(u)g'_j(u) du, \quad 1 \leq i, j \leq 2,$$

are introduced, and it is known that  $\mathcal{E}_{g_2}(X) = t$ . Then for all  $X \in CN$  we have

$$\mathcal{E}_{g_1}(X) \leq \frac{1}{G_{22}} \left( G_{12}t + \sqrt{(G_{11}G_{22} - G_{12}^2)(G_{22} - t^2)} \right), \quad (11)$$

and this bound is tight if the function

$$\tilde{x}(u) = \lambda_1 g_1'(1-u) + \lambda_2 g_2'(1-u),$$

where

$$\lambda_1 = \sqrt{\frac{G_{22} - t^2}{G_{11}G_{22} - G_{12}^2}}, \quad \lambda_2 = \frac{t - \lambda_1 G_{12}}{G_{22}}, \quad (12)$$

is nondecreasing on  $(0, 1)$ ; then  $\tilde{x}(u)$  defines the distribution on which the bound is attained.

Note that by Proposition 1 we have  $\mathcal{E}_{g_2}^2(X) \leq G_{22}$ , so the radicand is always nonnegative. The functions  $g_1'(1-u)$  and  $g_2'(1-u)$  are nondecreasing, and  $\lambda_1 \geq 0$ ; however, the nondecreasing condition for  $\tilde{x}(u)$  can be violated when  $\lambda_2 < 0$ .

For the sequel, it would be convenient to introduce the notation for the constant

$$p = \frac{\pi^2}{6} - 1 \approx 0.645.$$

**Corollary 2.** For all  $X \in CN$  we have

$$\mathcal{E}(X) \leq p \mathcal{CE}(X) + \sqrt{(1-p^2)(1-\mathcal{CE}^2(X))}, \quad (13)$$

and this bound is tight if  $\mathcal{CE}(X) \geq p$ .

By symmetry (2) of the entropies, we also have

$$\mathcal{CE}(X) \leq p \mathcal{E}(X) + \sqrt{(1-p^2)(1-\mathcal{E}^2(X))},$$

and this bound is tight if  $\mathcal{E}(X) \geq p$ . By inverting the inequality, we can also obtain a lower bound

$$\mathcal{E}(X) \geq p \mathcal{CE}(X) - \sqrt{(1-p^2)(1-\mathcal{CE}^2(X))}$$

in the range  $\mathcal{CE}(X) \geq \sqrt{1-p^2} \approx 0.764$  where this bound is nonnegative (but we cannot claim that it is tight). Similarly, a lower bound for  $\mathcal{CE}(X)$  can be found.

The question of what is the upper bound when  $\tilde{x}(u)$  is not nondecreasing remains open. In this case we deal with a problem of not the calculus of variations but optimal control (with an additional condition  $x'(u) \geq 0$ ), which is much more complicated. One can also apply an approach to establishing (not tight) bounds using special families of distributions, as was done in [11]. This approach is exploited in the proof of the following theorem

**Theorem 2.** For any  $0 < t < p$  we have

$$\max_{X \in CN, \mathcal{CE}(X)=t} \mathcal{E}(X) \geq \sqrt{\frac{1-a}{1+a}} (1 - \ln(1-a)),$$

where  $a$  is a unique solution on  $(0, 1)$  of the equation<sup>1</sup>

$$-\frac{a(\ln a - 1) - \text{Li}_2(1-a) + 1}{\sqrt{1-a^2}} = t.$$

By symmetry (2), an analogous estimate holds for  $\mathcal{CE}(X)$  given  $\mathcal{E}(X)$ , whence one can obtain a lower estimate for the maximum of  $\mathcal{E}(X)$  given  $\mathcal{CE}(X)$ .

Figure 1 represents plots of the obtained bounds for the entropies  $\mathcal{E}(X)$  and  $\mathcal{CE}(X)$ . In bold, we highlight the interval where the bound (13) is tight; the dotted line shows the bound of Theorem 2. Points of the bound marked by the triangle, star, and circle correspond to the distributions (5), (6), and (7). In the ranges  $\mathcal{CE}(X) < p$  and  $\mathcal{E}(X) < p$ , true bounds lie somewhere in between the solid and dotted lines. Establishing them deserves further investigation.

<sup>1</sup>Here,  $\text{Li}_m(z) = \sum_{n=1}^{\infty} z^n / n^m$  is the polylogarithm of order  $m$ .

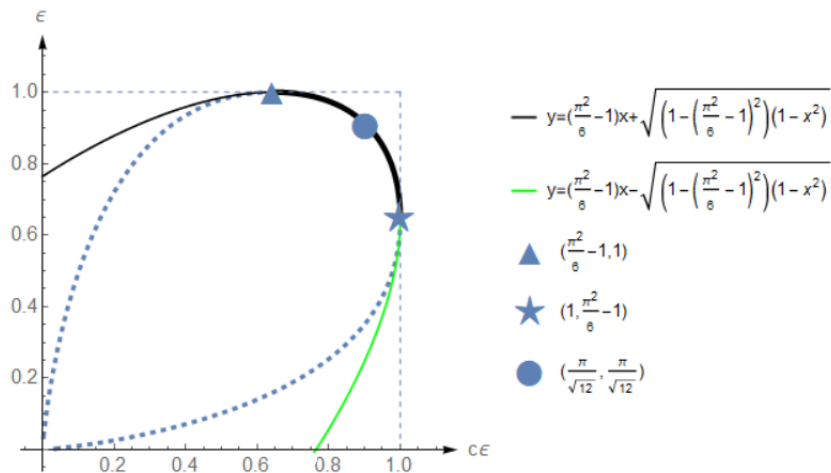


Figure 1: Plots of the bounds for the entropies  $\mathcal{E}(X)$  and  $C\mathcal{E}(X)$ .

### 3. PROOFS

**Proof of Corollary 1.** Let, for definiteness,  $\alpha < \beta$ ; then  $1/2 < \alpha \leq 1$ . Put  $g(u) = (u^\alpha - u^\beta) / (\beta - \alpha)$ ; then

$$g'(u) = \frac{\alpha u^{\alpha-1} - \beta u^{\beta-1}}{\beta - \alpha},$$

$$g''(u) = \frac{\alpha(\alpha - 1)u^{\alpha-2} - \beta(\beta - 1)u^{\beta-2}}{\beta - \alpha} < 0, \quad u \in (0, 1).$$

We obtain

$$\begin{aligned} G &= \int_0^1 \left( \frac{\alpha u^{\alpha-1} - \beta u^{\beta-1}}{\beta - \alpha} \right)^2 du \\ &= \frac{1}{(\beta - \alpha)^2} \int_0^1 (\alpha^2 u^{2\alpha-2} - 2\alpha\beta u^{\alpha+\beta-2} + \beta^2 u^{2\beta-2}) du \\ &= \frac{1}{(\beta - \alpha)^2} \left\{ \frac{\alpha^2}{2\alpha - 1} - \frac{2\alpha\beta}{\alpha + \beta - 1} + \frac{\beta^2}{2\beta - 1} \right\} \\ &= \frac{2\alpha\beta - \alpha - \beta + 1}{(2\alpha - 1)(2\beta - 1)(\alpha + \beta - 1)} \end{aligned} \tag{14}$$

and equations (9) and (10).

**Proof of Theorem 1.** By considering the Lagrangian

$$\mathcal{L} = \int_0^1 (\lambda_1 x(u)g'_1(1-u) + \lambda_2 x(u)g'_2(1-u) + \lambda_3 x(u) + \lambda_4 x^2(u)) du,$$

we obtain the Euler-Lagrange equation

$$\lambda_1 g'_1(1-u) + \lambda_2 g'_2(1-u) + \lambda_3 + 2\lambda_4 x(u) = 0,$$

where we may without loss of generality take  $\lambda_4 = -1/2$ .

Thus, we will seek for a function

$$\tilde{x}(u) = \lambda_1 g'_1(1-u) + \lambda_2 g'_2(1-u) + \lambda_3$$

satisfying the conditions

$$\int_0^1 \tilde{x}(u) du = 0, \quad \int_0^1 \tilde{x}^2(u) du = 1, \quad \int_0^1 \tilde{x}(u)g'_2(1-u) du = t.$$

The first condition, taking into account that  $g_i(0) = g_i(1) = 0$ ,  $i = 1, 2$ , gives  $\lambda_3 = 0$ ; the second and third yield a system of equations

$$\begin{cases} G_{11}\lambda_1^2 + 2G_{12}\lambda_1\lambda_2 + G_{22}\lambda_2^2 = 1, \\ G_{12}\lambda_1 + G_{22}\lambda_2 = t; \end{cases}$$

by solving this system with respect to  $\lambda_1$  and  $\lambda_2$ , we obtain (12).

Next, for any function  $x(u)$  corresponding to  $X \in CN$ , by the Cauchy-Bunyakovsky-Schwarz inequality we obtain

$$\int_0^1 x(u)\bar{x}(u) du = \lambda_1 \mathcal{E}_{g_1}(X) + \lambda_2 t \leq \left( \int_0^1 x^2(u) du \right)^{1/2} \left( \int_0^1 \bar{x}^2(u) du \right)^{1/2} = 1, \quad (15)$$

whence

$$\begin{aligned} \mathcal{E}_{g_1}(X) &\leq \frac{1 - \lambda_2 t}{\lambda_1} = \frac{G_{22} - (t - \lambda_1 G_{12})t}{\lambda_1 G_{22}} = \frac{\lambda_1 G_{12} t + G_{22} - t^2}{\lambda_1 G_{22}} \\ &= \frac{1}{G_{22}} \left( G_{12} t + \sqrt{(G_{11} G_{22} - G_{12}^2)(G_{22} - t^2)} \right). \end{aligned}$$

If  $\bar{x}(u)$  is nondecreasing and thus corresponds to some distribution, then with  $x(u) = \bar{x}(u)$  inequality (15) turns into equality, and the bound is attained.

**Proof of Corollary 2.** We apply Theorem 1 in the case of  $g_1(u) = -u \ln u$  and  $g_2(u) = -(1-u) \ln(1-u)$ ; then, as we have already obtained,  $G_{11} = G_{22} = 1$ , and we find

$$G_{12} = - \int_0^1 (\ln u + 1)(\ln(1-u) + 1) du = p;$$

plugging this into (11), we obtain (13). In this case we have

$$\bar{x}(u) = \lambda_1(-(\ln(1-u) + 1)) + \lambda_2(\ln u + 1),$$

where

$$\lambda_1 = \sqrt{\frac{1-t^2}{1-p^2}}, \quad \lambda_2 = t - p\lambda_1.$$

A necessary and sufficient condition for  $\bar{x}(u)$  to be nondecreasing on  $(0, 1)$  is  $\lambda_2 \geq 0$ , which happens to be equivalent to the inequality  $t \geq p$ .

**Proof of Theorem 2.** Consider a family of random variables  $X_a^0$ ,  $a \in [0, 1)$ , whose distribution is a mixture of zero (with probability  $a$ ) and the standard exponential distribution (with probability  $1-a$ ). Then the inverse CDFs take the form

$$x_a^0(u) = \begin{cases} 0, & 0 \leq u < a; \\ -\ln \frac{1-u}{1-a}, & a \leq u < 1. \end{cases}$$

We have

$$\mathbf{E}X_a^0 = 1 - a, \quad \mathbf{E}(X_a^0)^2 = 2(1 - a), \quad \mathbf{Var}X_a^0 = 2(1 - a) - (1 - a)^2 = 1 - a^2.$$

Put

$$X_a = \frac{X_a^0 - \mathbf{E}X_a^0}{\sqrt{\mathbf{Var}X_a^0}}.$$

Then  $X_a \in CN$ ,  $a \in [0, 1)$ ;  $X_0$  has distribution (5); and  $X_a \xrightarrow{d} 0$  as  $a \rightarrow 1 - 0$ .

Compute the corresponding entropies for  $0 < a < 1$ :

$$\begin{aligned} \mathcal{E}(X_a) &= \frac{\mathcal{E}(X_a^0)}{\sqrt{1-a^2}} = -\frac{1}{\sqrt{1-a^2}} \int_a^1 \ln \frac{1-u}{1-a} (\ln(1-u) + 1) du \\ &= \frac{(1-a)(1-\ln(1-a))}{\sqrt{1-a^2}} = \sqrt{\frac{1-a}{1+a}} (1-\ln(1-a)), \\ \mathcal{CE}(X_a) &= \frac{\mathcal{CE}(X_a^0)}{\sqrt{1-a^2}} = \frac{1}{\sqrt{1-a^2}} \int_a^1 \ln \frac{1-u}{1-a} (\ln u + 1) du \\ &= -\frac{a(\ln a - 1) - \text{Li}_2(1-a) + 1}{\sqrt{1-a^2}}, \end{aligned}$$

and  $\mathcal{CE}(X_a)$  strictly decreases in the interval  $0 < a < 1$ .

Thus, from the values of the entropies on the family  $X_a, a \in (0, 1)$ , we can obtain the estimate of Theorem 2.

## REFERENCES

- [1] Ahmadini, A. A. H., Hassan, A. S., Zaky, A. N., Alshqaq, S. S. (2020) Bayesian inference of dynamic cumulative residual entropy from Pareto II distribution with application to COVID-19. *AIMS Mathematics*. 6(3):2196–2216.
- [2] Balakrishnan, N., Bendre, S. M. (1993) Improved bounds for expectations of linear functions of order statistics. *Statistics*. 24(2):161–165.
- [3] Balakrishnan, N., Buono, F., Longobardi, M. (2022) On cumulative entropies in terms of moments of order statistics. *Methodol. Comput. Appl. Probab.* 24:345–359.
- [4] Di Crescenzo, A. and Longobardi, M. (2009) On cumulative entropies. *J. Stat. Plann. Inference*. 139:4072–4087.
- [5] Di Crescenzo, A. and Longobardi, M. (2013) Stochastic comparisons of cumulative entropies. In: Li, H., Li, X. (eds) *Stochastic Orders in Reliability and Risk*. Lecture Notes in Statistics, vol. 208. Springer, New York, NY. 167–182.
- [6] Gelfand, I. M., Kolmogorov, A. N., Yaglom A. M. (1993) Amount of information and entropy for continuous distributions. In: Shiriyayev, A.N. (eds) *Selected Works of A. N. Kolmogorov*. Mathematics and Its Applications, vol. 27. Springer, Dordrecht. 33–56.
- [7] Gumbel, E. J. (1954) The maxima of the mean largest value and of the range. *Ann. Math. Stat.* 25(1):76–84.
- [8] Hartley, H. O. and David, H. A. (1954) Universal bounds for mean range and extreme observation. *Ann. Math. Stat.* 25(1):85–99.
- [9] Kattumannil, S.K., Sreedevi, E. P., Balakrishnan, N. (2022) A generalized measure of cumulative residual entropy. *Entropy*. 24, Art. 444.
- [10] Klein, I. and Doll, M. (2020) (Generalized) maximum cumulative direct, residual, and paired  $\Phi$  entropy approach. *Entropy*. 22, Art. 91.
- [11] Ivanov, D. V. (2019) Conditional bounds of expected maxima of random variables and their reachability. *Systems and Means of Informatics*. 29(1):140–163. (in Russian)
- [12] Ivanov, D. V. (2023) On the bounds for the expected maxima of random samples with known expected maxima of two samples of smaller size. *Theory Probab. Appl.* 68(1):2–15.
- [13] Rao, M., Chen, Y., Vemuri, B., Wang, F. (2004) Cumulative residual entropy: A new measure of information. *IEEE Trans. Inf. Theory*. 50:1220–1228.
- [14] Shannon, C. E. (1948) A mathematical theory of communication. *Bell. Syst. Tech. J.* 27:379–423.

# SOME PROPERTIES OF TSALLIS ENTROPY BASED ON A DOUBLY TRUNCATED (INTERVAL) RANDOM VARIABLE

S. Jalayeri<sup>a</sup>, G.R. Mohtashami Borzadaran<sup>a\*</sup>, M. Khorashadizadeh<sup>b</sup>

•

<sup>a</sup> Ferdowsi University of Mashhad, Iran  
samirajalayeri@yahoo.com

<sup>a\*</sup> Ferdowsi University of Mashhad, Iran  
grmohtashami@um.ac.ir

<sup>b</sup> University of Birjand, Iran  
m.khorashadizadeh@birjand.ac.ir

## Abstract

*In this paper, we first study doubly truncated (interval) Tsallis entropy and suggest doubly truncated (interval) cumulative residual Tsallis entropy (ICRT), which is an extension of cumulative residual Tsallis entropy (CRT) and the dynamic CRT defined by the aid of Sati and Gupta and of Kumar, respectively. We investigate some properties and characterization of this measure, such as its relation with doubly truncated Shannon entropy, mean residual (past) life, and hazard rate (or reversed hazard rate). Also, the twin measure, doubly truncated (interval) cumulative past Tsallis entropy, is determined, and some of its properties are studied. Moreover, their monotonicity and related aging classes of distributions are expressed, and the upper (lower) bound for them is acquired. In the end, we propose four nonparametric estimators and compare their performance by utilizing simulation data. Also, being based on the best-proposed estimator, a real data set is additionally examined.*

**Keywords:** Doubly truncated (interval) Tsallis entropy, Doubly truncated (interval) cumulative residual Tsallis entropy (ICRT), Doubly truncated (interval) cumulative past Tsallis entropy (ICPT), Hazard rate, Reversed hazard rate, Mean residual life, Mean past life, Nonparametric estimators

## 1. INTRODUCTION

The notion of entropy, later generalized to information theory and statistical mechanics, was initially created by physicists in the area of equilibrium thermodynamics. The most famous one is due to [22], that plays an essential role in measuring the average uncertainty of a random variable. Entropy plays an important role in measuring the index of dispersion, volatility, or uncertainty related to a random variable  $X$ . Here and during this paper,  $X$  is an absolutely continuous nonnegative random variable, with probability density function (pdf)  $f(x)$  and survival function  $\bar{F}(x) = P(X > x)$ . Then the average amount of uncertainty associated with the random variable  $X$  as given by Shannon entropy, is

$$H(X) = - \int_0^{\infty} f(x) \ln f(x) dx.$$

Although, in certain situations, the Shannon entropy is not suitable where some generalized forms are of importance. Several generalized entropy measures are accessible in literature, which



have many huge properties consisting of smoothness, big dynamic range with respect to certain conditions, and many others, which lead them to greater flexibility in practice. One prevalent generalization is the Tsallis entropy, introduced by [24], determined as a generalization of the Boltzmann–Gibbs entropy. Inside the studying of statistical mechanics, Tsallis entropy gives a much broader view of how disorder emerges in macroscopic systems. For a continuous nonnegative random variable  $X$ , Tsallis entropy is determined as

$$T^\alpha(X) = \frac{1}{\alpha - 1} \left( 1 - \int_0^\infty f(x)^\alpha dx \right), \tag{1}$$

where  $0 < \alpha \neq 1$ . Clearly, when  $\alpha \rightarrow 1$ , we have  $T^\alpha(X) \rightarrow H(X)$ . Tsallis exploited its nonextensive features, and it has more and more extensive applications in science and technology. This entropy measure is extra flexible because of the parameter  $\alpha$ , and it increases the scope of application. Tsallis entropy preserves many significant characteristics of Shannon entropy except for the additivity property. From the years 2000 on, an increasingly wide spectrum of natural, artificial, and socially complicated systems were identified that verify the predictions and conclusions derived from this nonadditive entropy. Extensive or nonextensive statistical mechanics derive from the additivity or nonadditivity of the corresponding entropy measures. The Tsallis entropy is broadly utilized in physics to examine the distribution characterizing the movement of cold atoms in dissipative optical lattices [9] and signal processing [23]. More properties and applications of Tsallis entropy have been mentioned in [24, 25]. Considering the measures based on residual lifetime random variable,  $X_t = (X - t | X \geq t)$  has an essential role in many grounds, including reliability theory, survival analysis, and information theory. So, [10, 6] defined the residual Tsallis entropy (RT) based on the random variable  $X_t$  by

$$RT(X; t) = \frac{1}{\alpha - 1} \left( 1 - \int_t^\infty \left( \frac{f(x)}{\bar{F}(t)} \right)^\alpha dx \right).$$

The expected uncertainty involved in the remaining lifetime of a component is measured basically by RT. It is clear that  $RT(X; 0) = T^\alpha(X)$ . Lately, [10, 4] introduced an entropy-based measure of uncertainty in past lifetime distributions and denominated it past Tsallis entropy (PT). The uncertainty of the idle time of a component or system that is based on past lifetime random variable  $X_t^* = (t - X | X \leq t)$  is indicated by PT, and it is given by

$$PT(X; t) = \frac{1}{\alpha - 1} \left( 1 - \int_0^t \left( \frac{f(x)}{\bar{F}(t)} \right)^\alpha dx \right),$$

and also,  $PT(X; \infty) = T^\alpha(X)$ .

Currently, many researchers advanced new measures of uncertainty to overcome the limitations of traditional entropy measures and increase the applicability of information measures in diverse areas of science and engineering. With this motivation, [18] studied an alternative to Shannon differential entropy. The cumulative residual entropy (CRE) is obtained by replacing the pdf  $f(x)$  in  $H(X)$  with the survival function  $\bar{F}(x) = P(X > x)$ , given by  $H(X) = -\int_0^\infty \bar{F}(x) \ln \bar{F}(x) dx$ . The CRE is regarded to be greater stable due to the fact that the distribution function is greater regular than the pdf, and it owns more mathematical properties and special applications. Also, it is easily computable, always nonnegative, and its definition is valid in both the continuous and discrete cases. Additionally, the distribution exists despite the fact that the pdf does not.

In information theory, numerous attempts have been made by researchers, and an eminent amount of work has been done from both theoretical and application points of view for studying and extending the notion of CRE. Motivated by the extensive applicability of  $H(X)$ , a cumulative version of (1) studied by [19], is determined as the cumulative Tsallis entropy (CRT)

$$CRT(X) = \frac{1}{\alpha - 1} \left( 1 - \int_0^\infty \bar{F}(x)^\alpha dx \right).$$

Although [19] denoted that  $CRT(X)$  tends to  $CRE(X)$  when  $\alpha \rightarrow 1$ , where  $CRE(X) = - \int_0^\infty \bar{F}(x) \ln(\bar{F}(x)) dx$ , defined by [18], but [16] showed with a counter example that it is not true. The cumulative past Tsallis entropy (CPT) has also been introduced and studied by [16] as follows:

$$CPT(X) = \frac{1}{\alpha - 1} \left( 1 - \int_0^\infty F(x)^\alpha dx \right).$$

[19] gave the dynamic version of cumulative residual Tsallis entropy (DCRT), which is the CRT of the residual random variable  $X_t$  and it is given by

$$DCRT(X; t) = \frac{1}{\alpha - 1} \left( 1 - \int_t^\infty \left( \frac{\bar{F}(x)}{\bar{F}(t)} \right)^\alpha dx \right),$$

and  $DCRT(X; 0) = CRT(X)$ . Furthermore, [8] studied many properties of DCRT, and [16] introduced the dynamic version of cumulative past Tsallis entropy (DCPT) by

$$DCPT(X; t) = \frac{1}{\alpha - 1} \left( 1 - \int_0^t \left( \frac{F(x)}{F(t)} \right)^\alpha dx \right),$$

and  $DCPT(X; \infty) = CPT(X)$ . Occasionally, in many conditions, we just possess information between two points. Thus, we have to look at the statistical measures (particularly in information theory and reliability) under the case of doubly truncated random variables. For instance, in reliability, if  $X$  indicates the lifetime of a unit, then the random variable  $X_{t_1, t_2} = (X - t_1 | t_1 \leq X \leq t_2)$  is known as the doubly truncated residual lifetime. Note that the well-known random variable,  $X_t = (X - t | X \geq t)$ , is the particular case of  $X_{t_1, t_2}$  when  $t_2$  tends to  $\infty$ . Also, doubly truncated past lifetime is the random variable  $X_{t_1, t_2}^* = (t_2 - X | t_1 \leq X \leq t_2)$ , which in the specific case when  $t_1 = 0$ , it is the past lifetime random variable  $X_t^*$ . Another generalization of Tsallis entropy is based on a doubly truncated (interval) random variable [13], which reads as follows:

$$T^\alpha(X; t_1, t_2) = \frac{1}{\alpha - 1} \left( 1 - \int_{t_1}^{t_2} \left( \frac{f(x)}{F(t_2) - F(t_1)} \right)^\alpha dx \right), \tag{2}$$

where  $(t_1, t_2) \in D = \{(t_1, t_2) : F(t_1) < F(t_2)\}$  and  $T^\alpha(X; 0, \infty)$  is the Tsallis entropy  $T^\alpha(X)$ , and  $T^\alpha(X; t_1, \infty)$  is the residual entropy  $RT(X; t_1)$  and also  $T^\alpha(X; 0, t_2)$  is the past entropy  $PT(X; t_2)$ . Also, when  $\alpha \rightarrow 1$ , we have  $T^\alpha(X; t_1, t_2) \rightarrow H(X; t_1, t_2) = - \int_{t_1}^{t_2} \frac{f(x)}{F(t_2) - F(t_1)} \ln \left( \frac{f(x)}{F(t_2) - F(t_1)} \right) dx$ .

The distribution function estimation is not only an interesting problem by itself, but also it emerges naturally in actual problems of many scientific fields, consisting of seismology, hydrology, environmental sciences, and so on. Currently, in those disciplines, numerous methodologies have appeared for attacking statistical problems based on nonparametric ideas. With this motivation, the performance of four nonparametric estimators of ICPT is compared, and also a real-life data set is illustrated based on the best-proposed estimator.

In this paper, some properties of  $T^\alpha(X; t_1, t_2)$  are introduced. Additionally, we discuss the doubly truncated (interval) cumulative residual Tsallis entropy (ICRT) and doubly truncated (interval) cumulative past Tsallis entropy (ICPT), which can be general forms of the preceding findings. Some properties of ICRT and ICPT and their relationships with reliability measures, including hazard rate (or reversed hazard rate) and mean residual life (or mean past life), are studied. Finally, we consider four empirical and kernel-based estimators. Then, by using simulated data, we compare the behavior of the proposed estimators. In addition, a real data set from environmental monitoring is studied.

## 2. DOUBLY TRUNCATED TSALLIS ENTROPY

In this section, we express some properties and characterization results of  $T^\alpha(X; t_1, t_2)$ . First, for the  $T^\alpha(X; t_1, t_2)$ , an upper interval is acquired with respect to  $t_2$ , for any fixed  $t_1$ , in the next theorem. [13] proved a result similar to the following theorem, with respect to  $t_1$ , for any fixed

$t_2$ . Also, it should be noted that [11] introduced the generalized failure rate (GFR) based on the doubly truncated random variables by

$$h_1(t_1, t_2) = \lim_{h \rightarrow 0^+} \left[ \frac{P(t_1 \leq x \leq t_1 + h | t_1 \leq x \leq t_2)}{h} \right] = \frac{f(t_1)}{F(t_2) - F(t_1)} \tag{3}$$

and

$$h_2(t_1, t_2) = \lim_{h \rightarrow 0^-} \left[ \frac{P(t_2 \leq x \leq t_2 + h | t_1 \leq x \leq t_2)}{h} \right] = \frac{f(t_2)}{F(t_2) - F(t_1)}, \tag{4}$$

where their relationships with  $m(t_1, t_2) = E(X | t_1 \leq X \leq t_2) = \int_{t_1}^{t_2} x \frac{f(x)}{F(t_2) - F(t_1)} dx$  for  $(t_1, t_2) \in D$  are as follows:

$$h_1(t_1, t_2) = \frac{\frac{\partial m(t_1, t_2)}{\partial t_1}}{m(t_1, t_2) - t_1}, \tag{5}$$

$$h_2(t_1, t_2) = \frac{\frac{\partial m(t_1, t_2)}{\partial t_2}}{t_2 - m(t_1, t_2)}. \tag{6}$$

A lower (upper) bound for the  $ICRT(X; t_1, t_2)$  when increasing the ICRT property is acquired in the next theorem, for  $0 < \alpha < 1 (\alpha > 1)$ .

**Theorem 1.** The random variable  $X$  has increasing doubly truncated (interval) Tsallis entropy property if and only if the following inequalities are satisfied for all  $(t_1, t_2) \in D$  and  $0 < \alpha < 1 (\alpha > 1)$ :

$$\frac{1}{\alpha - 1} \left( 1 - \frac{1}{\alpha} \left( \frac{\frac{\partial m(t_1, t_2)}{\partial t_2}}{t_2 - m(t_1, t_2)} \right)^{\alpha - 1} \right) \leq (\geq) T^\alpha(X; t_1, t_2).$$

**Proof.** By differentiating  $T^\alpha(X; t_1, t_2)$  of the form (2) with respect to  $t_2$ , we have

$$\begin{aligned} \frac{\partial T^\alpha(X; t_1, t_2)}{\partial t_2} &= \frac{-1}{\alpha - 1} \left( \left( \frac{f(t_2)}{F(t_2) - F(t_1)} \right)^\alpha - \alpha \frac{f(t_2)}{F(t_2) - F(t_1)} \int_{t_1}^{t_2} \left( \frac{f(x)}{F(t_2) - F(t_1)} \right)^\alpha dx \right) \\ &= \frac{-1}{\alpha - 1} h_2^\alpha(t_1, t_2) + \frac{\alpha}{\alpha - 1} h_2(t_1, t_2) (1 - (\alpha - 1) T^\alpha(X; t_1, t_2)) \\ &= h_2(t_1, t_2) \frac{-1}{\alpha - 1} h_2^{\alpha - 1}(t_1, t_2) + \frac{\alpha}{\alpha - 1} (1 - (\alpha - 1) T^\alpha(X; t_1, t_2)). \end{aligned}$$

So, after suitable substitution of equation (6) and simplifying the equation we have,

$$T^\alpha(X; t_1, t_2) \leq (\geq) \frac{1}{\alpha - 1} \left( 1 - \frac{1}{\alpha} (h_2(t_1, t_2))^{\alpha - 1} \right),$$

the proof is complete. ■

We study the effect of increasing transformation on  $T^\alpha(Y; t_1, t_2)$ .

**Lemma 1.** Let  $X$  be a nonnegative continuous random variable with cumulative distribution function (cdf)  $F$ , and take  $Y = \phi(X)$ , where  $\phi(\cdot)$  is a strictly increasing differentiable function. Then

$$T^\alpha(Y; t_1, t_2) = \frac{1}{\alpha - 1} \left( 1 - \int_{\max\{0, \phi^{-1}(t_1)\}}^{\phi^{-1}(t_2)} \left( \frac{f(x)}{F(\phi^{-1}(t_2)) - F(\phi^{-1}(t_1))} \right)^\alpha \frac{1}{(\phi'(x))^{\alpha - 1}} dx \right).$$

If  $Z = aX + b$ , with  $a > 0$  and  $b \geq 0$ , so  $F_{aX+b}(z) = F_X\left(\frac{z-b}{a}\right)$ , then

$$T^\alpha(Z; t_1, t_2) = \frac{a^{\alpha - 1} - 1}{a^{\alpha - 1}(\alpha - 1)} + \left( \frac{a^{\alpha - 1} - 1}{a^{\alpha - 1}} \right) T^\alpha\left(X; \frac{t_1 - b}{a}, \frac{t_2 - b}{a}\right).$$

There are an identity and inequalities for doubly truncated (interval) Tsallis entropy based on the assumptions of the following proposition.

**Proposition 1.** Let  $X$  be a random variable with support in  $[0, r]$  where  $r > 0$  and symmetric with respect to  $\frac{r}{2}$ ; that is,  $F(x) = \bar{F}(r - x)$  for  $0 \leq x \leq r$ . Then

$$T^\alpha(X; t_1, t_2) = T^\alpha(X; r - t_1, r - t_2); \quad 0 \leq t_1, t_2 \leq r.$$

**Proof.** We have

$$\begin{aligned} T^\alpha(X; t_1, t_2) &= \frac{1}{\alpha - 1} \left( 1 - \int_{t_1}^{t_2} \left( \frac{f(x)}{F(t_2) - F(t_1)} \right)^\alpha dx \right) \\ &= \frac{1}{\alpha - 1} \left( 1 - \int_{t_1}^{t_2} \left( \frac{f(r - x)}{\bar{F}(r - t_2) - \bar{F}(r - t_1)} \right)^\alpha dx \right) \\ &= -\frac{1}{\alpha - 1} \left( 1 - \int_{r-t_1}^{r-t_2} \left( \frac{f(y)}{F(r - t_1) - F(r - t_2)} \right)^\alpha dy \right) \\ &= \frac{1}{\alpha - 1} \left( 1 - \int_{r-t_2}^{r-t_1} \left( \frac{f(y)}{F(r - t_1) - F(r - t_2)} \right)^\alpha dy \right) \\ &= T^\alpha(X; r - t_1, r - t_2). \end{aligned}$$

■

**Example 1.** If  $X$  is uniformly distributed in  $[0, r]$ , then for  $0 \leq t_1, t_2 \leq r$ , we have  $T^\alpha(X; t_1, t_2) = \frac{1}{\alpha - 1} (t_2 - t_1)^{1 - \alpha}$ , which is in agreement with Proposition 1.

**Proposition 2.** Let  $X$  be a nonnegative and absolutely continuous random variable. Then for  $\alpha > 1$  ( $0 < \alpha < 1$ ), we have

$$1 - (t_2 - t_1) \leq T^\alpha(X; t_1, t_2) \leq (t_2 - t_1) - 1. \tag{7}$$

**Proof.** The upper bound and lower bound given in (7) can be obtained from the well-known inequality  $\ln x \leq x - 1$ , where  $x > 0$ . Let  $x = \frac{f(x)}{F(t_2) - F(t_1)}$ . Then  $x^{\alpha - 1} > 0$  for  $\alpha > 1$  ( $0 < \alpha < 1$ ), and by using  $H(X; t_1, t_2) \leq (t_2 - t_1) - 1$  [15], the proof is complete. ■

**Proposition 3.** Let  $X$  be a nonnegative and absolutely continuous random variable with cdf  $F(x)$  and pdf  $f(x)$ . If  $f(x)$  is decreasing in  $x$ , then for  $0 < \alpha < 1$  ( $\alpha > 1$ ),

$$\frac{1 - h_1^\alpha(t_1, t_2)(t_2 - t_1)}{(\alpha - 1)} \geq (\leq) T^\alpha(X; t_1, t_2) \geq (\leq) \frac{1 - h_2^\alpha(t_1, t_2)(t_2 - t_1)}{(\alpha - 1)},$$

where  $h_1(t_1, t_2)$  and  $h_2(t_1, t_2)$  are defined in (3) and (4).

**Proof.** Let  $f(x)$  be decreasing in  $x$ . Then for  $t_1 \leq x \leq t_2$ , we have

$$\frac{f(t_1)}{F(t_2) - F(t_1)} \geq \frac{f(x)}{F(t_2) - F(t_1)} \geq \frac{f(t_2)}{F(t_2) - F(t_1)}.$$

So,

$$\int_{t_1}^{t_2} \left( \frac{f(t_1)}{F(t_2) - F(t_1)} \right)^\alpha dx \geq \int_{t_1}^{t_2} \left( \frac{f(x)}{F(t_2) - F(t_1)} \right)^\alpha dx \geq \int_{t_1}^{t_2} \left( \frac{f(t_2)}{F(t_2) - F(t_1)} \right)^\alpha dx.$$

Then

$$1 - h_1^\alpha(t_1, t_2)(t_2 - t_1) \leq 1 - \int_{t_1}^{t_2} \left( \frac{f(x)}{F(t_2) - F(t_1)} \right)^\alpha dx \leq 1 - h_2^\alpha(t_1, t_2)(t_2 - t_1).$$

Thus for  $0 < \alpha < 1$  ( $\alpha > 1$ ), after some calculations, the proof is complete. ■

**Example 2.** Let  $X$  be a nonnegative and absolutely continuous random variable with cdf  $F(x) = 1 - e^{-x}$  and pdf  $f(x) = e^{-x}$ . Then,  $T^\alpha(X; t_1, t_2) = \frac{1}{(\alpha - 1)} \left( 1 - \frac{1 - (e^{-\alpha t_1} - e^{-\alpha t_2})}{(e^{t_1} - e^{t_2})^\alpha} \right)$ , for all  $\alpha > 1$  ( $0 < \alpha < 1$ ) and  $t_1, t_2$  ( $t_1 < t_2$ ), which is in agreement with Proposition 2 and Proposition 3.

For increasing function  $f(x)$ , the above proposition can be similarly proved.

### 3. INTERVAL CUMULATIVE RESIDUAL AND PAST TSALLIS ENTROPY

Let  $X$  be an absolutely continuous random variable and let  $D = \{(x, y) : F(x) < F(y)\}$ . Then we define the  $ICPT$  and  $ICRT$  functions, respectively, as follows:

$$ICPT(X; t_1, t_2) = \frac{1}{\alpha - 1} \left( 1 - \int_{t_1}^{t_2} \left( \frac{F(x)}{F(t_2) - F(t_1)} \right)^\alpha dx \right) \tag{8}$$

and

$$ICRT(X; t_1, t_2) = \frac{1}{\alpha - 1} \left( 1 - \int_{t_1}^{t_2} \left( \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \right)^\alpha dx \right), \tag{9}$$

where  $(t_1, t_2) \in D$ . It is clear that,  $ICRT(X; 0, \infty)$  is  $CRT(X)$ , and  $ICRT(X; t_1, \infty)$  is  $DCRT(X; t_1)$ . Also,  $ICPT(X; 0, \infty)$  is  $CPT(X)$ , and  $ICPE(X; 0, t_2)$  is  $DCPT(X; t_2)$ . The applications of classes of life distributions can be demonstrated in different areas, including reliability, engineering, biological science, maintenance, and biometrics. Hence, statisticians and reliability analysts are interested in modeling survival information and classifications of life distributions based on a few aspects of aging. For instance, we refer the reader to [15, 1, 26]. So, the corresponding aging classes are defined as follows.

**Definition 1.** Consider the random variable  $X$ .

- $X$  is said to have decreasing interval cumulative residual Tsallis entropy (DICRT) property if and only if for any fixed  $t_2$ ,  $ICRT(X; t_1, t_2)$  is decreasing with respect to  $t_1$ .
- $X$  is said to have increasing interval cumulative past Tsallis entropy (IICPT) property if and only if for any fixed  $t_1$ ,  $ICPT(X; t_1, t_2)$  is increasing with respect to  $t_2$ .

An upper bound for  $ICRT(X; t_1, t_2)$  with the decreasing (increasing) ICRT property is acquired in the next theorems.

**Theorem 2.** The random variable  $X$  has decreasing (increasing) ICRT property if and only if the following inequality is satisfied for all  $(t_1, t_2) \in D$  and  $0 < \alpha < 1$  ( $\alpha > 1$ ):

$$ICRT(X; t_1, t_2) \leq (\geq) \frac{1}{\alpha - 1} \left( 1 - \frac{1}{\alpha} \left( \frac{\bar{F}(t_1)}{f(t_1)} \right)^\alpha \left( \frac{1 + \frac{\partial \mu(t_1, t_2)}{\partial t_1}}{\mu(t_1, t_2)} \right)^{\alpha - 1} \right).$$

**Proof.** By differentiating  $ICRT(X; t_1, t_2)$  of the form (9) with respect to  $t_1$ , we have

$$\begin{aligned} \frac{\partial ICRT(X; t_1, t_2)}{\partial t_1} &= \frac{1}{\alpha - 1} \left( \left( \frac{\bar{F}(t_1)}{\bar{F}(t_1) - \bar{F}(t_2)} \right)^\alpha \right. \\ &\quad \left. - \alpha \frac{f(t_1)}{\bar{F}(t_1) - \bar{F}(t_2)} \int_{t_1}^{t_2} \left( \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \right)^\alpha dx \right) \\ &= \frac{1}{\alpha - 1} \left( \frac{\bar{F}(t_1)}{f(t_1)} \right)^\alpha h_1^\alpha(t_1, t_2) \\ &\quad - \frac{\alpha}{\alpha - 1} h_1(t_1, t_2) (1 - (\alpha - 1) ICRT(X; t_1, t_2)). \end{aligned}$$

By the definition of the GFR in (3) and (4), their relationships with  $\mu(t_1, t_2) = E(X - t_1 | t_1 \leq X \leq t_2)$  and  $\mu^*(t_1, t_2) = E(t_2 - X | t_1 \leq X \leq t_2)$  are, respectively, as follows:

$$h_1(t_1, t_2) = \frac{1 + \frac{\partial \mu(t_1, t_2)}{\partial t_1}}{\mu(t_1, t_2)}, \tag{10}$$

$$h_2(t_1, t_2) = \frac{1 - \frac{\partial \mu^*(t_1, t_2)}{\partial t_2}}{\mu(t_1, t_2)}. \tag{11}$$

So, after suitable substitution of Eqs. (10) and (11) and simplifying the equations, we have

$$ICRT(X; t_1, t_2) \leq (\geq) \frac{1}{\alpha - 1} \left( 1 - \frac{1}{\alpha} \left( \frac{\bar{F}(t_1)}{f(t_1)} \right)^\alpha (h_1(t_1, t_2))^{\alpha-1} \right).$$

■

**Example 3.** Let  $X$  be distributed uniformly on  $(0, \beta)$ ,  $\beta > 0$ , then it can be easily verified that,

$$ICRT(X; t_1, t_2) = \frac{1}{\alpha - 1} \left( 1 - \frac{(\beta - t_1)^{\alpha+1} - (\beta - t_2)^{\alpha+1}}{(t_2 - t_1)^\alpha (1 + \alpha)} \right),$$

$$\mu(t_1, t_2) = \frac{t_2 - t_1}{2}.$$

the differentiation of  $ICRT$  with respect to  $t_1$  is negative for all  $(t_1, t_2) \in D$ , which shows that the uniform distribution has  $DICRT$  property and theorem2 is satisfied.

There exist no nonnegative random variables with increasing  $ICRT(IICRT)$  over the domain  $[0, \infty)$ , indicated in the following theorem.

**Theorem 3.** If  $X$  is a nonnegative nondegenerate random variable, then  $ICRT(X; t_1, t_2)$  cannot be an increasing function with respect to  $t_1$  for any real fixed  $t_2$ .

**Proof.** First note that, using lHopitals rule, we have

$$\begin{aligned} \lim_{t_1 \rightarrow t_2} ICRT(X; t_1, t_2) &= \lim_{t_1 \rightarrow t_2} \frac{1}{\alpha - 1} \left( 1 - \int_{t_1}^{t_2} \left( \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \right)^\alpha dx \right) \\ &= \frac{1}{\alpha - 1} \left( 1 - \lim_{t_1 \rightarrow t_2} \frac{\int_{t_1}^{t_2} (\bar{F}(x))^\alpha dx}{(\bar{F}(t_1) - \bar{F}(t_2))^\alpha} \right) \\ &= \frac{1}{\alpha - 1} \left( 1 - \lim_{t_1 \rightarrow t_2} \frac{(\bar{F}(t_1))^\alpha}{\alpha f(t_1) (F(t_2) - F(t_1))^{\alpha-1}} \right) \\ &= -\infty. \end{aligned}$$

Now, on the contrary, suppose that  $ICRT(X; t_1, t_2)$  is increasing in  $t_1$ . Then for all  $t_1 \leq t_2$ ,  $ICRT(X; t_1, t_2) \leq ICRT(X; t_2, t_2) = -\infty$ , which contradicts with the fact that  $ICRT(X; t_1, t_2) \in \mathfrak{R}$  for all  $(t_1, t_2) \in D$ . ■

In the following proposition, we obtain a lower bound, according to  $\mu(X) = \int_x^\infty \frac{F(x)}{F(t)} dt$ , for  $E(\mu(X)|t_1 \leq X \leq t_2)$ .

**Proposition 4.** Suppose that  $F$  is an absolutely continuous distribution function with  $ICRT(X; t_1, t_2) < \infty$ . Then, for  $0 < \alpha < 1$

$$E(\mu(X)|t_1 \leq X \leq t_2) \geq (\alpha - 1)ICRT(X; t_1, t_2) - 1.$$

**Proof.** By using  $E(\mu(X)|t_1 \leq X \leq t_2) \geq ICRT(X; t_1, t_2)$  [5], we have

$$\begin{aligned} &\int_{t_1}^{t_2} \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \log\left(\frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)}\right) dx \\ &\leq \int_{t_1}^{t_2} \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \left( \left( \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \right) - 1 \right) dx \\ &\leq \int_{t_1}^{t_2} \left( \left( \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \right) - 1 \right) dx \\ &\leq \int_{t_1}^{t_2} \left( \left( \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \right)^\alpha - 1 \right) dx \\ &= \int_{t_1}^{t_2} \left( \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \right)^\alpha - (t_2 - t_1) dx. \end{aligned}$$

Then

$$\begin{aligned} & - \int_{t_1}^{t_2} \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \log\left(\frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)}\right) dx \\ \geq & - \int_{t_1}^{t_2} \left(\frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)}\right)^\alpha + (t_2 - t_1) dx \\ \geq & - \int_{t_1}^{t_2} \left(\frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)}\right)^\alpha dx \\ = & (\alpha - 1)ICRT(X; t_1, t_2) - 1. \end{aligned}$$

■

The following theorem tries to clarify the problem, achieving when the interval entropy uniquely appoints the distribution function.

**Theorem 4.** Let  $X$  be a nonnegative and continuous random variable and let  $ICRT(X; t_1, t_2)$  be increasing with respect to  $t_1$  and decreasing with respect to  $t_2$ . Then  $ICRT(X; t_1, t_2)$  uniquely determines  $F(x)$ .

**Proof.** By differentiating  $ICRT(X; t_1, t_2)$  with respect to  $t_j (j = 1, 2)$ , we have

$$\begin{aligned} \frac{\partial ICRT(X; t_1, t_2)}{\partial t_2} &= \frac{1}{\alpha - 1} \left( - \left(\frac{\bar{F}(t_2)}{\bar{F}(t_1) - \bar{F}(t_2)}\right)^\alpha \right. \\ &\quad \left. + \alpha \frac{f(t_2)}{\bar{F}(t_1) - \bar{F}(t_2)} \int_{t_1}^{t_2} \left(\frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)}\right)^\alpha dx \right) \\ &= \frac{-1}{\alpha - 1} \left(\frac{\bar{F}(t_2)}{f(t_2)}\right)^\alpha h_2^\alpha(t_1, t_2) \\ &\quad + \frac{\alpha}{\alpha - 1} h_2(t_1, t_2) (1 - (\alpha - 1)ICRT(X; t_1, t_2)) \\ &= -h_2(t_1, t_2) \left( \frac{1}{\alpha - 1} \left(\frac{\bar{F}(t_2)}{f(t_2)}\right)^\alpha h_2^{\alpha-1}(t_1, t_2) \right. \\ &\quad \left. - \frac{\alpha}{\alpha - 1} (1 - (\alpha - 1)ICRT(X; t_1, t_2)) \right), \end{aligned}$$

and

$$\begin{aligned} \frac{\partial ICRT(X; t_1, t_2)}{\partial t_1} &= \frac{1}{\alpha - 1} \left( \left(\frac{\bar{F}(t_1)}{\bar{F}(t_1) - \bar{F}(t_2)}\right)^\alpha \right. \\ &\quad \left. - \alpha \frac{f(t_1)}{\bar{F}(t_1) - \bar{F}(t_2)} \int_{t_1}^{t_2} \left(\frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)}\right)^\alpha dx \right) \\ &= \frac{1}{\alpha - 1} \left(\frac{\bar{F}(t_1)}{f(t_1)}\right)^\alpha h_1^\alpha(t_1, t_2) \\ &\quad - \frac{\alpha}{\alpha - 1} h_1(t_1, t_2) (1 - (\alpha - 1)ICRT(X; t_1, t_2)) \\ &= h_1(t_1, t_2) \left( \frac{1}{\alpha - 1} \left(\frac{\bar{F}(t_1)}{f(t_1)}\right)^\alpha h_1^{\alpha-1}(t_1, t_2) \right. \\ &\quad \left. - \frac{\alpha}{\alpha - 1} (1 - (\alpha - 1)ICRT(X; t_1, t_2)) \right). \end{aligned}$$

Thus, for fixed  $t_2$  and arbitrary  $t_1$ ,  $h_1(t_1, t_2)$  is a positive solution to the following equation:

$$\begin{aligned} g(x_{t_2}) &= x_{t_2} \left( \frac{1}{\alpha - 1} \left(\frac{\bar{F}(t_1)}{\bar{F}(t_1)}\right)^\alpha x_{t_2}^{\alpha-1} - \frac{\alpha}{\alpha - 1} (1 - (\alpha - 1)ICRT(X; t_1, t_2)) \right) \\ &\quad - \frac{\partial ICRT(X; t_1, t_2)}{\partial t_1}. \end{aligned} \tag{12}$$

Similarly, for fixed  $t_1$  and arbitrary  $t_2$ , we have  $h_2(t_1, t_2)$  as a positive solution to the following equation:

$$\gamma(y_{t_1}) = y_{t_1} \left( \frac{1}{\alpha - 1} \left( \frac{\bar{F}(t_2)}{f(t_2)} \right)^\alpha y_{t_1}^{\alpha-1} - \frac{\alpha}{\alpha - 1} (1 - (\alpha - 1)ICRT(X; t_1, t_2)) \right) + \frac{\partial ICRT(X; t_1, t_2)}{\partial t_2}. \tag{13}$$

By differentiating  $g$  and  $\gamma$  with respect to  $x_{t_2}$  and  $y_{t_1}$ , we get

$$\frac{\partial g(x_{t_2})}{\partial x_{t_2}} = \frac{\alpha}{\alpha - 1} \left( \left( \frac{\bar{F}(t_1)}{f(t_1)} \right)^\alpha x_{t_2}^{\alpha-1} - (1 - (\alpha - 1)ICRT(X; t_1, t_2)) \right),$$

and

$$\frac{\partial \gamma(y_{t_1})}{\partial y_{t_1}} = \frac{\alpha}{\alpha - 1} \left( \left( \frac{\bar{F}(t_2)}{f(t_2)} \right)^\alpha y_{t_1}^{\alpha-1} - (1 - (\alpha - 1)ICRT(X; t_1, t_2)) \right).$$

Furthermore, the second-order derivatives of  $g$  and  $\gamma$  with respect to  $x_{t_2}$  and  $y_{t_1}$  are  $\alpha \left( \frac{\bar{F}(t_1)}{f(t_1)} \right)^\alpha x_{t_2}^{\alpha-2} > 0$  and  $\alpha \left( \frac{\bar{F}(t_2)}{f(t_2)} \right)^\alpha y_{t_1}^{\alpha-2} > 0$ , respectively. Then the functions  $g$  and  $\gamma$  are minimized at points  $x_{t_2} = \left( (1 - (\alpha - 1)ICRT(X; t_1, t_2)) \left( \frac{f(t_1)}{\bar{F}(t_1)} \right)^\alpha \right)^{\frac{1}{\alpha-1}}$  and  $y_{t_1} = \left( (1 - (\alpha - 1)ICRT(X; t_1, t_2)) \left( \frac{f(t_2)}{\bar{F}(t_2)} \right)^\alpha \right)^{\frac{1}{\alpha-1}}$ , respectively. In addition,

$$g(0) = -\frac{\partial ICRT(X; t_1, t_2)}{\partial t_1} < 0, \quad g(\infty) = \infty,$$

and

$$\gamma(0) = -\frac{\partial ICRT(X; t_1, t_2)}{\partial t_2} < 0, \quad \gamma(\infty) = \infty.$$

So, both functions  $g$  and  $\gamma$  first decrease and then increase with respect to  $x_{t_2}$  and  $y_{t_1}$ , respectively, which conclude that equations (12) and (13) have unique roots  $h_1(t_1, t_2)$  and  $h_2(t_1, t_2)$ , respectively. Now,  $ICRT(X; t_1, t_2)$  uniquely determines GFRs and the distribution function, with attention to Remark 3.1 [14]. ■

Similar to Theorems 2, 3, and 4 and Proposition 4, we have the following results:

- The random variable  $X$  has decreasing (increasing) ICRT property if and only if the following inequality is satisfied for all  $(t_1, t_2) \in D$  and  $0 < \alpha < 1$  ( $\alpha > 1$ ):

$$ICPT(X; t_1, t_2) \leq (\geq) \frac{1}{\alpha - 1} \left( 1 - \frac{1}{\alpha} \left( \frac{\bar{F}(t_2)}{f(t_2)} \right)^\alpha \left( \frac{1 - \frac{\partial \mu^*(t_1, t_2)}{\partial t_2}}{\mu(t_1, t_2)} \right)^{\alpha-1} \right).$$

- If  $X$  is a nonnegative nondegenerate random variable, then  $ICPT(X; t_1, t_2)$  cannot be a decreasing function with respect to  $t_2$  for any real fixed  $t_1$ .
- Suppose that  $F$  is an absolutely continuous distribution function with  $ICPT(X; t_1, t_2) < \infty$ , then

$$E(\mu^*(X) | t_1 \leq X \leq t_2) \geq (\alpha - 1)ICPT(X; t_1, t_2) - 1.$$

- Let  $X$  be a nonnegative and continuous random variable and let  $ICPT(X; t_1, t_2)$  be increasing with respect to  $t_1$  and decreasing with respect to  $t_2$ . Then  $ICPT(X; t_1, t_2)$  uniquely determines  $F(x)$ .

**Example 4.** Let  $X$  be distributed uniformly on  $(0, \beta)$ ,  $\beta > 0$ , then it can be easily verified that,

$$ICPT(X; t_1, t_2) = \frac{1}{\alpha - 1} \left( 1 - \frac{t_2^{\alpha+1} - t_1^{\alpha+1}}{(t_2 - t_1)^\alpha (1 + \alpha)} \right),$$



$$\mu^*(t_1, t_2) = \frac{t_2 - t_1}{2}.$$

As the *ICPT* is increasing with respect to  $t_2$ ,  $X$  has *IICPT* properties.

As in Lemma 1, the following theorem is proved by the same approach.

**Lemma 2.** Let  $X$  be a nonnegative continuous random variable with cdf  $F$ , and take  $Y = \phi(X)$ , where  $\phi(\cdot)$  is a strictly increasing differentiable function. Then

$$ICRT(Y; t_1, t_2) = \frac{1}{\alpha - 1} \left( 1 - \int_{\max\{0, \phi^{-1}(t_1)\}}^{\phi^{-1}(t_2)} \left( \frac{\bar{F}(x)}{\bar{F}(\phi^{-1}(t_1)) - \bar{F}(\phi^{-1}(t_2))} \right)^\alpha \phi'(x) dx \right).$$

**Proposition 5.** If  $Z = aX + b$ , with  $a > 0$  and  $b \geq 0$ , so  $\bar{F}_{aX+b}(z) = \bar{F}_X(\frac{z-b}{a})$ , then

$$ICRT(Z; t_1, t_2) = \frac{1-a}{\alpha-1} + aICRT(X; \frac{t_1-b}{a}, \frac{t_2-b}{a}).$$

There is an identity for doubly truncated (interval) CRT in the following theorem.

**Theorem 5.** Let  $X$  be a random variable with support in  $[0, r]$  and symmetric with respect to  $\frac{r}{2}$ , that is,  $\bar{F}(x) = F(r-x)$  for  $0 \leq x \leq r$ . Then

$$ICRT(X; t_1, t_2) = ICPT(X; r - t_2, r - t_1), \quad 0 \leq t_1, t_2 \leq r.$$

**Proof.** The theorem is proved by the following equation:

$$\begin{aligned} ICRT(X; t_1, t_2) &= \frac{1}{\alpha - 1} \left( 1 - \int_{t_1}^{t_2} \left( \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \right)^\alpha dx \right) \\ &= \frac{1}{\alpha - 1} \left( 1 - \int_{t_1}^{t_2} \left( \frac{F(r-x)}{F(r-t_1) - F(r-t_2)} \right)^\alpha dx \right) \\ &= -\frac{1}{\alpha - 1} \left( 1 - \int_{r-t_1}^{r-t_2} \left( \frac{F(y)}{F(r-t_1) - F(r-t_2)} \right)^\alpha dy \right) \\ &= \frac{1}{\alpha - 1} \left( 1 - \int_{r-t_2}^{r-t_1} \left( \frac{F(y)}{F(r-t_1) - F(r-t_2)} \right)^\alpha dy \right) \\ &= ICPT(X; r - t_2, r - t_1). \end{aligned}$$

■

**Example 5.** If  $X$  is uniformly distributed in  $[0, r]$ , then for  $0 \leq t_1, t_2 \leq r$ , we have  $ICRT(X; t_1, t_2) = ICPT(X; r - t_2, r - t_1) = \frac{1}{\alpha-1} \left( 1 - \frac{(r-t_1)^{\alpha+1} - (r-t_2)^{\alpha+1}}{(t_2-t_1)^\alpha (1+\alpha)} \right)$ , which is in agreement with Theorem 5.

Similar to Lemma 2, Proposition 5, and Theorem 5, we have the following results:

- Let  $X$  be a nonnegative continuous random variable with cdf  $F$ , and take  $Y = \phi(X)$ , where  $\phi(\cdot)$  is a strictly increasing differentiable function. Then

$$ICPT(Y; t_1, t_2) = \frac{1}{\alpha - 1} \left( 1 - \int_{\max\{0, \phi^{-1}(t_1)\}}^{\phi^{-1}(t_2)} \left( \frac{F(x)}{F(\phi^{-1}(t_2)) - F(\phi^{-1}(t_1))} \right)^\alpha \phi'(x) dx \right).$$

- If  $Z = aX + b$ , with  $a > 0$  and  $b \geq 0$ , so  $F_{aX+b}(z) = F_X(\frac{z-b}{a})$ , then

$$ICPT(Z; t_1, t_2) = \frac{1-a}{\alpha-1} + aICPT(X; \frac{t_1-b}{a}, \frac{t_2-b}{a}).$$

- Let  $X$  be a random variable with support in  $[0, r]$  and symmetric with respect to  $\frac{r}{2}$ , that is,  $F(x) = \bar{F}(r-x)$  for  $0 \leq x \leq r$ . Then

$$ICPT(X; t_1, t_2) = ICRT(X; r - t_2, r - t_1); \quad 0 \leq t_1, t_2 \leq r.$$

**Example 6.** If  $X$  is uniformly distributed in  $[0, r]$ , for  $0 \leq t_1, t_2 \leq r$ , we have  $ICPT(X; t_1, t_2) = ICRT(X; r - t_2, r - t_1) = \frac{1}{\alpha - 1} \left( 1 - \frac{t_2^{\alpha+1} - t_1^{\alpha+1}}{(t_2 - t_1)^\alpha (1 + \alpha)} \right)$ , which is in agreement with Remark 4 (part 1).

Let  $X$  and  $Y$  be two random variables. Also, the distribution function and density function of  $X$  are indicated by  $F(t)$  and  $f(t)$  and those of  $Y$  are denoted by  $G(t)$  and  $g(t)$ , separately. Now we compare the two random variables  $X$  and  $Y$  based on doubly truncated (interval) cumulative residual and past Tsallis entropy. So, we first need the following definitions, which can be seen in [20]

**Definition 2.**  $X$  is said to be less than or equal to  $Y$  in usual stochastic ordering, if  $\frac{f(x)}{g(x)}$  is decreasing in  $x > 0$ . We write  $X \leq^{lr} Y$ .

**Definition 3.**  $X$  is said to be less than or equal to  $Y$  in likelihood ratio ordering, if  $\bar{F}(x) \leq \bar{G}(x)$ , for all  $x > 0$ . We write  $X \leq^{st} Y$ .

Navarro and Rubio(2011) expressed that The two random variables  $X$  and  $Y$  satisfy  $X \leq^{lr} Y$  if, and only if,  $[X - t_1 | t_1 \leq X \leq t_2] \leq^{st} [Y - t_1 | t_1 \leq Y \leq t_2]$ , whenever  $(t_1 < t_2)$ . Also, we compare two random variables  $X$  and  $Y$  based on the properties of (interval) CRT and (interval) CPT in likelihood ratio ordering.

**Theorem 6.** Let  $X$  and  $Y$  be two nonnegative absolutely continuous random variables with survival functions  $\bar{F}(x)$  and  $\bar{G}(x)$ , respectively. If  $X \leq (\geq)^{lr} Y$  for all  $t_1, t_2 \geq 0$ , then  $ICRT(X; t_1, t_2) \leq (\geq) ICRT(Y; t_1, t_2)$ , for  $0 < \alpha < 1$ ; otherwise for  $\alpha > 1$ ,  $ICRT(X; t_1, t_2) \geq (\leq) ICRT(Y; t_1, t_2)$ .

**Proof.** The assumption  $X \leq (\geq)^{lr} Y$  implies that

$$\begin{aligned} \bar{F}_{X_{t_1, t_2}} &\leq (\geq) \bar{G}_{X_{t_1, t_2}}, \\ \left( \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \right)^\alpha &\leq (\geq) \left( \frac{\bar{G}(x)}{\bar{G}(t_1) - \bar{G}(t_2)} \right)^\alpha, \\ 1 - \int_{t_1}^{t_2} \left( \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \right)^\alpha dx &\geq (\leq) 1 - \int_{t_1}^{t_2} \left( \frac{\bar{G}(x)}{\bar{G}(t_1) - \bar{G}(t_2)} \right)^\alpha dx. \end{aligned}$$

For  $\alpha > 1$ , we have

$$\begin{aligned} \frac{1}{\alpha - 1} \left( 1 - \int_{t_1}^{t_2} \left( \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \right)^\alpha dx \right) &\geq (\leq) \frac{1}{\alpha - 1} \left( 1 - \int_{t_1}^{t_2} \left( \frac{\bar{G}(x)}{\bar{G}(t_1) - \bar{G}(t_2)} \right)^\alpha dx \right), \\ ICRT(X; t_1, t_2) &\geq (\leq) ICRT(Y; t_1, t_2). \end{aligned}$$

For  $0 < \alpha < 1$ , it follows that

$$\begin{aligned} \frac{1}{\alpha - 1} \left( 1 - \int_{t_1}^{t_2} \left( \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \right)^\alpha dx \right) &\leq (\geq) \frac{1}{\alpha - 1} \left( 1 - \int_{t_1}^{t_2} \left( \frac{\bar{G}(x)}{\bar{G}(t_1) - \bar{G}(t_2)} \right)^\alpha dx \right), \\ ICRT(X; t_1, t_2) &\leq (\geq) ICRT(Y; t_1, t_2). \end{aligned}$$

■

**Theorem 7.** Let  $X$  and  $Y$  be two nonnegative absolutely continuous random variables with cdfs  $F(x)$  and  $G(x)$ , respectively. If  $X \leq^{st} Y$  for all  $t_1, t_2 \geq 0$ , then  $ICPT(X; t_1, t_2) \geq ICPT(Y; t_1, t_2)$ , for  $0 < \alpha < 1$ ; otherwise for  $\alpha > 1$ ,  $ICPT(X; t_1, t_2) \leq ICPT(Y; t_1, t_2)$ .

**Proof.** The assumption that  $X \stackrel{st}{\leq} Y$  implies that

$$\begin{aligned} F_{X_{t_1, t_2}} &\geq G_{X_{t_1, t_2}}, \\ \left(\frac{F(x)}{F(t_2) - F(t_1)}\right)^\alpha &\geq \left(\frac{G(x)}{G(t_2) - G(t_1)}\right)^\alpha, \\ 1 - \int_{t_1}^{t_2} \left(\frac{F(x)}{F(t_2) - F(t_1)}\right)^\alpha dx &\leq 1 - \int_{t_1}^{t_2} \left(\frac{G(x)}{G(t_2) - G(t_1)}\right)^\alpha dx. \end{aligned}$$

for  $\alpha > 1$ , we have

$$\begin{aligned} \frac{1}{\alpha - 1} \left(1 - \int_{t_1}^{t_2} \left(\frac{F(x)}{F(t_2) - F(t_1)}\right)^\alpha dx\right) &\leq \frac{1}{\alpha - 1} \left(1 - \int_{t_1}^{t_2} \left(\frac{G(x)}{G(t_2) - G(t_1)}\right)^\alpha dx\right), \\ ICPT(X; t_1, t_2) &\leq ICPT(Y; t_1, t_2). \end{aligned}$$

For  $0 < \alpha < 1$ , it follows that

$$\begin{aligned} \frac{1}{\alpha - 1} \left(1 - \int_{t_1}^{t_2} \left(\frac{F(x)}{F(t_2) - F(t_1)}\right)^\alpha dx\right) &\geq \frac{1}{\alpha - 1} \left(1 - \int_{t_1}^{t_2} \left(\frac{G(x)}{G(t_2) - G(t_1)}\right)^\alpha dx\right), \\ ICPT(X; t_1, t_2) &\geq ICPT(Y; t_1, t_2). \end{aligned}$$

■

**Example 7.** Let

$$\bar{F}(x) = \begin{cases} \left(\frac{x_0}{x}\right)^{\beta_1}, & x > x_0, \\ 1, & x \leq x_0, \end{cases}$$

and

$$\bar{G}(x) = \begin{cases} \left(\frac{x_0}{x}\right)^{\beta_2}, & x > x_0, \\ 1, & x \leq x_0. \end{cases}$$

That is,  $X$  and  $Y$  have Pareto distributions with parameters  $\beta_1$  and  $\beta_2$ , respectively. If  $\beta_1 \geq \beta_2$  and  $0 < \beta_1, \beta_2 \leq \frac{1}{\alpha}$ , hence  $X \stackrel{lr}{\leq} Y$  for  $\alpha > 1$ , then  $ICRT(X; t_1, t_2) \geq ICRT(Y; t_1, t_2)$ . Also, the assumptions of the theorem hold, and therefore  $[X - t_1 | t_1 \leq X \leq t_2] \leq^{st} [Y - t_1 | t_1 \leq Y \leq t_2]$ , whenever  $(t_1 < t_2)$ .

#### 4. EMPIRICAL ESTIMATION OF ICPT

By utilizing various empirical estimators of the cdf, we suggest four non-parametric estimators  $ICPT(X; t_1, t_2)$  and also compare the implementation of the proposed estimators. For an actual-life fact set, we study the monotonicity of ICPT based totally on its kernel-smoothed estimator.

First, we introduce four nonparametric estimators, by mentioning the name  $ICPT^1(X; t_1, t_2)$ ,  $ICPT^2(X; t_1, t_2)$ ,  $ICPT^3(X; t_1, t_2)$  and  $ICPT^4(X; t_1, t_2)$ , of  $ICPT$  through utilizing empirical distribution function, mean empirical distribution function, median empirical distribution function, and kernel-smoothed function and their implementation by the Monte-Carlo simulation. Let  $X_1, X_2, \dots, X_n$  be an independent and identically distributed random sample drawn from a population having distribution function  $F(x)$  and survival function  $\bar{F}(x)$ . Now, the first nonparametric estimator of  $ICPT^1(X; t_1, t_2)$  may be written as

$$ICPT^1(X; t_1, t_2) = \frac{1}{\alpha - 1} \left(1 - \int_{t_1}^{t_2} \left(\frac{F_n^{(1)}(x)}{F_n^{(1)}(t_2) - F_n^{(1)}(t_1)}\right)^\alpha dx\right),$$

for  $0 < \alpha \neq 1$ , where  $F_n^{(1)}(x) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x)$ ,  $x \in R$ , is the empirical distribution function and

$$I(X_i \leq x) = \begin{cases} 1 & \text{if } X \leq x, \\ 0 & \text{otherwise,} \end{cases}$$

is the indicator function of the event  $X \leq x$ . Let  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  be the order statistics of random sample. Noting the sample values lying between  $t_1$  and  $t_2$  so that  $t_1 \leq x_{(j)}, x_{(j+1)}, \dots, x_{(k)} \leq t_2$ , then

$$\begin{aligned} ICPT^1(X; t_1, t_2) &= \frac{1}{\alpha - 1} \left( 1 - \sum_{i=j}^k \int_{x_{(i)}}^{x_{(i+1)}} \left( \frac{F_n^{(1)}(x)}{F_n^{(1)}(t_2) - F_n^{(1)}(t_1)} \right)^\alpha dx \right) \\ &= \frac{1}{\alpha - 1} \left( 1 - \frac{1}{(F_n^{(1)}(t_2) - F_n^{(1)}(t_1))^\alpha} \sum_{i=j}^k \int_{x_{(i)}}^{x_{(i+1)}} (F_n^{(1)}(x))^\alpha dx \right) \\ &= \frac{1}{\alpha - 1} \left( 1 - \frac{1}{(F_n^{(1)}(t_2) - F_n^{(1)}(t_1))^\alpha} \sum_{i=j}^k (x_{(i)} - x_{(i+1)}) (F_n^{(1)}(x))^\alpha \right). \end{aligned} \quad (14)$$

The second estimator of  $ICPT^2(X; t_1, t_2)$  can be acquired by replacing mean empirical distribution function  $F_n^{(2)}(x)$  in (14) as

$$ICPT^2(X; t_1, t_2) = \frac{1}{\alpha - 1} \left( 1 - \frac{1}{(F_n^{(2)}(t_2) - F_n^{(2)}(t_1))^\alpha} \sum_{i=j}^k (x_{(i)} - x_{(i+1)}) (F_n^{(2)}(x))^\alpha \right), \quad (15)$$

where the mean empirical distribution function is defined as

$$F_n^{(2)}(x) = \frac{1}{n+1} \sum_{i=1}^n I(X_i \leq x), \quad x \in R.$$

The third nonparametric estimator of  $ICPT^3(X; t_1, t_2)$  can be achieved by utilizing median empirical distribution function in (14) as follows:

$$ICPT^3(X; t_1, t_2) = \frac{1}{\alpha - 1} \left( 1 - \frac{1}{(F_n^{(3)}(t_2) - F_n^{(3)}(t_1))^\alpha} \sum_{i=j}^k (x_{(i)} - x_{(i+1)}) (F_n^{(3)}(x))^\alpha \right), \quad (16)$$

where  $F_n^{(3)}(x) = \sum_{i=1}^n \frac{I(X_i \leq x) - 0.3}{n + 0.4}$ ,  $x \in R$ , is the median empirical distribution function.

The fourth estimator can be defined by utilizing Kernel-smoothed estimator  $F_n^{(4)}(x)$  of the distribution function in (14) as follows:

$$ICPT^4(X; t_1, t_2) = \frac{1}{\alpha - 1} \left( 1 - \frac{1}{(F_n^{(4)}(t_2) - F_n^{(4)}(t_1))^\alpha} \sum_{i=j}^k (x_{(i)} - x_{(i+1)}) (F_n^{(4)}(x))^\alpha \right), \quad (17)$$

where  $F_n^{(4)}(x)$ , the kernel-smoothed estimator of distribution function, is defined as

$$F_n^{(4)}(x) = \frac{1}{n} \sum_{i=1}^n L\left(\frac{x - X_i}{h}\right),$$

where  $L$  is a distribution function of positive kernel  $K$ , that is,  $L(u) = \int_u^{-\infty} K(t) dt$  and  $h$  is the bandwidth of parameter. Now, we utilize the normal kernel function  $K(u) = \frac{1}{\sqrt{\pi}} \exp\left(-\frac{u^2}{2}\right)$ .

## 5. SIMULATION

It is widely recognized that the smoothed estimator has a better performance compared to a nonsmoothed estimator. To demonstrate the effectiveness of the empirical and kernel estimators, a Monte-Carlo simulation examination is accomplished. The estimated values are computed based on 1000 simulations from  $Exp(0.5)$  (exponential distribution) each of size  $n$  ( $n = 30, 35, 40, 50, 60$ ) for different truncation limits and  $\alpha = 0.2; 3.5$ . Bias and mean square error (MSE) are also calculated. In Tables 1 and 2, we present the exact value, bias, and the

MSE of the proposed estimators of ICPT. The MSE of the estimators corresponding to truncation limit (0.2, 4) and for  $\alpha = 0.2, 3.5$  is also displayed in Figure 1 for increasing sample size.

It is obvious that in nearly all cases  $ICPT^4(X; t_1, t_2)$  (17) performs way better with less MSE than the other estimators as determined in (14) (15) and (16). Further, for  $\alpha = 0.2$ ,  $ICPT^1(X; t_1, t_2)$  produces better result than  $ICPT^3(X; t_1, t_2)$ , while  $ICPT^2(X; t_1, t_2)$  yields poor estimates as MSE is higher in comparison with the other estimators of ICPT. Also, for  $\alpha = 3.5$ , it can be seen that there is a slight difference between the first, second and third estimators and The fourth estimator is significantly better estimator. It is expected, one can depict from Tables 1 and 2 that ICPT as a measure of uncertainty declines for a shrinking interval. Generally, we can conclude that kernel smoothed estimator gives better estimates of ICPT than the other proposed estimators in terms of MSE. Also, the values of MSE of the proposed estimators are reduced by increasing sample size, which is caused by dependence of the MSE of the empirical estimators to the sample size.

It is obvious that, in nearly, all cases  $ICPT^4(X; t_1, t_2)$  defined by (17) perform a way better with less MSE than the other estimators, as determined in (14), (15), and (16). Further more, for  $\alpha = 0.2$ ,  $ICPT^1(X; t_1, t_2)$  produces a better result than  $ICPT^3(X; t_1, t_2)$ , while  $ICPT^2(X; t_1, t_2)$  yields poor estimates as the MSE is higher in comparison with the other estimators of ICPT. Also, for  $\alpha = 3.5$ , it can be seen that there is a slight difference between the first, second and third estimators and The fourth estimator is significantly better estimator. It is expected that one can depict from Tables 1 and 2 that ICPT, as a measure of uncertainty, declines for a shrinking interval. Generally, we can conclude that the kernel-smoothed estimator gives better estimates of ICPT than the other proposed estimators in terms of the MSE. Also, the values of MSE of the proposed estimators are reduced by increasing sample size, which is caused by the dependence of the MSE of the empirical estimators on the sample size.

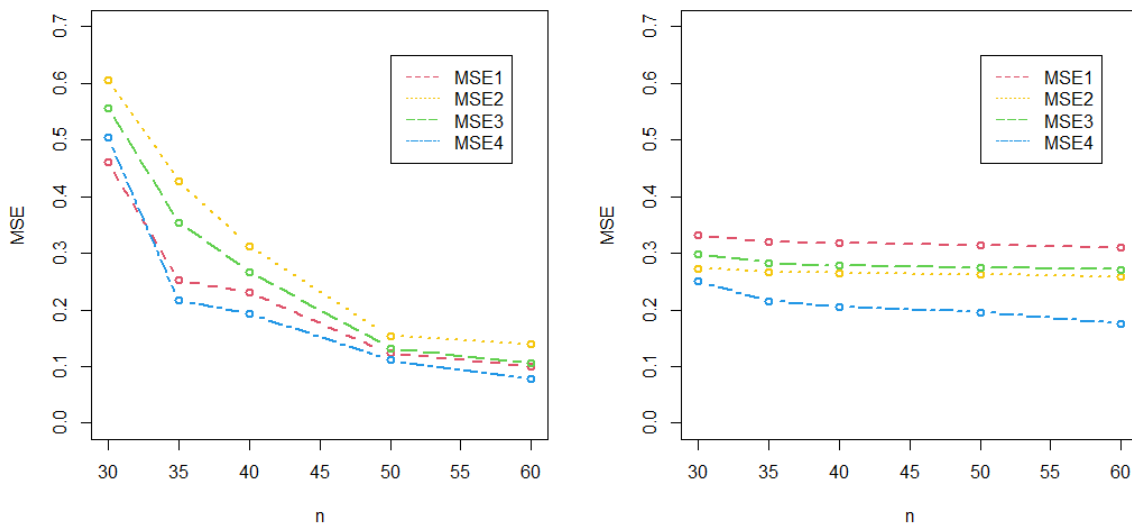
**Table 1:** Bias and MSE of  $ICPT^1(X; t_1, t_2)$ ,  $ICPT^2(X; t_1, t_2)$ ,  $ICPT^3(X; t_1, t_2)$  and  $ICPT^4(X; t_1, t_2)$  for  $\alpha = 3.5$  and different truncation limits ( $n = 30, 35, 40, 50, 60$ ).

$\alpha = 3.5$			$ICPT^1(X; t_1, t_2)$	$ICPT^2(X; t_1, t_2)$	$ICPT^3(X; t_1, t_2)$	$ICPT^4(X; t_1, t_2)$
$(t_1, t_2)$	$n$	Exact value	Bias1/ MSE1	Bias2/ MSE2	Bias3/ MSE3	Bias4/ MSE4
(0.1,4.5)	30		0.54832/0.325730	0.49378/0.28433	0.51636/ 0.30243	0.20438/0.17744
	35		0.549833/0.32429	0.50497/0.28408	0.52796/ 0.30195	0.23830/0.16294
	40	-0.50468	0.55001/0.32373	0.50943/ 0.28386	0.52995/0.29999	0.28788/0.16066
	50		0.55571/0.32236	0.51445/0.28234	0.53018/0.29798	0.35550/0.17280
	60		0.55720/0.32132	0.52253/ 0.28551	0.53218/0.29726	0.38246/0.17472
(0.2,4)	30		0.52237/0.33138	0.43733/0.27303	0.47260/0.29833	0.14750/0.25057
	35		0.53415/0.32091	0.45608/0.26707	0.48867/0.28337	0.22450/0.21577
	40	-0.53930	0.53631/0.31804	0.47334/0.26538	0.49120/0.27857	0.26394/0.20544
	50		0.53835/0.31399	0.48180/0.26242	0.49168/0.27427	0.32690/0.19580
	60		0.54001/ 0.31037	0.48255/ 0.25854	0.49838/0.27144	0.36282/0.17503
(0.3,3.9)	30		0.66016/0.59085	0.58444/0.46060	0.60264/0.49963	0.28883/0.38595
	35		0.67913/0.52852	0.60971/0.44713	0.61856/0.46831	0.37049/0.37528
	40	-0.71898	0.68008/ 0.50869	0.61383/ 0.43901	0.63064/0.45793	0.42447/0.34087
	50		0.68175/0.49875	0.62145/ 0.43391	0.64163/0.45061	0.49303/0.33731
	60		0.68447/0.49486	0.62355/0.42886	0.64253/0.44791	53037/0.33580

The nonparametric estimators of the distribution function are occasionally considered as plotting positions because they supply the ordinate values in plotting the distribution function.

**Table 2:** Bias and MSE of  $ICPT^1(X; t_1, t_2)$ ,  $ICPT^2(X; t_1, t_2)$ ,  $ICPT^3(X; t_1, t_2)$  and  $ICPT^4(X; t_1, t_2)$  for  $\alpha = 0.2$  and different truncation limits ( $n = 30, 35, 40, 50, 60$ ).

$\alpha = 0.2$			$ICPT^1(X; t_1, t_2)$	$ICPT^2(X; t_1, t_2)$	$ICPT^3(X; t_1, t_2)$	$ICPT^4(X; t_1, t_2)$
$(t_1, t_2)$	$n$	Exact value	Bias1/ MSE1	Bias2/ MSE2	Bias3/ MSE3	Bias4/ MSE4
(0.1,4.5)	30		0.01443/0.62585	0.20562/0.85397	0.15292/0.84328	0.31217/0.46912
	35		-0.08361/ 0.40997	0.10770/0.63329	0.03216/0.50496	0.19145/0.31041
	40	3.81827	-0.13264/0.38650	0.01795/0.46097	0.00052/0.41305	0.12438/0.27354
	50		-0.2265/0.21496	-0.12186/0.27259	-0.16442/0.23459	0.00011/0.15006
	60		-0.26605/0.17343	-0.19538/0.20058	-0.20950/0.18536	-0.10140/0.11435
(0.2,4)	30		0.0668/ 0.45989	0.21781/0.60482	0.13608/ 0.55571	0.33128/0.50470
	35		-0.02364/0.25232	0.091067/0.42669	0.07336/0.3542	0.17299/0.21602
	40	3.19537	-0.04258/0.23084	0.02478/0.31252	0.00027/0.26733	0.11015/0.19296
	50		-0.14634/ 0.12285	-0.05928/0.15386	-0.07653/0.13229	0.01410/0.11127
	60		-0.15977/0.09937	-0.10696/0.13898	-0.14774/0.10505	-0.053750/0.07863
(0.3,3.9)	30		0.06254/0.52412	0.15421/0.42303	0.10033/0.45151	0.21734/0.35308
	35		0.03341/0.27368	0.04638/ 0.23737	0.03507/0.26058	0.14486/0.22552
	40	3.04031	-0.09178/0.21888	0.02148/0.15290	0.01607/0.18283	0.07261/0.20835
	50		-0.15733/0.17706	-0.10894/0.11774	-0.11561/0.13794	0.04314/0.09471
	60		-0.18802/0.10640	-0.15383/0.09217	-0.15181/0.10345	-0.10009/0.07318



**Figure 1:** Graphical showing of the MSE of four estimators. Sample size for fixed truncation limit (0.2, 4). (I) Plot of the MSE for fixed truncation limit (0.2, 4) and  $\alpha = 0.2$  and (II) Plot of the MSE for fixed truncation limit (0.2, 4) and  $\alpha = 3.5$ .

## 6. REAL DATA

In this part, an actual life data set is examined to illustrate the applicability and usefulness of the best-proposed estimator of ICPT in actual status. For this purpose, we have taken into account the data set vinyl chloride acquired from clean upgradient groundwater monitoring wells [2]. Vinyl chloride is an organic compound that is unstable. In environmental investigations, this aspect is of extraordinary significance due to the fact that it is both anthropogenic and carcinogenic. Nonetheless, in lots of background monitoring wells, low levels of this component are determined. This compound low surface detections in clean upgradient background monitoring wells is because of cross pollution from air or gas or the analytic system itself. The data set is provided as follows. Data Set (g/l) : 5.1, 1.2, 1.3, 0.6, 0.5, 2.4, 0.5, 1.1, 8.0, 0.8, 0.4, 0.6, 0.9, 0.4, 2.0, 0.5, 5.3, 3.2, 2.7, 2.9, 2.5, 2.3, 1.0, 0.2, 0.1, 0.1, 1.8, 0.9, 2.0, 4.0, 6.8, 1.2, 0.4, 0.2 has been fitted with exponential distribution by [21]. They acclaimed that this data set follows  $Exp(0.5320814)$  (exponential distribution). To examine the behavior of the ICPT, we have calculated estimated values of  $ICPT^4(X; t_1, t_2)$  by means of the use of its best-proposed estimator for different trunca-

tion limits and  $\alpha = 0.2, 3.5$  as shown in Table 3. It has been determined that the estimated values are decreasing in  $t_1$  and increasing in  $t_2$  for  $\alpha = 0.2, 3.5$ . So, by increasing (decreasing) the fourth estimator of ICPT(doubly truncated CPT), the amount of the dispersion of vinyl chloride obtained from clean upgradient groundwater monitoring wells increases (decreases). As expected for  $0 < \alpha = 1$ , ICPT is an increasing function of the interval. It is worth noting that this result is according to the monotonicity of  $ICPT(X; t_1, t_2)$  for  $Exp(0.5320814)$  and  $\alpha = 0.3, 1.5$ .

**Table 3:** Kernel estimates of  $ICPT^4(X; t_1, t_2)$  for the Vinyl chloride data for different truncation limits  $(t_1, t_2)$  and  $\alpha = 0.2; 3.5$ .

$\alpha \setminus (t_1, t_2)$	(0.4,2.9)	(0.6,2.9)	(0.8,2.9)	(1,2.9)	(0.2,1.8)	(0.2,2)	(0.2,2.4)	(0.2,2.8)
0.2	2.035697	1.84515	1.658554	1.47661	1.108958	1.073237	1.576643	1.94168
3.5	-0.419520	-0.508537	-0.641847	-0.919710	-1.837256	-1.312966	-0.754232	-0.437168

## 7. CONCLUSION

In information theory and also in reliability, there are several uncertainty measures that play a central role. In this paper, we first studied the notion of doubly truncated (interval) Tsallis entropy and suggested the doubly truncated (interval) cumulative residual Tsallis entropy (ICRT) and doubly truncated (interval) cumulative past Tsallis entropy (ICPT) whose some of their properties and their relations with hazard rate (reversed hazard rate) and mean residual (past) life were studied. Also, we introduced ordering classes for ICRT and ICPT and gave some characterization. In the end, we have proposed four nonparametric estimators and compared their performance by utilizing simulation data. Also, based on the best-proposed estimator, an actual data set was additionally examined.

## REFERENCES

- [1] Barlow R.E. and Proschan F. Statistical theory of reliability and life testing: probability models, Florida State Univ Tallahassee,(1975).
- [2] Bhaumik, D. K., Kapur, K., and Gibbons, R. D. (2009). Testing parameters of a gamma distribution for small samples. *Technometrics*, 51(3):326–334.
- [3] Ebrahimi, N. (1996). How to measure uncertainty in the residual life time distribution. *Sankhy: The Indian Journal of Statistics, Series A*, 48–56.
- [4] Gupta, R. D. and Nanda, A. K. (2002).  $\alpha$ - and  $\beta$ -entropies and relative entropies of distributions. *Journal of Statistical Theory and Applications*, 1(3):177–190.
- [5] Khorashadizadeh, M., Rezaei Roknabadi, A. H. and Mohtashami Borzadaran, G. R. (2013). Doubly truncated (interval) cumulative residual and past entropy. *Statistics & Probability Letters*, 83(5):1464–1471.
- [6] Kumar, V. and Taneja, H. C. (2011). A generalized entropy-based residual lifetime distributions. *International Journal of Biomathematics*, 4(02):171–184.
- [7] Kundu, C. and Singh, S. (2020). On generalized interval entropy. *Communications in Statistics-Theory and Methods*, 49(8):1989–2007.
- [8] Kumar, V. (2017). Characterization results based on dynamic Tsallis cumulative residual entropy. *Communications in Statistics-Theory and Methods*, 46(17):8343–8354.
- [9] Lutz, E. (2003). Anomalous diffusion and Tsallis statistics in an optical lattice. *Physical Review A*, 67(5):051402.
- [10] Nanda, A. K. and Paul, P. (2006). Some results on generalized residual entropy. *Information Sciences*, 176(1):27–47.
- [11] Navarro, J. and Ruiz, J. M. (1996). Failure-rate functions for doubly-truncated random variables. *IEEE Transactions on Reliability*, 45(4):685–690.
- [12] Navarro, J., and Rubio, R. (2011). A note on necessary and sufficient conditions for ordering properties of coherent systems with exchangeable components. *Naval Research Logistics (NRL)*, 58(5):478–489.

- [13] Nourbakhsh M. and Yari G. Doubly truncated generalized entropy, In Proceedings of the 1st International Electronic Conference Conference on Entropy and its Applications, 3-21 November 2014,
- [14] Misagh, F. (2012). Some Properties of Interval Entropy Function and their Applications. *World Applied Sciences Journal*, 20(12):1666–1671.
- [15] Moharana, R. and Kayal, S. (2020). Properties of Shannon Entropy for Double Truncated Random Variables and its Applications. *Journal of Statistical Theory and Applications*, 19(2):261–273.
- [16] Mohamed, M. S. (2020). On Cumulative Tsallis Entropy and Its Dynamic Past Version. *Indian Journal of Pure and Applied Mathematics*, 51(4):1903–1917.
- [17] Moharana, R. and Kayal, S. (2019). On shift-dependent generalized entropies for doubly truncated random variable. *Journal of Statistics and Management Systems*, 22(5):923–942.
- [18] Rao, M., Chen, Y., Vemuri, B. C. and Wang, F. (2004). Cumulative residual entropy: a new measure of information. *IEEE Transactions on Information Theory*, 50(6):1220–1228.
- [19] Sati, M. M. and Gupta, N. (2015). Some characterization results on dynamic cumulative residual Tsallis entropy. *Journal of Probability and Statistics*, 8 pages, 287–294.
- [20] Shaked M. and Shanthikumar J.G. (Eds.). *Stochastic orders*, New York, NY: Springer New York, 2007.
- [21] Shanker, R., Hagos, F., and Sujatha, S. (2015). On modeling of Lifetimes data using exponential and Lindley distributions. *Biometrics & Biostatistics International Journal*, 2(5):1–9.
- [22] Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423.
- [23] Tong, S., Bezerianos, A., Paul, J., Zhu, Y. and Thakor, N. (2002). Nonextensive entropy measure of EEG following brain injury from cardiac arrest. *Physica A: Statistical Mechanics and its Applications*, 305(3-4):619–628.
- [24] Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, 52(1):479–487.
- [25] Tsallis, C. and Brigatti, E. (2004). Nonextensive statistical mechanics: A brief introduction. *Continuum Mechanics and Thermodynamics*, 16(3):223–235.
- [26] Zacks S. *Introduction to Reliability Analysis Probability Models and Methods*, Springer - Verlag, New York, 1992.



# GENERALIZED X-EXPONENTIAL BATHTUB SHAPED FAILURE RATE DISTRIBUTION AND ESTIMATION OF RELIABILITY OF MULTICOMPONENT STRESS- STRENGTH

Faryal Shabbir, Abdul Khalique

•

Department of Statistics National College  
of Business administration and Economics Lahore, Pakistan

[faryalshab4@gmail.com](mailto:faryalshab4@gmail.com)

[a.khalique57@gmail.com](mailto:a.khalique57@gmail.com)

## Abstract

*In an engineering setup, one is interested to know and determine the reliability of the system of different components. These components are usually subjected to different kinds of stress, and the reliability of the components needs to be estimated under stress. In this paper, we aim to estimate the reliability of a multicomponent stress-strength model assuming that the components of the system are working independently with a common life distribution. The system follows a comparatively new distribution named as; Generalized X-Exponential bathtub failure rate distribution. This paper studies the usefulness of this distribution in terms of estimating the maximum likelihood estimate of the reliability parameter and its asymptotic confidence intervals. Paper uses methods of parametric estimation and reliability estimation. Results are computed using Monte Carlo simulation for small samples. Real data set is presented to evaluate the performance of Generalized X Exponential Distribution (GXED) reliability estimator. Findings show that with the usage of proposed distribution, estimator of reliability parameter fits very well to the real-world situations*

**Key words:** Generalized X -Exponential distribution, Multicomponent stress-strength, Reliability, ML estimation, Average variance, Confidence intervals.

## I. Introduction

The X-Exponential distribution was introduced by Chacko [4], to add another model to the class of bathtub type failure rate distributions. When  $x$  is X-Exponential with parameters  $\alpha$  and  $\lambda$ . It has distribution function:  $F(x) = (1 - (1 + \lambda x^2)e^{-\lambda x})^\alpha$  with the corresponding density function:  $f(x) = \alpha e^{-\lambda x}(\lambda^2 x^2 - 2\lambda x + 1)(1 - (1 + \lambda x^2)e^{-\lambda x})^{\alpha-1}$ . Its properties and reliability applications were studied by the author. However, in order to get more flexibility to the model, Chacko and Deepthi [5] made a small change in the exponential part. The corresponding distribution is named as Generalized X-Exponential distribution. Basically, bathtub failure rate distribution's curve illustrates three phases of a product's life. First phase is known as early failure, next is a roughly prolonged intrinsic period and failure rate is approximately constant here. This stage is very important for reliability prediction of a product. And finally, there is a wear out failure phase, where failure rate increases. In the past several bathtub failure rate distributions have been studied by Kundu & Gupta, Srinivasa Rao [11] to carry out reliability testing by using single component stress

strength, as well as multi- component stress strength models. Since no substantial work has been done on reliability estimation of multicomponent stress strength by using a flexible distribution i.e., GXED, hence there was a need to study the reliability estimator of newly introduced Generalized X-Exponential distribution having distribution function,  $F(x) = \left(1 - (1 + \lambda x^2)e^{-\lambda(x^2+x)}\right)^\alpha$ ,  $x > 0, \lambda > 0$  and  $\alpha > 0$ .and the density function is:

$$f(x) = \alpha e^{-\lambda(x^2+x)} (\lambda(1 + \lambda x^2)(2x + 1) - 2\lambda x) \left(1 - (1 + \lambda x^2)e^{-\lambda(x^2+x)}\right)^{\alpha-1}; \alpha > 0, \lambda > 0 \tag{1}$$

$$\text{Failure rate} = \frac{\alpha e^{-\lambda(x^2+x)} (\lambda(1+\lambda x^2)(2x+1)-2\lambda x) \left(1 - (1+\lambda x^2)e^{-\lambda(x^2+x)}\right)^{\alpha-1}}{1 - \left(1 - (1+\lambda x^2)e^{-\lambda(x^2+x)}\right)^\alpha}; \quad x > 0, \alpha > 0, \lambda > 0 \tag{2}$$

The authors (Chacko and Deepthi) have investigated the properties and some reliability applications of the new model. Here we are interested in the reliability analysis of multicomponent system where the components are connected in parallel and function independently, with the same Generalized X -Exponential distribution GXED and stress too has the same distribution but with different parameters.

Let the random samples  $Y, X_1, X_2, X_3, \dots, X_K$  be independent,  $G(y)$  be the continuous distribution function of  $Y$ , and  $F(x)$  be the common distribution function of  $Y, X_1, X_2, X_3, \dots, X_K$ .The reliability in a multi component stress-strength model developed by Bhattacharyya and Johnson [2] is given by.

$R_{s,k} = P$  [at least  $s$  of the  $X_1, X_2, X_3, \dots, X_K$  exceed  $Y$ ]

$$= \sum_{i=s}^k \binom{k}{i} \int_{-\infty}^{+\infty} [1 - F(y)]^i [F(y)]^{k-i} dG(y) \tag{3}$$

Where  $X_1, X_2, X_3, \dots, X_K$  identically and independently distributed (iid) are with common distribution function  $F(x)$  and subjected to random stress  $Y$ . The probability in (3) is called ‘Reliability in a multicomponent stress –strength model’ Bhattacharyya and Johnson [2]. The survival probabilities of single component stress- strength version was considered by several authors for different distributions. Some of them are: Enis and Geisser [9], Downtown [8], Awad and Gharraf [1], McCool [18], Hanagal [12], Nandi and Aich [19], Surles and Padgett [27], Kundu and Gupta [15,16], Raqab et al. [26] and Kundu and Raqab [17]. More over Kotz & Pensky [14] studied the generalizations of stress strength model.

Reliability in a multicomponent stress-strength model was developed by Bhattacharyya and Johnson [2]. Pandey & Burhan [21] computed the estimation of reliability for a multicomponent model using Burr distribution. Zimmer et al [29] studied the reliability analysis for Burr X11 distribution. Estimation of reliability in models with correlated stress and strength has been studied by Balakrishnan & Lai [3]. Rao and Kantam [24] studied the estimation of reliability in a multicomponent stress- strength model for logistic distribution, Rao [23] also developed the procedure for the estimation of reliability in multicomponent stress-strength model based on Generalized exponential distribution. Ghitany et al. [10] studied the estimation of reliability of multicomponent model using Power Lindley distribution. Burr-X11 distribution for parametric and reliability estimation in a multicomponent stress-strength environment has been analyzed by Rao et al. [25]. Dey, S. et al [6] considered Bayesian and non-Bayesian estimation of multicomponent stress-strength reliability using Kumaraswami distribution.

Dey, Raheem & Mukherjee [7] derived the form of stress-strength reliability parameter for transmuted Rayleigh distribution. Hassan [13] developed the procedure for the estimation of stress-strength model using Lindley distribution. Estimation on Reliability in a multicomponent Stress-strength model with Power Lindley distribution is carried out by Abbas Pak et al [22]. Similarly, a recent study has been conducted on the estimation of stress strength reliability for Akash distribution by Akhila. K. Varghese & V. M. Chacko [28].

The aim of this paper is to estimate the reliability in a multi component stress-strength model based on  $X, Y$  being two independent random variables, where  $X \sim GXED, (\alpha_1, \lambda)$  and  $Y \sim GXED (\alpha_2, \lambda)$ . We use parametric estimation and estimation reliability. Suppose a system with  $k$  identical components, functions if at least  $s$  ( $1 \leq s \leq k$ ) components operate simultaneously. In its operating environment, the system is subjected to stress  $Y$  which is a random variable with distribution function  $G(\cdot)$ . The strengths of the components, that is the minimum stresses causing failure, are independently and identically distributed random variables with distribution function  $F(\cdot)$ . The reliability of the system can be obtained by (3). An attempt has been made here to study the estimation of reliability in a multicomponent stress-strength model with reference to two parameter  $GXED$ .

The remainder of the paper is organized as follows. In section 2, research methodology and procedure for expression of  $R_{s,k}$ . The asymptotic distribution and confidence interval of (3) are calculated using *MLE*. The results of small sample comparisons derived from Monte Carlo simulations and analysis of real data sets are described in section 3. Findings are discussed in section 4.

## 2. Maximum Likelihood Estimator of $R_{s,k}$

Let  $X \sim GXED (\alpha_1, \lambda)$  and  $Y \sim GXED (\alpha_2, \lambda)$  be independently distributed with unknown shape parameters  $(\alpha_2, \lambda)$  while common scale parameter  $\lambda$ . Using (3) the reliability in multicomponent stress-strength for two-parameter  $GXED$  distribution is as follows:

$$R_{s,k} = \sum_{i=s}^k \binom{k}{i} \int_0^{+\infty} [1 - F(y)]^i [F(y)]^{k-i} dG(y)$$

$$F(y) = \left(1 - (1 + \lambda x^2)e^{-\lambda(x^2+x)}\right)^\alpha; \quad x > 0, \alpha > 0, \lambda > 0$$

$$1 - F(y) = 1 - \left(1 - (1 + \lambda x^2)e^{-\lambda(x^2+x)}\right)^\alpha$$

$$dG(y) = \alpha e^{-\lambda(y^2+y)} (\lambda(1 + \lambda y^2)(2y + 1) - 2\lambda y) \left(1 - (1 + \lambda y^2)e^{-\lambda(y^2+y)}\right)^{\alpha-1} dy$$

$$R_{s,k} = \sum_{i=s}^k \binom{k}{i} \nu \int_0^1 (1-t)^i t^{k-1+\nu-1} dt$$

$$\text{where } t = 1 - \left(1 - (1 + \lambda x^2)e^{-\lambda(y^2+y)}\right)^\alpha \text{ and } \nu = \frac{\alpha_2}{\alpha_1}$$

After simplification we get

$$R_{s,k} = \sum_{i=s}^k \binom{k}{i} \nu B(i + 1, k - i + \nu) \tag{4}$$

The probability in (4) is termed reliability in a multicomponent stress-strength model. It is important to mention here that MLE of  $R_{s,k}$  depends on that of  $\alpha_1$  &  $\alpha_2$ . Hence, we need to calculate MLE of the latter to derive that of the former. Similarly, to find the MLE of  $\alpha_1$  &  $\alpha_2$  and we need to

find the MLE of  $\lambda$  as well. Here we assume that  $X_1, X_2, X_3 \dots \dots X_n$  is a random sample from GXED  $(\alpha_1, \lambda)$  and  $Y_1, Y_2, Y_3, \dots \dots Y_n$  is a random sample from GXED  $(\alpha_2, \lambda)$ .

The loglikelihood function *LLF* of these samples is expressed as:

$$L(\alpha_1, \alpha_2, \lambda) = m \ln \alpha_1 + n \ln \alpha_2 - (m + n) \lambda (x_i^2 + x_i + y_j^2 + y_j) + m \ln \sum (\lambda(1 + \lambda x_i^2)(2x_i + 1) - 2\lambda x_i) + (\alpha_1 - 1) \sum (\ln(1 - (1 + \lambda x_i^2)e^{-\lambda(x_i^2 + x_i)})) + n \ln \sum (\lambda(1 + \lambda y_j^2)(2y_j + 1) - 2\lambda y_j) + (\alpha_2 - 1) \sum (\ln(1 - (1 + \lambda y_j^2)e^{-\lambda(y_j^2 + y_j)})) \tag{5}$$

Thus, the *MLE* of  $\lambda$  is the solution of

$$\frac{\partial \log L(\alpha_1, \alpha_2, \lambda)}{\partial \lambda} = 0 \Rightarrow - \sum_{i=1}^m (x_i^2 + x_i) + \sum_{i=1}^m \frac{((2x_i+1)(1+2\lambda x_i^2)-2x_i)}{(\lambda(1+\lambda x_i^2)(2x_i+1)-2\lambda x_i)} + (\alpha_1 - 1) \sum_{i=1}^m \frac{(1+\lambda x_i^2)e^{-\lambda(x_i^2+x_i)}(x_i^2+x_i)-e^{-\lambda(x_i^2+x_i)}x_i^2}{(1-(1+\lambda x_i^2)e^{-\lambda(x_i^2+x_i)})} - \sum_{j=1}^n (y_j^2 + y_j) + \sum_{j=1}^n \frac{((2y_j+1)(1+2\lambda y_j^2)-2y_j)}{(\lambda(1+\lambda y_j^2)(2y_j+1)-2\lambda y_j)} + (\alpha_2 - 1) \sum_{j=1}^n \frac{(1+\lambda y_j^2)e^{-\lambda(y_j^2+y_j)}(y_j^2+y_j)-e^{-\lambda(y_j^2+y_j)}y_j^2}{(1-(1+\lambda y_j^2)e^{-\lambda(y_j^2+y_j)})} = 0 \tag{6}$$

Similarly, the *MLE* of  $\alpha_1$  can be obtained as the solution of

$$\frac{\partial \log L(\alpha_1, \alpha_2, \lambda)}{\partial \alpha_1} = 0 \Rightarrow \frac{m}{\alpha_1} + \sum_{i=1}^m \log(1 - (1 + \lambda x_i^2)e^{-\lambda(x_i^2 + x_i)}) = 0 \tag{7}$$

Also, for  $\alpha_2$

$$\frac{\partial \log L(\alpha_1, \alpha_2, \lambda)}{\partial \alpha_2} = 0 \Rightarrow \frac{n}{\alpha_2} + \sum_{j=1}^n \log(1 - (1 + \lambda y_j^2)e^{-\lambda(y_j^2 + y_j)}) = 0 \tag{8}$$

From (7) and (8) we obtain:

$$\alpha_1^{\wedge}(\lambda) = \frac{-m}{\sum_{i=1}^m \log(1 - (1 + \lambda x_i^2)e^{-\lambda(x_i^2 + x_i)})} \text{ and } \alpha_2^{\wedge}(\lambda) = \frac{-n}{\sum_{j=1}^n \log(1 - (1 + \lambda y_j^2)e^{-\lambda(y_j^2 + y_j)})} \tag{9}$$

Putting the values of  $\alpha_1^{\wedge}(\lambda)$  and  $\alpha_2^{\wedge}(\lambda)$  into equation (6), we got a function of  $\lambda$  which is nonlinear.

$$h(\lambda) = \lambda \tag{10}$$

$$\frac{\sum_{i=1}^m \frac{4\lambda x_i^3 + 2\lambda x_i^2 + 1}{2\lambda x_i^2 + 2x_i^2 + 1} + \sum_{j=1}^n \frac{4\lambda y_j^3 + 2\lambda y_j^2 + 1}{2\lambda y_j^2 + 2y_j^2 + 1}}{\sum_{i=1}^m (x_i^2 + x_i) + \sum_{j=1}^n (y_j^2 + y_j) + \frac{m}{\sum_{i=1}^m \log(1 - (1 + \lambda x_i^2)e^{-\lambda(x_i^2 + x_i)})} \sum_{i=1}^m \frac{x_i(1 + \lambda x_i^3 + \lambda x_i^2) \cdot e^{-\lambda(x_i^2 + x_i)}}{(1 - (1 + \lambda x_i^2)e^{-\lambda(x_i^2 + x_i)})} + \frac{n}{\sum_{j=1}^n \log(1 - (1 + \lambda y_j^2)e^{-\lambda(y_j^2 + y_j)})} \sum_{j=1}^n \frac{y_j(1 + \lambda y_j^3 + \lambda y_j^2) \cdot e^{-\lambda(y_j^2 + y_j)}}{(1 - (1 + \lambda y_j^2)e^{-\lambda(y_j^2 + y_j)})} + \frac{n}{\sum_{j=1}^n \log(1 - (1 + \lambda y_j^2)e^{-\lambda(y_j^2 + y_j)})} \sum_{i=1}^m \frac{x_i(1 + \lambda x_i^3 + \lambda x_i^2) \cdot e^{-\lambda(x_i^2 + x_i)}}{(1 - (1 + \lambda x_i^2)e^{-\lambda(x_i^2 + x_i)})}} \tag{11}$$

Here  $\lambda^{\wedge}$  is a fixed-point solution of nonlinear equation (10). It can be obtained using a simple iterative procedure:

$$h_{\lambda(j)} = \lambda(j + 1) \tag{12}$$

Where  $\lambda_j$  is the  $j^{th}$  iteration of  $\lambda^{\wedge}$ . During the simulation process, when the difference between  $\lambda_j$  and  $\lambda(j + 1)$  becomes sufficiently small; then we stop the iterative process. Once we obtain  $\lambda^{\wedge}$ , the parameters  $\alpha_1^{\wedge}$  and  $\alpha_2^{\wedge}$  can be obtained from (9) as respectively. To obtain the asymptotic confidence interval for  $R_{s,k}$  we proceed as follows.

### 2.1 Asymptotic Variance and Confidence Intervals

$$V(\alpha_1^{\wedge}) = [E(-\partial^2 L / \partial \alpha_1^2)]^{-1} = \frac{\alpha_1^2}{m} \text{ and } V(\alpha_2^{\wedge}) = [E(-\partial^2 L / \partial \alpha_2^2)]^{-1} = \frac{\alpha_2^2}{n} \tag{13}$$

The asymptotic variance AV of an estimate of  $R_{s,k}$  which is a function of two independent statistics  $\alpha_1^{\wedge}, \alpha_2^{\wedge}$  is established by Rao (1973):

$$AV(R_{s,k}^{\wedge}) = V(\alpha_1^{\wedge}) \left(\frac{\partial R_{s,k}}{\partial \alpha_1}\right)^2 + V(\alpha_2^{\wedge}) \left(\frac{\partial R_{s,k}}{\partial \alpha_2}\right)^2 \tag{14}$$

Thus from (14), asymptotic variance in  $R_{s,k}$  can be obtained for GXED.

We obtain  $R_{s,k}$  and their derivatives for  $(s, k) = (1, 3)$  and  $(2, 4)$  separately:

$$R_{1,3}^{\wedge} = \frac{3v^2 + 9v + 6}{(v + 1)(v + 2)(v + 3)} \text{ and } R_{2,4}^{\wedge} = \frac{12(v^2 + 3v + 2)}{(v + 1)(v + 2)(v + 3)(v + 4)}$$

$$\frac{\partial R_{1,3}^{\wedge}}{\partial \alpha_1} = \frac{3v(v^4 + 6v^3 + 13v^2 + 12v + 4)}{\alpha_1[(v + 1)(v + 2)(v + 3)]^2}$$

$$\frac{\partial R_{1,3}^{\wedge}}{\partial \alpha_2} = \frac{-3v(v^4 + 6v^3 + 13v^2 + 12v + 4)}{\alpha_1[(v + 1)(v + 2)(v + 3)]^2}$$

$$\frac{\partial R_{2,4}^{\wedge}}{\partial \alpha_1} = \frac{12v(2v^5 + 19v^4 + 68v^3 + 115v^2 + 92v + 28)}{\alpha_1[(v + 1)(v + 2)(v + 3)(v + 4)]^2} \text{ and}$$

$$\frac{\partial R_{2,4}^{\wedge}}{\partial \alpha_2} = \frac{-12(2v^5 + 19v^4 + 68v^3 + 115v^2 + 92v + 28)}{\alpha_1[(v + 1)(v + 2)(v + 3)(v + 4)]^2}$$

Therefore as  $n \rightarrow \infty$  and  $m \rightarrow \infty$ ,  $(R_{s,k}^{\wedge} - R_{s,k}) / AV(R_{s,k}^{\wedge}) \sim N(0, 1)$

$$AV(R_{1,3}^{\wedge}) = \frac{9v^2(v^4 + 6v^3 + 13v^2 + 12v + 4)^2(1/m + 1/n)}{[(v + 1)(v + 2)(v + 3)]^4}$$

$$\text{and } AV(R_{2,4}^{\wedge}) = \frac{144v^2(2v^5 + 19v^4 + 68v^3 + 115v^2 + 92v + 28)^2(1/m + 1/n)}{[(v + 1)(v + 2)(v + 3)(v + 4)]^4}$$

Where  $R_{s,k}^{\wedge} \mp 1.96\sqrt{AV(R_{s,k}^{\wedge})}$  is the asymptotic 95% confidence interval (C.I) of system reliability  $R_{s,k}$  and asymptotic 95% C.I for  $R_{1,3}$  is given by:

$$R_{1,3}^{\wedge} \mp 1.96 \frac{3v(v^4 + 6v^3 + 13v^2 + 12v + 4)\sqrt{1/m + 1/n}}{[(v + 1)(v + 2)(v + 3)]^2}$$

and the asymptotic 95% confidence interval (C.I) for  $R_{2,4}$  is given by:

$$R_{2,4}^{\wedge} \mp 1.96 \frac{12v(2v^5 + 19v^4 + 68v^3 + 115v^2 + 92v + 28)\sqrt{1/m + 1/n}}{[(v + 1)(v + 2)(v + 3)(v + 4)]^2}$$

## 3. Simulation Study

### 3.1 Results

5000 random samples are generated each of size 10(5)30 from stress and strength populations for different values of  $\alpha_1$  and  $\alpha_2$ : (2.0,2.5), (2.0,3.0), (2.0,3.5), (3.0,2.0), (3.0,2.5), (3.0,3.0). The MLE of scale parameter  $\lambda$  is estimated by the iterative method and using  $\lambda$  the shape parameters  $\alpha_1$  and  $\alpha_2$  are estimated from eq (8).

These ML estimators of  $\alpha_1$  and  $\alpha_2$  are then substituted in  $v$  to obtain the multicomponent reliability for  $(s, k) = (1,3)$  and  $(2,4)$ . The average bias and average MSE of reliability estimate over 5000 replications are presented in Table 1 and Table 2. Average length of confidence interval and coverage probability of the simulated 95% CIs of  $R_{s,k}$  are given in Table 3 and Table 4. The true values of reliability in multicomponent stress-strength with given combinations of  $\alpha_1, \alpha_2$  for  $(s, k) = (1,3)$  are 0.7058824, 0.6666667, 0.6315789, 0.8181074, 0.7826768, 0.75, 0.7142857 and for  $(s, k) = (2,4)$  are 0.5378151, 0.4848485, 0.4393593, 0.7011849, 0.6477772, 0.6, 0.5494505.

Here it is seen that the true value of reliability in multicomponent stress-strength decreases as  $\alpha_2$  is increased for a fixed value of  $\alpha_1$ , whereas reliability in multicomponent stress-strength also decreases as  $\alpha_1$  is increased for a fixed value of  $\alpha_2$ . Thus, the true value of reliability increases as  $v$  decreases and vice versa.

**Table 1:** Average bias of the simulated estimates of  $R_{s,k}(\alpha_1, \alpha_2)$

s, k	n, m	2.0,2.5	2.0,3.0	2.0,3.5	3.0,2.0	3.0,2.5	3.0,3.0
1,3	<b>10,10</b>	-.006581	-.0016484	-.007107	-.010848	-.017399	-.061099
	<b>15,15</b>	-.006425	-.0041552	-.003801	-.005277	-.005233	-.056924
	<b>20,20</b>	-.005644	-.0009021	-.002751	-.003377	-.003877	-.055159
	<b>25,25</b>	-.004301	-.0035697	-.002016	-.003189	-.003199	-.055473
	<b>30,30</b>	-.003075	-.0032642	-.001574	-.003726	-.002866	-.055240
2,4	<b>10,10</b>	-.003129	-.0009622	0.000633	-.011485	-.007666	-.008216
	<b>15,15</b>	-.005138	-.0022865	-.001908	-.003719	-.008930	-.005695
	<b>20,20</b>	0.000287	0.0005562	-.000367	-.006453	-.005077	-.004446
	<b>25,25</b>	-.000917	-.0003678	-.001444	-.001709	-.004725	-.005074
	<b>30,30</b>	-.000523	-.0020488	-.003957	-.002097	-.004196	-.003842

Results of Table 1 and Table 2 depicts that average bias and MSE decrease as sample size increases for both the cases of estimation of reliability. Bias is negative in all the combinations of parameters in both situations of  $(s, k)$ . This shows the consistency of MSE. Also, absolute bias increases as  $\alpha_1$  increases for a fixed value of  $\alpha_2$ . While MSE decreases as  $\alpha_1$  increases for a fixed value of  $\alpha_2$  for both the cases of  $(s, k)$ . Also, for fixed  $\alpha_1$  and increasing  $\alpha_2$  MSE increases for same sample.

**Table 2:** Average MSE of the simulated estimates of  $R_{s,k}(\alpha_1, \alpha_2)$

s, k	n, m	2.0,2.5	2.0,3.0	2.0,3.5	3.0,2.0	3.0,2.5	3.0,3.0
1,3	<b>10,10</b>	.008420	.008347	.0105324	.005115	.005915	.010999
	<b>15,15</b>	.005872	.006666	.0071572	.002870	.004008	.008089
	<b>20,20</b>	.0049068	.004907	.0054370	.002320	.003011	.006478
	<b>25,25</b>	.0036479	.004291	.0045213	.001871	.002486	.005984
	<b>30,30</b>	.002805	.003195	.0037037	.001478	.002033	.005489
2,4	<b>10,10</b>	.015654	.0154285	.0165471	.010602	.012716	.014210
	<b>15,15</b>	.010693	.010985	.0111963	.004762	.008500	.009423
	<b>20,20</b>	.0075629	.008428	.008489	.004305	.006470	.007210
	<b>25,25</b>	.006857	.006696	.0069927	.004016	.004969	.005663
	<b>30,30</b>	.005238	.005696	.0052105	.003354	.004067	.005053

**Table 3:** Average Length of the simulated 95% confidence intervals of  $R_{s,k}(\alpha_1, \alpha_2)$

s, k	n, m	2.0,2.5	2.0,3.0	2.0,3.5	3.0,2.0	3.0,2.5	3.0,3.0
1,3	10,10	.350894	.378572	.390008	.263240	.299103	.322061
	15,15	.290262	.311732	.323802	.214849	.242167	0.26764
	20,20.	.253883	.272129	.323815	.186494	.210426	.230430
	25,25	.228945	.243448	.254189	.166324	.188198	.206166
	30,30	.208672	.223215	.232022	.150368	.170951	.189016
2,4	10,10	.475256	.483035	.485837	.392269	.428761	.453479
	15,15	.395910	.404009	.405260	.322269	.351963	.373558
	20,20.	.346274	.357539	.354290	.280074	.308866	.327652
	25,25	.309512	.318022	.318480	.250861	.275602	.291288
	30,30	.285038	.291335	.292680	.230350	.252819	.269990

Table 3 and Table 4 findings show that as the sample size increases, length of CI also decreases and coverage probability in most the cases crossing 0.95 and for few it is 0.98, which shows the performance of CI using Generalized X- Exponential Distribution GXED is excellent and it covers most of the cases. Among the parameters, it is observed that length of CI increases for fixed value of  $\alpha_1$  for (1,3) while for fixed value of  $\alpha_2$  length of CI decreases.

**Table 4:** Average Coverage Probability of simulated 95% confidence intervals of  $R_{s,k}(\alpha_1, \alpha_2)$

s, k	n, m	2.0,2.5	2.0,3.0	2.0,3.5	3.0,2.0	3.0,2.5	3.0,3.0
1,3	10,10	.891333	.968000	.936667	.910667	.969333	.987333
	15,15	.972667	.944444	.880000	.950000	.925084	.905333
	20,20.	.905333	.980810	.914000	..968667	.912052	.956667
	25,25	.951333	.912300	.949333	.969333	.946000	.966667
	30,30	.953815	.965333	.926000	.952667	.896360	.936667
2,4	10,10	.964667	.957333	.9743178	.984000	.934667	.966677
	15,15	.962667	.966600	.953000	.970000	.942000	.967333
	20,20.	.963333	.955746	.952667	.969425	.954000	.953333
	25,25	.946000	.936667	.953333	.983333	.970883	.948007
	30,30	.902000	.946666	.937333	.956000	.960667	.960667

### 3.2 Data Analysis

In this section, we will deal with two real data sets, will show how reliability in a multicomponent stress-strength model can be applied for GXED. Both data sets were discussed by Zimmer et al. (1998) and Lio et al. (2010) for Burr-X11 reliability analysis. They showed that Burr-X11 distribution fits quite well. For both the data sets, here we are using GXED.

(X):0.19, 0.78, 0.96, 0.31, 2.78, 3.16, 4.15, 4.67, 4.85, 6.50, 7.35, 8.01, 8.27, 12.06, 31.75, 32.52, 33.91, 36.71 and 72.89

(Y):0.9, 1.5, 2.3, 3.2, 3.9, 5.0, 6.2, 7.5, 8.3, 10.4, 11.1, 12.6, 15.0, 16.3, 19.3, 22.6,

24.8, 31.5 And 53.0. Iterative procedure was used to calculate the value of  $\lambda$  using (8) and then  $\alpha_1$  and  $\alpha_2$  were obtained by substituting the MLE of  $\lambda$  in (10).

The final estimates of  $\alpha_1 = 0.844798$ ,  $\alpha_2 = 1.551717$  and  $\lambda = 0.04642891$ . Based on these estimates the MLE of  $R_{1,3}$  turned out to be 0.620246 and 95% CI (.4704636, .770028) while for  $R_{2,4}$ , came out to be 0.4250596; CI (.2752773,0.5748419).

#### 4. Discussion

In this paper, we analyzed the behavior of Generalized X-Exponential Distribution (GXED) in calculating the multicomponent stress-strength reliability estimates. We also calculated 95% CI & coverage probability for reliability estimates and results were excellent. Coverage probability touched up to 0.98, which shows GXED estimates, very accurately.

The simulation results indicated that average bias and MSE decreased as the sample size increased for both the cases of  $R_{s,k}$ . The real data sets also revealed GXED fits very well and provides quite close results. Hence, GXED can be used readily to calculate the reliability in a multicomponent stress- strength environment.

#### References

- [1] Awad, A.M., & Gharraf, M.K. (1986). Estimation of  $P(Y < X)$  in Burr case: A comparative study. *Communications in Statistics- Simulations and Computations*, 15, 389-403.
- [2] Bhattacharyya, G. K., & Johnson, R. A. (1974). Estimation of reliability in a multicomponent stress strength model. *Journal of the American Statistical Association*, 69(348), 966-970.
- [3] Balakrishnan, N., & Lai, C. D. (2009). *Continuous bivariate distributions*. Springer Science & Business Media.
- [4] Chacko, V. M. (2016). X-Exponential bathtub failure rate model. *Reliability: Theory & Applications*, 11(4 (43)), 55-65.
- [5] Chacko, V. M., & Deepthi, K. S. (2019). Generalized X-Exponential bathtub shaped failure rate distribution. *Journal of the Indian Society for Probability and Statistics*, 20(2), 157-171.
- [6] Dey, S., Mazucheli, J. & Anis, M. (2017). Estimation of reliability of multicomponent stress-strength for a Kumaraswamy distribution. *Communications in Statistics-Theory and Methods*, 46(4), 1560-1572.
- [7] Dey, S., Raheem, E., & Mukherjee, S. (2017). Statistical properties and different methods of estimation of Transmuted Rayleigh distribution. *Revista Colombiana de Estadística* , 40(1),165.
- [8] Downtown, F. (1973). The estimation of  $P(Y > X)$  in the normal case. *Technometrics*, 15, 551-558.
- [9] Enis, P., Geisser, S. (1971). Estimation of the probability that  $(Y < X)$ . *JASA*, 66, 162-168.
- [10] Ghitany, M.E., Al-Mutairi, D. K. & Aboukhamseen, S. M. (2015). Estimation of the reliability of a stress- strength system from Power Lindley distribution. *Communication in Statistics Simulation and Computation*, 44(1), 118-136.
- [11] Gupta, R.D., Kundu, D. (1999) Generalized exponential distributions. *Aust N Z. J Stat*, 41, 173– 188
- [12] Hanagal, D, D. (1997). Note on estimation of reliability under Bivariate Pareto Stress-strength model. *Statistical Papers*, 38, 453-459
- [13] Hassan, M. K. (2017). Estimation a stress-strength model for  $p(y_{r:n_1} < x_{k:n_2})$  using the Lindley distribution. *Revista Colombiana de Estadística*, 40(1), 105-121
- [14] Kotz, S. & Pensky, M. (2003). The stress-strength model and its generalization Theory and Applications. *World Scientific*.
- [15] Kundu, D., Gupta, R.D. (2005). Estimation of  $P(Y < X)$  for the Generalized Exponential distribution. *Metrika*, 61 (3), 291-308.
- [16] Kundu, D., Gupta, R.D. (2006). Estimation of  $P(Y < X)$  for Weibull distribution. *IEEE Transactions on Reliability*, 55 (2), 270-280.
- [17] Kundu, D., Raqab, M. Z. (2009). Estimation of  $R = P(Y < X)$  for three-parameter Weibull distribution., *Statistics and Probability Letters*, 79, 1839-1846.
- [18] McCool, J. I. (1991). Inference on  $P(Y < X)$  in the Weibull case. *Communications in Statistics-Simulations. & Computations*, 20, 129-148.
- [19] Nandi, SB. & Aich, S.B. (1994). A note on the estimation of  $P(X > Y)$  for some distributions useful in life testing. *IAPQR Transactions*, 19(1), 35-44
- [20] Nelson, W. (1982). *Applied Life Data Analysis*, John Wiley and Sons, NY.



- [21] Pandey, M., Uddin, Md. B. (1991). Estimation of reliability in multi-component stress strength model following Burr distribution. *Micro electronics Reliability*, 31 (1), 21-25.
- [22] Pak, A., Gupta, A. K., & Khoolejani, N. B. (2018). On reliability in a multicomponent stress- strength model with Power Lindley distribution. *Revista Colombiana de Estadística*, 41(2), 251-267.
- [23] Rao, G. S. (2012). Estimation of reliability in multicomponent stress-strength model based on Generalized Exponential distribution. *Colombian Journal of Statistic*, 35(1), 67-76.
- [24] Rao, G. S., Kantam, R. L. (2010). Estimation of reliability in multicomponent stress-strength model: Log-Logistic distribution. *Electronic Journal of Applied Statistical Analysis*, 3(2), 75-84.
- [25] Rao, G. S., M, Aslam. & Kundu. (2014). Burr Type X11 distribution parametric estimation of reliability in multicomponent stress-strength model. *Communication in Statistics-Theory and Methods*, 44(23), 4953-4961
- [26] Raqab, M. Z., Madi, M. T., Kundu, D. (2008). Estimation of  $P(Y < X)$  for the 3-parameter Generalized Exponential distribution. *Communication in Statistics-Theory and Methods*, 37 (18), 2854-2864.
- [27] Surles, J. G., Padgett, W. J. (1998). Inference for  $P(Y < X)$  in the Burr Type X model. *Journal of Applied Statistical Sciences*, 7, 225-238.
- [28] Varghese, A. K., & Chacko, V. M. (2022). Estimation of stress-strength reliability for Akash distribution. *Reliability: Theory & Applications*, 17(3 (69)), 52-58.
- [29] Zimmer, W. J., Keats, J. B., Wang, F. K. (1998). The Burr XII distribution in reliability analysis. *Journal of Quality Technology*, 30, 386-394.

# DEVELOPMENT OF AN INTEGRATED SAFETY SYSTEM FOR PRODUCTION FACILITIES: THE PROBLEM STATEMENT AND THE PROPOSED SOLUTION

Evgeny Gvozdev

•

Moscow State University of Civil Engineering  
[evgvozdev@mail.ru](mailto:evgvozdev@mail.ru)

## Abstract

*The article focuses on explosion and fire hazards at production facilities of enterprises where flammable liquids and gases, categorized by explosion and fire risks, are processed, handled, transported, and stored. The goal to be attained and the tasks to be solved towards this end are formulated in the article. Consolidated areas of knowledge, accumulating results of research into risk assessment within systems of integrated safety implemented at production facilities, are considered by the author. A model for development of a novel set of research and methodological instruments (methods, techniques, software and hardware) is presented for its further practical application. The problem of developing integrated safety systems for industrial facilities, posing explosion and fire hazards, as well as the solution, are presented by the author for the first time. The novelty of the solution lies in the computation of validity of the practical application of a novel set of research and methodological instruments. A reduction in damage from accidents and fires at production facilities is demonstrated. Ultimately, the socio-economic problem of reducing damage from accidents and fires is solved not only by Russian production facilities, but also by government agencies, including the EMERCOM of Russia (Ministry of the Russian Federation for Civil Defense, Emergencies and Elimination of Consequences of Natural Disasters), Ministry of Labor and Social Protection of Russia, and Federal Environmental, Industrial and Nuclear Supervision Service of Russia.*

**Keywords:** risk, explosion and fire hazard, integrated safety, integrated safety system, development model, target model

## I. Introduction

Analysis of statistics on accidents and fires at production facilities shows that the share of combined hazardous events (accidents and fires) reaches about 20% of the total number of accidents. Damage from combined events reaches about 46% of the total damage from accidents. Such events may cause injuries and fatalities to in-house personnel and third parties [1-3]. *Explosion and fire hazards arising at production facilities are understood as the state of a facility characterized by the possibility of an explosion or a fire or, alternatively, the occurrence of fire followed by an explosion*<sup>1</sup>. These are the conditions for several types of damage (material and economic damage, calculated in *rubles*; injuries and fatalities, calculated in *units*). Production facilities posing fire and explosion risks (hereinafter -

---

<sup>1</sup> Federal Law No. 123-FZ of 22.07.2008 Technical regulations on fire safety requirements.

PFPFER) are enterprises where flammable liquids and gases are processed, handled, transported, and stored. Such production facilities are categorized by explosion and fire risks<sup>2</sup>; they are categorized according to computations made for production premises and buildings.

Substantial damage deals with combined hazardous events (accidents and fires) resulting from conditions triggering a fire or an explosion at a hydrocarbon processing facility. Thereafter, a secondary factor of explosion or fire is in place, and eventually large volumes of hydrocarbons (hydrocarbon gases) cause destruction and spread over the territory of a production facility (a gas spill). [4]. The article focuses on several types of damage (*economic, material, and social damage*) to the following three subsystems: the occupational safety subsystem (hereinafter - OS); the industrial safety subsystem (hereinafter - InS); the fire safety subsystem (hereinafter - FS), included in the integrated safety system (hereinafter - ISS) at PFPFER. The ISS at PFPFER should be understood as a set of interacting industry-wide subsystems (OS, InS, FS) needed to protect personnel, property, equipment and environment from accidents and fires. Integrated safety (hereinafter - IS) at PFPFER should be understood as industry-wide subsystems (OS, InS, FS) characterized by preventability of hazardous events (accidents and fires) that can damage the assets to be protected. According to item 15 of Article 2 of Federal Law 123-FZ of 22.07.2008 titled Technical fire safety regulations, assets to be protected are products owned by natural persons or legal entities, government agencies or municipalities (including property items located in settlements, as well as buildings, structures, vehicles, process plants, equipment, assemblies, products and other property), that are subject to FS requirements for fire prevention and protection of people in case of fire.

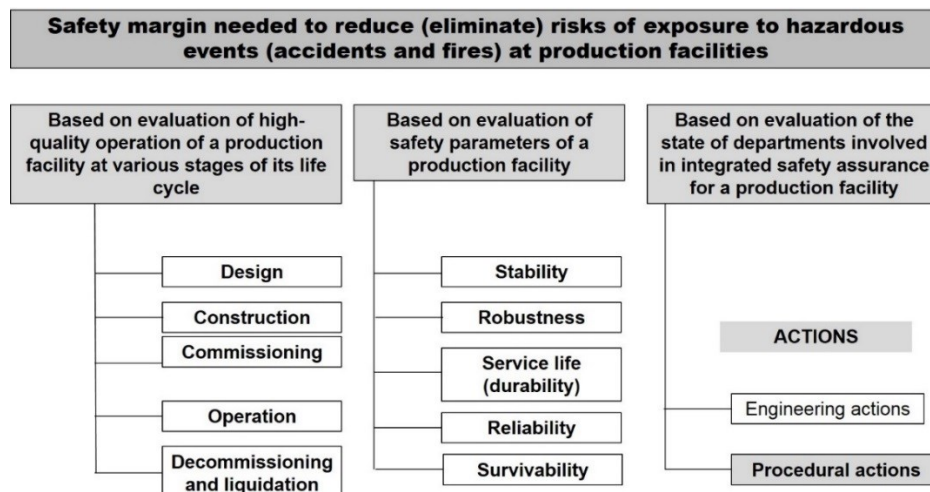
Reduction (elimination) of different types of damage depends on the availability of the required safety margin (figure 1), whose design value is determined using the following equation [5]

$$R_{(\tau)} = \frac{R_c(\tau)}{n_R}, \text{ where} \tag{1}$$

$R_c(\tau)$  is the boundary risk value (critical, threshold);

$n_R$  is the safety margin value required to reduce (eliminate) the risk.

The *safety margin* is understood as a set of factors characterized by the sufficiency of actions required to solve problems arising as a result of dangerous events (accidents and fires).



**Figure 1:** Clustered focus areas for evaluation of factors affecting the safety margin needed to reduce (eliminate) risks

<sup>2</sup> Code of Regulations CR 12.13130.2009 Categorization of premises, buildings and outdoor installations by explosion and fire hazards

The key idea aimed at reducing damage from hazardous events is to ensure the availability of a well-grounded safety margin designed to reduce (eliminate) various risks. Obviously, various applications and relationships determine the nature of the safety margin.

The focus areas listed above are highly relevant. The author conducted an analytical study to develop well-grounded solutions pre-compared with other well-known actions<sup>3</sup> aimed at improving the ISS at PPFER to meet the principal requirement of point 10, Resolution 842, issued by the Government of the Russian Federation on September 24, 2013 "On Procedure for Awarding Academic Degrees".

## II. Analysis of fundamental areas for improving and developing integrated safety at industrial enterprises

Fundamental documents governing the vital activity of Russia, including its essential industrial infrastructure, include *National Security Strategies of the Russian Federation, approved by Decrees of the President of the Russian Federation*<sup>4</sup>, which govern the development of comprehensive actions towards their implementation.

The outcome of a research project on IS problems, solved using the risk-oriented approach at industrial facilities, is consolidated areas of research that demonstrate valuable research findings (figure 2).



**Figure 2:** Findings used to solve problems of integrated safety management at industrial facilities

Research Area 1 (see Figure 2) considers theoretical fundamentals and their connection with the risk-oriented methodology and its implementation to ensure the IS of industrial enterprises; the following fundamental principles are formulated:

- using fundamental principles of risk analysis  $R(\tau)$  in the three principal areas of vital activity (social ( $N$ ), natural ( $S$ ) and technogenic ( $T$ ) activities), conducted as a single complex socio-natural-technogenic system of humans-nature-infrastructure during time  $\tau$  [6]

$$R_{(\tau)} = F_R\{R_N(\tau), R_S(\tau), R_T(\tau)\}; \quad (2)$$

- developing a generalized model of risk assessment at industrial facilities that demonstrates changes, triggered by the factor values of risks  $R(\tau)$ , or probabilities  $P(\tau)$  of dangerous events (accidents, fires, emergencies) and respective damage (economic damage, assessed in *rubles*; social damage assessed in the *number of people injured, killed, also known as casualties*). These types of damage are related to the main spheres of life, including the social sphere ( $N$ ), the natural sphere ( $S$ ), and the

<sup>3</sup> URL:<https://docs.cntd.ru/document/499047147> (Date of access: July 1, 2023)

<sup>4</sup> Decree of the President of the Russian Federation № 1666 issued on 02.07.2021 On National Security Strategy of the Russian Federation; Decree of the President of the Russian Federation № 400 issued on 19.12.2012 On the Strategy of the State National Policy of the Russian Federation for the Period through 2025.

technogenic sphere ( $T$ ), that make up a single complex system, consisting of humans-nature-infrastructure, during time  $t$  [6].

$$R_{(\tau)} = F_R\{P(\tau), U(\tau)\}; \tag{3}$$

$$R_{(\tau)} = F_R\{P(\tau), U(\tau)\}; \tag{4}$$

$$U_{(\tau)} = F_U\{U_N(\tau), U_S(\tau), U_T(\tau)\}; \tag{5}$$

- drafting scenarios of events, occurring in a complex system, and making a quantitative assessment of risks  $R(\tau)$ , using parameters of principal triggers and destructive factors of dangerous energies  $E(\tau)$ , substances  $W(\tau)$ , and information flows  $I(\tau)$  [7]

$$R_{(\tau)} = F_R\{E(\tau), W(\tau), I(\tau)\}; \tag{6}$$

- complying with the fundamental requirement concerning the non-exceedance of acceptable risks by calculated values of risks (formulas 2-6) in the process of implementing a risk-oriented approach [8].

$$R_{(\tau)} \leq [R_{(\tau)}], \text{ where} \tag{7}$$

$[R_{(\tau)}]$  is the parameter that has a limit value of an assessed acceptable risk. Applicable regulations (RLA, or regulatory legal acts, and RD, or regulatory documents) of the Russian Federation set the limit value of an assessed risk.

Researchers from the Russian Academy of Sciences formulated the fundamental substantiation of acceptable risks  $[R_{(\tau)}]$ , whose calculated value is identified using the following equation [9]

$$R_{(\tau)} = \frac{R_c(\tau)}{n_R}, \text{ where} \tag{8}$$

$R_c(\tau)$  is the threshold value of risk (critical, limit risk);

$n_R$  is the value of the safety margin used to reduce (eliminate) the risk considered above. The principle of choosing the reasonable *rational safety margin is sufficiency of compensatory actions aimed at reducing (eliminating) risks* (Figure 3).

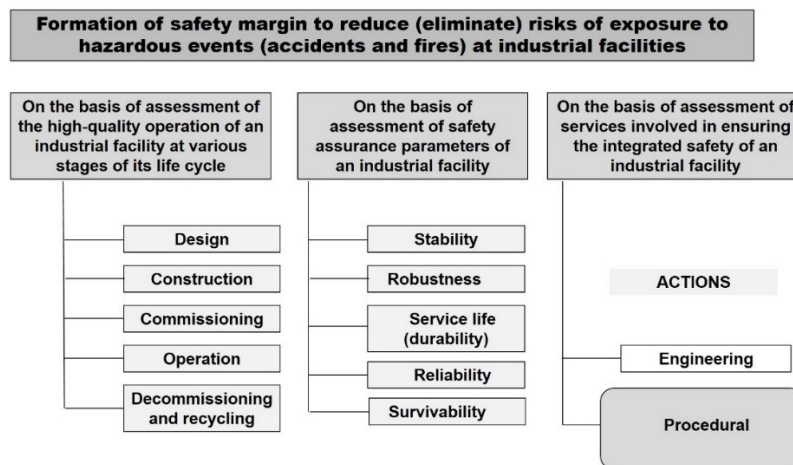


Figure 2: Assessment of factors, affecting the safety margin needed to reduce (eliminate) risks

The key approach to reducing damage from the impact of hazardous events encompasses a reasonable safety margin designed to reduce (eliminate) various risks. Obviously, the safety margin to be analyzed will be based on its application in different areas and considered using versatile methods of analysis in relation to risks (Figure 3).

### III. Existing and new proposed methods of risk assessment in the field of integrated safety of industrial enterprises

In the course of solving the problem, it was necessary to build awareness of approaches and techniques used in practice, as well as to provide more information about Research Areas 2 and 3 (see Figure 1). Methods<sup>5</sup>, including risk assessment procedures (recommendations) applicable by production facilities, were used to assess the risks arising within subsystems (InS, FS, OT). Information about the results of analytical comparison is provided in *Table 1*.

**Table 1.** Comparison between risk assessment methods, used to ensure practical integrated safety of industrial enterprises, and new methods, proposed and substantiated by the author of the article

<b>Methods of risk assessment within the framework of integrated safety of industrial facilities</b>			
In the field of industrial safety	In the field of fire safety	In the field of occupational safety	Original methods proposed by E.V. Gvozdev
<i>Methods belonging to the group of logical-graphical methods</i>			
Event Tree Analysis; Failure Tree Analysis; "What - If" method	Logical event trees	Scenario Analysis; Decision Tree Analysis; Structured What-If Method (Swift)	Bayesian Trust Networks (BTN) method
<i>Methods belonging to the group of expert analysis methods</i>			
Check-List; Hazard and Operability Analysis (HOA); HAZID (Hazard Identification) method	-	Checklists; Bow-tie analysis; HAZOP (Hazard and Operability Study) method.	Analysis of hierarchies and pairwise comparisons method (MAI)
<i>Methods belonging to the group whereby characteristics are calculated using individual weighting coefficients</i>			
Failure Type and Consequence Analysis (FTCA); Safety actions analysis; quantitative accident risk assessment	Determining the time of blocking evacuation routes in case of fire; determining the estimated evacuation time	Cause-effect analysis; matrix method based on scoring; LOPA layers of protection analysis; HRA (Hyman Reliability Assessment); occupational disease risk assessment; cost effectiveness analysis (cost-benefit analysis)	Method of complex numbers (Symb method)

<sup>5</sup> Order № 387 issued by Federal Environmental, Industrial and Nuclear Supervision Service of Russia on 03.11.2022 On Approval of Safety Guidelines Titled Methodological Fundamentals for Hazard Analysis and Accident Risk Assessment at Hazardous Production Facilities; Order № 404 issued by the EMERCOM of Russia (Ministry of the Russian Federation for Civil Defence, Emergencies and Elimination of Consequences of Natural Disasters) on 10.07.2009 On Approval of the Methodology for Determining Estimated Fire Risk Values at Production Facilities; Order № 929 issued by the Ministry of Labor of Russia on 28.12.2021 On Approval of Recommendations for Selecting Methods of Assessing Occupational Risk Levels and Reducing Levels of Such Risks.

The table presents consolidated groups of methods used in the subsystems (InS, FS, OS). Their practical application allows obtaining results in the form of final (qualitative or quantitative) estimated risk values.

Comparative results of the practical application of methods were obtained in the format of final (qualitative or quantitative) estimated risk values (Table 2).

**Table 2.** Comparative results of final estimated risk values obtained in the course of risk assessment within the framework of integrated safety of industrial facilities

<b>Results of risk assessment within the framework of integrated safety of industrial facilities</b>			
In the field of industrial safety	In the field of fire safety	In the field of occupational safety	Results obtained using methods proposed by E.V. Gvozdev
<i>Results presented as qualitative values</i>			
Risk prioritization based on categorization of hazards from accidents, <i>risk priority value (1;2;3)</i>	-	+	Prioritization of risk based on a general ranked list, <i>the value of the risk priority (1;2;...; n) depends on damage</i>
Risk values ranging from negligible to higher than acceptable risk, <i>risk value (A; B; C; D)</i>		+	
Risk values with criticality of deviations, <i>risk value (high; medium; low)</i>		+	
<i>Results presented as quantitative values</i>			
Risk values of the frequency of depressurization of engineering pipelines, <i>risk value <math>10^{-n}</math>/year, where n is a power value</i>	+	-	Risk values for the value of cause and effect relationships, <i>risk value of the probability of implementation (1-100%)</i>
Risk values of damage to people, <i>risk value of the probability of implementation (1-100%), risk value <math>10^{-n}</math>/year, where n is a power value.</i>	Risk values needed to determine the estimated evacuation time, <i>risk value (min.)</i>	-	Risk values based on the calculation of the impact factor of services, <i>risk value (0,001-0,475)</i>

#### IV. Purpose of the study

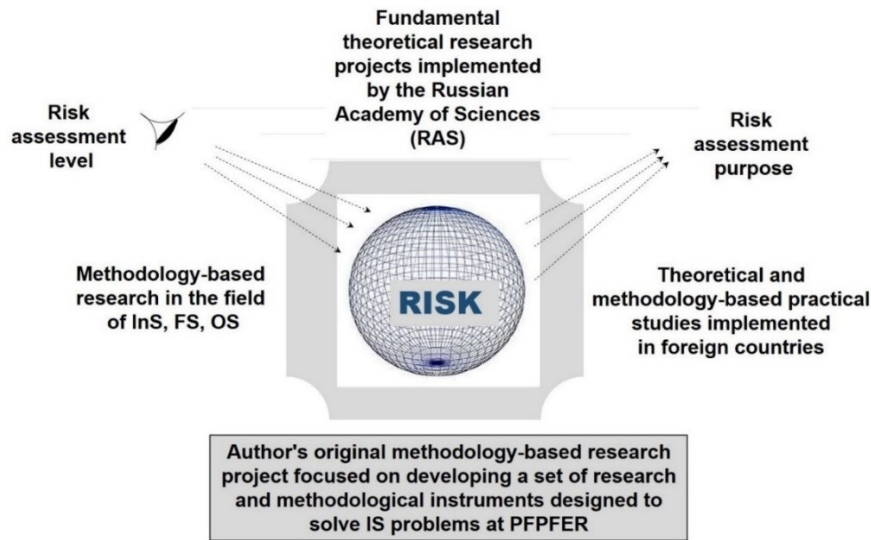
The purpose of the study is to substantiate the adequate practical use of a novel set of research and methodological instruments developed by the author and to make sure that the socio-economic effect of its application is higher than that of the ISS that are currently used by PPFER, in other words, to confirm the feasibility of new methods (groups of methods) to be used to assess risks arising within this system. In this case, the assessment process will be based on the practical data backed by the experiments.

To achieve this purpose, the author employed a methodology comprising the awareness of procedures, whose core elements are *methods and methodology, contributed to the set of research and methodological instruments*, used to solve problems of research and practice [10]. The following tasks

were to be solved:

1. Presenting the statement of and the proposed solution to the problem of the future development of ISS at PFPFER.
2. Presenting the case justifying the adequacy of this solution to confirm the practical applicability of the proposed set of research and methodological instruments.

Below is a model for selecting a new risk assessment methodology required to develop ISS at PFPFER (figure 4).



**Figure 4:** Development model required for risk assessment at PFPFER

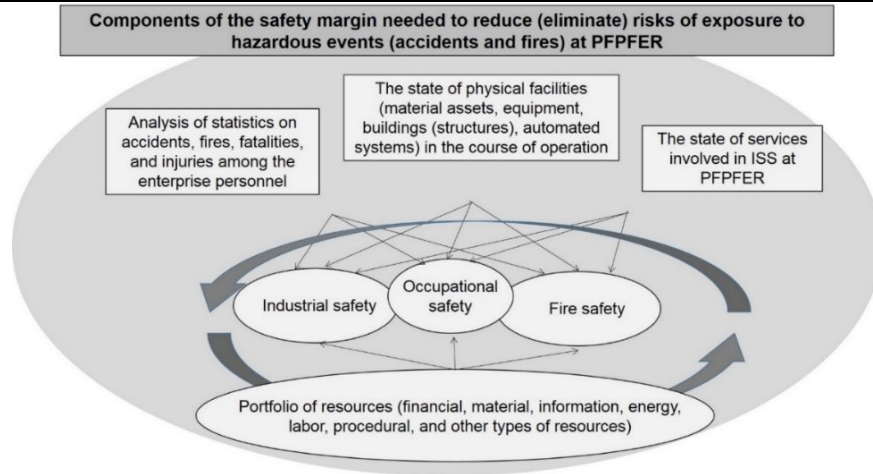
The proposed model (see figure 4) is not considered as a model with the same assessment criteria. It allows looking beyond the horizon of unexplored risks, developing a new set of research and methodological instruments (models, methods, techniques) for practical use. Each new solution to a research problem has new features added to research and methodological instruments (for example, a solution to the new formulation of a problem can result in a new solution to this problem) [11]. It will open the way for a transition to development of ISS that are currently in operation at PFPFER enterprises.

## V. Formulating and solving the problem of ISS development at PFPFER enterprises

The prospective development of ISS at PFPFER enterprises requires a safety margin composed of various resources (financial, material, information, energy, labor, and other types of resources) to be contributed to subsystems (OS; InS; FS) to reduce (eliminate) risks [12]. Figure 5 has a block scheme of the safety margin.

Given that material, economic, labor, time, information, and other resources are the main constituents of the corporate safety margin for enterprises under consideration, identification of the nature and extent of risks and coordination of subsystems (OS, InS, FS) support, adjustable to ensure the highest socio-economic effect, is a challenging task [13 – 21].





**Figure 5.** The block scheme of the safety margin needed to reduce (eliminate) risks of damage from hazardous events (accidents and fires) at PPFER enterprises

**Problem formulation.** Assume that management teams of the enterprises under consideration make a decision to develop ISS at PPFER or to bring its operation to a new qualitative level within a pre-set period of time. The management of these enterprises identifies a transition period, including the initial point of reference ( $t_0$ ) and the point of goal achievement ( $t_1$ ), included in the following model:

$$\sum_{i=1}^N P_{\varphi_{n1}}(t_0) \rightarrow \sum_{i=1}^N P_{\varphi_{n2}}^*(t_1), \tag{9}$$

where  $P_{\varphi_{n1}}(t_0), P_{\varphi_{n2}}^*(t_1)$  are values, describing the state of subsystems (OS, InS, FS) at the beginning and at the end of the transition period;  
 $n_1; n_2$  are values of resources calculated for the initial point of reference ( $t_0$ ) and the point of goal achievement ( $t_1$ ).

If these values change and became equal to  $P_{\varphi_{n2}}^*(t_1)$  during period ( $t_1$ ), total changes will be calculated as follows:

$$\sum_{i=1}^N |P_{\varphi_{n2}}^*(t_1) - P_{\varphi_{n1}}(t_0)| = \Delta(t_1), \tag{10}$$

where  $\Delta(t_1)$  is the total difference in changes for all values over period  $t_1$ ;

|...| is the sign showing the modulus of a number.

**The task is to substantiate** calculations of efficiency of the practical application of a set of research and methodological instruments and to demonstrate a reduction in damage to subsystems under consideration.

**Solution.** If the total difference  $\Delta(t_1)$  showing changes in all values during period  $t_1$  is available, the value of the development change  $Y(t_1)$  at the point of goal achievement ( $t_1$ ) can be calculated using the following formula:

$$Y(t_1) = \frac{\Delta(t_1)}{E_n}, \tag{11}$$

where  $E_n$  is the efficiency of the volumetric contribution of resources to  $(1 - N)$  industry-focused subsystems (OS, InS, FS). It is the value whose calculation needs hundreds of different parameters. The following calculations must be made to find effective contributions of resources to the  $(1 - N)$  industry-specific subsystem (OS, InS, FS).

Assume that actual ( $P_{\varphi_{n1}}^{Risk}(t_0)$ ) boundary and ( $P_{\varphi_{n1}}^{Cp}(t_0)$ ) mean values of damage to subsystems (OS,

InS, FS) are available for the previous period. Further, the target damage reduction value is identified for the forecast period ( $t_0 - t_1$ ) with account taken of the deviation of the actual damage values from the mean ones

$$\sum_{i=1}^n |P_{\varphi_{n1}}^{Cp}(t_0) - P_{\varphi_{n1}}^{Risk}(t_0)|. \quad (12)$$

Further, the rational target value  $P_{\varphi_{n1}}^{Prog}(t_1)$  of a reduction in damage to subsystems (OS, InS, FS) is found; it is subject to comparison as a ratio of values for current and projected periods ( $t_0 - t_1$ ), taking into account a deviation of mean values of damage from planned values of damage:

$$\sum_{i=1}^n |P_{\varphi_{n1}}^{Prog}(t_1) - P_{\varphi_{n1}}^{Cp}(t_0)|, \quad (13)$$

where  $n$  is the total number of values used in the calculations.

$\lambda_n$ , the coefficient affecting a reduction in damage from accidents and fires at PPFER, calculated for current values, can be written as follows:

$$\lambda_n(t_0) = \frac{\sum_{i=1}^n |P_{\varphi_{n1}}^{Cp}(t_0) - P_{\varphi_{n1}}^{Risk}(t_0)|}{\sum_{i=1}^n |P_{\varphi_{n1}}^{Cp}(t_0)|}. \quad (14)$$

$\lambda_n$ , the coefficient affecting a reduction in damage from accidents and fires at PPFER, calculated for the ISS development period, can be formulated as follows:

$$\lambda_n(t_1) = \frac{\sum_{i=1}^n |P_{\varphi_{n1}}^{Prog}(t_1) - P_{\varphi_{n1}}^{Cp}(t_0)|}{\sum_{i=1}^n |P_{\varphi_{n1}}^{Cp}(t_0)|}. \quad (15)$$

The achieved target value of the ISS development at PPFER (conventional period)  $Y(t_1 - t_0)$  will be calculated using the difference between coefficients  $\lambda_n$ , affecting the reduction in damage from accidents and fires at these enterprises

$$Y(t_1 - t_0) = \lambda_n(t_1) - \lambda_n(t_0). \quad (16)$$

The proposed theoretical models of ISS at PPFER, designed for the present and future systems, are characterized by numerous parameters requiring computations to be made for all risks. Hence, there is a need to develop the ISS designed for PPFER, meaning that there is a need to develop *a set of research and methodological instruments* to ensure the availability of the safety margin to ensure the maintenance and development of subsystems (OS, InS, FS), and to improve the efficiency of ISS at the enterprises considered in this article.

## VI. The case substantiating ISS improvement at PPFER through the use of a set of research and methodological instruments

Let's analyze the calculation made within the framework of an experiment to make verifications using formula (8) together with the data obtained using methods contributed to the software registered with the Federal Service for Intellectual Property<sup>6</sup> (figure 6).

<sup>6</sup> Certificates of state registration of computer software:

№ 2022614215 RF Calculator for evaluation of industrial and fire safety actions at oil and gas enterprises of Russia; published 17.03.2022, by E.V. Gvozdev, B.S. Sadovsky, N.R. Ruppa, P.A. Butovchenko;

RF № 2023611653. Rater for assessment of industrial and fire safety at oil and gas enterprises of Russia; published 24.01.2023 E.V. Gvozdev, N.M. Migalchinsky, T.E. Koldin, D.S. Sinyakin.

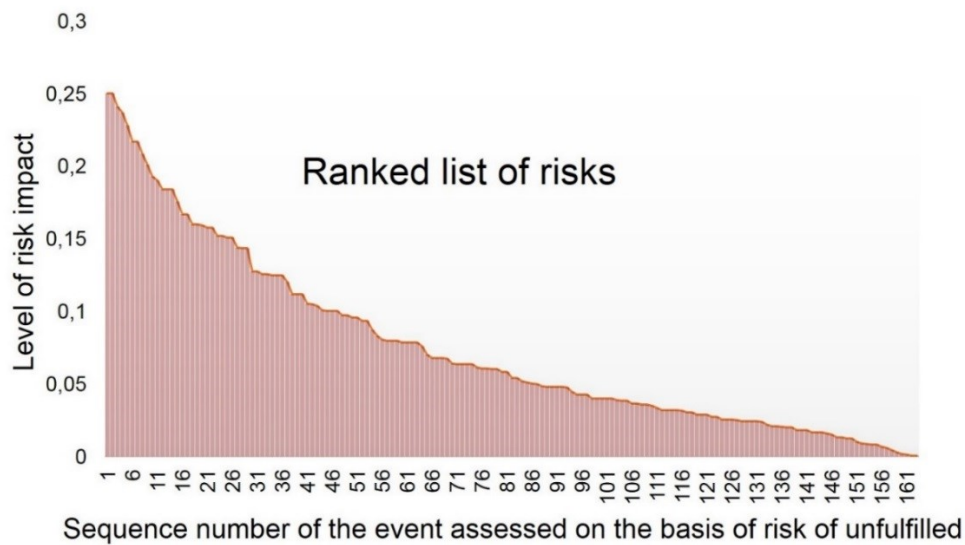


Figure 6. Ranked list of unfulfilled activities assessed by risk

The sampling has minimum (from 0,001) and maximum (0,250) limit values of risk calculated for unimplemented actions, extracted from statistics of accidents and fires for the period of 9 years<sup>7</sup>.

In the present-day environment Corporate Decision Makers (hereinafter - DMs) use their practical experience to distribute resources between ISS subsystems (OS, InS, FS). Thus, DMs rise the probability of errors in prioritizing factors of damage from accidents and fires on the ranked list of risks.

In the course of experiments some factors were randomly disregarded (about 25% in total) according to the following procedure: every 2nd factor was disregarded for the 1st experiment; every 3rd factor was disregarded for the 2nd experiment; every 4th factor was disregarded for the 3rd experiment (figure 6). As for the sampling analyzed using a set of research and methodological instruments, factors 121 to 161 (about 25% in total) were disregarded. Calculation formulas are presented for each experiment in the form of a system of equations:

$$\begin{cases} A_1 = \sum_{0,001}^{0,250} (n_1 + n_2 + \dots + n_{161}) - (n_1 + n_3 + \dots + n_{79}) \in (25\%) = \lambda_{n_{A_1}}(t_0) \\ A_2 = \sum_{0,001}^{0,250} (n_1 + n_2 + \dots + n_{161}) - (n_1 + n_4 + \dots + n_{118}) \in (25\%) = \lambda_{n_{A_2}}(t_0) \Rightarrow A_{PE3}(t_0) \\ A_3 = \sum_{0,001}^{0,250} (n_1 + n_2 + \dots + n_{161}) - (n_1 + n_5 + \dots + n_{157}) \in (25\%) = \lambda_{n_{A_3}}(t_0) \end{cases} \quad (17)$$

$$B = \sum_{0,001}^{0,250} (n_1 + n_2 + \dots + n_{161}) - (n_{121} + n_{122} + \dots + n_{161}) \in (25\%) = B(t_1)$$

Experimental results for present and future ISS at PFPFER are shown in figure 7.

The graph shows that the ISS can be improved at PFPFER, if a set of research and methodological instruments are employed. In other words, risks of damage from accidents and fires can be reduced by 18% during the period of the ISS development.

<sup>7</sup> Information about accidents and fires is available on the website of the Federal Environmental, Industrial and Nuclear Supervision Service of Russia at [https://www.gosnadzor.ru/industrial/oil/lessons/index.php?sphrase\\_id=2569631](https://www.gosnadzor.ru/industrial/oil/lessons/index.php?sphrase_id=2569631), accessed 15.08.2023

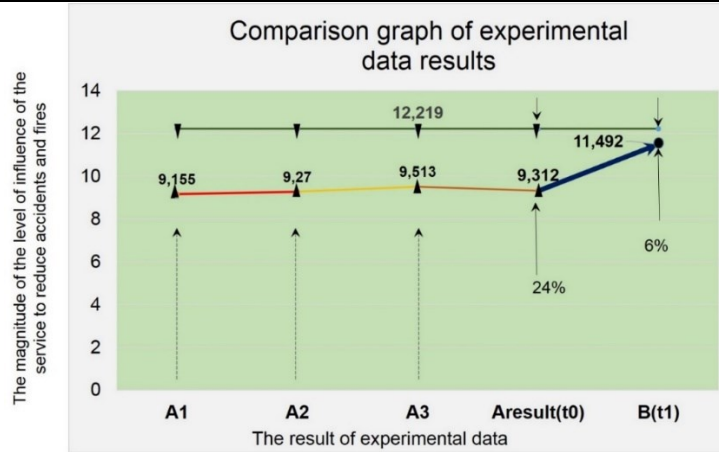


Figure 7. Experimental data for present and future ISS at PPFER

Figure 5 shows the ratio of the present-day target value of ISS at PPFER, equaling  $Y(t_0)$ , to the achieved target development value of ISS at PPFER, equaling  $Y(t_1)$ . A positive trend towards the reduction in damage from accidents and fires can be formulated as follows:

$$m\lambda_{(1-n)} = B(t_1) - A_{PE3}(t_0) \approx 2,2, \quad (18)$$

where  $m\lambda_{(1-n)}$  is the mathematical expectation of the total value, affecting the reduction in damage from accidents and fires, calculated using a set of research and methodological instruments;  $A_{RES}(t_0)$ ;  $B(t_1)$  are the final results based on the experimental data for the current and future ISS at PPFER (figure 7).

The total calculated value, affecting the reduction in damage from accidents and fires and immediately related to all damage from accidents and fires reported on the website of the Federal Environmental, Industrial and Nuclear Supervision Service for the period of 9 years, is shown in table 3. The value applies to all Russian oil and gas companies.

Table 3. Ratio of one unit of damage to total damage from accidents and fires at Russian oil and gas enterprises in 2014-2022

Category	Total damage from accidents and fires according to reports	Ratio of total damage to one unit of damage according to experimental data
Fatalities, number of persons	49	4
Injuries, number of persons	122	10
Economic damage, billion rubles	19,7	1,615

Values of socio-economic damage can be reduced to a conventional unit based on the experimental data (table 1) using the ratio of total calculated damage from accidents and fires to different categories of assets to be protected (table 1), as well as to the total calculated value  $\lambda_n const \approx 12,2$  (figure 5).

The socio-economic effect  $E_{(1-n)}$  can be calculated as follows to rise the ISS at PPFER in the course of the development period:

$$E_{(1-n)} = \frac{m\lambda_{(1-n)}}{y_{(t_1-t_0)}^*}, \quad (19)$$

where  $Y_{(t_1-t_0)}^*$  is the time frame (3 years and more) prescribed by the management of these enterprises for the development of ISS at PFPFER. A conceptual solution to the problem of the present-day ISS development at PFPFER is found. It will bring socio-economic benefits in the future in case of a reduction in the number of accidents and fires at these enterprises.

This case confirms the feasibility of the ISS development through the practical application of research and methodological instruments at PFPFER. The proposed approach to development of the current ISS at PFPFER, presented by the author in a formal form, can be applied at any other production facilities of the Russian Federation.

### VII. ISS development at Russian production facilities: the proposal to be made to the management team

Executives of Russian production facilities can consider different ISS development periods for PFPFER. Below is the projected socio-economic effect attainable during one development period equaling one year (figure 8).

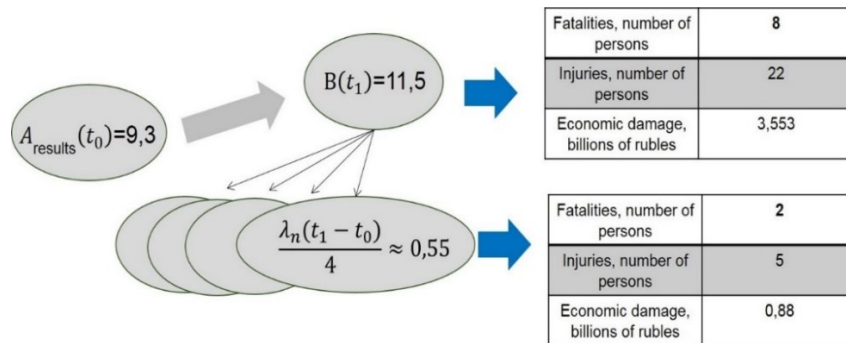


Figure 8. Target model describing the achievement of the socio-economic effect during one year

Schedules outline the time horizon needed to achieve the strategic objective, which sets the time frame for achieving sub-objectives at tactical and operational levels. The case of time horizons is presented for forecasting purposes (figure 9).

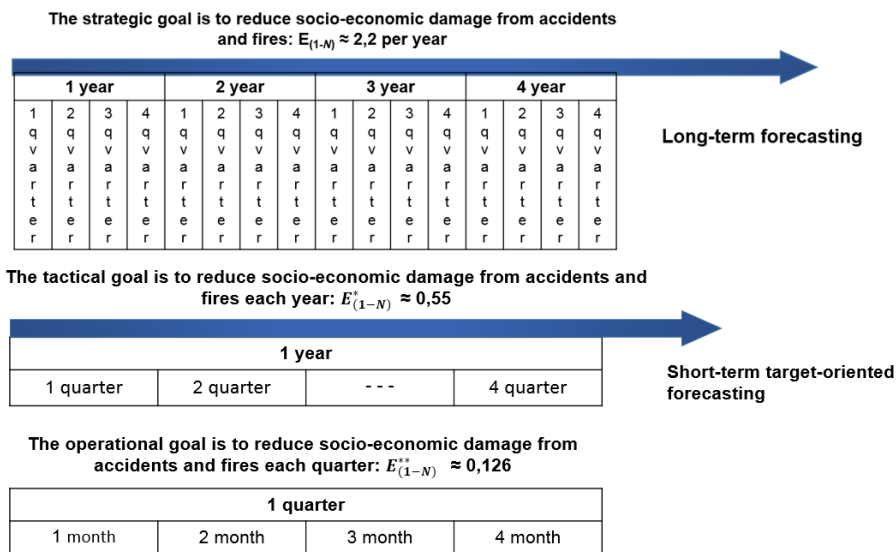


Figure 9. Time horizons for ISS development at Russian production facilities

*The strategic objective* is formulated by the management team for a long period of time, but its achievement should be broken down into steps (monthly or quarterly goals).

*At the tactical level*, the horizon for scheduling future activities is limited to one year.

*At the operational level*, all actions implemented within a month (a quarter) are taken and registered in the documents of target-oriented operational scheduling.

Production facilities should take advantage of the main items (models, methods, techniques, etc.) from the set of research and methodological instruments, developed by the author of the article, to achieve their targets at the strategic level, including a major reduction in damage from accidents and fires at Russian production facilities.

## Conclusion

1. The relevance of the ISS development at PFPFER is substantiated by the author. In the future, ISS will be able to reach a qualitatively new level through the assessment of risks of damage from combined hazardous events (accidents and fires). Consolidated improvement areas are identified, and research results are available for practical application in this area of research.
2. The idea of a new category of combined risks of hazardous events (accidents and fires) is presented. Its originality lies in the fact that the adjustment of the required safety margin should take into account the state of subsystems and the effect of services (OS, InS, FS) on the state of ISS at PFPFER. The author of the article presents a new research area that requires expanding the range of methods used in practice to assess the state of subsystems (OS, InS, FS).
3. The formulation of and a solution to the problem of the future development of ISS at PFPFER is presented. Its uniqueness lies in the fact that this solution can present an individual risk as a quantitative value as a result of assessment of each action left unimplemented in the field of OS, InS, and FS.
4. The case substantiating the adequacy of results is presented. It proves the practical usability of a new set of research and methodological instruments, which (1) have a socio-economic effect for Russian production facilities, and (2) lead a positive trend towards a smaller damage from accidents and fires registered and reported by different government authorities, such as the Ministry of the Russian Federation for Civil Defense, Emergencies and Elimination of Consequences of Natural Disasters, the Ministry of Labor, and the Federal Environmental, Industrial and Nuclear Supervision Service of Russia.

## References

- [1] Gvozdev, E., 2021. Visualization and diagnostics of reliability of an object with changing functioning conditions. E3S Web of Conferences, Moscow, April 22–24. DOI 10.1051/e3sconf/202126302009
- [2] Gvozdev, E.V., Matvienko, Y.G., 2021. About the evaluation of the state of integrated safety of industrial enterprises. Technosphere safety technologies 4(94), 76-95. DOI 10.25257/TTS.2021.4.94.76-95
- [3] Gvozdev, E.V., 2022. About the evaluation of the state of integrated safety at enterprises of oil and gas industry of Russia. Fire and Explosion Safety, vol. 31, 1, 49-64. DOI 10.22227/0869-7493.2022.31.01.49-64
- [4] Gvozdev, E.V., 2023. Formulating the problem of rational allocation of resources to ensure integrated safety. Real Estate: Economics, Management, 2, 50-55
- [5] Makhutov, N.A., Pokrovskii, A.M., Dubovitskii, E.I., 2019. Analysis of Crack Resistance of an Oil Trunk Pipeline Considering the Varying Failure Viscosity in the Neighborhood of a Welded

Joint. Journal of Machinery Manufacture and Reliability, vol. 48, 1, 35-42. DOI 10.3103/S1052618819010126

[6] N. A. Makhutov, V. N. Permyakov, R. S. Akhmetkhanov, Risk analysis and security of critical facilities of the oil and gas industry (Tyumen, TyumNGU, 560 p., 2013)

[7] Gadolina, I. V. Varied approaches to loading assessment in fatigue studies / I. V. Gadolina, N. A. Makhutov, A. V. Erpalov // International Journal of Fatigue. – 2021. – Vol. 144. – P. 106035. – DOI 10.1016/j.ijfatigue.2020.106035

[8] Makhutov, N. A. Assessment of extreme thermo-mechanical states of engineering systems under operating loading conditions / N. A. Makhutov, M. M. Gadenin, D. O. Reznikov // Acta Mechanica. – 2021. – Vol. 232, No. 5. – P. 1829-1839. – DOI 10.1007/s00707-020-02920-3

[9] Makhutov, N. A. Analysis and management of strength parameters, resource and risks of safe operation of power plants with various types of energy resources / N. A. Makhutov, M. M. Gadenin // Problems of mechanical engineering and machine reliability. – 2022. – No. 1. – pp. 47-56. – DOI 10.31857/S0235711922010060

[10] Dissertation presentation alert. <https://www.safety.ru/about/dissertation/obyavozaschite/031022-gvozdev-evgeniy-vladimirovich> (August 14, 2023)

[11] Dolgov, A.I., 2013. Methodology of scientific research. DSTU, 161 p

[12] Abrosimov N.V., Aksyutin O.E., Aleshin, A.V., et al., 2019. Security of Russia. Legal, socio-economic and scientific-technical projects. Scientific foundations of industrial safety. Moscow, Znanie, 824 p. ISBN 978-5-87633-183-0

[13] Mileshko, L.P., 2023. Economics and safety management. Yurait, 99 p. ISBN 978-5-534-13764-4

[14] Fedortsov, A.B., Silivanov, M.O., 2021. Physiologically based control of the process of gentle light awakening. Journal of Physics. Conference Series 16. Moscow. DOI 10.1088/1742-6596/2127/1/012064

[15] Odintsov, B.E., Urintsov A.I., Mamedova, N.A., 2021. Business process performance management of enterprises in terms of inverse calculations AIP Conference Proceeding, Volgograd. DOI 10.1063/5.0067990

[16] Mkrtychyan, L., Straub, U., Giachino, M., Kocher, T., Sansavini, G., 2022. Insurability risk assessment of oil refineries using Bayesian Belief Networks. Journal of Loss Prevention in the Process Industries, 74. <https://doi.org/10.1016/j.jlp.2021.104673>

[17] Massel, A.G., Pesterev, D.V., 2017. Transformation of cognitive models into knowledge base of production expert system. Proceedings of the 19th International Workshop on Computer Science and Information Technologies. Germany, Baden-Baden. USATU, vol. 1, 121-124. ISBN 978-5-1030-8, ISBN 978-4-4221-1031-5

[18] Vladimirovich, G.E., Mikhailovna, C.V., 2019. The Modern Strategy to the Process of Managing Complex Security of the Enterprise on the Basis of Rational Centralization. International Journal of Innovative Technology and Exploring Engineering, 9(1), 4614-4620. <https://doi.org/10.35940/ijitee.A4944.119119>

[19] Qingsong, X., Leiguang, J., Xuedong, L., 2020. Decision-making optimization of complex product design task change perturbation based on prospect theory. Journal of Industrial Engineering and Engineering Management, 34(1). <https://doi.org/10.13587/j.cnki.jieem.2020.01.020>

[20] Bochkov, A. Reflections on Dual nature of risk. Toward a formalism / A. Bochkov // Reliability: Theory & Applications. – 2023. – Vol. 18, No. S5(75). – P. 44-74. – DOI 10.24412/1932-2321-2023-575-44-74.

[21] Arefyeva, E. Managing the risk of emergencies caused by groundwater flooding of the historical built-up areas / E. Arefyeva // Reliability: Theory & Applications. – 2023. – Vol. 18, No. S5(75). – P. 125-135. – DOI 10.24412/1932-2321-2023-575-125-135.

# TRANSIENT AND METAHEURISTIC COST SCRUTINY OF $M^X/G(A, B)/1$ RETRIAL QUEUE WITH RANDOM FAILURE UNDER EXTENDED BERNOULLI VACATION WITH IMPATIENT CUSTOMERS

RANI R<sup>1</sup> AND INDHIRA K\*



<sup>1,\*</sup>Vellore Institute of Technology, Vellore - 632 014, Tamil Nadu, India.  
kindhira@vit.ac.in.

## Abstract

*The transient and metaheuristic cost analysis of a  $M^X/G(a, b)/1$  retrial queue with random failure during an extended Bernoulli vacation with impatient clients is covered in this study. Any batch that arrives and discovers the server is busy, down, or on vacation joins an orbit. In the alternative, only one new customer from the group joins the service right away, while the others join the orbit. After providing each service, the server either waits to serve the following customer with probability  $(1 - \theta)$  or goes on vacation with probability  $\theta$ . It has been found that these systems express steady-state solutions and are dependent on time probability generating functions in consideration of their Laplace transforms. We also discuss a few exceptional and particular instances. After that, the impact of different parameters on the system's effectiveness is evaluated. We are also talking about ANFIS. Additional approaches employed in this study to swiftly determine the system's optimum cost include genetic algorithms (GA), artificial bee colonies (ABC), and particle swarm optimization (PSO). We also examined the graph-based convergence of several optimization algorithms.*

**Keywords:** Batch arrival, Retrial queues, Feedback, Extended Bernoulli Vacation, ANFIS, Cost Optimization.

## 1. INTRODUCTION

For the development, capacity planning, performance assessment, and optimization of numerous real-world systems, queueing theory offers a potent tool. Chaudhry and Templeton [1] provided a comprehensive analysis of bulk queuing. Bulk arrival analysis, a condensed form of customer examination, is a great place to start with customised models. Bulk service queuing models were created by Bailey [2]. He invented the process known as "fixed-batch service". The server continuously offers a specific batch of services to each set of users in fixed-batch service queuing systems (QS).

The "retrial queueing" system, which is used when a customer enters and the server is occupied, requires the customer to leave the appropriate area and repeat his request after a certain period of time. This property is essential for network technologies, cognitive networks, online computing systems, manufacturing systems, and other systems.

Sumitha and Udaya Chandrika [3] investigated a retrial queueing system with starting failure, single vacation, and orbital search. In batch arrival retrial queues, Radha et al. [4] studied some system performance measures are evaluated using the supplementary variable technique (SVT) and the steady-state (SS) probability generating function (PGF) for system size.

Gomez-Corral has talked a lot about a retrial QS with FCFS discipline and typical retrial



periods. The  $M/G/1$  retrial queue with feedback and starting failures was described by Krishna Kumar et al. [5]. Yang, Tao, and Hui Li [6] investigated an  $M/G/1$  retrial queue with a starting failure-prone server. An analysis of a feedback retrial queueing system with starting failures and a single vacation was studied by Mokaddis et al. [7].

In a Vacation, queueing system the server could be temporarily unavailable for a number of reasons, including maintenance monitoring, tending to other queues, or simply taking a break. When the server is unavailable to users, that time period is referred to as a "vacation". A single server batch arrival Bernoulli feedback QS with a waiting server, K-variant vacations, and anxious clients was examined by Bouchentouf et al. [8]. The transient behaviour of a batch arrival feedback retrial queue with starting failure and Bernoulli vacation (BV) was investigated by Ayyappan and Sathiy a [9]. Assuming that repair, service, and vacation times are randomly distributed, the time-dependent PGF are also computed in relation to their Laplace transforms(LT).

Numerous academics who have studied queueing techniques with interruptions have as their primary tenet that, in the event of a failure, the service channel will be promptly repaired. A transient analysis of the  $M^{[X_1]}, M^{[X_2]}/G1, G2/1$  retrial QS's with priority services, working breakdown, start up/close down time, BV, renegeing, and balking was studied by Ayyappan et al. [10]. Kulkarni et al. [11] established a retrial queue with a server prone to failures and maintenance. Ayyappan and Shyamala [12] created an  $M^{[X]}/G/1$  with Bernoulli schedule, server vacation, random breakdown and second optional repair. And also calculate the typical length of the line and the typical wait period in closed form. When the repair is finished a number of consumers who had previously used the services wait for the remainder to be provided. Jau-Chuan Ke et al. [13] demonstrated a waiting line with customers complaining and providing feedback the servers malfunctioned. Furthermore, if all servers are already in use when a customer arrives, he will either join a retrial orbit or decline. When a service is finished the client can exit the system or rejoin the retrial group to receive more services. They can also design a cost function to determine the system's ideal parameter settings under the stability condition. Computer telecommunication systems is a example of application for these types.

A consumer may try again until they are happy if they are not satisfied with the service they received. Takacs [14] investigates this at first allowing the consumer who has finished the service to provide feedback to the rear of the line. An  $M/(G1, G2)/1$  feedback retrial queue with two phase service, variant vacation policy under delaying repair for impatient Customers was analysed by Rajadurai et al. [15].

Many real-world systems have impatient customers as a built-in feature, particularly when the customer is a human, a perishable product, or some moving object that can depart the service area and their waiting period in the queue reaches certain pre-defined threshold values. This clearly explains why queueing literature frequently discusses the impatience phenomenon. Accounting impatience is crucial in the setting of lines for group service because a client could spend a large amount of time in the system while waiting for the accumulation of a sufficient number of customers.

More focus has been placed on the numerous retrial lineups with non-persistent (impatient) consumers. A discussion about the study of a retrial queue with group service of impatient clients involved D'rienzo et al. [16]. A batch arrival retrial queueing model with starting failures and customer impatience was addressed by Nila and Sumitha [17]. Customers arrive in batches in line with the Poisson process. In certain situations, the clients refuse and break their promises. The analysis of a retrial QS with priority services, working breakdown, BV, admission control, and balking was explained by Ayyappan et al. [18]. Ayyappan and Nirmala [19] explored an analysis of customer's impatience on bulk service QS's with an unreliable server, setup time and two types of multiple vacations. Sethi.R et al. [20] investigated the cost optimization and ANFIS computing of an unreliable  $M/M/1$  queueing system with customers' impatience under n-policy. The ideal Cost Analysis for Discrete-Time Recurrent Queue with Bernoulli Feedback and Emergency Vacation was described by M. Vaishnavi [21]. In order to calculate costs, PSO, ABC, and GA are also used. To ensure the best deal, these methods compare and contrast the outputs.

The paper’s structure is as follows: Section 2 provides a detailed explanation of the mathematical model. Section 3 discusses the ideas and formulae governing our system as well as how to obtain the time-dependent solution of our model. The PGF for the queue length at each given epoch and the SS performance of the system are explicitly determined in Section 4. In Section 5, the pertinent stability condition has been uncovered. In Section 6, we precisely estimate the mean queue size, mean queue waiting time, and efficiency features for each state of the system. In Section 7, we present a practical illustration. We offer a numerical study and associated graphs in Section 8. Furthermore, an ANFIS was provided in Section 9. The Cost optimization is offered by Section 10. The conclusion is presented in Section 11.

## 2. MODEL DESCRIPTION AND ANALYSIS

We suppose that the underlying queueing model is as follows:

**Arrival process:** Customers enter a poisson stream, and bulk service is offered on an FCFS basis. Considering that a batch of “i” customers enters the system,  $\Lambda > 0$  represents the average batch arrival rate, and  $\Lambda c_i dt (i \geq 1)$  represents the first order probability during the short interval of time  $(\omega, \omega + d\omega]$ . We define a batch arrival and a bulk service as having a smallest batch size of “a” and a highest batch size of “b”.

**Retrial process:** When a customer arrives and discovers that the server is busy, unavailable, or broken, the customer has two options: (1) leave the service area with a probability of  $d$  and join a pool of blocked customers known as an orbit; or (2) balk the system with a probability of  $\bar{d}$  in accordance with FCFS, which implies that only the customer at the head of the orbit queue is permitted access to the server.

When the server is idle, the customer at the head of the retrial queue engages with potential primary customers to see who can cancel their service request and, with prob.,  $g$ , either move up in the retrial queue or leave the system with prob.,  $(1 - g)$ .

A general (arbitrary) distribution with the distribution function  $A(u)$  and the density function  $a(u)$  determines the retrial interval.

Let  $g(\zeta)d\zeta$  be the conditional prob., density of completing the retrial within the range  $(\zeta, \zeta + d\zeta]$ , where  $\zeta$  is the elapsed retrial time.

$$g(\zeta) = \frac{a(\zeta)}{1 - A(\zeta)}$$

and therefore,

$$a(u) = g(u)e^{-\int_0^u g(\zeta)d\zeta}$$

Inter-retrial times have an arbitrary dist.,  $A(\zeta)$  with corresponding Laplace-Stieltjes transforms (LST)  $A^*(u)$ .

**Service process:** The server enters an idle state whenever a fresh or returning user comes before quickly resuming regular operations for the newcomers. A generic (arbitrary) distance with the distance function  $B(\zeta)$  and the density function  $b(\zeta)$  follows the service time.

Given the elapsed retrial time  $\zeta$ , define  $\phi(\zeta)d\zeta$  as the conditional probability of service completion within the range  $(\zeta, \zeta + d\zeta]$ .

$$\phi(\zeta) = \frac{b(\zeta)}{1 - B(\zeta)}$$

and therefore,

$$b(\omega) = \phi(\omega)e^{-\int_0^\omega \phi(\zeta)d\zeta}$$

The random variable  $B$  with the dist., function  $B(\zeta)$  and LST  $B^*(\omega)$  denotes the service time.

**Random failure:** Failures are anticipated to occur sporadically throughout the system and ought to follow a poisson stream with an average failure rate of  $\tau > 0$ . The repair times follow a general dist., which is represented by the random variable  $D$  and the dist., function  $D(\zeta)$ , with the LST  $D^*(\omega)$ .

The length of repairs is determined by a general (arbitrary) dist., with a dist., function  $D(\zeta)$  and a density function  $d(\zeta)$ . Given an elapsed repair time of  $\zeta$ , define  $\alpha(\zeta)d\zeta$  as the conditional probability of completing repairs within the range  $(\zeta, \zeta + d\zeta]$ .

$$\alpha(\zeta) = \frac{d(\zeta)}{1 - D(\zeta)}$$

and therefore,

$$d(\omega) = \alpha(\omega)e^{-\int_0^\omega \alpha(\zeta)d\zeta}$$

**Extended Bernoulli vacation:** If there are any unfinished parts of the service, the server has two options: either accept the BV with a probability of  $\theta$  or keep serving them with a probability of  $(1 - \theta)$ . After the vacation is over, the server either undertakes the second type of optional extended Bernoulli vacation with a prob., of  $\mu$  or continues to serve the remaining batches with a prob., of  $(1 - \mu)$ .

The random variable  $F$  with the distance function  $F(\zeta)$  and LST  $F^*(\omega)$  is employed to represent the server's leisure time. This arbitrary variable  $F$  follows a general distribution.

The server's vacation time follows a general(arbitrary) dist., function  $F(\omega)$  and density function  $f(\omega)$ . Let  $\beta(\zeta)d\zeta$  be the conditional prob., of a completion of a vacation during the interval  $(\zeta, \zeta + d\zeta]$ , given that the elapsed repair time is  $\zeta$ , so that

$$\beta(\zeta) = \frac{f(\zeta)}{1 - F(\zeta)}$$

and therefore,

$$f(\omega) = \beta(\omega)e^{-\int_0^\omega \beta(\zeta)d\zeta}$$

The system's stochastic processes are all considered to be independent of one another.

**Feedback Rule:** Clients who are unhappy with their offerings can re-join the line once they've been completed, give feedback to receive another service with minimal difficulty, or both  $p$  ( $0 \leq p \leq 1$ ), otherwise the system must be terminated with complement prob.  $q = (1 - p)$

### 3. DEFINITIONS:

We define

1.  $P_n(\zeta, \omega)$  = Prob., that the server will be idle at time  $\omega$  with  $n(n \geq 0)$  customers in the orbit and  $\zeta$  for the customer's elapsed retrial time.
2.  $Q_n(\zeta, \omega)$  = Prob., that the server will be busy at time  $\omega$  with  $n(n \geq 0)$  customers in the orbit and  $\eta$  for the customer's elapsed retrial time.
3.  $R_n(\zeta, \omega)$  = Prob., that at time  $\omega$ , there are  $n(n \geq 0)$  customers in the orbit and the server is offline due to system repair and waiting for repairs to start with elapsed repair time  $\zeta$ .

4.  $V_n(\zeta, \omega)$  = Prob., that there are  $n(n \geq 0)$  consumers in orbit at time  $\omega$  and the server is on vacation with elapsed vacation time  $\zeta$ .
5. There are no customers in the orbit at time  $\omega$ , and the server is inactive but still available in the system, according to the probability  $P_0(\omega)$ .

The following differential-difference equations regulate the model:

$$\frac{d}{d\omega} P_0(\omega) = -\Lambda P_0(\omega) + (1 - \theta) \bar{d} \int_0^\infty Q_0(\zeta, \omega) \phi(\zeta) d\zeta + (1 - \mu) \int_0^\infty V_0(\zeta, \omega) \beta(\zeta) d\zeta \quad (1)$$

$$\frac{\partial}{\partial \zeta} P_n(\zeta, \omega) + \frac{\partial}{\partial \omega} P_n(\zeta, \omega) = -[\Lambda + g(\zeta)] P_n(\zeta, \omega), n \geq 1 \quad (2)$$

$$\frac{\partial}{\partial \zeta} Q_0(\zeta, \omega) + \frac{\partial}{\partial \omega} Q_0(\zeta, \omega) = -[\Lambda + \tau + \phi(\zeta)] Q_0(\zeta, \omega) \quad (3)$$

$$\frac{\partial}{\partial \zeta} Q_n(\zeta, \omega) + \frac{\partial}{\partial \omega} Q_n(\zeta, \omega) = -[\Lambda + \tau + \phi(\zeta)] Q_n(\zeta, \omega) + \Lambda \sum_{k=1}^n C_k Q_{n-k}(\zeta, \omega), n \geq 1 \quad (4)$$

$$\frac{\partial}{\partial \zeta} R_0(\zeta, \omega) + \frac{\partial}{\partial \omega} R_0(\zeta, \omega) = -[\Lambda + \alpha(\zeta)] R_0(\zeta, \omega), n = 0 \quad (5)$$

$$\frac{\partial}{\partial \zeta} R_n(\zeta, \omega) + \frac{\partial}{\partial \omega} R_n(\zeta, \omega) = -[\Lambda + \alpha(\zeta)] R_n(\zeta, \omega) + \Lambda \sum_{k=1}^n C_k R_{n-k}(\zeta, \omega), n \geq 1 \quad (6)$$

$$\frac{\partial}{\partial \zeta} V_0(\zeta, \omega) + \frac{\partial}{\partial \omega} V_0(\zeta, \omega) = -[\Lambda + \beta(\zeta)] V_0(\zeta, \omega), n = 0 \quad (7)$$

$$\frac{\partial}{\partial \zeta} V_n(\zeta, \omega) + \frac{\partial}{\partial \omega} V_n(\zeta, \omega) = -[\Lambda + \beta(\zeta)] V_n(\zeta, \omega) + \Lambda \sum_{k=1}^n C_k V_{n-k}(\zeta, \omega), n \geq 1 \quad (8)$$

The following boundary conditions must be met in order to answer the given equation:

$$P_n(0, \omega) = (1 - \theta) \bar{d} \int_0^\infty Q_n(\zeta, \omega) \phi(\zeta) d\zeta + (1 - \theta) d \int_0^\infty Q_{n-1}(\zeta, \omega) \phi(\zeta) d\zeta + \int_0^\infty R_n(\zeta, \omega) \alpha(\zeta) d\zeta + (1 - \mu) \int_0^\infty V_n(\zeta, \omega) \beta(\zeta) d\zeta, n \geq 1 \quad (9)$$

$$Q_0(0, \omega) = \Lambda p (1 - g) \sum_{r=a}^b \sum_{k=0}^{a-1} C_k \int_0^\infty P_{n-k+b}(\zeta, \omega) d\zeta + (1 - \theta) p \sum_{r=a}^b \int_0^\infty P_r(\zeta, \omega) g(\zeta) d\zeta + \sum_{r=a}^b \int_0^\infty V_r(\zeta, \omega) \beta(\zeta) d\zeta \quad (10)$$

$$Q_n(0, \omega) = \Lambda p (1 - g) \sum_{k=0}^{a-1} C_k \int_0^\infty P_{n-k+b}(\zeta, \omega) d\zeta + p \int_0^\infty P_{n+b}(\zeta, \omega) g(\zeta) d\zeta + \Lambda g \int_0^\infty P_{n+b}(\zeta, \omega) d\zeta + \int_0^\infty V_{n+b}(\zeta, \omega) \beta(\zeta) d\zeta \quad (11)$$

$$R_0(\zeta, 0, \omega) = \tau Q_0(\zeta, \omega), n = 0 \quad (12)$$

$$R_n(\zeta, 0, \omega) = \tau Q_n(\zeta, \omega), n \geq 1 \quad (13)$$

$$V_n(0, \omega) = \theta \int_0^\infty Q_n(\zeta, \omega) \phi(\zeta) d\zeta, n \geq 1 \quad (14)$$

We presume that the system is initially empty of users and that the server is idle. Thus, the initial conditions are

$$V_n(0) = R_n(0) = Q_n(0) = 0, n \geq 0$$

$$P_0(0) = 1, P_n^i(0) = 0, n \geq 1 \quad (15)$$

Generating functions of the queue length (The time-dependent solution):

$$\begin{aligned}
 P(\zeta, \Psi, \omega) &= \sum_{n=0}^{\infty} \Psi^n P_n^i(\zeta, \omega); P(\Psi, \omega) = \sum_{n=0}^{\infty} \Psi^n P_n(\omega) \\
 Q(\zeta, \Psi, \omega) &= \sum_{n=0}^{\infty} \Psi^n Q_n(\zeta, \omega); Q(\Psi, \omega) = \sum_{n=0}^{\infty} \Psi^n Q_n(\omega) \\
 R(\zeta, \iota, \Psi, \omega) &= \sum_{n=0}^{\infty} \Psi^n R_n(\zeta, \iota, \omega); R(\zeta, \Psi, \omega) = \sum_{n=0}^{\infty} \Psi^n R_n(\zeta, \omega) \\
 V(\zeta, \Psi, \omega) &= \sum_{n=0}^{\infty} \Psi^n V_n(\zeta, \omega); V(\Psi, \omega) = \sum_{n=0}^{\infty} \Psi^n V_n(\omega) \\
 C(\Psi) &= \sum_{n=1}^{\infty} C_n \Psi^n; Q(\Psi) = \sum_{r=0}^{a-1} Q_r \Psi^r
 \end{aligned} \tag{16}$$

which define the LT of a function  $f(\omega)$  as it converges within the circle defined by  $z \leq 1$ .

$$\tilde{f}(s) = \int_0^{\infty} e^{-s\omega} f(\omega) d\omega, \mathcal{R}(s) \geq 0 \tag{17}$$

Using (15) and the LT from equations (1) through (14), we arrive at

$$(s + \Lambda) \bar{p}_0(s) = 1 + (1 - \theta) \bar{d} \int_0^{\infty} \bar{Q}_0(\zeta, s) \phi(\zeta) d\zeta + (1 - \mu) \int_0^{\infty} \bar{V}_0(\zeta, s) \beta(\zeta) d\zeta \tag{18}$$

$$\frac{\partial}{\partial \zeta} \bar{P}_n(\zeta, s) + [s + \Lambda + g(\zeta)] \bar{P}_n(\zeta, s) = 0, n \geq 1 \tag{19}$$

$$\frac{\partial}{\partial \zeta} \bar{Q}_0(\zeta, s) + [s + \Lambda + \phi(\zeta)] \bar{Q}_0(\zeta, s) = 0 \tag{20}$$

$$\frac{\partial}{\partial \zeta} \bar{Q}_n(\zeta, s) + [s + \Lambda + \phi(\zeta)] \bar{Q}_n(\zeta, s) = \Lambda \sum_{k=1}^n C_k \bar{Q}_{n-k}(\zeta, s), n \geq 1 \tag{21}$$

$$\frac{\partial}{\partial \zeta} \bar{R}_0(\zeta, \iota, s) + [s + \Lambda + \alpha(\zeta)] \bar{R}_0(\zeta, s) = 0 \tag{22}$$

$$\frac{\partial}{\partial \zeta} \bar{R}_n(\zeta, \iota, s) + [s + \Lambda + \alpha(\zeta)] \bar{R}_n(\zeta, s) = \Lambda \sum_{k=1}^n C_k \bar{R}_{n-k}(\zeta, s), n \geq 1 \tag{23}$$

$$\frac{\partial}{\partial \zeta} \bar{V}_0(\zeta, s) + [s + \Lambda + \beta(\zeta)] \bar{V}_0(\zeta, s) = 0 \tag{24}$$

$$\frac{\partial}{\partial \zeta} \bar{V}_n(\zeta, s) + [s + \Lambda + \beta(\zeta)] \bar{V}_n(\zeta, s) = \Lambda \sum_{k=1}^n C_k \bar{V}_{n-k}(\zeta, s), n \geq 1 \tag{25}$$

$$\begin{aligned} \bar{P}_n(0, s) &= (1 - \theta)\bar{d} \int_0^\infty \bar{Q}_n(\zeta, s)\phi(\zeta)d\zeta + (1 - \theta)d \int_0^\infty \bar{Q}_{n-1}(\zeta, s)\phi(\zeta)d\zeta \\ &+ \int_0^\infty \bar{R}_n(\zeta, s)\alpha(\zeta)d\zeta + (1 - \mu) \int_0^\infty \bar{V}_n(\zeta, s)\beta(\zeta)d\zeta, n \geq 1 \end{aligned} \quad (26)$$

$$\begin{aligned} \bar{Q}_0(0, s) &= \Lambda p(1 - g) \sum_{r=a}^b \sum_{k=0}^{a-1} C_k \int_0^\infty \bar{P}_{n-k+b}(\zeta, s)d\zeta \\ &+ (1 - \theta)p \sum_{r=a}^b \int_0^\infty \bar{P}_r(\zeta, s)g(\zeta)d\zeta + \sum_{r=a}^b \int_0^\infty \bar{V}_r(\zeta, s)\beta(\zeta)d\zeta \end{aligned} \quad (27)$$

$$\begin{aligned} \bar{Q}_n(0, s) &= \Lambda p(1 - g) \sum_{k=0}^{a-1} C_k \int_0^\infty \bar{P}_{n-k+b}(\zeta, s)d\zeta + p \int_0^\infty \bar{P}_{n+b}(\zeta, s)g(\zeta)d\zeta \\ &+ \Lambda g \int_0^\infty \bar{P}_{n+b}(\zeta, s)d\zeta + \int_0^\infty \bar{V}_{n+b}(\zeta, s)\beta(\zeta)d\zeta \end{aligned} \quad (28)$$

$$\bar{R}_0(\zeta, 0, s) = \tau\bar{Q}_0(\zeta, s), n = 0 \quad (29)$$

$$\bar{R}_n(\zeta, 0, s) = \tau\bar{Q}_n(\zeta, s), n \geq 1 \quad (30)$$

$$\bar{V}_n(0, s) = \theta \int_0^\infty \bar{Q}_n(\zeta, s)\phi(\zeta)d\zeta, n \geq 1 \quad (31)$$

By multiplying equations (19) through (31) by  $\Psi^n$  and adding the results over  $n$ , we can obtain using the generating function mentioned in equation (16).

$$\frac{\partial}{\partial \zeta} \bar{P}(\zeta, \Psi, s) + [s + \Lambda + g(\zeta)]\bar{P}(\zeta, \Psi, s) = 0 \quad (32)$$

$$\frac{\partial}{\partial \zeta} \bar{Q}(\zeta, \Psi, s) + [s + \Lambda(1 - C(\Psi)) + \phi(\zeta)]\bar{Q}(\zeta, \Psi, s) = 0 \quad (33)$$

$$\frac{\partial}{\partial \zeta} \bar{R}(\zeta, \Psi, s) + [s + \Lambda(1 - C(\Psi)) + \alpha(\zeta)]\bar{R}(\zeta, \Psi, s) = 0 \quad (34)$$

$$\frac{\partial}{\partial \zeta} \bar{V}(\zeta, \Psi, s) + [s + \Lambda(1 - C(\Psi)) + \beta(\zeta)]\bar{V}(\zeta, \Psi, s) = 0 \quad (35)$$

$$\begin{aligned} \bar{P}(0, \Psi, s) &= (1 - \theta)(\bar{d} + d\Psi) \int_0^\infty \bar{Q}(\zeta, \Psi, s)\phi(\zeta)d\zeta + \int_0^\infty \bar{R}(\zeta, \Psi, s)\alpha(\zeta)d\zeta \\ &+ (1 - \mu) \int_0^\infty \bar{V}(\zeta, \Psi, s)\beta(\zeta)d\zeta - \bar{d}(1 - \theta) \int_0^\infty \bar{Q}_0(\zeta, s)\phi(\zeta)d\zeta \\ &- (1 - \mu) \int_0^\infty \bar{V}_0(\zeta, s)\beta(\zeta)d\zeta, n \geq 1 \end{aligned} \quad (36)$$

$$\begin{aligned} \Psi^b \bar{Q}(0, \Psi, s) &= \Lambda(1 - g)pC(\Psi) \int_0^\infty \bar{P}(\zeta, \Psi, s)d\zeta + p \int_0^\infty \bar{P}(\zeta, \Psi, s)g(\zeta)d\zeta \\ &+ \Lambda g \int_0^\infty \bar{P}(\zeta, \Psi, s)d\zeta + \int_0^\infty \bar{V}(\zeta, \Psi, s)\beta(\zeta)d\zeta \end{aligned} \quad (37)$$

$$\bar{R}(\zeta, 0, \Psi, s) = \tau\bar{Q}(\zeta, \Psi, s), n \geq 1 \quad (38)$$

$$\bar{V}(0, \Psi, s) = \theta \int_0^\infty \bar{Q}(\zeta, \Psi, s)\phi(\zeta)d\zeta, n \geq 1 \quad (39)$$

Equation (18) in (36) gives us

$$\begin{aligned} \bar{P}(0, \Psi, s) &= [1 - (s + \Lambda)\bar{P}_0(s)] + (1 - \theta)(\bar{d} + d\Psi) \int_0^\infty \bar{Q}(\zeta, \Psi, s)\phi(\zeta)d\zeta \\ &+ \int_0^\infty \bar{R}(\zeta, \Psi, s)\alpha(\zeta)d\zeta + (1 - \mu) \int_0^\infty \bar{V}(\zeta, \Psi, s)\beta(\zeta)d\zeta \end{aligned} \quad (40)$$

Equation (32), when integrated between 0 and  $\zeta$ , yields

$$\bar{P}(\zeta, \Psi, s) = \bar{P}(0, \Psi, s)e^{-(s+\Lambda)\zeta - \int_0^\zeta g(\omega)d\omega} \tag{41}$$

Once more, integrating equation (41) by parts with respect to  $\zeta$  yields,

$$\bar{P}(\Psi, s) = \bar{P}(0, \Psi, s) \left[ \frac{1 - \bar{A}(s + \Lambda)}{s + \Lambda} \right] \tag{42}$$

where,

$$\bar{A}(s + \Lambda) = \int_0^\infty e^{-(s+\Lambda)\zeta} dA(\zeta)$$

When integrating equations (33) to (35) from 0 to  $\zeta$ , similar outcomes are found.

$$\bar{Q}(\zeta, \Psi, s) = \bar{Q}(0, \Psi, s)e^{-\zeta(\Psi, s)\zeta - \int_0^\zeta \phi(\omega)d\omega} \tag{43}$$

$$\bar{R}(\zeta, \iota, \Psi, s) = \bar{R}(\zeta, 0, \Psi, s)e^{-\zeta(\Psi, s)\zeta - \int_0^\zeta \alpha(\omega)d\omega}$$

$$\bar{R}(\zeta, \Psi, s) = \bar{R}(\zeta, 0, \Psi, s) \left[ \frac{1 - \bar{D}(\zeta(\Psi, s))}{\zeta(\Psi, s)} \right] \tag{44}$$

$$\bar{V}(\zeta, \Psi, s) = \bar{V}(0, \Psi, s)e^{-\zeta(\Psi, s)\zeta - \int_0^\zeta \beta(\omega)d\omega} \tag{45}$$

where the values of  $\bar{P}(0, \Psi, s), \bar{Q}(0, \Psi, s), \bar{R}(0, \Psi, s)$  and  $\bar{V}(0, \Psi, s)$  are given by (37) to (40). Taking into account  $\zeta$  yields, integrate equations (43) to (45) by parts once more.

$$\bar{Q}(\Psi, s) = \bar{Q}(0, \Psi, s) \left[ \frac{1 - \bar{B}(\zeta(\Psi, s))}{\zeta(\Psi, s)} \right] \tag{46}$$

$$\bar{R}(\Psi, s) = \tau \bar{Q}(0, \Psi, s) \left[ \frac{1 - \bar{B}(\zeta(\Psi, s))}{\zeta(\Psi, s)} \right] \left[ \frac{1 - \bar{D}(\zeta(\Psi, s))}{\zeta(\Psi, s)} \right] \tag{47}$$

$$\bar{V}(\Psi, s) = \bar{V}(0, \Psi, s) \left[ \frac{1 - \bar{F}(\zeta(\Psi, s))}{\zeta(\Psi, s)} \right] \tag{48}$$

Where,

$$\bar{B}(\zeta(\Psi, s)) = \int_0^\infty e^{-\zeta(\Psi, s)\zeta} dB(\zeta)$$

$$\bar{D}(\zeta(\Psi, s)) = \int_0^\infty e^{-\zeta(\Psi, s)\zeta} dD(\zeta)$$

$$\bar{F}(\zeta(\Psi, s)) = \int_0^\infty e^{-\zeta(\Psi, s)\zeta} dF(\zeta)$$

are, in order, the LST of the following values: retrial time  $A(\zeta)$ , service time  $B(\zeta)$ , repair time  $D(\zeta)$ , and vacation time  $F(\zeta)$ .

Now, multiplying both side of equations (41),(43) to (45) by  $g(\zeta), \phi(\zeta), \alpha(\zeta)$  and  $\beta(\zeta)$  and integrating over  $\zeta$ , we obtain

$$\int_0^\infty \bar{P}(\zeta, \Psi, s)g(\zeta)d\zeta = \bar{P}(0, \Psi, s)\bar{A}(s + \Lambda) \tag{49}$$

$$\int_0^\infty \bar{Q}(\zeta, \Psi, s)\phi(\zeta)d\zeta = \bar{Q}(0, \Psi, s)\bar{B}(\zeta(\Psi, s)) \tag{50}$$

$$\int_0^\infty \bar{R}(\zeta, \iota, \Psi, s)\alpha(\zeta)d\zeta = \bar{R}(\zeta, 0, \Psi, s)\bar{D}(\zeta(\Psi, s)) \tag{51}$$

$$\int_0^\infty \bar{V}(\zeta, \Psi, s)\beta(\zeta)d\zeta = \bar{V}(0, \Psi, s)\bar{F}(\zeta(\Psi, s)) \tag{52}$$

Using equations (50) in (39)

$$\bar{V}(0, \Psi, s) = \theta \bar{Q}(0, \Psi, s) \bar{B}(\zeta(\Psi, s)) \tag{53}$$

Using equations (49) in (37) and (38), we get

$$\bar{Q}(0, \Psi, s) = \frac{\bar{P}(0, \Psi, s)}{\Psi^b - \theta \bar{F}(\zeta(\Psi, s)) \bar{B}(\zeta(\Psi, s))} \left[ \Lambda(1-g)pC(\Psi) \left( \frac{1 - \bar{A}(s+\Lambda)}{s+\Lambda} \right) + p\bar{A}(s+\Lambda) + \Lambda g \left( \frac{1 - \bar{A}(s+\Lambda)}{s+\Lambda} \right) \right] \tag{54}$$

$$\bar{R}(\zeta, 0, \Psi, s) = \tau \bar{Q}(0, \Psi, s) \left( \frac{1 - \bar{B}(\zeta(\Psi, s))}{(\zeta(\Psi, s))} \right) \tag{55}$$

Using equation (50) to (52) in (40) we get

$$\bar{P}(0, \Psi, s) = \frac{Nr(\Psi)}{Dr(\Psi)} \tag{56}$$

$$\begin{aligned} Nr(\Psi) &= [1 - (s + \Lambda)\bar{P}_0(s)][\Psi^b - \theta \bar{F}(\zeta(\Psi, s)) \bar{B}(\zeta(\Psi, s))] \\ Dr(\Psi) &= \Psi^b - \theta \bar{F}(\zeta(\Psi, s)) \bar{B}(\zeta(\Psi, s)) \\ &\quad - \left[ \Lambda(1-g)pC(\Psi) \left( \frac{1 - \bar{A}(s+\Lambda)}{s+\Lambda} \right) + p\bar{A}(s+\Lambda) + \Lambda g \left( \frac{1 - \bar{A}(s+\Lambda)}{s+\Lambda} \right) \right] \\ &\quad \left[ (1 - \theta)(\bar{d} + d\Psi) \bar{B}(\zeta(\Psi, s)) + \tau \bar{D}(\zeta(\Psi, s)) \left( \frac{1 - \bar{B}(\zeta(\Psi, s))}{(\zeta(\Psi, s))} \right) \right. \\ &\quad \left. + \theta(1 - \mu) \bar{F}(\zeta(\Psi, s)) \bar{B}(\zeta(\Psi, s)) \right] \end{aligned}$$

where e,

$$\zeta(\Psi, s) = s + \Lambda(1 - C(\Psi))$$

Subs/-  $\bar{P}(0, \Psi, s)$  from equation (56) into equation (53) to (55)

$$\bar{Q}(0, \Psi, s) = \left[ \frac{\Lambda(1-g)pC(\Psi) \left( \frac{1 - \bar{A}(s+\Lambda)}{s+\Lambda} \right) + p\bar{A}(s+\Lambda) + \Lambda g \left( \frac{1 - \bar{A}(s+\Lambda)}{s+\Lambda} \right)}{\Psi^b - \theta \bar{F}(\zeta(\Psi, s)) \bar{B}(\zeta(\Psi, s))} \right] \left[ \frac{Nr(\Psi)}{Dr(\Psi)} \right] \tag{57}$$

$$\bar{R}(\zeta, 0, \Psi, s) = \tau \left( \frac{1 - \bar{B}(\zeta(\Psi, s))}{(\zeta(\Psi, s))} \right) \left[ \frac{Nr(\Psi)}{Dr(\Psi)} \right] \left[ \frac{\Lambda(1-g)pC(\Psi) \left( \frac{1 - \bar{A}(s+\Lambda)}{s+\Lambda} \right) + p\bar{A}(s+\Lambda) + \Lambda g \left( \frac{1 - \bar{A}(s+\Lambda)}{s+\Lambda} \right)}{\Psi^b - \theta \bar{F}(\zeta(\Psi, s)) \bar{B}(\zeta(\Psi, s))} \right] \tag{58}$$

$$\bar{V}(0, \Psi, s) = \theta \bar{B}(\zeta(\Psi, s)) \left[ \frac{Nr(\Psi)}{Dr(\Psi)} \right] \left[ \frac{\Lambda(1-g)pC(\Psi) \left( \frac{1 - \bar{A}(s+\Lambda)}{s+\Lambda} \right) + p\bar{A}(s+\Lambda) + \Lambda g \left( \frac{1 - \bar{A}(s+\Lambda)}{s+\Lambda} \right)}{\Psi^b - \theta \bar{F}(\zeta(\Psi, s)) \bar{B}(\zeta(\Psi, s))} \right] \tag{59}$$

Updating equations (56) to (59) in (42), (46) to (48) We determine the PGF of various conditions in the system under a transient condition.



#### 4. THE STEADY STATE'S FINDINGS:

To define the SS prob., we disregard the argument  $\omega$  whenever it appears in the time-dependent analysis.

$$\lim_{s \rightarrow 0} s\bar{f}(s) = \lim_{\omega \rightarrow \infty} f(\omega)$$

$$P(\Psi) = P(0, \Psi) \left( \frac{1 - \bar{A}(\Lambda)}{\Lambda} \right) \tag{60}$$

$$Q(\Psi) = \left( \frac{1 - \bar{B}(\zeta(\Psi))}{\zeta(\Psi)} \right) P(0, \Psi) \tag{61}$$

$$\left[ \frac{\Lambda(1 - g)pC(\Psi) \left( \frac{1 - \bar{A}(\Lambda)}{\Lambda} \right) + p\bar{A}(\Lambda) + \Lambda g \left( \frac{1 - \bar{A}(\Lambda)}{\Lambda} \right)}{\Psi^b - \theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))} \right] \tag{62}$$

$$R(\Psi) = \tau \left( \frac{1 - \bar{B}(\zeta(\Psi))}{\zeta(\Psi)} \right) \left( \frac{1 - \bar{D}(\zeta(\Psi))}{\zeta(\Psi)} \right) \\ P(0, \Psi) \left[ \frac{\Lambda(1 - g)pC(\Psi) \left( \frac{1 - \bar{A}(\Lambda)}{\Lambda} \right) + p\bar{A}(\Lambda) + \Lambda g \left( \frac{1 - \bar{A}(\Lambda)}{\Lambda} \right)}{\Psi^b - \theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))} \right] \tag{63}$$

$$V(\Psi) = \theta\bar{B}(\zeta(\Psi)) \left( \frac{1 - \bar{F}(\zeta(\Psi))}{\zeta(\Psi)} \right) \\ P(0, \Psi) \left[ \frac{\Lambda(1 - g)pC(\Psi) \left( \frac{1 - \bar{A}(\Lambda)}{\Lambda} \right) + p\bar{A}(\Lambda) + \Lambda g \left( \frac{1 - \bar{A}(\Lambda)}{\Lambda} \right)}{\Psi^b - \theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))} \right] \tag{64}$$

where,

$$P(0, \Psi) = \frac{Nr(\Psi)}{Dr(\Psi)} \\ Nr(\Psi) = [1 - \Lambda\bar{P}_0][\Psi^b - \theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))] \\ Dr(\Psi) = \Psi^b - \theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi)) \\ - \left[ \Lambda(1 - g)pC(\Psi) \left( \frac{1 - \bar{A}(\Lambda)}{\Lambda} \right) + p\bar{A}(\Lambda) + \Lambda g \left( \frac{1 - \bar{A}(\Lambda)}{\Lambda} \right) \right] \\ \left[ (1 - \theta)(\bar{d} + d\Psi)\bar{B}(\zeta(\Psi)) + \tau\bar{D}(\zeta(\Psi)) \left( \frac{1 - \bar{B}(\zeta(\Psi))}{\zeta(\Psi)} \right) \right] \\ + \theta(1 - \mu)\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi)) \tag{65}$$

### 4.1. Queue sizes distribution at a certain epoch:

The PGF is a of the queue size dist., at a random interval, is obtained by adding (60) to (63) with the idle term.

$$\begin{aligned}
 K(\Psi) &= \frac{Nr(\Psi)}{Dr(\Psi)} \tag{66} \\
 Nr(\Psi) &= \Lambda P_0 \zeta(\Psi) \left( \Psi^b - \theta \bar{F}(\zeta(\Psi)) \bar{B}(\zeta(\Psi)) - [(1-g)pC(\Psi)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda)] \right. \\
 &\quad \left. + g(1-\bar{A}(\Lambda)) \right) \left[ (1-\theta)(\bar{d} + d\Psi) \bar{B}(\zeta(\Psi)) + \tau \bar{D}(\zeta(\Psi)) \left( \frac{1-\bar{B}(\zeta(\Psi))}{(\zeta(\Psi))} \right) \right. \\
 &\quad \left. + \theta(1-\mu) \bar{F}(\zeta(\Psi)) \bar{B}(\zeta(\Psi)) \right] - (1-\bar{A}(\Lambda)) \zeta(\Psi) [\Psi^b - \theta \bar{F}(\zeta(\Psi)) \bar{B}(\zeta(\Psi))] \\
 &\quad + \Lambda [(1-g)pC(\Psi)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))] \\
 &\quad [(1-\bar{B}\zeta(\Psi)) + \tau(1-\bar{B}\zeta(\Psi))(1-\bar{D}\zeta(\Psi)) + \theta \bar{B}\zeta(\Psi)(1-\bar{F}\zeta(\Psi))] \\
 &\quad + (1-\bar{A}(\Lambda)) \zeta(\Psi) [\Psi^b - \theta \bar{F}(\zeta(\Psi)) \bar{B}(\zeta(\Psi))] \\
 &\quad + \Lambda [(1-g)pC(\Psi)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))] \\
 &\quad [(1-\bar{B}\zeta(\Psi)) + \tau(1-\bar{B}\zeta(\Psi))(1-\bar{D}\zeta(\Psi)) + \theta \bar{B}\zeta(\Psi)(1-\bar{F}\zeta(\Psi))] \\
 Dr(\Psi) &= \zeta(\Psi) \Lambda \left\{ \Psi^b - \theta \bar{F}(\zeta(\Psi)) \bar{B}(\zeta(\Psi)) - [(1-g)pC(\Psi)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda)] \right. \\
 &\quad \left. + g(1-\bar{A}(\Lambda)) \right) \left( (1-\theta)(\bar{d} + d\Psi) \bar{B}(\zeta(\Psi)) + \tau \bar{D}(\zeta(\Psi)) \left( \frac{1-\bar{B}(\zeta(\Psi))}{(\zeta(\Psi))} \right) \right. \\
 &\quad \left. + \theta(1-\mu) \bar{F}(\zeta(\Psi)) \bar{B}(\zeta(\Psi)) \right) \left. \right\}
 \end{aligned}$$

### 5. STABILITY CONDITION

The PGF needs to meet  $P(1)=1$ . Applying the L'Hopital rules and equating the expression to 1 results in the result that satisfies the requirement.

$$\begin{aligned}
 &b - [(1-g)pE(I)(1-\bar{A}(\Lambda))][(1-\theta)(d+\bar{d}) + \theta(1-\mu)] + p(1-g)(1-\bar{A}(\Lambda)) \\
 &+ p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))[(1-\theta)(d+\bar{d})(1-\Lambda E(I)E(B)) + \tau E(B)] \\
 &- \theta \Lambda(1-\mu)E(I)A_1 + \Lambda \theta E(I)A_1
 \end{aligned}$$

Now we can determine the prob., that are unknown.  $P(1)=1$  is therefore fulfilled if

$$\left[ \Psi^b - \theta \bar{F}(\zeta(\Psi)) \bar{B}(\zeta(\Psi)) - [(1-g)pC(\Psi)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))] \right. \\
 \left. \left[ (1-\theta)(\bar{d} + d\Psi) \bar{B}(\zeta(\Psi)) + \tau \bar{D}(\zeta(\Psi)) \left( \frac{1-\bar{B}(\zeta(\Psi))}{(\zeta(\Psi))} \right) + \theta(1-\mu) \bar{F}(\zeta(\Psi)) \bar{B}(\zeta(\Psi)) \right] \right] > 0$$

$$\rho = \frac{[(1-g)pE(I)(1-\bar{A}(\Lambda))][(1-\theta)(d+\bar{d}) + \theta(1-\mu)] + [p(1-g)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g((1-\bar{A}(\Lambda))][(1-\theta)(d+\bar{d})(1-\Lambda E(I)E(B)) + \tau E(B)] - \theta \Lambda(1-\mu)E(I)A_1] + \Lambda \theta E(I)A_1}{b} \tag{67}$$

then  $\rho < 1$  is the condition to be satisfied for the existence of the SS for the model under consideration.

### 6. PERFORMANCE EVALUATION:

This section includes system performance metrics, a model stability study, and some unique system prob., while the system is in various states.

We obtain the following prob., if the system fulfills the stability requirement  $\rho < 1$ .

- Let  $P$  be the SS Prob., that the server is idle during the retrial time.

$$P = \lim_{\Psi \rightarrow 1} P(\Psi) = P(1) = \frac{(1 - \theta)(1 - \Lambda p_0)(1 - \bar{A}(\Lambda))}{\Lambda(1 - \theta) - [p(1 - g)(1 - \bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1 - \bar{A}(\Lambda))]} \frac{1}{[(1 - \theta)(d + \bar{d}) + \theta(1 - \mu)]}$$

- If the server is busy, let  $Q$  be the SS Prob.,

$$Q = \lim_{\Psi \rightarrow 1} Q(\Psi)$$

$$Q(1) = E(B) \times \left\{ \frac{(1 - \Lambda p_0)[p(1 - g)(1 - \bar{A}(\Lambda))E(I)]}{b + \Lambda\theta E(I)A_1 - [(1 - g)pE(I)(1 - \bar{A}(\Lambda))][(1 - \theta)(d + \bar{d}) + \theta(1 - \mu)] + [p(1 - g)(1 - \bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1 - \bar{A}(\Lambda))]}{(1 - \theta)(d + \bar{d})(1 - \Lambda E(I)E(B)) + \tau E(B) - \Lambda\theta(1 - \mu)E(I)A_1} \right\}$$

- $R$  ought to indicate the SS Prob., that the server is being repaired.

$$R = \lim_{\Psi \rightarrow 1} R(\Psi)$$

$$R(1) = \tau E(B)E(D) \times \left\{ \frac{(1 - \Lambda p_0)[p(1 - g)(1 - \bar{A}(\Lambda))E(I)]}{b + \Lambda\theta E(I)A_1 - [(1 - g)pE(I)(1 - \bar{A}(\Lambda))][(1 - \theta)(d + \bar{d}) + \theta(1 - \mu)] + [p(1 - g)(1 - \bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1 - \bar{A}(\Lambda))]}{[(1 - \theta)(d + \bar{d})(1 - \Lambda E(I)E(B)) + \tau E(B) - \Lambda\theta(1 - \mu)E(I)A_1]} \right\}$$

- Using  $V$  as the SS Prob., we may assume that the server is on vacation.

$$V = \lim_{\Psi \rightarrow 1} V(\Psi)$$

$$V(1) = \theta E(F)E(I) \times \left\{ \frac{(1 - \Lambda p_0)(-\Lambda E(B)p(1 - g)(1 - \bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1 - \bar{A}(\Lambda)) + p(1 - g)(1 - \bar{A}(\Lambda)))}{b + \Lambda\theta E(I)A_1 - [(1 - g)pE(I)(1 - \bar{A}(\Lambda))][(1 - \theta)(d + \bar{d}) + \theta(1 - \mu)] + [p(1 - g)(1 - \bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1 - \bar{A}(\Lambda))][(1 - \theta)(d + \bar{d})]}{(1 - \Lambda E(I)E(B)) + \tau E(B) - \Lambda\theta(1 - \mu)E(I)A_1} \right\}$$

### 6.1. Average queue length:

Computing at  $\Psi = 1$  and differentiating (65) with regard to  $\Psi$  yields the mean number of users in the queue ( $L_q$ ) under SS conditions.

$$L_q = \lim_{\Psi \rightarrow 1} \frac{d}{d\Psi} P(\Psi)$$

$$P'(1) = \frac{Nr''(1)Dr'(1) - Dr''(1)Nr'(1)}{2(Dr'(1))^2}$$

$$\begin{aligned}
 D'(1) &= -\Lambda^2 E(I) \left\{ \Psi^b - \theta \bar{F}(\zeta(\Psi)) \bar{B}(\zeta(\Psi)) - [(1-g)pC(\Psi)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda)] \right. \\
 &\quad \left. + g(1-\bar{A}(\Lambda)) \left[ (1-\theta)(\bar{d} + d\Psi) \bar{B}(\zeta(\Psi)) + \tau \bar{D}(\zeta(\Psi)) \left( \frac{1-\bar{B}(\zeta(\Psi))}{\zeta(\Psi)} \right) \right] \right. \\
 &\quad \left. + \theta(1-\mu) \bar{F}(\zeta(\Psi)) \bar{B}(\zeta(\Psi)) \right\} \\
 D''(1) &= -\Lambda^2 \{ E(I(I-1)) [1-\theta - [p-g(p-1)(1-\bar{A}(\Lambda))] [1-\theta\mu] \\
 &\quad + 2E(I) [b + \Lambda\theta E(I)A_1 - ((1-g)pE(I)(1-\bar{A}(\Lambda))] [(1-\theta)(\bar{d} + \bar{d}) + \theta(1-\mu)] \\
 &\quad + [p(1-g)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))] [(1-\theta)(\bar{d} + \bar{d})(1-\Lambda E(I)E(B)) \\
 &\quad + \tau E(B) - \Lambda\theta(1-\mu)E(I)A_1] \} \\
 N'(1) &= -\Lambda^2 E(I) \{ 1-\theta - [p(1-g)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))] [(1-\theta)(\bar{d} + \bar{d}) \\
 &\quad + \theta(1-\mu)] \} + (1-\Lambda) \left\{ -\Lambda E(I)(1-\theta)(1-\bar{A}(\Lambda)) + \Lambda^2 E(I)E(B) \right. \\
 &\quad \left. (p(1-g)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))) + \theta\Lambda E(I)E(F) \right\} \\
 N''(1) &= -\Lambda^2 \{ E(I(I-1)) (1-\theta - A_4(1-\theta\mu)) + 2E(I) (b + \theta\Lambda E(I)A_1 - A_2(1-\theta\mu)) \\
 &\quad + A_4[(1-\theta)(1-\Lambda E(I)E(B))] + \tau E(B) - \theta\Lambda(1-\mu)E(I)A_1 \} \\
 &\quad + (1-\Lambda) \{ (1-\bar{A}(\Lambda)) [-(1-\theta)\Lambda E(I(I-1)) - \Lambda E(I)(b + \theta\Lambda E(I)A_1)] \\
 &\quad + \Lambda^2 E(I)E(B)A_2 + \Lambda^2 [E(I(I-1))E(B) + E(I)E^2(B)] [p(1-g)((1-\bar{A}(\Lambda))) \\
 &\quad + p\bar{A}(\Lambda) + g((1-\bar{A}(\Lambda)))] - \theta\Lambda^2 E(B)E(F)E(I)^2 \\
 &\quad + \Lambda\theta [E(I(I-1))E(F) + E(I)E^2(F)] \}
 \end{aligned}$$

where e,

$$\begin{aligned}
 A_1 &= E(B) + E(F) \\
 A_2 &= p(1-g)E(I)((1-\bar{A}(\Lambda))) \\
 A_4 &= p-g(p-1)(1-\bar{A}(\Lambda))
 \end{aligned}$$

- The Little's formula ( $W_q$ ) is used to determine how long an average customer waits in queue.

$$W_q = \frac{L_q}{\Lambda E(I)}$$

### 7. Practical application of the model:

The field of telecommunications networks may be able to use the suggested model. This system manages a lot of consumer telephone communications. Call takers are referred to as servers and callers as customers in this context. A consumer may elect to exit the system if he calls and discovers that all the servers are occupied (impatience). Customers wait in orbit while the server is overloaded, out of commission, or under going maintenance. If a server has any questions or concerns that fall outside of their area of expertise, they may need to refer them to other servers who are available or speak with a senior in order to acquire the answers. A service failure can be used to represent this circumstance. The speed at which the agent receives responses from the expert in this case is known as the repair rate. Additionally, the server may do various maintenance procedures known as "vacations." Additionally, after each customer's service is finished dissatisfied customers may re-join the line and be classified as feedback consumers.

### 8. Numerical Results

In this section, we'll use MATLAB to demonstrate how different parameters affect observations of system behavior. The batch size distance of the arrivals in this section is geometry; with a mean

of 2. Here, the exponential distance is followed by the service, vacation, and repair stages. By creating erroneous assumptions about the parameters, we make sure that the stability criterion is satisfied. Tables 1 to 3 present estimated values for our queueing system's utilization factor ( $\rho$ ), average queue length ( $L_q$ ), and average waiting time ( $W_q$ ).

**Table 1:** The effects of arrival rate ( $\Lambda$ ) on  $\rho$ ,  $L_q$ , and  $W_q$

$g = 0.5, p = 1.5, E = 0.6, G = 2.2, \theta = 3, d = 3,$   
 $e = 0.6, \mu = 0.9, B = 1.5, D = 1, F = 0.7, z = 1, b = 2, \tau = 1.8$

Arrival rate ( $\Lambda$ )	$\rho$	$L_q$	$W_q$
0.30	0.022680	3.505127	5.841879
0.31	0.096336	4.616984	7.446748
0.32	0.169992	6.010404	9.391256
0.33	0.243648	7.742148	11.730527
0.34	0.317304	9.878047	14.526540
0.35	0.390960	12.494169	17.848813
0.36	0.464616	15.678118	21.775164

**Table 2:** The effects of the service rate  $\phi(\zeta)$  on  $\rho$ ,  $L_q$ ,  $W_q$

$g = 7.8, p = 0.7, E = 0.8, G = 6, \theta = 1, d = 3, e = 4.6,$   
 $\mu = 0.7, D = 1, F = 0.7, \Lambda = 0.3, z = 1, b = 2, \tau = 1$

service rate ( $B$ )	$\rho$	$L_q$	$W_q$
0.50	0.737200	0.223375	0.372292
0.51	0.687360	0.189247	0.315412
0.52	0.637520	0.160488	0.267480
0.53	0.587680	0.135989	0.226649
0.54	0.537840	0.114943	0.191572
0.55	0.488000	0.096748	0.161247
0.56	0.438160	0.080947	0.134912

**Table 3:** The effects of the Breakdown rate ( $\tau$ ) on  $\rho$ ,  $L_q$ ,  $W_q$

$g = 0.2, p = 0.7, E = 2.9, G = 9, \theta = 1,$   
 $d = 7, e = 8.6, \mu = 0.2, B = 7, D = 2, F = 0.7, \Lambda = 0.4, z = 2, b = 4$

breakdown rate ( $\tau$ )	$\rho$	$L_q$	$W_q$
1.0	0.264592	5.619523	7.024404
1.1	0.303092	6.095933	7.619916
1.2	0.341592	6.575234	8.219042
1.3	0.380092	7.057428	8.821784
1.4	0.418592	7.542516	9.428145
1.5	0.457092	8.030500	10.038125
1.6	0.495592	8.521381	10.651727

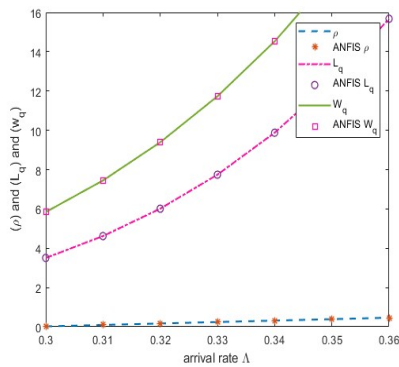
The two-dimensional graph that represents the system measurement of performance is shown in Figure 1 (a – c).

- The figure 1 (a) demonstrates how the utilization factor ( $\rho$ ), estimated queue length ( $L_q$ ), and expected waiting time ( $W_q$ ) all increase as the arrival rate ( $\Lambda$ ) does.
- The figure 1 (b) shows that while the utilization factor ( $\rho$ ) decreases, the service rate  $\phi(\zeta)$  rises. Expected waiting time ( $W_q$ ) and queue length ( $L_q$ ) decrease.
- The breakdown rate ( $\tau$ ), utilization factor ( $\rho$ ), expected queue size ( $L_q$ ), and expected waiting time ( $W_q$ ) all show increasing trends in the figure 1 (c).

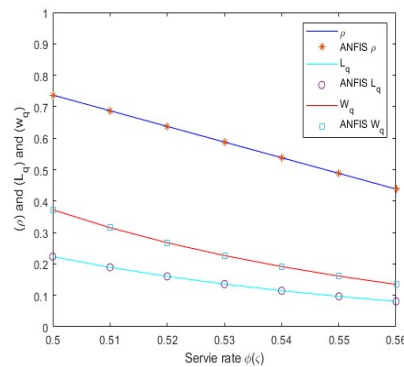
The three-dimensional graph of the system indicators of performance is shown in Figure 2 (a – c).

- The surface in figure 2 (a) shows the growth of the arrival rate ( $\Lambda$ ), estimated length of the line ( $L_q$ ), and estimated wait time ( $W_q$ ).
- Figure 2 (b) shows that as the service rate  $\phi(\zeta)$  rises, the estimated queue size ( $L_q$ ) and waiting time ( $W_q$ ) both decrease.
- Figure 2 (c) shows that as the breakdown rate  $\tau$  rises, expected queue lengths ( $L_q$ ) and waiting times ( $W_q$ ) also rise.

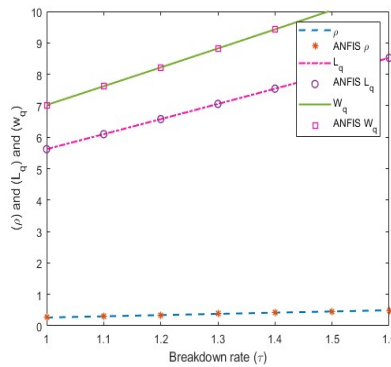
The numerical results above allow us to determine the influence of attributes on the system's evaluation criteria, and we can be assured that they are representative of realistic conditions.



(a)  $\rho, L_q, W_q$  versus arrival rate  $\Lambda$



(b)  $\rho, L_q, W_q$  versus Service rate  $\phi(\zeta)$



(c)  $\rho, L_q, W_q$  versus Breakdown rate  $\tau$

**Figure 1:** 2D representation effects

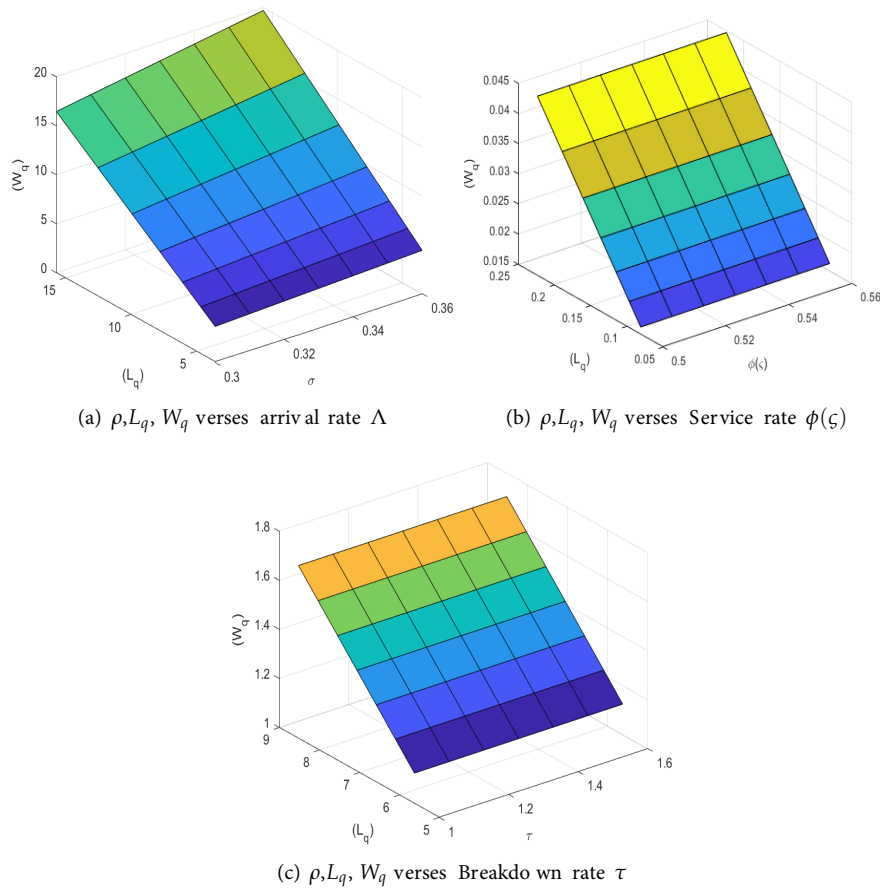


Figure 2: 3D representation effects

## 9. Adaptive Neuro-Fuzzy Inference System (ANFIS)

The ANFIS modal is actually applicable in a variety of fields such as modes of transport, congestion, telecommuting, atmospheric research, etc. Artificial neural networks are used in communications networks to accomplish a variety of goals, including an increase in customers, expense reduction, shorter wait times, etc. With variations in arrival rates while on vacation, service rates, repair rates, and repair to busy rates, the current modal allows us to examine the impatience of the client while they wait for the service.

A very helpful approach for ANFIS is created by combining soft computing methods, artificial neural networks (ANNs), and fuzzy systems (FS). We are showing a simplified idea of the ANFIS architecture by using the fuzzy parameters. We can implement an ANFIS input-output function and input-output data pairs as fuzzy if-then logic. The fuzzy toolbox of MATLAB software can be utilized for contrasting the computational finding with the implementation of an ANFIS network.

The input parameters and the membership function are assumed to be the  $\Lambda$ ,  $\phi(\zeta)$ , and  $\tau$  Gaussian functions in order to produce computational results based on ANFIS. (see Fig. 3a, b, c). It is assumed that the linguistic values are low, moderate, or high. Tick marks are placed over the curves made for the results obtained analytically in Figure 1a, 1b and 1c to indicate the results produced by the ANFIS approach for the queue size. The figures show that the numerical outcomes produced using the Runge-Kutta method and the ANFIS results are nearly identical.

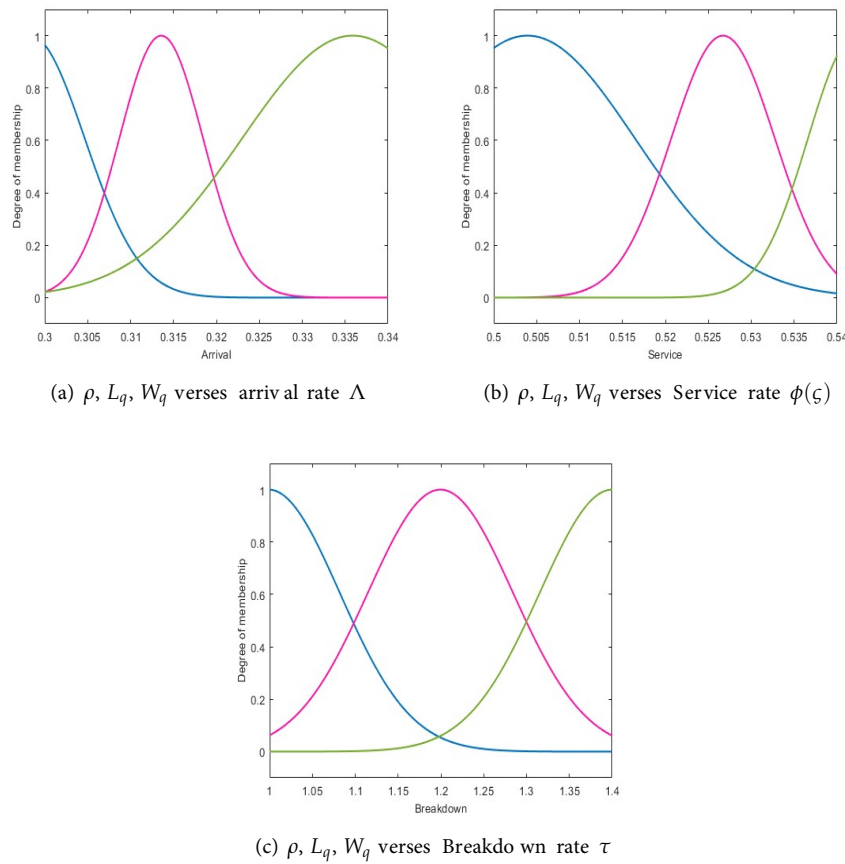


Figure 3: ANFIS representation effects

### 10. Cost Optimization:

The term "optimization" describes the method of determining the set of parameters for an objective function that produces the highest or lowest outcome. The continual, business-oriented activity known as "cost optimization" aims to reduce expenditures and costs while raising the organization's value. Standardizing, streamlining, and rationalizing platforms, application development, procedures, and services are all part of this process, along with establishing the most competitive possible terms and prices for all business transactions. The operating cost and profit of a system are closely tied in real-world situations. Therefore, the system's designers or managers place a lot of emphasis on reducing operational expenses per unit of time in order to enhance the system's earnings. Our objective is to identify the best cost per unit of time (TC) characteristics. In order to do this and increase the cost-effectiveness of our developed approach, we will build our competence in this field

$C_h$  - Holding expense for every user in the system per unit of time.

$C_b$  - The cost for each unit of time the server is turned on and used.

$C_v$  - The cost imposed on the server in vacation mode per unit of time.

$C_r$  - The cost to repair the server after its failure, calculated per unit of time.

$C_1$  - The cost per unit over a busy time.

$C_2$  - Cost for each unit of time used over the vacation period.

$$TC = C_h L_q + C_v V + C_b Q + C_r R + C_1 \gamma_b + C_2 \gamma_v$$



The TC problem is solved using metaheuristic optimisation methods including PSO, ABC, and GA. In view of the importance of cost optimisation, this study was conducted using the global search optimisation algorithms particle swarm optimisation (PSO), artificial bee colony (ABC), and genetic algorithms (GA), each of which is separately described in three different subsections of this section. If the algorithm's assumptions are correct, local search techniques frequently offer the level of computer efficiency required to find the global optimal. Tables 5 to 7 display the effects of  $\Lambda$ ,  $\tau$ , and  $\phi$  on  $TC^*$  using PSO, ABC, and GA.

**Table 4: Cost sets for optimal policy**

Cost sets	$C_h$	$C_v$	$C_b$	$C_r$	$C_1$	$C_2$
1	10	9	7	6	7	8
2	8	4	6	4	8	9
3	7	6	8	3	9	6

### 10.1. Particle Swarm Optimization (PSO)

One of the meta-heuristic methods used to solve optimization issues is the particle swarm optimization (PSO) technique, which has been employed successfully in a number of single objective optimization problems. Kennedy and Eberhart first proposed this algorithm. The PSO algorithm has the benefit of being simple to implement and apply for solving different function optimization problems, which can be categorized as function minimization or maximization problems.

**Table 5: Effect of  $\Lambda$ ,  $\tau$ ,  $\phi(\zeta)$  on  $TC^*$  using PSO**

$$g = 0.2, p = 0.7, G = 9, \theta = 0.95, d = 7, e = 8.6, \\ c = 0.2, B = 7, D = 2, \tau = 1.6, b = 4$$

Cost sets	$TC^*$			
	Cost set 1	Cost set 2	Cost set 3	
$\Lambda$	0.4	149.1752	133.0711	127.2781
	0.5	162.6882	143.5244	136.3697
	0.6	173.5857	152.1798	143.7058
$\tau$	1.6	149.1752	133.0711	127.2781
	1.7	161.5959	141.8755	134.7960
	1.8	175.6139	151.7985	143.2466
$\phi(\zeta)$	7	149.1752	133.0711	127.2781
	8	184.5033	159.0594	151.0219
	9	230.2302	192.6975	181.7546

### 10.2. Artificial Bee Colony(ABC)

One of Dervis Karaboga's most recent algorithms—created in 2005—is called the Artificial Bee Colony and was modeled after the cunning behaviour of honey bees. Basic process indicators like colonies and highest levels are essentially all that are used. Like PSO and differential evolutionary approaches, it is equally simple to comprehend. The search for huge areas of nectar-containing

food sources, and ultimately the one with the most nectar, is the bees' main goal. This population-based search approach is the main one used by ABC. The cost of the suggested structure is decreased through a process known as ABC.

**Table 6: Effect of  $\Lambda, \tau, \phi(\zeta)$  on  $TC^*$  using ABC**

$$g = 0.2, p = 0.7, G = 9, \theta = 0.95, d = 7, e = 8.6, \\ c = 0.2, \Lambda = 0.4, D = 2, \tau = 1.6, b = 4$$

Cost sets	$TC^*$			
	Cost set 1	Cost set 2	Cost set 3	
$\Lambda$	0.4	108.0030	108.8585	109.6965
	0.5	112.6458	113.7397	114.0666
	0.6	115.8805	117.2631	116.9601
$\tau$	1.6	108.0030	108.8585	109.6965
	1.7	112.4344	113.2737	114.1395
	1.8	117.4656	117.8425	118.7674
$\phi(\zeta)$	7	108.0030	108.8585	109.6965
	8	120.7245	120.9982	122.5394
	9	138.2173	134.5400	136.4184

### 10.3. Genetic Algorithm (GA)

The genetic algorithm, created in the 1960s and 1970s by Bremermann, Holland, and their colleagues, is a technique for addressing optimization problems brought on by natural selection, the mechanism that promotes evolution in biology. They are frequently employed to deliver superior solutions to stochastic search issues. The full procedure serves as a representation of the criteria for choice that were used to select the people who would make the best parents for the coming human generation.

**Table 7: Effect of  $\Lambda, \tau, \phi(\zeta)$  on  $TC^*$  using GA**

$$g = 0.2, p = 0.7, G = 9, \theta = 0.95, d = 7, e = 8.6, \\ c = 0.2, \Lambda = 0.4, D = 2, B = 7, b = 4$$

Cost sets	$TC^*$			
	Cost set 1	Cost set 2	Cost set 3	
$\Lambda$	0.4	152.5341	131.8364	124.6592
	0.5	163.2414	141.3111	131.8853
	0.6	170.0165	148.0781	136.4720
$\tau$	1.6	152.5341	131.8364	124.6592
	1.7	167.5959	142.6723	133.7200
	1.8	184.4744	154.7915	143.8279
$\phi(\zeta)$	7	152.5341	131.8364	124.6592
	8	192.8582	162.3320	151.7762
	9	243.6750	200.7627	185.9493

#### 10.4. Analogy of PSO, ABC and GA

This section compares the three approaches—particle swarm optimization (PSO), artificial bee colony (ABC), and genetic algorithm (GA)—to determine which has the least expense using the corresponding MATLAB programs. Then, one by one, the MATLAB programs for each of the aforementioned algorithms are run. We found that all three programs generated values that were nearly identical. Because of this, the three solutions are nearly comparable in terms of their optimum results and the fewest associated costs. It proves the reliability (local) and potency of these three simple techniques. Any technique can be used to calculate the optimal cost; however, PSO outperforms all others in comparison to our model. Because PSO has so many advantages, we have found that it is the best approach out of all of them. It performs well in global queries, requires a small number of arguments, is easy to configure, and is unaffected by design variable scalability. In addition to suffering sluggish convergence in a concentrated searching region, PSO has a tendency to lead to swift and early convergence in mid-optimal locations (being able to impair local search capabilities).

#### 10.5. Convergence in PSO, ABC and GA

After employing an optimization methodology like PSO, ABC, or GA, it is crucial to comprehend whether a particle recovers to normal or not and when it will roam around in search of a better solution. As a result, convergence is a significant component of cost evaluation. A statistical analysis (Fig. 4) of the outcomes demonstrates that ABC exceeds the PSO approach. For the whole standard optimization, ABC had fewer functional evaluations overall than PSO. The findings demonstrate that PSO converges more quickly. ABC cannot be employed if a speedy result is required for time-sensitive applications.

The study shows the applicability of our concept to real-world situations. Some of the analysts' financial issues will be partially overcome once they know how much the system will cost overall. The current situation may heavily rely on the cost-benefit assessment that was produced, which serves to illustrate the logic of our strategy and aid network administrators and specialists in lowering the issue of communications services that explicitly deal with blocking.

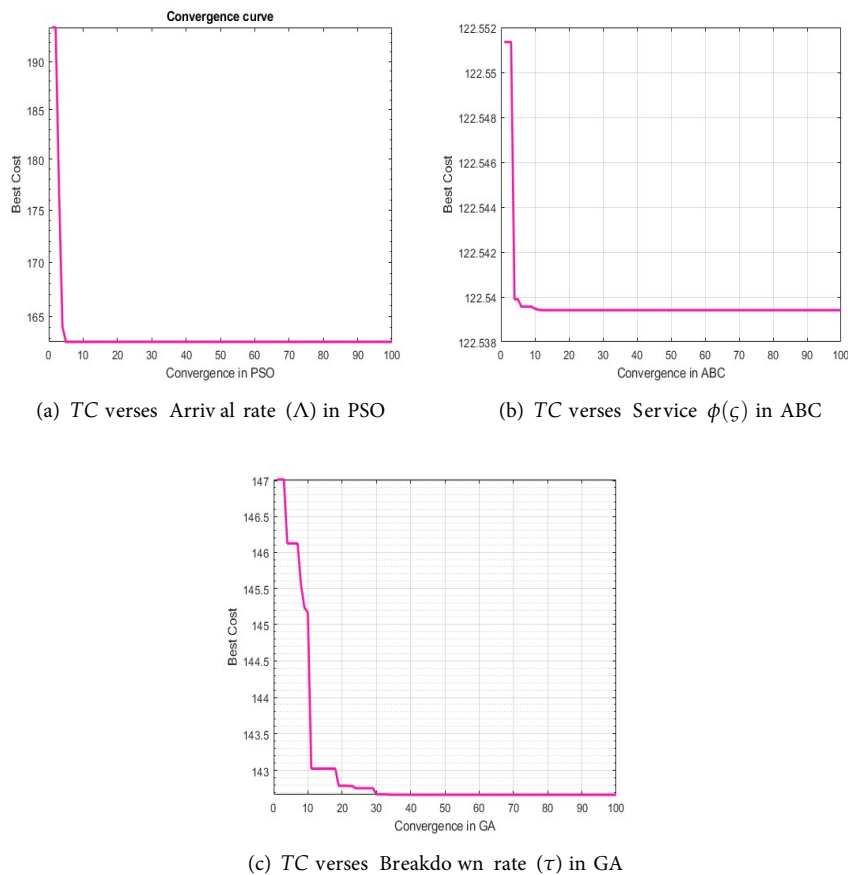
### 11. Conclusion:

This paper investigates the  $M^X/G(a,b)/1$  retrial queue with random failure and feedback under extended Bernoulli vacation with impatient customers. The SVT is utilized to determine indicators of efficiency for the various system stages. The efficiency of the system is then evaluated after considering the effects of various parameters. Finally, we gave a thorough explanation of the ANFIS. PSO, ABC, and GA are also used to compute the total cost. In an effort to find the best offer, these techniques compare and contrast the outcomes. The impetus for this study came from the prospective applications for the developed model, such as call centres, wireless networks, or telecommunication infrastructures, which might be powered by controlled precision test queueing systems to provide outstanding service at low prices. The simple mail transfer protocol utilizes a way to convey the messages between the mail servers. The recommended approach might be used in an email system's transfer model.

#### DECLARATIONS:

**Acknowledgments:** Not applicable

**Funding information:** This research did not receive any specific grant from funding agencies in



**Figure 4:** Cost Optimization effects

the public, commercial, or not-for-profit sectors.

**Conflicts of interest:** The authors declare no conflict of interest.

**Data availability:** Not applicable

**Authors contribution:** All the authors made substantial contributions to the conception or design of the work.

**Competing Interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## REFERENCES

- [1] Chaudhry, M. and Templeton, J. (1983). A first course in bulk queues. *John Wiley & Sons*. <http://dx.doi.org/10.1016/j.asej.2016.08.025>.
- [2] Bailey, N.T. (1954). On queueing processes with bulk service. *J R Stat Soc Ser A Stat Soc., Series B (Methodological)*. 80-87. <https://doi.org/10.1111/j.2517-6161.1954.tb00149.x>
- [3] Sumitha, D. and Udaya Chandrika, K. (2012). Retrial queueing system with starting failure, single vacation and orbital search. *Int.J. Comput. Appl.* 40.13, 29-33.
- [4] Radha, J., Rajadurai, P., Indhira, K. and Chandrasekaran, V.M. (2014). A batch arrival retrial queue with K-optional stages of service, Bernoulli feedback, single vacation and random breakdown. *Glob. J. Pure Appl. Math.* 10.2, 265-283.
- [5] Krishna Kumar, B., Pavai Madheswari, S. and Vijayakumar, A. (2002). The M/G/1 retrial queue with feedback and starting failures. *Appl. Math. Model.* 26.11, 1057-1075. [https://doi.org/10.1016/S0307-904X\(02\)00061-6](https://doi.org/10.1016/S0307-904X(02)00061-6)

- [6] Yang, Tao, Hui Li. (1994). The M/G/1 retrial queue with the server subject to starting failure. *Queueing Syst.* 16.1, 83-96. <https://doi.org/10.1007/BF01158950>
- [7] Mokaddis, G.S., Metwally, S.A. and Zaki, B.M. (2007). A feedback retrial queueing system with starting failures and single vacation. *J. Appl. Sci. Eng.* 10.3, 183-192. <https://doi.org/10.6180/jase.2007.10.3.01>
- [8] Bouchentouf, Amina Angelika, Abdelhak Guendouzi. (2021). Single server batch Arrival Bernoulli feedback queueing system with waiting server, K-variant vacations and impatient customers. *Oper. Res. Forum.* Vol.2. No.1. Springer International Publishing. <https://doi.org/10.1007/s43069-021-00057-0>.
- [9] Ayyappan, G. and Sathiy a, K. (2013). Transient analysis of batch arrival feedback retrial queue with starting failure and Bernoulli vacation. *Math. Model. Anal.* 3.8, 60-67.
- [10] Ayyappan, Govindhan, Udayageetha, J. (2020). Transient Analysis of M [X1], M [X2]/G1, G2/1 retrial queueing system with priority services, working breakdown, start up/close down time, Bernoulli vacation, reneging and balking. *Pak. J. Stat. Oper.* 203-216. <https://doi.org/10.18187/pjsor.v16i1.2181>.
- [11] Kulkarni, Vidyadhar, G., Bong Dae Choi. (1990). Retrial queues with server subject to breakdowns and repairs. *Queueing syst.* 7.2, 191-208. <https://doi.org/10.1007/BF01158474>
- [12] Ayyappan, G. and Shyamala, S. (2013). M[X]/G/1 with Bernoulli Schedule Server Vacation Random Break Down and second optional Repair. *J. Comp., & Model.* 3, 159-175.
- [13] Ke, Jau-Chuan, Tzu-Hsin Liu, Siping Su, Zhe-George Zhang. (2022). On retrial queue with customer balking and feedback subject to server breakdowns. *Commun. Stat.* 51.17, 6049-6063. <https://doi.org/10.1080/03610926.2020.1852432>
- [14] Takacs, L. (1963). A single-server queue with feedback. *Bell Syst. Tech. J.* 2, 505-519.
- [15] Rajadurai, P., Saravananarajan, M., Chandrasekaran, V.M. and Indhira, K. (2015). An M/(G1, G2)/1 Feedback Retrial Queue with Two Phase Service, Variant Vacation Policy Under Delaying Repair for Impatient Customer. *Int. J. Fuzzy Math. Arch* 6, 45-55.
- [16] 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. *J Ambient Intell Humaniz Comput.* 11.6, 2591-2599. <https://doi.org/10.1007/s12652-019-01318-x>
- [17] Nila, M. and Sumitha, D. Batch Arrival Retrial Queueing Model with Starting Failures, Customer Impatience, Multi Optional Second Phase and Orbital Search.
- [18] Ayyappan, G., Thamizhselvi, P., Somasundaram, B. and Udayageetha, J. (2021). Analysis of an retrial queueing system with priority services, working breakdown, Bernoulli vacation, admission control and balking. *Int. j. stat. manag. syst.* 24.4, 685-702. <https://doi.org/10.1080/09720529.2020.1744812>
- [19] Ayyappan, G. and Nirmala, M. (2021). Analysis of customer's impatience on bulk service queueing system with unreliable server, setup time and two types of multiple vacations. *Int. J. Ind. Syst.* 38.2, 198-222. <https://doi.org/10.1504/IJISE.2021.115321>
- [20] Sethi, R., Jain, M., Meena, R. K. and Garg, D. (2020). Cost optimization and ANFIS computing of an unreliable M/M/1 queueing system with customers' impatience under n-policy. *Int. J. Appl. Comput. Math.* 6, 1-14. <https://doi.org/10.1007/s40819-020-0802-0>
- [21] Miss Vaishnavi, Shweta Upadhyaya and Rakhee Kulshrestha (2022). Optimal Cost Analysis for Discrete-Time Recurrent Queue with Bernoulli Feedback and Emergency Vacation. *Int. J. Appl. Comput. Math.* 8.5, 254. <https://doi.org/10.1007/s40819-022-01445-8>

# REPORTING METHODOLOGY AND ALGORITHM OF MODES OF COMPLEX ENERGY SYSTEMS WITH PHASE COORDINATES

Huseyngulu Guliyev<sup>1</sup> and Famil Ibrahimov<sup>2</sup>

•

<sup>1,2</sup>Azerbaijan Technical University, Baku, Azerbaijan  
AZ1073, H. Javid avenue 25

<sup>1</sup>huseyngulu@mail.ru, <sup>2</sup>amfanet@mail.ru

## Abstract

*A mathematical model, algorithm and program have been developed to study any types of complex asymmetric steady-state modes and transient processes of a multi-machine power system with a renewable energy source in phase coordinates, the results of which can be used in the operational control of power system operating modes with any type of emergency automation. The developed methodology and software package can also be used in industry to check the possibility of long-term operation in the considered asymmetrical mode from the point of view of the operating conditions of the system generators and electrical receivers, to determine the need to use baluns, to select their parameters and installation locations, to ensure the efficiency of asymmetrical modes, as well as for conducting various tests and analyzing accidents that have occurred.*

**Keywords:** power system, asymmetrical steady-state modes, transient processes, emergency automation, relay protection

## I. Introduction

At present, methods and algorithms for calculating steady-state and transition modes of complex power systems are used in the replacement scheme of a symmetrical three-phase system. When modeling switching processes in non-symmetrical short-circuit and incomplete phase modes, instead of asymmetry, it is performed by adding a shunt or additional resistance. In modern conditions, where the integration of renewable energy sources and digital technologies into the energy system takes place, solving the mentioned problem with traditional methods and algorithms becomes significantly more complicated, and sometimes in complex asymmetric modes, when the transposition of electric transmission lines is not considered, when the parameters of the line and other elements of the system differ in phase, in substations when three-phase transformers are connected with special schemes, it is quite difficult and sometimes impossible to solve. At the same time, carrying out non-symmetric settled and transition modes in a fictitious two-axis coordinate system using transformation formulas significantly increases modeling and reporting errors and makes adequate decision-making difficult for mode control. Taking into account the above, the issue of expressing mode parameters in phase coordinates ( $a, b, c$ ) during the calculation and study of steady and transition modes appears as an actual

solution [1-4].

It should be noted that the advantage of the  $(a,b,c)$  coordinate system over other calculation systems, especially the  $d,q,0$  system, is that all mode quantities correspond to real-time values, and re-transformation and calculation of the results are not required to obtain the phase quantities.

The solution of the given problem in phase coordinates is quite universal, as the modeling of various types of non-symmetrical short-circuits and settled symmetric modes is relatively easy, and the simplicity of the reporting algorithm allows the use of modern high-performance computing systems. In this case, the main difficulty is to design a three-phase replacement scheme for the elements of the power system in non-symmetrical quasi-steady and transition modes. Therefore, the replacement schemes and mathematical models of the main elements of the power system (generator, transformer, power transmission line, load) for the calculation of non-symmetric modes were presented in [5-11], respectively.

In order to solve the problem posed in the conditions of integration of renewable energy sources, first, the calculation of the non-symmetrical settled mode is carried out in the phase coordinates, and here the pre-accident mode will differ significantly from the linear scheme of the mode due to the reasons we mentioned above. In the next stage, generators in a multi-machine complex energy system are combined with the algorithm of reporting transition processes in phase coordinates, taking into account the complete Park-Gorev equations [12-15].

In the work under review, the equations for the stator windings of conventional and wind generator machines are used in the  $(a,b,c)$  coordinate system, and for the rotor quantities, the  $d,q,0$  system is used. The periodic coefficients in the equations of synchronous and asynchronous machines are calculated as the angle between the stator and rotor axes in each interval of the mathematical solution of the equations. It should be noted that since their periodic coefficients are expressed as  $\sin \gamma$  and  $\cos \gamma$ , their calculation does not cause any difficulties.

In order to express each three-phase element in phase coordinates, their description with a suitable three-phase replacement scheme was used in the calculation methodology. In this case, the operation of the complex transformation coefficient in different branching cases is taken into account for voltage regulation in power transformers and autotransformers. For the purpose of reporting, a matrix of nodal equations is established for separate elements of the energy system, and based on it, the results of the report are used not only for tuning relay protection and automation devices but also for more complex issues, in other words, mode symmetrization.

## II. Solving the system of nodal equations in phase coordinates for the study of symmetric and non-symmetric regimes of the complex energy system

The steady-state mathematical model of a three-phase network is analogous to the model of a single-line network and is a system of nonlinear mathematical equations with complex coefficients and variables. All known methods can be used to solve it [16-18].

The well-known Gauss-Seidel method was used to solve the system of nodal voltage equations in the form of a current balance and the system of nodal equations written in the form of a matrix.

$$| \dot{Y} || \dot{U} | = | \dot{I} |$$

Data for nodes is given as  $P_L + jQ_L$  load power for each phase, and for generators as  $P_G + jQ_G$  corresponding to each phase or  $P_G, | \dot{U}_G |$ . As in a single-line circuit, a node is taken as a balancing node, for which the emf's are assumed to be  $120^\circ$  from each other in the three phases. The voltage

at the balancing node is determined according to the following expression:

$$\begin{pmatrix} \dot{Y}_{11} & \dot{Y}_{12} & \dot{Y}_{13} & -\dot{Y}_0 \\ \dot{Y}_{21} & \dot{Y}_{22} & \dot{Y}_{23} & -\dot{Y}_0 \\ \dot{Y}_{31} & \dot{Y}_{32} & \dot{Y}_{33} & -\dot{Y}_0 \\ -\dot{Y}_0 & -\dot{Y}_0 & -\dot{Y}_0 & 3\dot{Y}_0 + \dot{Y}_{N0} \end{pmatrix} \begin{pmatrix} \dot{U}_1 \\ \dot{U}_2 \\ \dot{U}_3 \\ 0 \end{pmatrix} = \begin{pmatrix} \dot{S}_1/\dot{U}_1 + \dot{Y}_1 E_a^{\delta a1} \\ \dot{S}_2/\dot{U}_2 + a^2 \dot{Y}_1 E_a^{\delta a1} \\ \dot{S}_3/\dot{U}_3 + a^2 \dot{Y}_1 E_a^{\delta a1} \\ 0 \end{pmatrix} \quad (1)$$

here  $a = 1 \angle 120^\circ = e^{j2\pi/3}$ ;  $\dot{Y}_1, \dot{Y}_2, \dot{Y}_0$  – forward, reverse and zero sequence conductors;  $U_1, U_2, U_3$  – forward, reverse and zero sequence voltages;  $S_1, S_2, S_3$  – the full powers of individual phases;  $E_a^{\delta a1}$  – is the EMF of the balancing node.

In this case, the following restrictions are taken into account according to voltage and power:

$$\begin{aligned} P_{i\min} &\leq P_{Gi} \leq P_{i\max}, \\ Q_{i\min} &\leq Q_{Gi} \leq Q_{i\max}, \\ \dot{U}_{i\min} &\leq |\dot{U}_i| \leq \dot{U}_{i\max}, \\ &i = 1, 2, 3, \dots, n \end{aligned}$$

The allowable limits characterize the change of active power  $P_i$ , reactive power  $Q_i$  and voltage modulus  $|\dot{U}_i|$  at node  $i$ . The algorithm uses the procedure of accelerated accumulation of the iteration process, where the new acceleration coefficient  $\omega_{new}$  is recalculated depending on the given number of iterations of the old coefficient  $\omega_{start}$ .

$$\omega^{new} = \frac{2}{1 + \sqrt{1 - \frac{(\omega^{start} + \lambda - 1)^2}{\omega^{start} - \lambda}}} \quad (3)$$

here

$$\lambda = \sqrt{\frac{\sum_{i=1}^n |\Delta \dot{U}_1^{(k+1)}|^2}{\sum_{i=1}^n |\Delta \dot{U}_1^{(k)}|^2}} \quad (4)$$

The value of the acceleration coefficient is taken in the range  $1 \leq \omega \leq 2$ . The iteration process ends after the given precision is met.

Considering the given expression, an algorithm and software were developed for the calculation of asymmetric modes in multi-machine complex systems with renewable energy sources, according to which the equations of the energy system elements are expressed in phase coordinates.

### III. Determination of currents and voltages in phase coordinates in asymmetric regimes of power systems

Calculation of short-circuit currents and single-phase modes can be performed based on the results of calculating the previous short-circuit pre-emergency mode. In this case, you can simultaneously simulate any type of short circuit, including short circuit through impedance. Short circuits and phase breaks are taken into account directly when drawing up nodal equations. At the



same time, all the necessary information is entered into the computer, taking into account the capacitive conductivities of the corresponding component lines, the resistance of the component transformers (autotransformers); resistances included in the neutral of transformers, generators, etc. The phase discontinuity of a branch can also be replaced by including an infinitely large resistance in it [19,20].

In the conductivity matrix  $| \dot{Y} |$ , the short circuit is quite simply taken into account through transition resistance. To do this, it is enough to set the value of the contact resistance in the source data. When modeling a fault in the nodes of the system circuit, it is necessary to set a pre-provided code for the fault type. To calculate short-circuit currents and open-phase modes, the Gauss-Seidel method was used.

#### IV. Modeling in phase coordinates of transient processes in complex regulated power systems.

The equations of synchronous machines are modeled using the full Park–Gorev equations and simplified Lebedev–Zhdanov equations, taking into account electromagnetic transient processes in the rotor circuits [21-26]

The initial equations for calculating the modes of a synchronous machine in coordinates  $a, b, c$  are the following differential equations for the stator winding voltages:

$$\left. \begin{aligned} p\psi_a &= e_a - i_a r_a \\ p\psi_b &= e_b - i_b r_b \\ p\psi_c &= e_c - i_c r_c \end{aligned} \right\}, \quad (5)$$

where  $\psi_a, \psi_b, \psi_c$  – flux linkage of the stator winding phases;  $i_a, i_b, i_c$  – stator winding phase currents;  $r_a, r_b, r_c$  – active resistance of stator winding phases;  $e_a, e_b, e_c$  – voltage at the terminals of the generator stator phase windings;  $p = \frac{d}{d\tau}$  – differentiation operator with respect to synchronous time  $\tau = 2\pi ft$ .

To this system of equations one should add the stress equations for the rotor circuits and the rotor motion equations:

$$\left. \begin{aligned} p\psi_f &= e_f - i_f r_f \\ p\psi_{kd} &= -i_{kd} r_{kd} \\ p\psi_{kq} &= -i_{kq} r_{kq} \\ pS &= \frac{1}{H} (M_m + M_e) \\ p\theta &= S \end{aligned} \right\}, \quad (6)$$

where  $\left. \begin{aligned} \psi_f, \psi_{kd}, \psi_{kq} \\ i_f, i_{kd}, i_{kq} \\ r_f, r_{kd}, r_{kq} \end{aligned} \right\}$  – flux linkage of current and active resistance of the excitation winding and

dampers along the longitudinal and transverse axes;  $e_f$  – voltage applied to the excitation winding;  $S$  – slip;  $H$  – inertial constant in el. rad;  $M_m$  – load torque on the shaft of a synchronous machine;  $M_e$  – electromagnetic torque of synchronous machine;  $\theta$  – working angle (angle between

the transverse axis of the rotor and the representing vector of phase voltages.

To solve systems of equations (4) and (5) on a PC using any of the well-known numerical methods of Runge–Kutta, Adams Euler, etc. [3] it is necessary that the number of variables equals the number of equations. Experience shows that it is advisable to express all currents through the flux linkage of the circuits. For this purpose, well-known relationships obtained from calculations of symmetric modes using the Park–Gorev equations are used [3].

$$\left. \begin{aligned} i_d &= a\psi_d - b\psi_f - c\psi_{kd} \\ i_q &= g\psi_q - h\psi_{kq} \\ i_f &= -b\psi_d + d\psi_f - e\psi_{kd} \\ i_{kd} &= -c\psi_d - e\psi_f + f\psi_{kd} \\ i_{kq} &= -h\psi_q + k\psi_{kq} \\ i_0 &= \frac{\psi_0}{x_0} \end{aligned} \right\}, \quad (7)$$

where the coefficients  $a, b, c, d, e, f, g, h, k$  are expressed through the machine parameters as follows:

$$\left. \begin{aligned} a &= \frac{X_f X_{kd} - x_{ad}^2}{\Delta}; f = \frac{X_d X_f - x_{ad}^2}{\Delta}; \\ b &= \frac{x_{ad} X_{kd} - x_{ad}^2}{\Delta}; g = \frac{X_{kq}}{X_q X_{kq} - x_{aq}^2}; \\ c &= \frac{x_{ad} X_f - x_{ad}^2}{\Delta}; h = \frac{x_{aq}}{X_q X_{kq} - x_{aq}^2}; \\ d &= \frac{X_d X_{kd} - x_{ad}^2}{\Delta}; k = \frac{X_q}{X_q X_{kq} - x_{aq}^2}; \\ e &= \frac{X_d x_{ad} - x_{ad}^2}{\Delta}; \end{aligned} \right\} \quad (8)$$

$$\Delta = X_d (X_f X_{kd} - x_{ad}^2) - x_{ad} (x_{ad} X_{kd} - x_{ad}^2) - x_{ad} (x_{ad} X_f - x_{ad}^2).$$

The parameters included in these expressions represent the mutual or complete reactivity of the circuits.

$$\begin{aligned} X_f &= x_{ad} + x_f; & X_{kd} &= x_{ad} + x_{kd}; & X_{kq} &= x_{aq} + x_{kq}; \\ X_d &= x_{ad} + x_e; & X_q &= x_{aq} + x_l; \end{aligned}$$

To transition from stator currents  $i_d, i_q, i_0$  to phase values  $i_a, i_b, i_c$  we use the known relations.

$$\left. \begin{aligned} i_a &= i_0 + i_d \cos \gamma - i_q \sin \gamma; \\ i_b &= i_0 + i_d \cos(\gamma - \rho) - i_q \sin(\gamma - \rho); \\ i_c &= i_0 + i_d \cos(\gamma + \rho) - i_q \sin(\gamma + \rho); \end{aligned} \right\} \quad (9)$$

$$\left. \begin{aligned} \psi_0 &= \frac{1}{3}(\psi_a + \psi_b + \psi_c); \\ \psi_d &= \frac{2}{3}[\psi_a \cos \gamma + \psi_b \cos(\gamma - \rho) + \psi_c \cos(\gamma + \rho)]; \\ \psi_q &= \frac{2}{3}[\psi_a \sin \gamma + \psi_b \sin(\gamma - \rho) + \psi_c \sin(\gamma + \rho)]; \end{aligned} \right\} \quad (10)$$

where  $\rho = \frac{2\pi}{3} = 120^\circ$  for a machine with symmetrically arranged three-phase windings,  $\gamma = \tau + \theta + \frac{\pi}{2}$  – the angle between the stationary axis of phase a and the rotating longitudinal axis of the rotor.

## V. Modeling in phase coordinates of transient processes in complex regulated power systems.

The block diagram of the algorithm for calculating asymmetric modes and transient processes in phase coordinates is shown in Figure 1.

As a result of calculating symmetrical or asymmetrical modes for each phase, the following are determined: modules and voltage angles in nodes, flows of active and reactive power along lines and transformers, losses in each element and in the system as a whole, generation of reactive power in those nodes where voltage modules are specified and other information if necessary.

When calculating short circuit (SC) modes, the output information is also displayed on the display screen in tabular form and includes: currents at the short circuit point, residual voltages in the circuit nodes and their phases, currents or power flows along the branches, including losses in each element and in the system generally.

In this work, it is possible to perform calculations during a short circuit at an intermediate point of a branch without introducing additional nodes into the design diagram.

To illustrate the performance of the developed methodology and program for calculating symmetrical and complex-asymmetrical modes in complex multi-machine power systems, let us consider several examples for a specific circuit shown in Figure 2. As can be seen, a 10 MW wind turbine is integrated into the system through a T3 transformer.

All necessary data for the system under study are presented in Tables 1 ÷ 3. The values are presented in p. u. and reduced to  $S_b = 100MVA$ . Transformers T1-T3 have a connection diagram Y0/Δ. Transformers T4 and T5 have three-phase-two-phase and star-zigzag connection schemes, respectively.

Tables 4 and 5 present the results of calculating symmetrical and asymmetrical modes in phase coordinates, where the load was represented by  $P_L + jQ_L = \text{const}$ . Note that the voltage values in the secondary windings of a three-phase-two-phase transformer are distributed as follows:

$$\begin{aligned} \dot{U}_{a2-a1} &= (\dot{U}_{41} - \dot{U}_{43}), \quad \dot{U}_{\phi2-\phi1} = (\dot{U}_{42} - \dot{U}_{44}), \quad \text{where} \\ \dot{U}_{a2-a1} &= 0,855 \angle 83,15^\circ - 0,197 \angle -76,996^\circ = 1,0424 \angle 86,83^\circ; \\ \dot{U}_{\phi2-\phi1} &= 0,836 \angle -5,011^\circ - 0,209 \angle -175,166^\circ = 1,043 \angle -3,05^\circ. \end{aligned}$$

A comparison of the calculation results presented in Tables 4 and 5 indicates that the levels of voltage values in the nodes are different due to the different representations of loads in the nodes of the circuit.

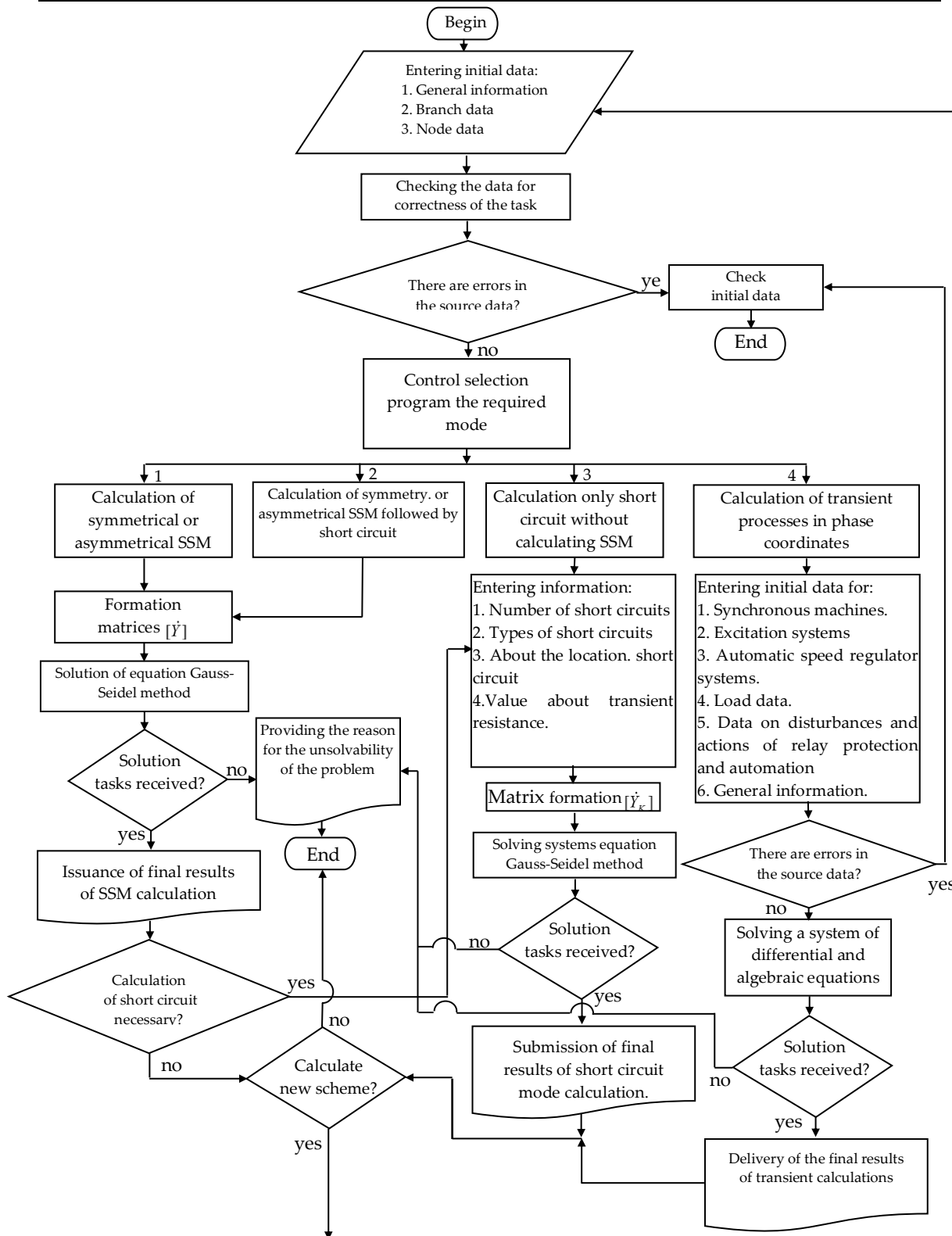


Figure 1: Block diagram of the algorithm for calculating power system modes in phase coordinates

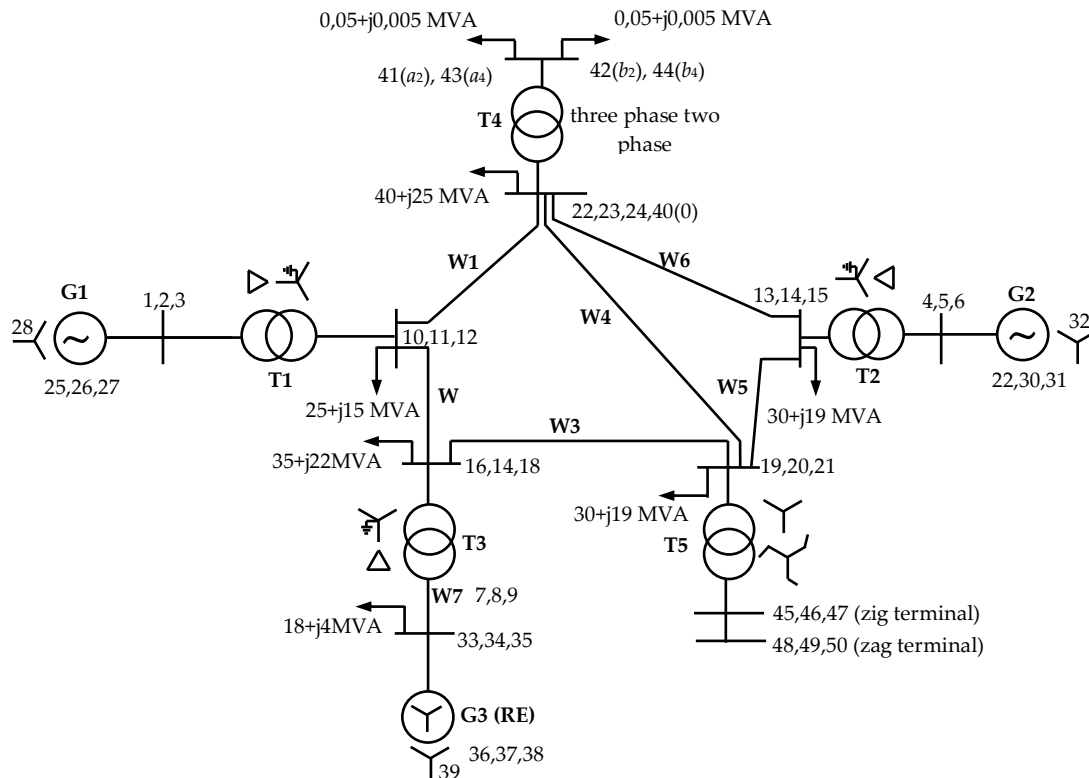


Figure 2: Scheme of the studied multi-machine power system with renewable energy sources

Table 1: Initial data of power transmission lines of power systems (p.u.)

art node	End node	Positive sequence resistors			Zero sequence resistances		
		$R_1$	$X_1$	$B_1$	$R_0$	$X_0$	$B_0$
10, 11, 12	22, 23, 24	0,0145	0,0660	0,0108	0,0456	0,1944	0,0058
10, 11, 12	16, 17, 18	0,0110	0,0496	0,0080	0,0342	0,1458	0,0050
16, 17, 18	19, 20, 21	0,0091	0,0413	0,0068	0,0285	0,1215	0,0043
19, 20, 21	22, 23, 24	0,0056	0,0248	0,0041	0,0171	0,0729	0,0025
19, 20, 21	13, 14, 15	0,0035	0,0155	0,0021	0,0114	0,0486	0,0027
13, 14, 15	22, 23, 24	0,0013	0,0330	0,0054	0,0228	0,0972	0,0034
7, 8, 9	33, 34, 35	0,0073	0,0330	0,0054	0,0228	0,0972	0,0034

Table 2: Initial data of system generators (p.u.)

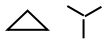


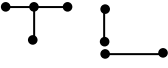
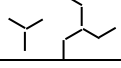
Start node	End node	Positive sequence resistors		Zero sequence resistances	
		$R_1$	$X_1$	$R_0$	$X_0$
1, 2, 3	25, 26, 27	0,0	0,0967	0,0	0,0467
4, 5, 6	29, 30, 31	0,0	0,17	0,0	0,085
33, 34, 35	36, 37, 38	0,0	0,17	0,0	0,085

The obtained results of calculating the asymmetric mode indicate that the noted violations of the symmetric mode do not cause deep violations of the level of asymmetry of the mode parameters in the network circuit and such a mode is acceptable.

The calculation results for a single-phase short circuit at the generator terminals are presented in Figure 3, where a complete coincidence with the experimentally taken curves of the transition

process was obtained.

**Table 3:** Initial data of power transformers (p.u.)

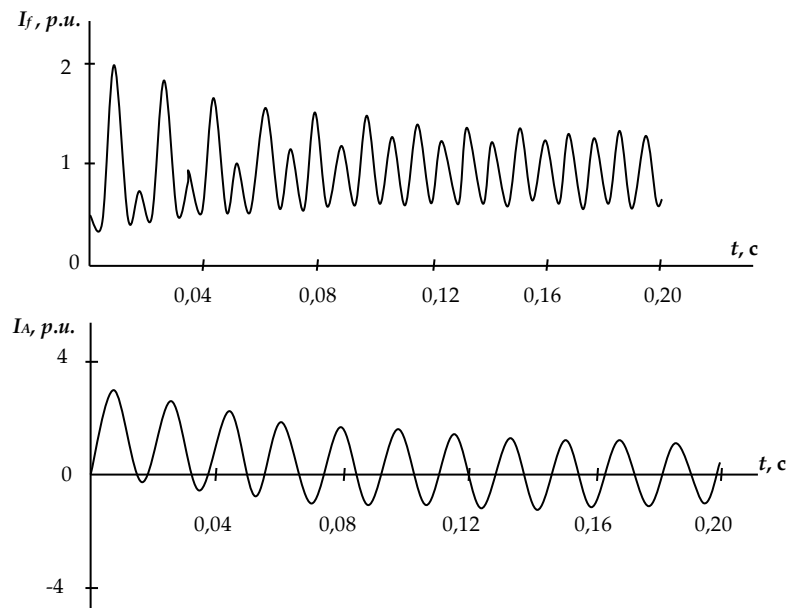
Start node	End node	Connection type	X
1, 2, 3	10, 11, 12		0,0533
4, 5, 6	13, 14, 15		0,12
7, 8, 9	16, 17, 18		0,16
22, 23, 24	41, 42, 43, 44		0,09( $X_m$ ) 0,16( $X_r$ )
19, 20, 21	45, 46, 47 48, 49, 50		0,17

**Table 4:** Calculation results of the symmetrical mode in phase coordinates

Start node	End node	P (MW)	Q (MVar)	Start node	End node	P (MW)	Q (MVar)
1	25	-32,373	-15,122	25	1	32,373	16,230
2	26	-32,306	-14,879	26	2	32,306	15,976
3	27	-32,497	-14,971	27	3	32,497	16,081
4	29	-19,985	-9,405	29	4	19,985	10,151
5	30	-20,105	-9,534	30	5	20,105	10,292
6	31	-19,912	-9,582	31	6	19,912	10,329
10	16	12,079	3,889	16	10	-12,063	-4,690
11	17	12,106	3,965	17	11	-12,090	-4,768
12	18	11,943	4,030	18	12	-11,926	-4,833
10	22	12,317	5,404	22	10	-12,292	-6,470
11	23	12,230	5,382	23	11	-12,207	-6,452
12	24	12,103	5,536	24	12	-12,077	-6,610
13	19	6,511	1,803	19	13	-6,509	-2,090
14	20	6,437	1,662	20	14	-6,436	-1,950
15	21	6,466	1,749	21	15	-6,465	-2,037
13	22	4,002	0,986	22	13	-4,001	-1,568
14	23	3,912	0,880	23	14	-3,911	-1,463
15	24	3,974	0,972	24	15	-3,973	-1,555
16	19	4,617	3,748	19	16	-4,614	-4,470
17	20	4,489	3,649	20	17	-4,487	-4,373
18	21	4,444	3,785	21	18	-4,440	-4,508
19	22	0,996	0,186	22	19	-0,996	-0,626
20	23	0,924	0,139	23	20	-0,924	-0,579
21	24	0,988	0,203	24	21	-0,988	-0,643
7	33	-3,966	-6,288	33	7	3,969	5,701
8	34	-4,024	-6,332	34	8	4,028	5,745
9	35	-3,837	-6,442	35	9	3,841	5,856
33	36	-10,047	-8,021	36	33	10,047	8,272
34	37	-10,074	-8,038	37	34	10,075	8,292
35	38	-9,881	-8,178	38	35	9,880	8,428

**Table 5:** Calculation results for the asymmetric mode in phase coordinates

Start node	End node	P (MW)	Q (MVAR)	Start node	End node	P (MW)	Q (MVAR)
1	25	-38,839	-21,470	25	1	38,839	23,199
2	26	-32,511	-20,268	26	2	32,511	21,553
3	27	-37,071	-15,517	27	3	37,072	16,920
4	29	-24,543	-10,624	29	4	24,543	11,722
5	30	-16,229	-12,413	30	5	16,229	13,057
6	31	-19,228	-4,594	31	6	19,228	5,183
10	16	12,622	4,220	16	10	-12,605	-5,008
11	17	15,593	5,171	17	11	-15,566	-5,913
12	18	13,322	7,676	18	12	-13,295	-5,419
10	22	11,936	5,932	22	10	-11,913	-6,993
11	23	14,615	6,308	23	11	-14,580	-7,324
12	24	12,881	8,829	24	12	-12,844	-9,824
13	19	8,625	1,825	19	13	-8,523	-2,105
14	20	4,315	2,852	20	14	-4,314	-3,139
15	21	5,362	-1,541	21	15	-5,361	1,260
13	22	5,114	1,095	22	13	-5,111	-1,671
14	23	1,412	1,713	23	14	-1,411	-2,293
15	24	2,381	-2,027	24	15	-2,381	1,458
16	19	3,321	4,121	19	16	-3,319	-4,942
17	20	5,276	3,489	20	17	-5,272	-4,203
18	21	4,866	5,752	21	18	-4,861	-6,446
19	22	1,068	0,315	22	19	-1,067	-0,753
20	23	-0,993	0,455	23	20	0,994	-0,892
21	24	-0,400	-1,603	24	21	0,400	1,175
7	33	-3,408	-7,157	33	7	3,412	6,575
8	34	-2,302	-5,661	34	8	2,305	5,068
9	35	-4,268	-5,472	35	9	4,270	4,882
33	36	-10,082	-9,145	36	33	10,083	9,429
34	37	-8,970	-7,701	37	34	8,970	7,914
35	38	-10,947	-7,463	38	35	10,947	7,732



**Figure 3:** Current in the field winding (a) and in phase A (b) with a single-phase short circuit

## VI. Conclusions

1. A mathematical model of a power system with a renewable energy source, an algorithm and a program have been developed for studying any types of longitudinal-transverse, complex asymmetric steady-state modes, short-circuit modes and transient processes in phase coordinates and a dialogue complex created on its basis, which can be used in operational dispatch control operating modes of power systems for any type of emergency automation.

2. The developed methodology and complex program can be used in industry to solve the following practical problems: to test the possibility of long-term operation in the considered asymmetrical mode from the point of view of the operating conditions of EPS generators and power receivers; to determine the need to use baluns, to select their parameters and installation locations, to ensure the efficiency of asymmetrical modes; for carrying out various tests, analyzing accidents that have occurred; for selecting response parameters and assessing the sensitivity of relay protection devices, parameters of automation devices.

3. The obtained comparative results of calculating the symmetrical and asymmetrical modes of power systems indicate that the levels of voltage values in the nodes are different due to the different representations of loads in the nodes of the circuit. The noted violations of the symmetrical mode do not cause deep violations of the level of asymmetry of the mode parameters in the network circuit and such a mode is acceptable.

## References

- [1] Merkuryev G. V., Shargin Yu. M. Formation of a mathematical model of the energy system for calculations of electromechanical transient processes // *Electricity*, 2008, No. 12, pp. 3-7
- [2] Vainshtein R. A., Lozinsky K. S., Ivanov V. P., Kobytsev M. I. Improvement of calculations of asymmetric modes in programs for calculating electromechanical transient processes // *Electricity*, 2008, No. 7, pp. 19-23



[3] Guseinov A. M. Assessment of the degree of influence of various factors on the synchronous dynamic stability of the electric power system // Problems of Energy, part 1 No. 3, 2007, pp. 13-25.

[4] Hashimov A. M., Rakhmanov N. R., Guliyev H. B. Assessment of the risk of disruption of the stability of the energy system at various shares of integration of renewable sources with variable power generation. Methodological issues in studying the reliability of large energy systems: Vol. 73. Reliability of energy systems in the context of the energy transition. Rep. ed. Academician of the Russian Academy of Sciences V. A. Stennikov, Alushta, Irkutsk: ISEM SB RAS. 2022. 667 p., pp. 592-601

[5] Guliyev H. B. Method for determining the critical parameters of the sustainability of an energy system with integrated renewable energy sources. Rudenko International Conference "Methodological Problems in Reliability Study of Large Energy Systems" (RSES 2022), E3S Web of Conf., Volume 384, 2023, pp.1-5

[6] Guliyev H. B., Ibrahimov F. Sh. Estimation of critical parameters of the states of the power system with renewable energy sources at random shutdowns of its main elements. Electroenergetics, Electrotechnics, Electromechanics + Control (EEEC), Scientific – industrial journal, Vol.12, No.2, 2022, pp.26-33.

[7] Guliyev H. B. Neuro-fuzzy algorithm for controlling voltage and reactive power in network with distributed generation based on renewable energy sources. COIA 2022, Proceeding of the 8th International Conference on Control and Optimization with Industrial Applications, Volume 1, 24-26 August, 2022, Baku, Azerbaijan, p.357-359

[8] Bian, J.; Wang, H.; Wang, L.; Li, G.; Wang, Z. Probabilistic optimal power flow of an AC/DC system with a multiport current flow controller. *CSEE J. Power Energy Syst.* 2021, 7, 744–752.

[9] Bernal-Romero, D. L.; Montoya, O. D.; Arias-Londoño, A. Solution of the Optimal Reactive Power Flow Problem Using a Discrete-Continuous CBGA Implemented in the DigSILENT Programming Language. *Computers* 2021, 10, 151.

[10] Zhao, J.; Gómez-Expósito, A.; Netto, M.; Mili, L.; Abur, A.; Terzija, V.; Kamwa, I.; Pal, B.; Singh, A.K.; Qi, J.; et al. Power System Dynamic State Estimation: Motivations, Definitions, Methodologies, and Future Work. *IEEE Trans. Power Syst.* 2019, 34, 3188–3198.

[11] Chen, Y.; Chen, H.; Jiao, Y.; Ma, J.; Lin, Y. Data-driven Robust State Estimation Through Off-line Learning and On-line Matching. *J. Mod. Power Syst. Clean Energy* 2021, 9, 897–909.

[12] Zhao, J.; Zhang, G.; Dong, Z. Y.; La Scala, M. Robust Forecasting Aided Power System State Estimation Considering State Correlations. *IEEE Trans. Smart Grid* 2018, 9, 2658–2666.

[13] Caro, E.; Conejo, A.J. State estimation via mathematical programming: A comparison of different estimation algorithms. *IET Gener. Transm. Distrib.* 2012, 6, 545–553.

[14] Zamora-Cárdenas, E. A.; Pizano-Martínez, A.; Lozano-García, J. M.; Gutiérrez-Martínez, V.J.; Cisneros-Magaña, R. Computational development of a practical educational tool for state estimation of power systems using the MATLAB optimization toolbox. *Int. J. Electr. Eng. Educ.* 2019, 56, 105–123.

[15] Madani, R.; Sojoudi, S.; Lavaei, J. Convex Relaxation for Optimal Power Flow Problem: Mesh Networks. *IEEE Trans. Power Syst.* 2015, 30, 199–211.

[16] Chen, J.; Abur, A. Placement of PMUs to Enable Bad Data Detection in State Estimation. *IEEE Trans. Power Syst.* 2006, 21, 1608–1615.

[17] Zhao, J.; Netto, M.; Huang, Z.; Yu, S.S.; Gómez-Expósito, A.; Wang, S.; Kamwa, I.; Akhlaghi, S.; Mili, L.; Terzija, V.; et al. Roles of Dynamic State Estimation in Power System Modeling, Monitoring and Operation. *IEEE Trans. Power Syst.* 2021, 36, 2462–2472.

- [18] Chen, L.; Li, Y.; Huang, M.; Hui, X.; Gu, S. Robust Dynamic State Estimator of Integrated Energy Systems Based on Natural Gas Partial Differential Equations. *IEEE Trans. Ind. Appl.* 2022, 58, 3303–3312.
- [19] Zhao, H.; Tian, B. Robust Power System Forecasting-Aided State Estimation With Generalized Maximum Mixture Correntropy Unscented Kalman Filter. *IEEE Trans. Instrum. Meas.* 2022, 71, 1–10.
- [20] Sheng, T.; Yin, G.; Wang, B.; Guo, Q.; Dong, J.; Sun, H.; Pan, Z. State estimation approach for combined heat and electric networks. *CSEE J. Power Energy Syst.* 2022, 8, 225–237
- [21] Liu, C.; Deng, R.; He, W.; Liang, H.; Du, W. Optimal Coding Schemes for Detecting False Data Injection Attacks in Power System State Estimation. *IEEE Trans. Smart Grid* 2022, 13, 738–749
- [22] Basu, S.; Lavrova, O.; Ranade, S. Analysis of Bad Data Processing Methodologies in Power System State Estimation. In Proceedings of the 2021 North American Power Symposium (NAPS), College Station, TX, USA, 11–14 April 2021; pp. 1–6.
- [23] Gay, D.M. The AMPL Modeling Language: An Aid to Formulating and Solving Optimization Problems. *Numer. Anal. Optim.* 2015, 134, 95–116
- [24] Bedoya, J.C.; Xie, J.; Wang, Y.; Zhang, X.; Liu, C.C. Resiliency of Distribution Systems Incorporating Asynchronous Information for System Restoration. *IEEE Access* 2019, 7, 101471–101482.
- [25] Florez, H. A. R.; Marujo, D.; López, G. P.; López-Lezama, J. M.; Muñoz-Galeando, N. State Estimation in Electric Power Systems Using an Approach Based on a Weighted Least Squares Non-Linear Programming Modeling. *Electronics* 2021, 10, 2560.
- [26] Nagy, I.; Suzdaleva, E.; Mlynářová, T. Comparison of state estimation using finite mixtures and hidden Markov models. In Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems, Prague, Czech Republic, 15–17 September 2011, Vol. 2, pp. 527–531.

# BAYES ESTIMATION OF CAPABILITY INDEX USING THREE-PARAMETER WEIBULL DISTRIBUTION

SONAM GUBRELEY<sup>1</sup>, ANKITA GUPTA<sup>2</sup>, AND SATYANSHU K. UPADHYAY<sup>1</sup>



<sup>1</sup>Department of Statistics

<sup>2</sup>Statistics Section, Mahila Mahavidyalay

Banaras Hindu University, Varanasi - 221 005, India.

sonamgubreley05@gmail.com

## Abstract

*The process capability index is an important tool used in quality control and process improvement. Generally, the index is estimated under the assumption of a normal distribution, although some other distributions are also recommended in the literature. This paper instead considers a three-parameter Weibull distribution and obtains an estimate of the process capability index under the Bayesian framework. Bayesian development is based on the use of non-informative priors and the posterior sample-based inferences are drawn using an important Markov Chain Monte Carlo technique, namely, the Gibbs sampler algorithm. Finally, a numerical illustration based on two real datasets is provided.*

**Keywords:** Process capability index, Gibbs sampler, Three-parameter Weibull distribution

## 1. INTRODUCTION

With the advancement of technology, there is an ever-increasing demand for high-quality products and services. Smart manufacturing process employing various advanced technologies facilitate automation, enhance productivity, improve maintenance and monitoring and reduce scope of human error. However, associated software products need to be examined for quality assurance.

The quality and reliability of the product can be assessed through various statistical tools, among which, process capability index (PCI) has been found propitious by the manufacturers as it is useful in assisting decision-making and boosting efforts in process performance. PCI is a measuring tool for accurately analysing the potential of a process and its performance. For quality control engineers, it is extremely important since it quantifies the relationship between the process's actual performance and the product's predetermined parameters. The index ascertains whether the process meets the defined manufacturing prerequisites. In this regard many capability indices have been developed so far (see, for example, [31], [11], [14] and [5]). The first index put forward in the literature was  $C_p$ , which simply calculates the span of the specifications relative to the six-sigma spread in the process (see [31]). As per this index, the process mean is centred between the lower and the upper specification limits. One of the major issues with this index is that it does not take into account the location of the process mean relative to the specifications. Moreover, if the process is not centred on the specification region, it would be possible to have a substantial percentage of the products with characteristics outside the specification limit although  $C_p$  may be high. In order to overcome this problem, [11] introduced another capability index,  $C_{pk}$ , which takes process centring into account in addition to the spread of the specifications relative to the six-sigma spread in the process. In other words, it measures the distance between the specification limits closest to the average from the quality characteristic of interest. Mathematically,  $C_p$  and  $C_{pk}$  can be defined as

$$C_p = \frac{USL - LSL}{6\sigma_p}, \tag{1}$$

$$C_{pk} = \min(C_{pu}, C_{pl}), \tag{2}$$

where

$$C_{pu} = \frac{USL - \mu_p}{3\sigma_p}, \tag{3}$$

$$C_{pl} = \frac{\mu_p - LSL}{3\sigma_p}, \tag{4}$$

USL and LSL are the upper and lower specification limits, respectively,  $\mu_p$  denotes the process mean and  $\sigma_p$  represents the process standard deviation.

Both of these PCIs are defined under two important assumptions, that is, the process is under statistical control and the quality characteristic of the process of interest is normally distributed (see [31]). Perhaps, because of these assumptions, a bulk of literature is available on the estimation of PCIs under the assumption of normality (see, for example, [1], [2], [13] and [23]). However, industrial processes are often not normally distributed and, for such scenarios, the values of conventional PCIs may be absurd and possibly misrepresent the quality of the product. For example, one may refer to [10], [27] and [24] for a systematic and detailed coverage. In order to remove this discrepancy, [3] proposed the quantile-based measure to estimate the capability index for non-normal distributions, which is given as under.

$$C_{pk} = \min\left(\frac{USL - M}{U_p - M}, \frac{M - LSL}{M - L_p}\right), \tag{5}$$

where  $U_p$ ,  $L_p$  and  $M$  are the 99.865th, 0.135th, and 50th percentiles of the target distribution, respectively, USL and LSL indicate upper and lower specification limits. A value of  $C_{pk} < 1$  is unfavourable and indicates that the process is incapable, whereas, a value of  $1 \leq C_{pk} \leq 1.33$  indicates that the process is barely capable and  $C_{pk} \geq 1.33$  shows that the process is capable to meet the consumers' requirements.

Besides normality assumption, several developments can be seen in literature on non-normal assumptions as well. [3], [14], [17], [16], [22], [12], [9], [26] and [20] are some of the important among other references where capability indices are estimated under the assumption of non-normal distributions. A thorough literature review on the estimation of PCIs for non-normal datasets reveals that most of the developments are done using classical framework and only a few of them considered Bayesian approach for estimating capability index. Further, in statistical process control, most of the datasets lie at a particular location, generally far from zero, and, therefore, it becomes imperative to assess capability index by considering a model which has a location parameter even if one is dealing with non-normal data. To the best of our knowledge, there is no reference in the literature that entertains a non-normal model with location parameter for estimating the capability index. To bridge this gap, this paper considers a three-parameter Weibull distribution for estimating the capability index and performs a Bayes analysis of the distribution.

The Weibull distribution is an important distribution that has received enough attention in the field of reliability and quality control. Its versatility stems from the fact that it incorporates increasing, decreasing and stable hazard rates for different values of its shape parameter (see [18] and [15], etc). The literature on the analysis of Weibull distribution has considered both two-parameter and three-parameter form of model where the former model is defined without a threshold parameter. The two-parameter Weibull distribution is comparatively easier to deal with as compared to three-parameter model form and, therefore, the literature on both classical and Bayes analysis of two-parameter Weibull distribution is available in bulk (see, for example, [19], [28], [15], [25], among others). On the other hand, the three-parameter Weibull distribution is much richer because of the involvement of a threshold parameter although its analysis is slightly

more challenging due to sometime unusual behaviour of the likelihood function, especially when the shape parameter is less than unity (see also [30] and [32]). As a result, this model is comparatively less entertained in the literature. [30], [32] and [28] are some of the important references among others where this form of the model is explored.

As mentioned, this paper is an attempt to provide Bayes analysis of the three-parameter Weibull distribution with ultimate objective of finding the estimate of PCI. The entire development is done using non-informative priors for the model parameters. It is seen that the resulting posterior is analytically intractable to draw exact posterior based inferences and, therefore, the paper utilizes an important Markov Chain Monte Carlo (MCMC) procedure, namely the Gibbs sampler algorithm, to simulate posterior samples and draw the sample based inferences including those of PCI. Finally, the proposed methodology is numerically illustrated on the basis of two real datasets from a juice manufacturing company.

The plan of the paper is as follows. The next section briefly describes the three-parameter Weibull model and its Bayesian formulation. Section 3 provides numerical illustration based on two real datasets. Finally, a brief conclusion is provided in the last section.

## 2. MODEL FORMULATION

### 2.1. Likelihood function

The probability density function (pdf) of the three-parameter Weibull distribution is

$$f(x|\theta, \beta, \mu) = \frac{\beta}{\theta} \left( \frac{x - \mu}{\theta} \right)^{\beta-1} \exp \left[ - \left( \frac{x - \mu}{\theta} \right)^\beta \right], \quad x > \mu; \quad \theta, \beta, \mu > 0 \quad (6)$$

where  $\theta$ ,  $\beta$  and  $\mu$  are the scale, shape and location parameters, respectively. The distribution exhibits increasing hazard rate for  $\beta > 1$ , decreasing hazard rate for  $\beta < 1$  and, for  $\beta = 1$ , the distribution reduces to two-parameter exponential model possessing constant hazard rate. Let us use the notation  $W(\theta, \beta, \mu)$  to denote the three-parameter Weibull distribution given in (6). The reliability function and the hazard function of  $W(\theta, \beta, \mu)$  at time  $t$  are, respectively, given by

$$R(t) = \exp \left[ - \left( \frac{t - \mu}{\theta} \right)^\beta \right], \quad (7)$$

and

$$h(t) = \frac{\beta}{\theta} \left( \frac{t - \mu}{\theta} \right)^{\beta-1}. \quad (8)$$

Similarly, the expressions for  $U_p$ ,  $L_p$  and  $M$  for the model  $W(\theta, \beta, \mu)$  can be written as

$$U_p = \theta [2.86967]^{1/\beta} + \mu, \quad (9)$$

$$L_p = \theta [0.00058]^{1/\beta} + \mu, \quad (10)$$

and

$$M = \theta [\ln 2]^{1/\beta} + \mu, \quad (11)$$

respectively.

Let us now assume that an experiment consisting of  $n$  units is being conducted and let  $\underline{x} = (x_i; i = 1, 2, \dots, n)$  be the resulting observations. Then, the likelihood function for the dataset  $\underline{x}$  can be expressed as

$$L(\underline{x}|\theta, \beta, \mu) = \left( \frac{\beta}{\theta} \right)^n \prod_{i=1}^n \left( \frac{x_i - \mu}{\theta} \right)^{\beta-1} \exp \left[ - \sum_{i=1}^n \left( \frac{x_i - \mu}{\theta} \right)^\beta \right]. \quad (12)$$

## 2.2. Bayesian formulation

To conduct Bayesian analysis, it is essential to specify prior distribution for the parameters of the entertained model. Several types of priors are proposed in the literature for the Weibull parameters. The paper, however, considers joint non-informative prior as suggested by [32] and the same is given as

$$g(\theta, \beta, \mu) \propto \frac{1}{\theta\beta}. \quad (13)$$

Obviously, the parameter  $\mu$  is assigned a constant prior over the positive real space.

The updated belief in the form of posterior distribution can be obtained by combining the prior distribution as specified in (13) with the likelihood function given in (12) via Bayes theorem. The joint posterior up to proportionality can, therefore, be written as

$$p(\theta, \beta, \mu | \underline{x}) \propto \frac{\beta^{n-1}}{\theta^{n\beta+1}} \prod_{i=1}^n (x_i - \mu)^{\beta-1} \exp \left[ - \sum_{i=1}^n \left( \frac{x_i - \mu}{\theta} \right)^\beta \right]; \quad \theta > 0, \beta > 0, \mu < \min(\underline{x}). \quad (14)$$

Obviously, the posterior given in (14) is analytically intractable and, therefore, one has to proceed with some approximation or simulation based alternative approaches for drawing the desired inferences from the posterior. As mentioned, this paper considers Gibbs sampler algorithm, an important MCMC procedure, because of its straightforwardness and ease of implementation. The algorithm requires specification of low-dimensional full conditionals for simulating the high dimensional posterior where both full conditionals and the posterior need to be specified up to proportionality only. The algorithm starts with the appropriately chosen initial values for the variates and then simulates the full conditionals one by one in a cyclic fashion with most recent available values for all the given variates at every stage. Obviously, the appropriately chosen initial values are updated after the first cycle of iteration from all the full conditionals. The process is continued for a large number of cycles until some systematic pattern of convergence is achieved among the generating variates. Moreover, it can be easily seen that the posterior (14) results in three one-dimensional full conditionals corresponding to  $\theta$ ,  $\beta$  and  $\mu$  and these full conditionals can be easily simulated resulting in an easy implementation of the Gibbs sampler algorithm. For further details on the algorithm, one can refer to [7], [6] and [32], among others.

Coming on to the full conditionals derived from (14), it can be seen that the full conditional for  $\theta$  happens to be the kernel of gamma distribution after appropriate transformation and, hence,  $\theta$  can be easily generated from a gamma generating routine (see [4]). The full conditional of  $\beta$  can be seen to be log concave and, therefore,  $\beta$  can be simulated using adaptive rejection sampling procedure (see [8]). The generation of  $\mu$  from its full conditional is based on the rejection algorithm using the envelope density  $g_1(\mu | \beta, x_1) = \left( \frac{\beta}{x_1^\beta} \right) (x_1 - \mu)^{(\beta-1)}$ ;  $x_1 > \mu$ , where  $x_1$  is minimum of  $(x_i; i = 1, 2, \dots, n)$  (see [32] for further details).

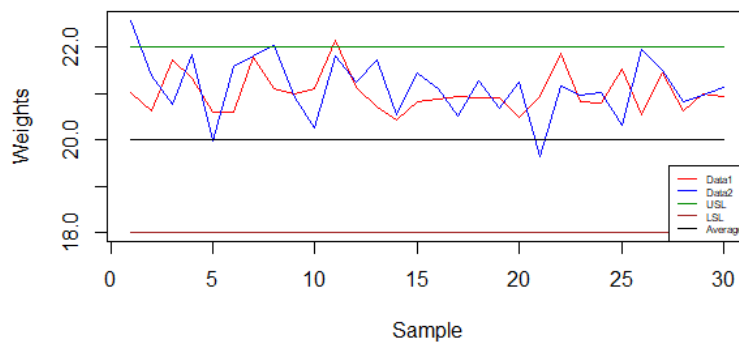
## 3. NUMERICAL ILLUSTRATION

For numerical illustration of the proposed formulation, the paper considers two real datasets on the weights (in grams) of thirty juice packs of grape and strawberry flavours. In the discussion that follows, the dataset on weights of juice packs of grape flavour is referred to as the *Data1* whereas that of strawberry flavour is referred to as the *Data2*. The two datasets are presented in Table 1 and these are actually collected to assess the process of filling powdered juice bags. The two datasets were first reported by [21] where the authors analysed the datasets under the assumption of normal distribution and evaluated  $C_{pk}$  by considering the specification limits as: LSL= 18.0 and USL= 22.0. These specification limits were specified in accordance with the guidelines provided by the National Institute of Metrology, Quality and Technology (INMETRO), the Brazilian organisation responsible for the quality control.

Before proceeding with the analysis of datasets, let us plot the control charts with the specification limits of 18.0 and 22.0. The control charts are presented in Figure 1 where the red line corresponds to *Data1* and the blue line corresponds to *Data2*. Moreover, the specification limits 18.0 and 22.0 suggest that the process must hover around the mean of these specification limits although the Figure 1 clearly suggests that the process is not centred around its mean. In fact, there are certain values that lie outside the provided range, which ultimately suggest that the process is out of control.

**Table 1:** Data on weights (in grams) of juice packs

Data1				
21.011	20.635	21.732	21.333	20.587
20.587	21.784	21.088	20.997	21.100
22.155	21.116	20.707	20.413	20.822
20.883	20.930	20.908	20.897	20.486
20.935	21.867	20.814	20.795	21.520
20.537	21.438	20.621	20.975	20.919
Data2				
22.572	21.376	20.768	21.833	19.970
21.583	21.813	22.025	20.892	20.241
21.816	21.232	21.730	20.529	21.435
21.106	20.519	21.263	20.684	21.233
19.624	21.150	20.962	21.024	20.316
21.942	21.495	20.819	20.973	21.115



**Figure 1:** Control chart for the two datasets.

Further, before carrying out the Bayes analysis of the considered datasets, let us check the compatibility of two datasets with the assumed model (6). The compatibility was examined based on Kolmogorov-Smirnov (KS) test statistic which was evaluated using maximum likelihood (ML) estimates of the model parameters. It may be noted that the ML estimates for  $\theta, \beta$  and  $\mu$  were found to be 0.693, 1.475 and 20.391, respectively, for *Data1* and 2.635, 4.244 and 18.737, respectively, for *Data2*. Finally, for *Data1*, the KS statistic was found to be 0.110 with the corresponding p-value as 0.860 while for *Data2*, the KS statistic was 0.066 with the corresponding p-value as 0.998. Obviously, the two datasets provide good compatibility with the model  $W(\theta, \beta, \mu)$ .

For performing the Bayes analysis, the Gibbs sampler algorithm was implemented on the posterior (14) as per details given in subsection 2.2. Convergence monitoring was done using

ergodic averages, obtained separately for each of the three variates, using a single long run of the iterating chain. It was found that 50K iterations were good enough for getting stationarity behaviour of the ergodic averages. Once the convergence was assessed, equally spaced observations at a gap of 10 were chosen to make auto correlation negligibly small. In this way, a posterior sample of size 1K was taken from the marginal posterior of each of  $\theta$ ,  $\beta$  and  $\mu$  (see also [29] and [32]). Once the samples of  $\theta$ ,  $\beta$  and  $\mu$  are obtained, the same can be used in (9)-(11) by substitution to get the corresponding samples of size 1K from the posterior of each of  $U_p$ ,  $L_p$  and  $M$ . Finally, the samples of  $U_p$ ,  $L_p$  and  $M$  so obtained can be used to get the posterior samples of size 1K corresponding to  $C_{pk}$  given in (5).

**Table 2:** Estimated posterior summaries for  $\theta$ ,  $\beta$ ,  $\mu$  and  $C_{pk}$

Datasets	Parameters	Estimated Posterior Summaries				
		Mean	Median	Mode	0.95 HPDI	
Data1	$\theta$	0.701	0.700	0.698	0.587	0.816
	$\beta$	1.493	1.490	1.483	1.215	1.794
	$\mu$	20.384	20.386	20.391	20.350	20.412
	$C_{pk}$	1.227	1.221	1.211	0.931	1.532
Data2	$\theta$	2.663	2.589	2.439	1.770	3.725
	$\beta$	4.259	4.139	3.901	2.522	6.242
	$\mu$	18.706	18.780	18.928	17.715	19.561
	$C_{pk}$	0.870	0.869	0.867	0.695	1.053

Table 2 provides a few important posterior based summaries of different posterior characteristics corresponding to various entertained model parameters, each estimated on the basis of corresponding 1K posterior samples. These summaries are shown in the form of estimated posterior mean, median, mode and the highest posterior density intervals with 0.95 coverage probability (0.95 HPDI) for each of the two datasets. It can be observed from Table 2 that the estimated posterior mean, median and mode corresponding to each parameter for both the datasets are quite close to each other, implying that the posterior distributions are approximately symmetric. Furthermore, the width of 0.95 HPDIs for all the parameters are quite small indicating less variability in the estimated values of the parameters and, hence, ensuring the consistency of the estimated values. An important finding presented in Table 2 is that  $1 \leq C_{pk} \leq 1.33$  for *Data1*, indicating that the process is barely capable whereas for *Data2*  $C_{pk} < 1$  implying that the process is incapable and requires further improvement. A similar conclusion was drawn on the basis of control charts shown in Figure 1.

#### 4. CONCLUSION

Technological advancements have typically led to an expansion of the industry, wherein the need for high-quality goods and services is reinforced by a competitive environment. From this vantage point, industries that deal with manufacturing are always susceptible to manufacturing process failures leading to the products that may not meet the desired specifications. The manufacturing sector has made extensive use of PCIs, providing a numerical gauge of a process's ability to produce goods that satisfy the factory-set quality standards. In estimating PCIs, more often the assumption is made that the data are generated randomly using a normal model. Nonetheless, asymmetric data are found in many circumstances. This paper has successfully demonstrated the utility of the three-parameter Weibull model in estimating the aforesaid index. Further, the Bayesian methodology developed in the paper is also found to offer the intended inferences in a routine manner. The inferential results show that the process pertaining to *Data1* is barely capable while that of *Data2* is incapable to offer the desired quality assurance.



## REFERENCES

- [1] Albing, M. (2006). *Process capability analysis with focus on indices for one-sided specification limits*. PhD thesis, Luleå Tekniska Universitet.
- [2] Chou, Y. and Owen, D. B. (1989). On the distributions of the estimated Process capability indices. *Communications in Statistics-Theory and Methods*, 18(12):4549–4560.
- [3] Clements, J. A. (1989). Process capability calculations for non-normal distributions. *Quality Progress*, 22:95–100.
- [4] Devroye, L. (1986). *Non-Uniform Random Variate Generation*. Springer-Verlag, New York.
- [5] Feigenbaum, A. V. (1951). *Quality Control: Principles, practice and administration: An industrial management tool for improving product quality and design and for reducing operating costs and losses*. McGraw-Hill.
- [6] Gelfand, A. E., Hills, S. E., Racine-Poon, A., and Smith, A. F. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association*, 85(412):972–985.
- [7] Gelfand, A. E. and Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409.
- [8] Gilks, W. R. and Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 41(2):337–348.
- [9] Gupta, P. K. and Singh, A. K. (2017). Classical and Bayesian estimation of Weibull distribution in presence of outliers. *Cogent Mathematics*, 4(1):1300975.
- [10] Hosseinifard, Z., Abbasi, B., and Niaki, S. (2014). Process capability estimation for leukocyte filtering process in blood service: A comparison study. *IIE Transactions on Healthcare Systems Engineering*, 4(4):167–177.
- [11] Kane, V. E. (1986). Process capability indices. *Journal of quality technology*, 18(1):41–52.
- [12] Kashif, M., Aslam, M., Al-Marshadi, A. H., and Jun, C. (2016). Capability indices for non-normal distribution using gini's mean difference as measure of variability. *IEEE Access*, 4:7322–7330.
- [13] Kocherlakota, S. and Kocherlakota, K. (1991). Process capability index: bivariate normal distribution. *Communications in Statistics-Theory and Methods*, 20(8):2529–2547.
- [14] Kotz, S. and Johnson, N. L. (2002). Process capability indices—A Review, 1992–2000. *Journal of quality technology*, 34(1):2–19.
- [15] Lawless, J. F. (2011). *Statistical models and methods for lifetime data*. John Wiley & Sons, New York.
- [16] Leiva, V., Marchant, C., Saulo, H., Aslam, M., and Rojas, F. (2014). Capability indices for Birnbaum–Saunders processes applied to electronic and food industries. *Journal of Applied Statistics*, 41(9):1881–1902.
- [17] Lin, G., Pearn, W., and Yang, Y. (2005). A Bayesian approach to obtain a lower bound for the  $C_{pm}$  capability index. *Quality and Reliability Engineering International*, 21(6):655–668.
- [18] Mann, N. R., Schafer, R. E., and Singpurwalla, N. D. (1974). *Methods for Statistical Analysis of Reliability and Life Data*. John Wiley & Sons, New York.
- [19] Martz, H. F. and Waller, R. A. (1982). *Bayesian reliability analysis*. Wiley series in probability and mathematical statistics. John Wiley & Sons, New York.
- [20] Meng, F., Yang, J., and Huang, S. (2021). Hypothesis testing of Process capability index  $C_{pk}$  from the perspective of generalized fiducial inference. *Quality and Reliability Engineering International*, 37(4):1578–1598.
- [21] Molina, R. (2018). Técnicas de avaliação de medidas e verificação dos índices de capacidade. *Presidente Prudente-São Paulo*.
- [22] Panichkitkosolkul, W. (2016). Confidence intervals for the Process capability index  $C_p$  based on confidence intervals for variance under non-normality. *Malaysian Journal of Mathematical Sciences*, 10(1):101–115.
- [23] Pearn, W. (1998). New generalization of process capability index  $C_{pk}$ . *Journal of Applied Statistics*, 25(6):801–810.

- [24] Piña-Monarez, M. R., Ortiz-Yañez, J. F., and Rodríguez-Borbón, M. I. (2016). Non-normal capability indices for the weibull and lognormal distributions. *Quality and Reliability Engineering International*, 32(4):1321–1329.
- [25] Ramos, P. L., Almeida, M. H., Louzada, F., Flores, E., and Moala, F. A. (2022). Objective Bayesian inference for the Capability index of the Weibull distribution and its generalization. *Computers & Industrial Engineering*, 167:108012.
- [26] Saha, M., Dey, S., Yadav, A. S., and Ali, S. (2021). Confidence intervals of the index  $C_{pk}$  for normally distributed quality characteristics using classical and Bayesian methods of estimation. *Brazilian Journal of Probability and Statistics*, 35(1):138–157.
- [27] Sennaroğlu, B. and Şenvar, Ö. (2015). Performance comparison of Box-cox transformation and weighted variance methods with Weibull distribution. *Journal of Aeronautics and Space Technologies*, 8(2):49–55.
- [28] Singpurwalla, N. D. (2006). *Reliability and risk: A Bayesian Perspective*. John Wiley & Sons, New York.
- [29] Smith, A. F. and Roberts, G. O. (1993). Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Methodological)*, 55(1):3–23.
- [30] Smith, R. L. and Naylor, J. (1987). A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 36(3):358–369.
- [31] Sullivan, L. P. (1984). Reducing variability: A new approach to quality. *Quality Progress*, 17(7):15–21.
- [32] Upadhyay, S., Vasishta, N., and Smith, A. (2001). Bayes inference in life testing and reliability via Markov chain Monte Carlo simulation. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 63(1):15–40.

# A NEW ALGORITHM TO SOLVE MULTI-OBJECTIVE TRANSPORTATION PROBLEM WITH GENERALIZED TRAPEZOIDAL FUZZY NUMBERS

RAMAKANT SHARMA<sup>1</sup>, SOHAN LAL TYAGI<sup>2\*</sup>



1. Research Scholar, Department of Mathematics, SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, Ghaziabad, 201204, India  
rs5364@srmist.edu.in
- 2\*. Assistant Professor, Department of Mathematics, SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, Ghaziabad, 201204, India  
\* Corresponding Author  
drsohanlyagi@gmail.com

## Abstract

*Transportation Problem is a specific type of linear programming problem (LPP). Today, in the real world, the decision maker handles the multi-objectives at the same time. Fuzzy Concepts are used in LPP to handle the uncertainty and vagueness of data. This paper presents a new algorithm to solve a special type of fuzzy transportation problem (FTP) with the generalized trapezoidal fuzzy numbers (GTpFN) in which the decision maker is not certain about the exact value of transportation charge and the availabilities and requirements are the real numbers. In this Proposed Algorithm first, the fuzzy multi-objective transportation problem (FMOTP) is converted into a Crisp multi-objective transportation problem (MOTP) by the Proposed ranking function, and then the Crisp MOTP is transformed into a single objective transportation problem using the sum of objective functions values. The proposed algorithm gives an efficient compromise solution of FMOTP. To elaborate the proposed algorithm, one numerical example is solved.*

**Keywords:** Ranking function, Multi-Objective Transportation Problem, Generalized Trapezoidal fuzzy number.

## 1. INTRODUCTION

The transportation problem (TP) is a classical optimization problem in operations research and logistics. To satisfy requirements and availabilities, it involves determining the most cost-effective way to distribute a product from various providers to various consumers. TP aims to minimize the total transportation cost. Traditional methods, including the Vogel approximation method, the Matrix Minima approach, and the North West Corner method, are used to solve the TP. In the real-world scenario, nowadays the decision maker can handle multiple objectives at a single time in which the decision maker is unsure about the precise value of transportation cost, requirements, and availabilities. The multi-objective transportation problem (MOTP) is a linear optimization problem with several variable objectives and equality Constraints. Fuzzy concepts often deal with such types of uncertainty and vagueness in the exact cost of transportation, availabilities, and requirements. The Concept of fuzzy transportation problems (FTP) was developed to find the

solution to the TP's unpredictable parameters, such as fuel prices, weather conditions, product supply, demands, etc. Trapezoidal fuzzy numbers (TpFN) are useful when modeling uncertain parameters in transportation problems, such as requirements and availabilities quantities or transportation costs, which are not precisely known but have a range of potential values.

The TP was developed by F.L. Hitchcock [1] originally in 1941. The TP was represented by a standard LPP form that can be solved by the simplex method. Lotfi A. Zadeh [2] was given the concept of fuzziness in 1965. Charnes and Cooper [3] developed The Stepping Stone approach offers an alternative approach for obtaining information from the Simplex Method. Zimmermann H.J. [4] was the first to use an appropriate membership function to solve an LP problem with multiple objectives. Ringuest et al. [5] gave two interactive algorithms for solving MOTP. Bit et al. [6] Solved TP problems with several criteria by using Fuzzy Programming. Chanas et al. [7] Proposed a model based on fuzzy linear programming to solve TPs in which cost coefficients are crisp values and supplies and demands are fuzzy values. Liu et al. [8] developed a method that is based on the extension principle to solve FTPs. Kiruthiga, M., et al. [9] used Interval arithmetic based on Alpha-cut to solve non-linear programming problems (NLP). M Afwat et al. [10] introduced the Product Approach to find an efficient solution for MOTP. Gani and Razak [11] presented a parametric approach for two-stage fuzzy cost-minimizing TP that has supplies and demands in the form of a trapezoidal fuzzy number. Bagheri. M. et al. [12] presented the DEA approach to solving FMOTP. Maity. G. et al. [13] studied the MOTP under uncertain environments. Dinagar and Palanivel [14] studied the FTP with trapezoidal fuzzy numbers. Pandian et.al [15] developed the zero-point method to find the fuzzy solution for the FTP. Hamiden Abd El-Waheed Khalifa et.al. [16] Presented a fuzzy geometric programming approach to find an optimal compromise solution for two-stage multi-objective TP. Srikanth Gupta et. al. [17] Investigated the MOTP with capacitated restrictions that have some linear objective functions and some that are fractional. Murshid Kamal et.al. [18] Studied the MOTP, where the objective function is type-2 TpFN in which supply and demand follow various types of probabilistic distributions. They used the fuzzy goal programming method to find an optimal solution. M.A. Sayed et.al. [19] Developed a novel approach to solve intuitionistic Fuzzy fractional MOTP. H. Adb E. Khalifa [20] proposed a signed distance ranking function method to obtain the set of efficient solution fuzzy MOTP. Yi-Mang et al. [21] adopted two fuzzy ranking methods based on their mean graded values and distance from the mean ranking function and proposed a novel ripple-spreading algorithm to solve FMOLPP. Y Kacher et al. [22] presented a novel two-step generalized parametric approach to solving different fuzzy parametric-based MOTP. SG Bodke [23] introduced a method to solve fuzzy MOTP after converting it into Crisp MOTP which is based on Zimmerman technique using the exponential membership function.

This paper presents a new algorithm for solving fuzzy MOTP with cost values as generalized trapezoidal fuzzy numbers and requirements availabilities are the real numbers. In this algorithm, firstly, the fuzzy MOTP is converted into Crisp MOTP by the proposed Ranking function. After converting the fuzzy MOTP into Crisp MOTP, the Crisp MOTP is changed into the single objective crisp transportation problem. The algorithm is based on row/column maximum and minimum. Our proposed method directly obtained a unique, efficient solution, which leads to a Compromise Solution of crisp and fuzzy MOTP.

## 2. ABBREVIATIONS

1. Linear Programming Problem - LPP
2. Transportation Problem - TP
3. Fuzzy Transportation Problem - FTP
4. Trapezoidal Fuzzy Number - TpFN
5. Generalized Trapezoidal Fuzzy Number - GTpFN

6. Multi-Objective Transportation Problem - MOTP
7. Fuzzy Multi-Objective Transportation Problem - FMOTP
8. Decision Maker- DM
9. Fuzzy Transportation Cost - FTC
10. Minimum Transportation Cost - MTC
11. Single objective transportation problem - SOTP

### 3. BASIC DEFINITIONS

1. Fuzzy Number: A fuzzy set  $\tilde{A}$  is said to be fuzzy number if its membership function  $\tilde{A} : \mathbb{R} \rightarrow [0, 1]$  has satisfy the following conditions:  
 $\tilde{A}(\lambda x_1 + (1 - \lambda)x_2) \geq \min\{\tilde{A}(x_1), \tilde{A}(x_2)\}$   
 there exist a  $x \in \mathbb{R}$  such that  $\tilde{A}(x) = 1$   
 $\tilde{A}$  is piece-wise continuous
2. Generalized Trapezoidal Fuzzy Numbers (GTpFN): A fuzzy number  $\tilde{A} = (p_1, p_2, p_3, p_4; w)$  where  $p_1 < p_2 < p_3 < p_4$  and  $0 < w \leq 1$  with membership function defined as:

$$\mu_{\tilde{A}(x)} = \begin{cases} w\{1 - \frac{p_2-x}{p_2-p_1}\} & \text{if } p_1 \leq x \leq p_2 \\ w & \text{if } p_2 \leq x \leq p_3 \\ w\{1 - \frac{x-p_3}{p_4-p_3}\} & \text{if } p_3 \leq x \leq p_4 \\ 0 & \text{Otherwise} \end{cases}$$

3. Properties of Trapezoidal Fuzzy Numbers (TpFN): let  $\tilde{A} = (p_1, p_2, p_3, p_4; w_1)$  and  $\tilde{B} = (q_1, q_2, q_3, q_4; w_2)$  be any two GTpFNs. then  
 $\tilde{A} + \tilde{B} = (p_1, p_2, p_3, p_4; w_1) + (q_1, q_2, q_3, q_4; w_2) = (p_1 + q_1, p_2 + q_2, p_3 + q_3, p_4 + q_4; \min(w_1, w_2))$   
 $\tilde{A} - \tilde{B} = (p_1, p_2, p_3, p_4; w_1) - (q_1, q_2, q_3, q_4; w_2) = (p_1 - q_1, p_2 - q_2, p_3 - q_3, p_4 - q_4; \min(w_1, w_2))$   
 $\tilde{A} \times \tilde{B} = (p_1, p_2, p_3, p_4; w_1) \times (q_1, q_2, q_3, q_4; w_2) = \{\min(p_1q_1, p_1q_4, p_4q_1, p_4q_4), \min(p_2q_2, p_2q_3, p_3q_2, p_3q_3), \max(p_2q_2, p_2q_3, p_3q_2, p_3q_3), \max(p_1q_1, p_1q_4, p_4q_1, p_4q_4)\}$   
 $\sigma \tilde{A} = (\sigma p_1, \sigma p_2, \sigma p_3, \sigma p_4)$ , where  $\sigma$  is any constant.

### 4. PROPOSED RANKING METHOD

The ranking method is used to compare the fuzzy numbers. Assuming that the natural order is preserved, The ranking function  $\mathfrak{R} : T(\mathbb{R}) \rightarrow \mathbb{R}$  defined on set of real numbers maps each fuzzy number into a real number where  $T(\mathbb{R})$  is set of the fuzzy numbers.

the proposed ranking function for the Trapezoidal number  $\tilde{A} = (p_1, p_2, p_3, p_4; w)$  is given as

$$\mathfrak{R}(\tilde{A}) = \frac{2p_1 + 5w(p_2 + p_3) + 2p_4}{14}$$

### 5. PROPERTIES OF RANKING FUNCTIONS

$\tilde{A} = (p_1, p_2, p_3, p_4; w_1)$  and  $\tilde{B} = (q_1, q_2, q_3, q_4; w_2)$  be any two GTpFNs. then the properties of the ranking function is given as:

$$\tilde{A} \leq \tilde{B} \text{ iff } \mathfrak{R}(\tilde{A}) \leq \mathfrak{R}(\tilde{B})$$

$$\begin{aligned} \tilde{A} &\equiv \tilde{B} \text{ iff } \mathfrak{R}(\tilde{A}) \equiv \mathfrak{R}(\tilde{B}) \\ \tilde{A} &\geq \tilde{B} \text{ iff } \mathfrak{R}(\tilde{A}) \geq \mathfrak{R}(\tilde{B}) \end{aligned}$$

## 6. MATHEMATICAL MODEL FOR FMOTP WITH GTpFN

The FMOTP with k objectives in mathematical form is given as:

$$\text{Min } \tilde{Z}_k(x) = \sum_{i=1}^m \sum_{j=1}^n \tilde{a}_{ij}^{(k)} x_{ij} \text{ for } k = 1, 2, \dots$$

Subject to

$$\begin{aligned} \sum_{i=1}^m x_{ij} &= d_j: \text{ for fixed } j=1, 2, \dots, n \\ \sum_{j=1}^n x_{ij} &= s_i: \text{ for fixed } i=1, 2, \dots, m \\ x_{ij} &\geq 0 \end{aligned}$$

Where,

$s_i$  = the product's availability at the i-th source

$d_j$  = the product's requirements at the j-th destinations

$\tilde{a}_{ij}^{(k)}$  = the fuzzy cost for transporting one unit of the given product from i-th source to j-th destination of k-th objective

$x_{ij}$  = Product's quantity transported from i-th source to j-th destination.

$\tilde{a}_{ij}^{(k)}$  are the GTpFNs.

## 7. EFFICIENT SOLUTION

A feasible solution  $X^0 = \{x_{ij}^0, i = 1, 2, \dots, m, j = 1, 2, \dots, n\}$  is called an efficient solution to the problem (T) if there does not exist any feasible solution Y of MOTP such that  $Z_1(X) \leq Z_1(X^0)$  and  $Z_2(X) \leq Z_2(X^0)$ .

## 8. OUR PROPOSED ALGORITHM

The Compromise efficient fuzzy solution of fuzzy MOTP is obtained by the proposed algorithm. The proposed algorithm's steps are as follows:

Step I: In this step first, the fuzzy MOTP is converted into crisp MOTP by the proposed Ranking Method. The proposed ranking method converted the fuzzy quantities into crisp quantities. The crisp MOTP in mathematical form can be given as:

$$\text{Min } Z_k(x) = \sum_{i=1}^m \sum_{j=1}^n a_{ij}^{(k)} x_{ij} \text{ for } k = 1, 2, \dots$$

Subject to

$$\begin{aligned} \sum_{i=1}^m x_{ij} &= d_j: \text{ for fixed } j=1, 2, \dots, n \\ \sum_{j=1}^n x_{ij} &= s_i: \text{ for fixed } i=1, 2, \dots, m \\ x_{ij} &\geq 0 \end{aligned}$$

Where

$s_i$  = the product availability at the i-th source

$d_j$  = the product requirements at the j-th destinations

$a_{ij}^{(k)}$  = the crisp cost for transporting one unit quantity of product from i-th source to j-th destination of k-th objective,

$x_{ij}$  = quantity of product transported from i-th source to j-th destination.

the crisp MOTP is represented in tabular form in Table 1.

**Table 1:** Tabular representation of Crisp MOTP

Destination → source ↓	$A_1$	$A_2$	.....	$A_n$	Availability $\{s_i\}$
$B_1$	$a_{11}^{(1)}$ $a_{11}^{(2)}$ :	$a_{12}^{(1)}$ $a_{12}^{(2)}$ :	.....	$a_{1n}^{(1)}$ $a_{1n}^{(2)}$ :	$s_1$
$B_2$	$a_{21}^{(1)}$ $a_{21}^{(2)}$ :	$a_{22}^{(1)}$ $a_{22}^{(2)}$ :	.....	$a_{2n}^{(1)}$ $a_{2n}^{(2)}$ :	$s_2$
:	:	:	:	:	:
$B_m$	$a_{m1}^{(1)}$ $a_{m1}^{(2)}$ :	$a_{m2}^{(1)}$ $a_{m2}^{(2)}$ :	.....	$a_{mn}^{(1)}$ $a_{mn}^{(2)}$ :	$s_m$
Requirement ( $d_j$ )	$d_1$	$d_2$	.....	$d_n$	

Step II: In this step the sum of the objectives is calculated.

$$t_{ij} = \sum_{v=1}^k C_{ij}^{(r)}$$
 for  $1 \leq i \leq m$  and  $1 \leq v \leq k$ .

then the crisp Single-Objective Transportation Problem (SOTP) in tabular form is represented in table 2.

**Table 2:** Tabular representation of Crisp SOTP

Destination → source ↓	$A_1$	$A_2$	.....	$A_n$	Availability $\{s_i\}$
$B_1$	$t_{11}$	$t_{12}$	.....	$t_{1n}$	$s_1$
$B_2$	$t_{21}$	$t_{22}$	.....	$t_{2n}$	$s_2$
:	:	:	:	:	:
$B_m$	$t_{m1}$	$t_{m2}$	.....	$t_{mn}$	$s_m$
Requirement ( $d_j$ )	$d_1$	$d_2$	.....	$d_n$	

Step III: Penalties of each row and columns

The Penalties for each row  $B_i : 1 \leq i \leq m$  is calculated as:

Row penalties  $\mu_i = [\text{maximum } (t_{1r}) - \text{minimum } (t_{1r})]$  for  $1 \leq r \leq n, \forall i : 1 \leq i \leq m$

Similarly, the penalties for each column  $A_p : 1 \leq p \leq n$  is calculated as:

Column Penalties  $\rho_j = [\text{maximum } (t_{s1}) - \text{minimum } (t_{s1})]$  for  $1 \leq s \leq m, \forall j : 1 \leq j \leq n$ .

The Crisp MOTP with penalties in given in table 3.

**Table 3:** Tabular representation of Crisp SOTP with penalties

Destination → source ↓	$A_1$	$A_2$	.....	$A_n$	Availability ( $s_i$ )	Row penalties ( $\mu_i$ )
$B_1$	$t_{11}$	$t_{12}$	.....	$t_{1n}$	$s_1$	$\mu_1$
$B_2$	$t_{21}$	$t_{22}$	.....	$t_{2n}$	$s_2$	$\mu_2$
:	:	:	.....	:	:	:
$B_m$	$t_{m1}$	$t_{m2}$	.....	$t_{mn}$	$s_m$	$\mu_m$
Requirement ( $d_j$ )	$d_1$	$d_2$	.....	$d_n$		
Column Penalties ( $\rho_j$ )	$\rho_1$	$\rho_2$	.....	$\rho_n$		

Step IV: In this step, the maximum penalty ( $\delta$ ) is calculated as:

$$\delta = \max \{ \mu_i, \rho_j; 1 \leq i \leq m, 1 \leq j \leq n \}$$

Select that Row/column which has a maximum penalty  $\delta$

Step V: In this step, select the cell that has minimum objective value in the row/column, that is selected in step IV.

Step VI: The maximum requirement/availability is allocated to the cell that is selected in Step V, and ignore that row/column that requirement/availability is satisfied.

Step VII: Repeat the process until all the requirements/availabilities are not fulfilled.

The flowchart of the proposed algorithm is given in figure 1.

To elaborate the proposed algorithm, a numerical example is considered.

Example: A fuzzy MOTP with three objective functions is considered.

Coe-efficient Matrix for a first objective function

$$C_1 = \begin{bmatrix} \tilde{a}_{11}^{(1)} & \tilde{a}_{12}^{(1)} & \tilde{a}_{13}^{(1)} & \tilde{a}_{14}^{(1)} & \tilde{a}_{15}^{(1)} \\ \tilde{a}_{21}^{(1)} & \tilde{a}_{22}^{(1)} & \tilde{a}_{23}^{(1)} & \tilde{a}_{24}^{(1)} & \tilde{a}_{25}^{(1)} \\ \tilde{a}_{31}^{(1)} & \tilde{a}_{32}^{(1)} & \tilde{a}_{33}^{(1)} & \tilde{a}_{34}^{(1)} & \tilde{a}_{35}^{(1)} \\ \tilde{a}_{41}^{(1)} & \tilde{a}_{42}^{(1)} & \tilde{a}_{43}^{(1)} & \tilde{a}_{44}^{(1)} & \tilde{a}_{45}^{(1)} \end{bmatrix}$$

$$= \begin{bmatrix} (9, 10, 11, 12; 0.8) & (12, 13, 14, 18; 0.8) & (8, 9.5, 11.5, 13; 0.8) & (3, 7, 9, 8; 0.8) & (5, 9, 12, 16; 0.8) \\ (6, 7, 8, 13; 0.8) & (2, 3, 3.5, 6; 0.8) & (5, 6.5, 8.5, 14; 0.8) & (5, 6, 9, 14; 0.8) & (4, 5, 6, 9; 0.8) \\ (4, 6, 7, 12; 0.8) & (3, 4, 7, 10; 0.8) & (7, 9, 12, 14; 0.8) & (3, 14, 15, 16; 0.8) & (2, 2.5, 4, 6; 0.8) \\ (2, 6, 9, 10; 0.8) & (6, 9, 10, 12; 0.8) & (3.5, 13, 16, 16.5; 0.8) & (1, 1.5, 2.5, 5; 0.8) & (1, 2, 2.5, 4; 0.8) \end{bmatrix}$$

Coe-efficient Matrix for the second objective function

$$C_2 = \begin{bmatrix} \tilde{a}_{11}^{(2)} & \tilde{a}_{12}^{(2)} & \tilde{a}_{13}^{(2)} & \tilde{a}_{14}^{(2)} & \tilde{a}_{15}^{(2)} \\ \tilde{a}_{21}^{(2)} & \tilde{a}_{22}^{(2)} & \tilde{a}_{23}^{(2)} & \tilde{a}_{24}^{(2)} & \tilde{a}_{25}^{(2)} \\ \tilde{a}_{31}^{(2)} & \tilde{a}_{32}^{(2)} & \tilde{a}_{33}^{(2)} & \tilde{a}_{34}^{(2)} & \tilde{a}_{35}^{(2)} \\ \tilde{a}_{41}^{(2)} & \tilde{a}_{42}^{(2)} & \tilde{a}_{43}^{(2)} & \tilde{a}_{44}^{(2)} & \tilde{a}_{45}^{(2)} \end{bmatrix}$$

$$= \begin{bmatrix} (1, 2.5, 3.5, 4; 0.6) & (10, 12, 13, 15.5; 0.6) & (9.5, 10, 11, 15; 0.6) & (0.5, 1, 1.5, 2.75; 0.6) & (2.5, 5, 6, 9; 0.6) \\ (0.25, 1, 1.5, 3) & (9, 11, 14, 16.5; 0.6) & (8.5, 10, 15, 17; 0.6) & (2.5, 7, 8, 10; 0.6) & (1.5, 2, 3, 5; 0.6) \\ (4.5, 12, 13, 14; 0.6) & (0.75, 1, 1.5, 2.5; 0.6) & (4, 11, 14, 14.5; 0.6) & (2, 4, 7, 9.5; 0.6) & (3, 6, 9, 9.5; 0.6) \\ (1.5, 2, 3, 5; 0.6) & (7, 8, 13, 17; 0.6) & (2.5, 9, 10, 11; 0.6) & (8, 10, 15, 17.5; 0.6) & (6, 8, 13, 18.5; 0.6) \end{bmatrix}$$



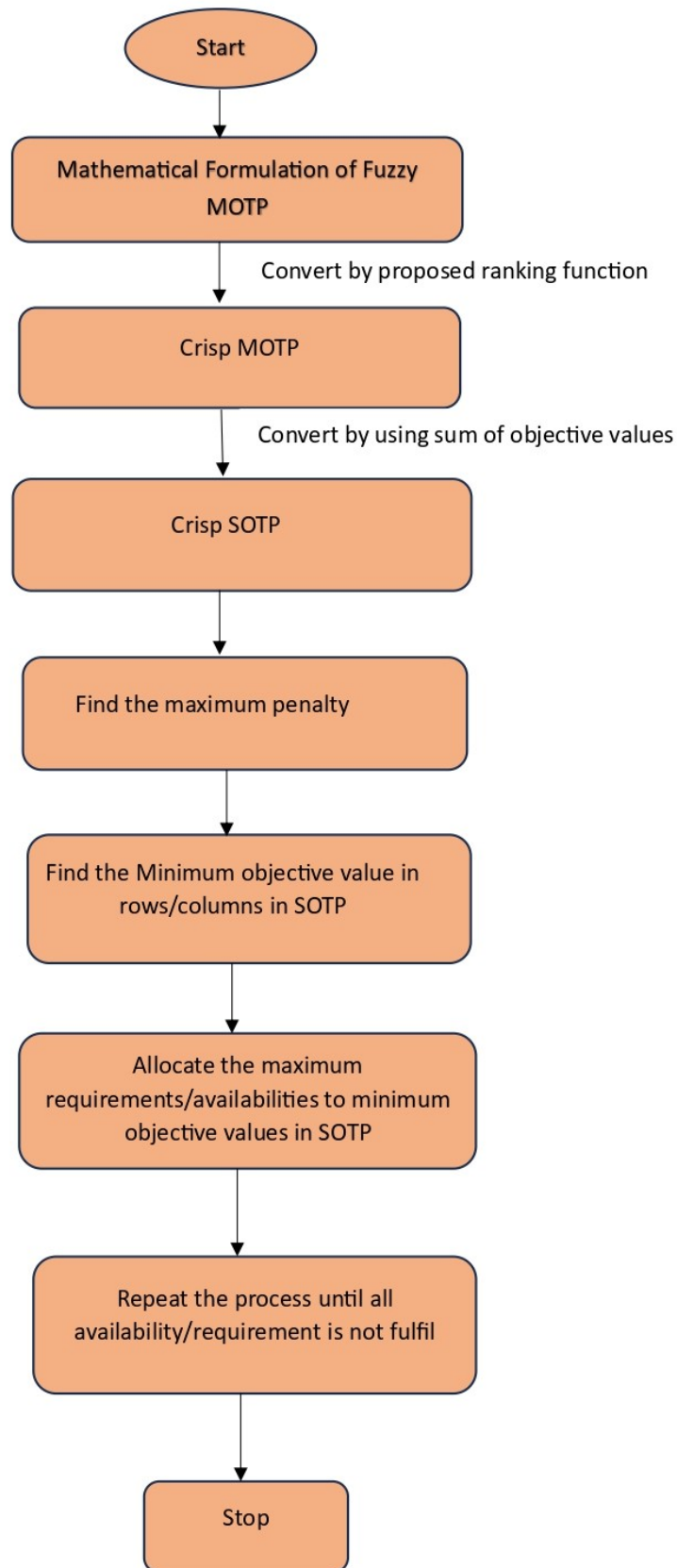


Figure 1: Flowchart for the Proposed Algorithm

Coe-efficient Matrix for the third objective function

$$C_3 = \begin{bmatrix} \tilde{a}_{11}^{(3)} & \tilde{a}_{12}^{(3)} & \tilde{a}_{13}^{(3)} & \tilde{a}_{14}^{(3)} & \tilde{a}_{15}^{(3)} \\ \tilde{a}_{21}^{(3)} & \tilde{a}_{22}^{(3)} & \tilde{a}_{23}^{(3)} & \tilde{a}_{24}^{(3)} & \tilde{a}_{25}^{(3)} \\ \tilde{a}_{31}^{(3)} & \tilde{a}_{32}^{(3)} & \tilde{a}_{33}^{(3)} & \tilde{a}_{34}^{(3)} & \tilde{a}_{35}^{(3)} \\ \tilde{a}_{41}^{(3)} & \tilde{a}_{42}^{(3)} & \tilde{a}_{43}^{(3)} & \tilde{a}_{44}^{(3)} & \tilde{a}_{45}^{(3)} \end{bmatrix}$$

$$= \begin{bmatrix} (0.5, 1, 2, 6; 1) & (2.5, 3, 4, 8; 1) & (4.5, 6, 10; 1) & (1, 2, 4, 5; 1) & (3, 4, 7, 11.5; 1) \\ (2, 2.5, 4.5, 8.5; 1) & (6, 7, 8, 12.5; 1) & (1, 2, 5, 9.5; 1) & (7.5, 8, 9, 13; 1) & (0.25, 1, 2, 6.25; 1) \\ (3.5, 4, 5, 9; 1) & (1, 2.5, 3.5, 5; 1) & (3, 4, 5, 9.5; 1) & (2, 2.5, 3.5, 4; 1) & (4, 4.5, 6.5, 10.5; 1) \\ (3.5, 5, 6, 11; 1) & (7, 7.5, 9.5, 13.5; 1) & (2.5, 4, 7, 12; 1) & (1.5, 2.5, 3.5, 4.5; 1) & (0.25, 0.5, 1.5, 1.75; 1) \end{bmatrix}$$

Availabilities:  $s_1 = 5, s_2 = 4, s_3 = 2, s_4 = 9$ .

Requirements:  $d_1 = 4, d_2 = 4, d_3 = 6, d_4 = 2, d_5 = 4$

Step I: In this step, the fuzzy MOTP is transformed into crisp MOTP using the Ranking function given above:

$$\tilde{a}_{11}^{(1)} = (9, 10, 11, 12; 0.8)$$

Here  $p_1 = 9, p_2 = 10, p_3 = 11, p_4 = 12, w = 0.8$

$$\Re(\tilde{a}_{11}^{(1)}) = \frac{2 \times 9 + 5 \times 0.8(10 + 11) + 2 \times 12}{14} = 9 = a_{11}^{(1)}$$

Similarly, all the fuzzy values  $\tilde{a}_{ij}^{(v)}$  for  $1 \leq j \leq 5, 1 \leq i \leq 4$  and  $1 \leq v \leq 3$ . can be converted in crisp values by using ranking function.

The crisp MOTP is represented in tabular form in table 4.

**Table 4:** Tabular representation of Crisp MOTP

Destination → source ↓	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	A <sub>4</sub>	A <sub>5</sub>	Availability {s <sub>i</sub> }
B <sub>1</sub>	9	12	9	6	9	5
	2	9	8	1	4	
	2	4	6	3	6	
B <sub>2</sub>	7	3	7	7	5	4
	1	9	9	5	2	
	4	8	4	9	2	
B <sub>3</sub>	6	5	9	11	3	2
	8	1	8	4	5	
	5	3	5	3	6	
B <sub>4</sub>	6	8	11	2	2	9
	2	8	6	9	8	
	6	9	6	3	1	
Requirement (d <sub>j</sub> )	4	4	6	2	4	

Step II: In this step, the sum of objectives value is calculated.

$$t_{11} = \sum_{v=1}^3 a_{11}^{(v)} = a_{11}^1 + a_{11}^2 + a_{11}^3 = 9 + 2 + 3 = 13$$

Similarly,

$$t_{12} = \sum_{v=1}^3 a_{12}^{(v)} = 25, t_{13} = \sum_{v=1}^3 a_{13}^{(v)} = 23, t_{14} = \sum_{v=1}^3 a_{14}^{(v)} = 10, t_{15} = \sum_{v=1}^3 a_{15}^{(v)} = 19$$

$$t_{21} = \sum_{v=1}^3 a_{21}^{(v)} = 12, t_{22} = \sum_{v=1}^3 a_{22}^{(v)} = 20, t_{23} = \sum_{v=1}^3 a_{23}^{(v)} = 20, t_{24} = \sum_{v=1}^3 a_{24}^{(v)} = 21, t_{25} = \sum_{v=1}^3 a_{25}^{(v)} = 9$$

$t_{31} = \sum_{v=1}^3 a_{31}^{(v)} = 18, t_{32} = \sum_{v=1}^3 a_{32}^{(v)} = 9, t_{33} = \sum_{v=1}^3 a_{33}^{(v)} = 22, t_{34} = \sum_{v=1}^3 a_{34}^{(v)} = 18, t_{35} = \sum_{v=1}^3 a_{35}^{(v)} = 14$   
 $t_{41} = \sum_{v=1}^3 a_{41}^{(v)} = 14, t_{42} = \sum_{v=1}^3 a_{42}^{(v)} = 25, t_{43} = \sum_{v=1}^3 a_{43}^{(v)} = 23, t_{44} = \sum_{v=1}^3 a_{44}^{(v)} = 14, t_{45} = \sum_{v=1}^3 a_{45}^{(v)} = 11$   
 Then, Crisp SOTP is given in Table 5.

**Table 5:** Tabular representation of Crisp SOTP

Destination → source ↓	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	A <sub>4</sub>	A <sub>5</sub>	Availability {s <sub>i</sub> }
B <sub>1</sub>	13	25	23	10	19	5
B <sub>2</sub>	12	20	20	21	9	4
B <sub>3</sub>	18	9	22	18	14	2
B <sub>4</sub>	14	25	23	14	11	9
Requirement (d <sub>j</sub> )	4	4	6	2	4	

Step III: the penalties for each Row and Columns is calculated as:

Rows Penalties:

$$\begin{aligned} \mu_1 &= [\text{maximum}(t_{1r}) - \text{minimum}(t_{1r})] \text{ for } 1 \leq r \leq 5 \\ &= [\text{maximum}(t_{11}, t_{12}, t_{13}, t_{14}, t_{15}) - \text{minimum}(t_{11}, t_{12}, t_{13}, t_{14}, t_{15})] \\ &= [\text{maximum}(13, 25, 23, 10, 19) - \text{minimum}(13, 25, 23, 10, 19)] \\ &= 25 - 10 = 15 \end{aligned}$$

Similarly, the remaining row penalties is

$$\mu_2 = 12, \mu_3 = 13, \mu_4 = 14$$

Columns Penalties:

$$\begin{aligned} \rho_1 &= [\text{maximum}(t_{s1}) - \text{minimum}(t_{s1})] \text{ for } 1 \leq s \leq 4 \\ &= [\text{maximum}(t_{11}, t_{21}, t_{31}, t_{41}) - \text{minimum}(t_{11}, t_{21}, t_{31}, t_{41})] \\ &= [\text{maximum}(13, 12, 18, 14) - \text{minimum}(13, 12, 18, 14)] \\ &= 18 - 12 = 6 \end{aligned}$$

Similarly, the remaining column penalties is

$$\rho_2 = 16, \rho_3 = 3, \rho_4 = 11, \rho_5 = 10$$

The tabular representation of SOTP with row and column penalties is given in Table 6.

**Table 6:** Tabular representation of Crisp SOTP With Penalties

Destination → source ↓	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	A <sub>4</sub>	A <sub>5</sub>	Availability {s <sub>i</sub> }	row penalties $\mu_i$
B <sub>1</sub>	13	25	23	10	19	5	15
B <sub>2</sub>	12	20	20	21	9	4	12
B <sub>3</sub>	18	9	22	18	14	2	13
B <sub>4</sub>	14	25	23	14	11	9	14
Requirement (d <sub>j</sub> )	4	4	6	2	4		
Column penalties (ρ <sub>j</sub> )	6	16	3	11	10		

Step IV: In this Step, the maximum penalty δ is calculated.

$$\begin{aligned} \delta &= \max \{ \mu_i, \rho_j; 1 \leq i \leq 4, 1 \leq j \leq 5 \} \\ &= \max \{ \mu_1, \mu_2, \mu_3, \mu_4, \rho_1, \rho_2, \rho_3, \rho_4, \rho_5 \} \\ &= \max \{ 15, 12, 13, 14, 6, 16, 3, 11, 10 \} \\ &= 16 = \rho_2 \end{aligned}$$

Here, Column A<sub>2</sub> has a maximum penalty.

Step V: In the table 6, in Column  $A_2$ , the cell  $a_{32}$  ( $= 9$ ) has the minimum objective value.

Step VI: Now, allocate  $\min(2,4) = 2$  to cell  $t_{32}$  and delete the Row  $B_3$  whose availability is fulfilled.

Step VII : Apply the same procedure from Step II to Step VIII for making the possible allocation in the remaining rows and columns, hence the 2<sup>nd</sup>, 3<sup>rd</sup>, 4<sup>th</sup>, 5<sup>th</sup>, 6<sup>th</sup>, 7<sup>th</sup>, and 8<sup>th</sup> allocations as 2,4,3,1,2,2,4 at cells  $t_{14}, t_{45}, t_{11}, t_{41}, t_{22}, t_{23}, t_{43}$  positions respectively. the optimum allocation of MOTP is given in Table 7.

**Table 7:** Final allocation table

Destination → source ↓	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$	Availability { $s_i$ }
$B_1$	13(3)	25	23	10(2)	19	5
$B_2$	12	20(2)	20(2)	21	9	4
$B_3$	18	9(2)	22	18	14	2
$B_4$	14	25	23(4)	14	11(4)	9
Requirement ( $d_j$ )	4	4	6	2	4	

## 9. RESULT ANALYSIS

In this section, the results obtained from the example are analyzed. The final Solution table for the example is shown in Table 8

**Table 8:** Final Solution table

Obtained Allocations	$x_{11} = 3, x_{12} = 0, x_{13} = 0, x_{14} = 2, x_{15} = 0$ $, x_{21} = 0, x_{22} = 2, x_{23} = 2, x_{24} = 0, x_{25} = 0, x_{31} = 0$ $, x_{32} = 2, x_{33} = 0, x_{34} = 0, x_{35} = 0, x_{41} = 1,$ $x_{42} = 0, x_{43} = 4, x_{44} = 0, x_{45} = 4$
Fuzzy Compromise efficient solution of MOTP	$[(71, 137, 170, 206; 0.8), (76, 123.5, 169.5, 212.5; 0.6),$ $(34, 53, 87, 148 : 1)]$
Crisp Compromise efficient solution of MOTP	$(127, 104, 76)$
Nature of Crisp Compromise Crisp Solution	Non-Degenerate

Physical Interpretation of the results: The obtained solution, as presented in Table 8, can be physically interpreted as follows:

(I) For First Objective Function: using the proposed algorithm the minimum fuzzy transportation cost (FTC) is  $[(71,137,170,206;0.8)]$ . It has the following physical interpretation:

(i) In the decision-maker's estimation, The minimum transportation cost (MTC) will be greater than Rs. 71 and less than Rs. 206 units.

(ii) The decision-maker is 80% satisfied overall with the statement that transportation costs will be 137-180.

(iii) The following values of the remaining minimum transportation cost can be used to determine the decision-maker's overall level of satisfaction:

If  $x$  is the MTC, then the overall decision-maker satisfaction level for  $x = \mu_{\hat{A}}(x) \times 100$

Where

$$\mu_{A(x)} = \begin{cases} 0.8\{1 - \frac{137-x}{66}\} & \text{if } 71 \leq x \leq 137 \\ 0.8 & \text{if } 137 \leq x \leq 170 \\ 0.8\{1 - \frac{x-170}{36}\} & \text{if } 170 \leq x \leq 206 \\ 0 & \text{Otherwise} \end{cases}$$

(II) For Second Objective Function: using the proposed algorithm the minimum FTP is [(76,123.5,169.5,212.5;0.6)] It has the following physical interpretation:

(i) The MTC, in the decision-maker's estimation, will be greater than Rs. 76 and less than Rs. 212.5 units.

(ii) The decision-maker is 60% satisfied overall with the statement that transportation costs will be 123.5-169.5.

(iii) The following values of the remaining minimum transportation cost can be used to determine the decision-makers overall level of satisfaction:

If  $x$  is the MTC, then the overall decision-maker satisfaction level for  $x = \mu_{\tilde{A}}(x) \times 100$

Where

$$\mu_{A(x)} = \begin{cases} 0.6\{1 - \frac{123.5-x}{47.5}\} & \text{if } 76 \leq x \leq 123.5 \\ 0.6 & \text{if } 123.5 \leq x \leq 169.5 \\ 0.6\{1 - \frac{x-169.5}{43}\} & \text{if } 169.5 \leq x \leq 212.5 \\ 0 & \text{Otherwise} \end{cases}$$

(III) For Third Objective Function: using the proposed algorithm the minimum FTC is [(34,53,87,148;1)]. It has the following physical interpretation:

(i)The MTC in the decision-makers estimation, will be greater than Rs. 34 and less than Rs. 148 units.

(ii) The decision-maker is 100% satisfied overall with the statement that transportation costs will be 53-87.

(iii) The following values of the remaining minimum transportation cost can be used to determine the decision-makers overall level of satisfaction:

If  $x$  is the MTC, then the overall decision-maker satisfaction level for  $x = \mu_{\tilde{A}}(x) \times 100$

Where

$$\mu_{A(x)} = \begin{cases} 1.0\{1 - \frac{53-x}{19}\} & \text{if } 34 \leq x \leq 53 \\ 1.0 & \text{if } 53 \leq x \leq 87 \\ 1.0\{1 - \frac{x-87}{61}\} & \text{if } 87 \leq x \leq 148 \\ 0 & \text{Otherwise} \end{cases}$$

## 10. CONCLUSION

This paper presents a new algorithm for solving the fuzzy multi-objective transportation problem (FMOTP) with objective function values as generalized trapezoidal fuzzy numbers, availabilities, and requirements are given as real numbers. The proposed algorithm first converts the fuzzy

MOTP into Crisp MOTP by ranking function after converting, the multi-objective MOTP is converted into Single objective TP. The proposed algorithm gives an efficient compromise solution and also provides a satisfaction level to the decision-maker in real-life situations. the proposed algorithm is less time-consuming and simple to use.

## REFERENCES

- [1] Hitchcock, F. L. The distribution of a product from several sources to numerous localities. *Journal of mathematics and physics*. 1941 Apr;20(1-4):224-30.
- [2] Charness A, Copper WW. The stepping stone method for explaining linear Programming calculation in transportation problem. *Management Science*. 1954; 1:49-69.
- [3] Zadeh, L. A. Fuzzy sets. *Information and control*. 1965; 8(3):338-353.
- [4] Zimmermann, H. J. Fuzzy programming and linear programming with several objective functions. *Fuzzy sets and systems*. 1978 Jan 1;1(1):45-55.
- [5] Chanas S., Kolodziejczyk, W., Machaj A. A fuzzy approach to the transportation problem. *Fuzzy sets and Systems*. 1984 Aug 1;13(3):211-21.
- [6] Ringuest, J. L. and Rinks, D. B. Interactive Solution for linear multi-objective transportation problem. *European Journal of Operational Research*. 1987 Oct; 32(1):96-106.
- [7] Bit, A. K, Biswal, M. P and Alam S. Fuzzy programming approach to multicriteria decision making transportation problem. *Fuzzy sets and systems*. 1992 Sep 10;50(2):135-141.
- [8] Saad, and Abass, S. A. A parametric study on transportation problem under fuzzy environment. 2002.
- [9] Gani, A. N and Razak, K. A. Two stage fuzzy transportation problem. *Journal of Physical Sciences*.2006;10:63-69.
- [10] Dinagar, D. S and Palanivel K. The transportation problem in fuzzy environment. *International journal of algorithms, computing and mathematics*. 2009 Aug;2(3):65-71.
- [11] Pandian P and Natarajan G. A new algorithm for finding a fuzzy optimal solution for fuzzy transportation problems. *Applied mathematical sciences*. 2010 May;4(2):79-90.
- [12] Kiruthiga, M. and Loganathan C. Fuzzy multi-objective linear programming problem using membership function. *International journal of science, engineering, and technology, applied sciences*. 2015;5(8):1171-8.
- [13] Maity. G, Roy, S. K, and Verdegay, J. L. Multi-objective transportation problem with cost reliability under uncertain environment. *International Journal of Computational Intelligence Systems*. 2016 Sep 2;9(5):839-49.
- [14] Gupta. S, Ali. I and Ahmed A. Multi-choice Multi-objective Capacitated transportation problem, a Case Study of uncertain demand and supply. *Journal of Statistics and Management Systems*. 2018 May 4;21(3):467-91.
- [15] Khalifa, A. E. A signed distance method for solving multi-objective transportation problems in fuzzy environment. *International journal of research in industrial engineering*. 2019 Sep 1;8(3):274-82.
- [16] El Sayed, M. A and Abo-Sinna, M. A. A novel approach for fully intuitionistic fuzzy multi-objective fractional transportation problem. *Alexandria Engineering Journal*. 2021 Feb 1;60(1):1447-63.
- [17] Ahmed, J. S, Mohammed, H. J ad Chaloob, I. Z. Application of a fuzzy multi-objective defuzzification method to solve a transportation problem. *Mater. Today Proc*. 2021.
- [18] Kamal M, Alarjani A, Haq A, Yusufi, F. N and Ali I. Multi-objective transportation problem under type-2 trapezoidal fuzzy numbers with parameters estimation and goodness of fit. *Transport*. 2021 Nov 24;36(4):317-38.
- [19] Khalifa HA, Kumar P, Alharbi MG. On characterizing solution for multi-objective fractional two-stage solid transportation problem under fuzzy environment. *Journal of intelligent systems*. 2021 Apr 21;30(1):620-35.

- [20] Bagheri M, Ebrahimnejad A, Razavyan S, Hosseinzadeh Lotfi F and Malekmohammadi N. Fuzzy arithmetic DEA approach for fuzzy multi-objective transportation problem. *Operational Research*. 2022 Apr 1:1-31.
- [21] Ma, Y. M, Hu, X. B and Zhou H. A deterministic and nature-inspired algorithm for the fuzzy multi-objective path optimization problem. *Complex and Intelligent Systems*. 2023 Feb;9(1):753-65.
- [22] Kacher Y, Singh P. A generalized parametric approach for solving different fuzzy parameter-based multi-objective transportation problems. *Soft Computing*. 2023 Oct 10:1-20.
- [23] Bodkhe, S. G. Multi-objective transportation problem using fuzzy programming techniques based on exponential membership functions. *International Journal of Statistics and Applied Mathematics* 2023; 8(5): 20-24

# USE OF THERMAL IMAGING METHOD OF CONTROL FOR INSPECTION OF BUILDING STRUCTURES FOR TIGHTNESS

Sofia Skachkova, Anton Avgutsevichs

•

Federal State Budgetary Institution "All-Russian Research Institute for Civil Defense and Emergencies of the Ministry of Emergency Situations of Russia" (Federal Center for Science and High Technologies)

sofiakovaleva55@yandex.ru, ministr.82@mail.ru

## Abstract

*In any industry related to the construction of buildings and structures we have heard about the need to assess the technical condition of various objects to assess and analyze the risks associated with the possible collapse of buildings (structures), loss of life and high costs to eliminate these consequences. Since many objects fail over time, and in general to determine the wear and tear and the possible term of further safe operation, it is necessary to conduct a technical survey. The article describes the principle of operation of thermal imaging devices for determining the reliability of building structures in residential premises, and also raises problems, the solution of which can simplify the use of thermal imaging devices in the inspection of buildings and structures and reduce the economic costs of damage compensation in case of timely detection and elimination of any defects.*

**Keywords:** thermal imager, thermal imaging device, reliability of building structures, risk analysis, building collapse, economic costs, inspection of buildings and structures, hidden defects, energy audit, temperature, energy, air leakage

## I. Introduction

In the Russian Federation, the construction industry is developing quite rapidly, and new buildings and structures are constantly being constructed. Companies that build and monitor the safe operation of various facilities are obliged to assess and determine the category of the technical condition of buildings and structures, which assesses the current state of the object under study, its operational properties, including the condition of the foundation soils, based on a comparison of the actual values of the parameters under study with the values of the same parameters established by the project or regulatory document.

## II. Methods

During the inspection and assessment of the technical condition of buildings and structures we mainly fix only visible defects, but quite often, many different damages in the building structures of buildings and structures are hidden and directly affect the load-bearing capacity of objects. For the fastest detection of any damage, especially hidden damage, where the integrity of structural



elements has been violated, namely cracks, cavities, etc. in closed building structures, thermal imaging methods of control of buildings and structures are used [1]. Such methods involve the use of devices called thermal imaging cameras. This device is based on non-contact and remote scanning of various structural elements of buildings and structures, and is used to visually determine and evaluate the thermal insulation characteristics of a building or room in real time [2]. Thermal imagers differ from each other by their technical characteristics, such as matrix parameters, ergonomics parameters, range of infrared radiation capture. Such devices are used in various spheres of human activity, for example, in medicine, veterinary medicine, engineering, housing and utilities, firefighting, etc. This article deals with the application of thermal imaging in the field of construction and examines the inspection of window glazing with the help of a thermal imaging device Testo 875 2i (Figure. 1).



**Figure 1:** General view of the device - Testo 875 2i thermal imager [1].

The thermal imaging method is used to observe the temperature distribution of the surface being inspected, which is displayed as a color picture where different colors correspond to different temperatures [2]. There is a temperature difference between the interior and exterior surfaces, and with the help of different colors on the thermal imager screen it is possible to identify defects where there is an increased heat loss due to the violation of thermal insulation [3].

Thermal imaging devices began to be used due to the fact that in 2009 came into force the federal law № 261 "On energy saving and energy efficiency" [4], and the subsequent introduction of thermal imaging devices in the field of energy saving and energy efficiency [4], and the subsequent order of the Ministry of Energy of the Russian Federation № 182 "On approval of requirements to the energy passport, compiled on the basis of the results of mandatory energy inspection" [5].

Thanks to the use of thermal imaging technology for inspection of the technical condition of structural elements of buildings and structures, it becomes possible to detect hidden defects at the early stages of their development, which makes it possible to prevent the transition of structural elements from "workable" technical condition to "limited-workable" or even "emergency" [6].

If we consider only the sphere of construction and architecture, it is possible with the help of thermal imaging devices:

- detect hidden defects in building structures [7];
- visualize energy loss [7];
- detect missing or deformed insulation layers;
- locate air leaks;

- detect mold, rotten or poorly insulated areas;
- identify temperature bridges;
- locate moisture penetration in flat roofs;
- find irregularities in piping and heating risers;
- find faults in electrical supply lines.

We have carried out energy audit of the installed double-glazed windows to determine their quality and tightness in the apartment of a residential building located in the Moscow region, Balashikha, Balashikha, Prospekt Lenina, 73 (Fig.2). Conducting an energy audit in the house is necessary to check the technical characteristics of the residential premises, its energy efficiency and cost-effectiveness. The main purpose of the survey is to identify the causes of high costs for electricity, gas, heat and water, as well as to determine possible ways to reduce energy consumption [8,9].



**Figure 2:** Schematic of the location of the object under study

The research was conducted using the following methods:

- method of comparison, comparison of the data of submitted documents and normative acts;
- analysis of normative and technical sources.
- carrying out thermal imaging inspection.

## Thermal inspection of window structures of a residential building



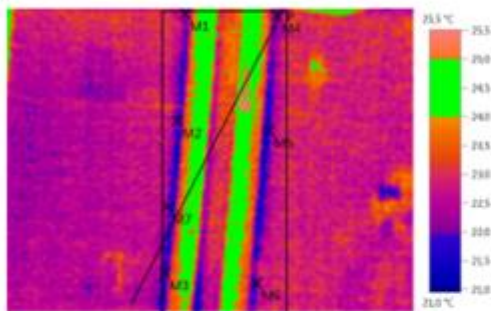
**Firm** Regional Center for Forensic Expertise  
 Balashikha, Prospekt Lenina, 73

**Customer** Pyatainina E.V.  
 Balashikha, Meshchery str. 4, kv.13

**Controller**

**Instrument** Testo 875-2i      **Serial №** 60342865

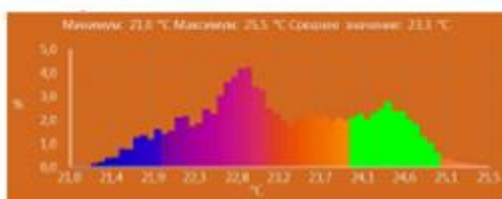
### Order



**Graphical data**      **Date** 09/07/2023      **Radiation coefficient** 0,95  
**Time** 8:37:56      **Reflective temperature [°C]** 22,0  
**File** IV-001.BMT

### Image selection

Measured objects	t, [°C]	Radiation	Reflective t, [°C]	Remarks
Measuring point 1	22,3	0,95	22	-
Measuring point 2	21,7	0,95	22	-
Measuring point 3	21,6	0,95	22	-
Measuring point 4	22,1	0,95	22	-
Measuring point 5	21,6	0,95	22	-
Measuring point 6	21,8	0,95	22	-
Measuring point 7	22,2	0,95	22	-



**Histogram**



**Profile line**

**Figure 3:** Analysis of the conducted thermal imaging inspection of the apartment window units

## Thermal inspection of window structures of a residential building



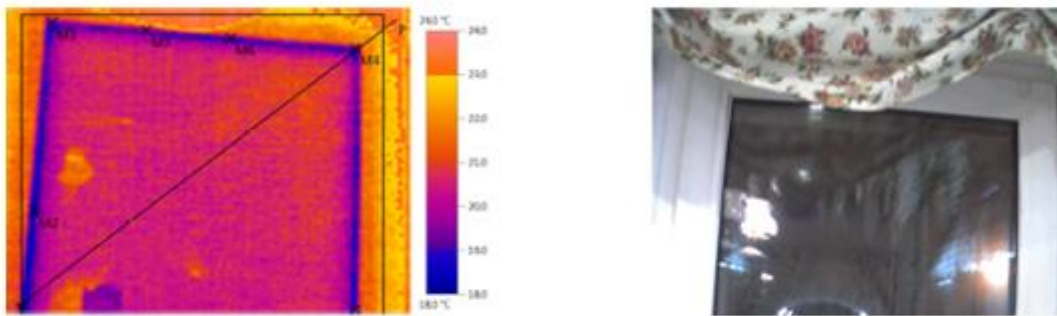
**Firm** Regional Center for Forensic Expertise  
 Balashikha, Prospekt Lenina, 73

**Customer** Pyatainina E.V.  
 Balashikha, Meshchery str. 4, kv.13

**Controller**

**Instrument** Testo 875-2i      **Serial №** 60342865

### Order

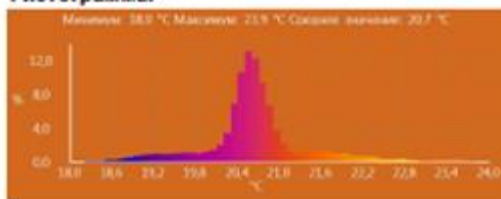


<b>Graphical data</b>	<b>Date</b> 09/07/2023	<b>Radiation coefficient</b> 0,95
	<b>Time</b> 8:38:32	<b>Reflective temperature [°C]</b> 22,0
	<b>File</b> IV-00110.BMT	

### Image selection

Measured objects	t, [°C]	Radiation	Reflective t, [°C]	Remarks
Measuring point 1	19,3	0,95	22	-
Measuring point 2	18,4	0,95	22	-
Measuring point 3	18,9	0,95	22	-
Measuring point 4	19,2	0,95	22	-
Measuring point 5	19,0	0,95	22	-
Measuring point 6	19,3	0,95	22	-
Measuring point 7	19,6	0,95	22	-

### Гистограмма:



**Histogram**

### Линия профиля:



**Profile line**

**Figure 4:** Analysis of the conducted thermal imaging inspection of the balcony door of the apartment

## Thermal inspection of window structures of a residential building



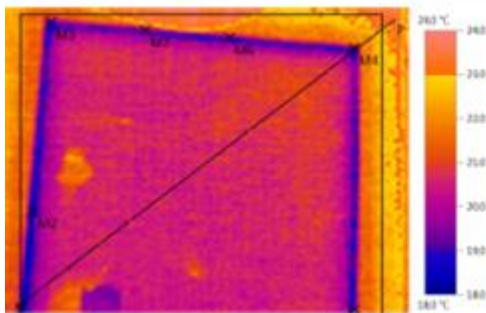
**Firm** Regional Center for Forensic Expertise  
Balashikha, Prospekt Lenina, 73

**Customer** Pyatainina E.V.  
Balashikha, Meshchery str. 4, kv.13

**Controller**

**Instrument** Testo 875-2i      **Serial №** 60342865

### Order



**Graphical data**      **09/07/2023**      **Radiation coefficient**      **0,95**  
**8:39:32**      **Reflective temperature [°C]**      **22,0**  
**IV-00114.BMT**

### Image selection

Measured objects	t, [°C]	Radiation	Reflective t, [°C]	Remarks
Measuring point 1	18,0	0,95	22	-
Measuring point 2	18,3	0,95	22	-
Measuring point 3	18,2	0,95	22	-
Measuring point 4	18,5	0,95	22	-
Measuring point 5	18,4	0,95	22	-
Measuring point 6	18,2	0,95	22	-

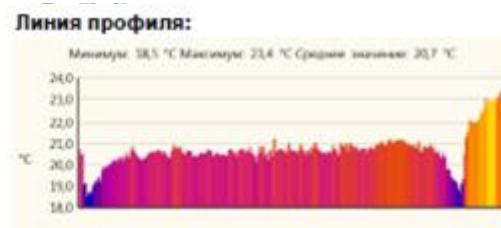
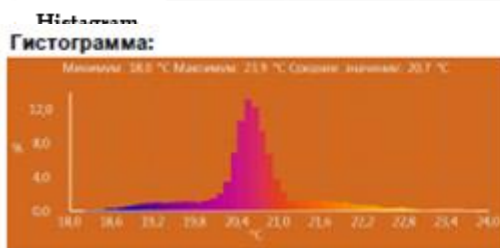


Figure 5: Analysis of the conducted thermal imaging inspection of the apartment window units

### III. Results

According to the results of the data obtained by taking thermal imaging and instrumental parameters of window structures, the following conclusions were obtained:

- the presence of voids (under the sill) was revealed;
- defects in filling the installation gaps during the construction of joints;
- thermal imaging inspection revealed air leakage both from the inside of the premises and from the outside;
- traces of condensation and mold formation on the slopes were revealed.

The reason of occurrence of available defects and shortcomings is low quality of manufacture of products and their installation. Translucent structures made of PVC profile installed in the apartment do not meet the basic requirement for windows of any design (tightness), it is not possible to use them for their intended purpose (operation) [10].

#### IV. Discussion

In conclusion, it can be concluded that non-destructive thermal imaging methods are an indispensable tool for locating hidden defects in structural elements of buildings and structures, because they can be effective even in areas inaccessible for other diagnostic methods, and they help to identify the causes of high costs for electricity, gas, heat and water, as well as to identify possible ways to reduce energy consumption.

#### References

- [1] Yeschenko D.V. Practical application of thermal imaging analysis and control methods / Yeschenko D.V., Nikitin A.T., Belov O.A. // Scientific Journal "Vestnik KamchatGTU", 2020, № 54.- S.6.- URL: <https://kamchatgtu.ru/wp-content/uploads/2022/12/Вестник-54.pdf>, free.
- [2] User Manual "Thermal Imager Testo-875 2i" // URL: <https://averus-pribor.ru/upload/instructions/testo-875i.pdf>.
- [3] GOST R 54852-2011 "Buildings and structures. Method of thermal imaging quality control of thermal insulation of building envelopes".
- [4] On Energy Saving and on Increasing Energy Efficiency, and on Amending Certain Legislative Acts of the Russian Federation: Federal Law of 23.11.2009 № 261-FZ [adopted by the State Duma on 11.11.2009] // Collection of Legislation of the Russian Federation. - 2009. № 48. St. 5711.
- [5] On approval of requirements to the energy passport, compiled as a result of mandatory energy inspection, and energy passport, compiled on the basis of project documentation, and rules for sending a copy of the energy passport, compiled as a result of mandatory energy inspection: Order of the Ministry of Energy of the Russian Federation of 19.04.2010 № 182. Federation from 19.04.2010 № 182.
- [6] GOST R 53778-2010 Buildings and structures. Rules of inspection and monitoring of technical condition. Moscow: Standardinform, 2008. 90 c.
- [7] Anikina, I.D. Thermal imaging inspection of buildings / Anikina, I.D., Porshnev, G.P., Sergeyev, V.V. // Scientific and Technical Bulletins of SPbSPU "Science and Education", 2012, № 2-1.-S.94.- URL: <https://cyberleninka.ru/article/n/teplovizionnoe-obsledovanie-zdaniy/viewer>, free.
- [8] Pro Gorod Samara // Samara News: every day. Internet portal. May 2023.12. URL.: [https://progorodsamara.ru/interesnoe/view/cto-takoe-energoaudit-doma#:~:text=Energy audit%20dome%20-%20this%20complex, fund%20i%20 solutions%20ecological%20problems/](https://progorodsamara.ru/interesnoe/view/cto-takoe-energoaudit-doma#:~:text=Energy%20audit%20dome%20-%20this%20complex,fund%20i%20solutions%20ecological%20problems/) (accessed: 20/10/2023).
- [9] Karpov D.F. Thermal imaging method for determining humidity fields of surfaces of building structures of buildings and structures // Scientific and theoretical journal "Bulletin of BSTU", 2019, No. 6.-P.28.- URL: <https://cyberleninka.ru/article/n/prakticheskoe-primeneniye-metodov-teplovizionnogo-analiza-i-kontrolya/viewer>, free.
- [10] GOST 23166-2021 "Window and balcony translucent enclosing structures. General technical conditions".

# ON THE CHARACTERIZATION AND APPLICATIONS OF A THREE-PARAMETER IMPROVED WEIBULL-WEIBULL DISTRIBUTION

A. S. MOHAMMED<sup>1</sup>, B. ABBA<sup>2,3\*</sup>, I. ABDULLAHI<sup>2</sup>, Y. ZAKARI<sup>1</sup>, A. I. ISHAQ<sup>1</sup>



<sup>1</sup>Department of Statistics, Ahmadu Bello University, Zaria.

<sup>2</sup>Department of Mathematics, Yusuf Maitama Sule University, Kano-Nigeria.

<sup>3</sup>School of Mathematics and Statistics, Central South University, Changsha, Hunan Province, China.

mohammedas@abu.edu.ng

badamasiabba@gmail.com

ibraabdul@googlemail.com

yzakari@abu.edu.ng

binishaq05@gmail.com

\*Corresponding author: Email: babba@yumsuk.edu.ng

## Abstract

*Parametric modeling of complex lifetime data characterized with nonmonotone hazard rate (NMHR) has in recent years attract the interest of many researchers and practitioners. The three-parameter improved Weibull-Weibull distribution introduced in 2022 has demonstrated a better NMHR modeling potential in the analysis of several failure times identified with bathtub hazard rate (BHR). In this study, we present the characterization, properties and two data sets' applications of the distribution. Various properties of the distribution obtained, include moment generating function, moments, skewness, kurtosis, and some types of entropy. Numerical results for mean, variance, skewness, and kurtosis are computed using simulation studies. Estimation of the distribution parameters is performed using the method of maximum likelihood, and the estimation method is assessed by Monte Carlo simulation experiments. The two illustrations further ascertain the capability of the model for modeling lifetime data from different scientific investigation areas.*

**Keywords:** Improved Weibull-Weibull distribution, non-monotone failure rate, characterization, maximum likelihood method, failure time data.

## 1. INTRODUCTION

Weibull distribution is one of the leading and widely used classical distributions. It has played a vital role in solving many problems in applied areas, such as reliability engineering, renewable energy, weather forecast, and biological studies analysis. For example, the Weibull model is used in modelling the failure time of devices [1], analysis of wind speed data to determine the wind power density [2], further more, [3] applied the distribution to describe soil particle-size and was recently used for fatigue life prediction of mechanical parts [4]. Because of its positively and negatively skewed density shapes, the distribution may be the first choice when modeling monotone hazard rates. One drawback with Weibull is its inability to accommodate non-monotone failure rates, such as the bathtub and unimodal failure rates [5]. For instance, the unimodal-shaped failure rate can be observed in the course of a successful surgery, where the patient is at high risk initially due to infection and other complications. The bathtub-shaped failure rate can be observed

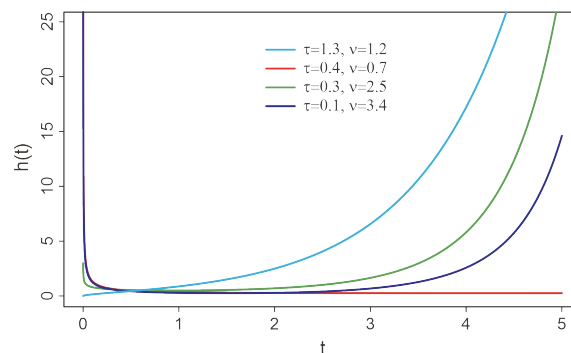
in the course of a population followed from birth or manufactured items with early failure due to faulty parts. Different new classes of distributions were developed based on modifications of the Weibull distribution to cope with the non-monotonic failure rates. Among others, are the exponentiated Weibull by [6], modified Weibull by [7], odd Weibull by [8], beta modified Weibull by [9], Weibull-Weibull by [10], Weibull-exponential by [11], odd generalized exponential-Weibull by [12], new Weibull-Weibull by [13], exponentiated additive Weibull distribution by [14], on monotonic and non-monotonic failure rates by [15], Four-Parameter Weibull distribution by [16] and flexible additive Chen-Gompertz by [17]. These distributions have one or more additional parameter(s) compared to Weibull distribution that makes them more flexible for modeling datasets with both monotone and non-monotone failure rates. The improved Weibull-Weibull (IWW3) distribution is a three-parameter model recently established by [18]. The distribution was proposed as an enhanced version of Weibull-Weibull (WW) distribution by [10] introduced in an earlier study to widen its applicability in modeling different complex lifetime data distinguished by various monotone and non-monotone hazard rate (HR) shapes. The IWW3 distribution is expressed by the following survival/reliability function as;

$$S(x) = \exp \left\{ - \int_0^{\infty} h(t) dx \right\} = \exp \left\{ - \left( e^{\zeta^\eta} - 1 \right)^\varphi \right\}, \quad t > 0, \quad (1)$$

where  $\zeta = \frac{x}{\tau}$ ,  $\eta > 0$  and  $\varphi > 0$  are the two shape parameters,  $\tau > 0$  is the scale parameter of the model, and  $h(x)$  represents the associated HR function (HRF) of the model, defined as

$$h(x) = \frac{\eta\varphi}{\tau} \zeta^{\eta-1} e^{\zeta^\eta} \left( e^{\zeta^\eta} - 1 \right)^{\varphi-1}. \quad (2)$$

The HRF of IWW3 distribution was characterized to have an increasing shape when  $\varphi, \eta \geq 1$ , and a decreasing pattern when  $\varphi, \eta < 1$ . Depending on the chosen values within the ranges of  $\eta > 1$  and  $\varphi\eta < 1$ , the HRF exhibits various bathtub curves. The HRF shapes are displayed in Figure 1, which visually explains the three main forms of the HFR shapes.



**Figure 1:** Curves describing various shapes of IWW3 hazard rate function at different values.

The probability density function (PDF) of the model is expressed as

$$f(x) = \frac{\eta\varphi}{\tau} \zeta^{\eta-1} e^{\zeta^\eta} \left( e^{\zeta^\eta} - 1 \right)^{\varphi-1} \exp \left[ - \left( e^{\zeta^\eta} - 1 \right)^\varphi \right]. \quad (3)$$

The distribution was demonstrated to provide a better fit in practice among several other two to five-parameter Weibull and non-Weibull distributions. More specifically, the censored failure and running times of 30 devices by [19] was identified to exhibit bathtub-shaped HR. The findings established the superiority of IWW3 distribution over some well-known Weibull extensions, including the exponentiated Weibull by [6], modified Weibull by [7], exponential Weibull by [20], alongside the WW [10] and other distributions. Motivated by the IWW3 distribution flexibility, and the original study by [18] only proposed the model, discussed its failure rate function and estimation methods. In this study, we present other important aspects of the



distribution, including the model's characterizations and some of its properties. Other real-life data applications of the distributions are also demonstrated.

The rest of the paper is arranged as follows. Section 2 discuss some of the properties and entropies of the distribution. The characterization of the distribution by two truncated moments and based on hazard function are given in section 3. The estimation of the proposed distribution parameters via the maximum likelihood method is presented in section 4. In sections 5 and 6, we assessed the estimators numerically by simulation studies and applications of the model to two lifetime data, respectively. We conclude the paper in section 7.

## 2. DISTRIBUTION PROPERTIES

Here, we discuss the moment generating function, moments, Rényi entropy and Mathai-Houbold entropy. We obtained an approximation for the values of the mean, variance, skewness and kurtosis of  $X$  using Monte Carlo simulation technique.

### 2.1. Moment generating function and moments

**Definition 1:** Let  $X$  be a random variable with the IWW3 density function (3). Then the moment generating function of  $X$  is given by

$$M_X(s) = \varphi \sum_{i \geq 0} \sum_{j \geq 0} \frac{(-1)^{i+j} \Gamma(\varphi(i+1))}{i!j! \Gamma(\varphi(i+1)-j)} \sum_{p \geq 0} \frac{\tau^p s^p \Gamma(1+p/\eta)}{p! (j - \varphi(i+1))^{1+\frac{p}{\eta}}} \quad (4)$$

The  $r^{th}$  moment about origin of  $X$  is generated from (4), and is defined as follows.

**Definition 2:** Let  $X$  be a random variable with the IWW3 density function (3). Then the  $r^{th}$  ( $r > 0$ , real) generalized ordinary moment of  $X$  is  $\mu_r = \int_{-\infty}^{\infty} x^r f(x) dx$ . For  $X \sim IWW3(\eta, \varphi, \tau)$ , the  $r^{th}$  moment from (4), is given as

$$\mu_r = E(X^r) = \sum_{i \geq 0} \sum_{j \geq 0} \frac{(-1)^{i+j} \Gamma(\varphi(i+1)) \Gamma(1+r/\eta) \tau^r \varphi}{i!j! (j - \varphi(i+1))^{1+r/\eta} \Gamma(\varphi(i+1)-j)}, \quad r \in N \quad (5)$$

In particular, the first four moments (for  $r = 1, \dots, 4$ ) can be used to calculate the mean ( $\mu_1$ ), variance ( $\sigma^2$ ), skewness ( $\sqrt{\beta_1}$ ) and kurtosis ( $\beta_2$ ) based on some well-known results.

The Monte Carlo simulation was performed for  $N = 1000$  samples each of size  $n = 200$  from the  $IWW3(\eta, \varphi, \tau)$  distribution, with  $\Psi = (\eta_0, \varphi_0, 1.5)^T$  - the vector of parameters, where  $\eta = 0.7, 0.9, 1.5, 2.0$  and  $3.0$ , and  $\varphi = 0.5, 1.0, 1.5, 2.0$  and  $3.0$ . Table 1 listed the numerical results for the mean, variance, skewness and kurtosis with their standard deviations (SDs) in parenthesis. We can notice from the Table, that the estimates of these properties varies for various combinations of the distribution parameters with a consistent decrease in the SDs for the mean and variance. The distribution shifted from right to left-skewed distribution when  $\eta\varphi \geq 2$ . The skewness and kurtosis plots are displayed in Figure 2 as functions of  $\eta$  and  $\varphi$ . A decrease in the values of the skewness and kurtosis are observed from the plots as values of the  $\eta$  and  $\varphi$  increases.

### 2.2. Rényi and Mathai-Houbold entropies

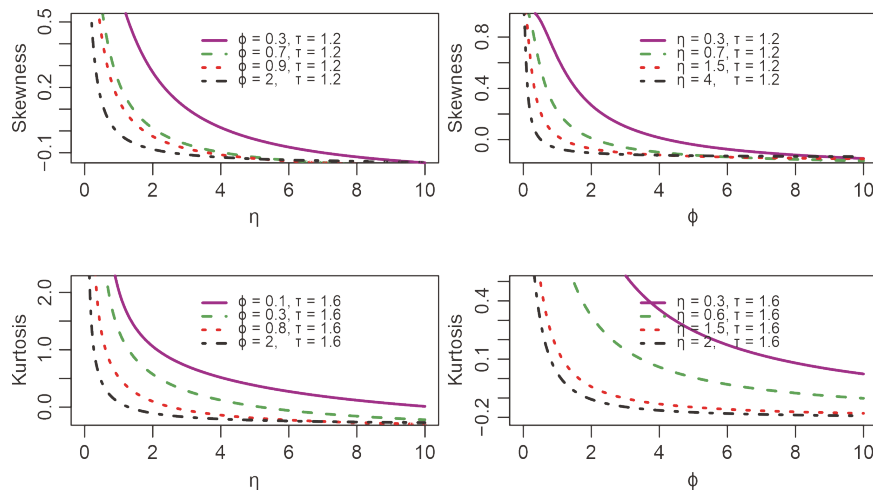
**Definition 3:** Let the random variable  $X$  have the density function given by (3). Then the Rényi entropy of  $X$  is given by

$$I_T(\alpha) = \frac{1}{1-\alpha} \log(M_{i,j}(\alpha, \Psi)) \quad (6)$$

where,  $M_{i,j}(\alpha, \Psi) = \frac{\eta^{\alpha-1} \varphi}{\tau^{\alpha-1}} \sum_{i \geq 0} \sum_{j \geq 0} \frac{(-1)^{i+j} \Gamma(\varphi(i+1) - (\alpha-1)) \Gamma((\eta-1)(\alpha-1)\eta^{-1} + 1)}{i!j! (j - \varphi(i+1))^{(\eta-1)(\alpha-1)\eta^{-1} + 1} \Gamma(\varphi(i+1) - (\alpha+j-1))}$ ,  $\alpha > 0$  and  $\alpha \neq 1$ .

**Table 1:** Mean, variance, skewness and kurtosis with standard deviations between parentheses;  $\tau = 1.5$  and some values of  $\eta$  and  $\phi$ .

$\eta$	$\phi$	Mean ( $\mu'_1$ )	Variance ( $\sigma^2$ )	Skewness( $\sqrt{\beta_1}$ )	Kurtosis( $\beta_2$ )
0.7	0.5	1.195 (0.128)	3.021 (0.635)	2.045 (0.408)	7.729 (3.268)
	1.0	0.830 (0.057)	0.609 (0.091)	1.277 (0.273)	4.563 (1.598)
	1.5	0.778 (0.038)	0.284 (0.033)	0.785 (0.197)	3.264 (0.833)
	2.0	0.772 (0.030)	0.173 (0.018)	0.466 (0.163)	2.784 (0.510)
	3.0	0.785 (0.021)	0.089 (0.008)	0.074 (0.143)	2.617 (0.275)
0.9	0.5	1.066 (0.094)	1.660 (0.263)	1.535 (0.257)	5.099 (1.506)
	1.0	0.875 (0.048)	0.450 (0.053)	0.870 (0.186)	3.295 (0.774)
	1.5	0.862 (0.034)	0.229 (0.022)	0.451 (0.151)	2.690 (0.439)
	2.0	0.871 (0.027)	0.144 (0.013)	0.177 (0.139)	2.560 (0.293)
	3.0	0.894 (0.019)	0.075 (0.007)	-0.160 (0.142)	2.711 (0.227)
1.5	0.5	1.007 (0.061)	0.710 (0.071)	0.780 (0.139)	2.736 (0.400)
	1.0	1.001 (0.036)	0.259 (0.022)	0.231 (0.122)	2.376 (0.231)
	1.5	1.030 (0.027)	0.140 (0.013)	-0.096 (0.127)	2.524 (0.190)
	2.0	1.054 (0.021)	0.088 (0.009)	-0.304 (0.139)	2.782 (0.234)
	3.0	1.086 (0.015)	0.045 (0.005)	-0.546 (0.164)	3.246 (0.372)
2.0	0.5	1.032 (0.050)	0.492 (0.041)	0.448 (0.116)	2.238 (0.217)
	1.0	1.077 (0.031)	0.188 (0.016)	-0.062 (0.116)	2.376 (0.162)
	1.5	1.116 (0.022)	0.101 (0.010)	-0.349 (0.135)	2.766 (0.232)
	2.0	1.142 (0.018)	0.063 (0.007)	-0.522 (0.154)	3.126 (0.334)
	3.0	1.173 (0.012)	0.031 (0.004)	-0.711 (0.182)	3.624 (0.500)
3.0	0.5	1.098 (0.039)	0.305 (0.023)	0.041 (0.104)	2.064 (0.117)
	1.0	1.178 (0.024)	0.115 (0.011)	-0.416 (0.127)	2.728 (0.231)
	1.5	1.219 (0.017)	0.059 (0.007)	-0.646 (0.158)	3.306 (0.392)
	2.0	1.243 (0.013)	0.036 (0.004)	-0.772 (0.180)	3.702 (0.521)
	3.0	1.269 (0.009)	0.017 (0.002)	-0.895 (0.206)	4.140 (0.683)



**Figure 2:** Plots of the Skewness and Kurtosis of the IWW3 as a function of  $\eta$  and  $\phi$ , respectively.

**Definition 4:** Let  $X$  have the PDF given by (3). Then the Mathai-Houbold (M-H) entropy of  $X$  is given by

$$J_{MH}(\delta) = \frac{\left(\frac{\eta}{\tau}\right)^{1-\delta} \varphi^{2-\delta} M_{i,j}^*(\delta, \Psi) - 1}{\delta - 1} \quad (7)$$

where  $e, M_{i,j}^*(\delta, \Psi) = \sum_{i \geq 0} \sum_{j \geq 0} \frac{(-1)^{i+j} (2-\delta)^i \Gamma(\varphi(2-\delta+i) - (1-\delta)) \Gamma((2-\delta) - (1-\delta)\eta^{-1})}{i! j! (j - \varphi(2-\delta+i))^{(2-\delta) - (1-\delta)\eta^{-1}} \Gamma(\varphi(2-\delta+i) - (1-\delta+j))}$ ,  $\delta \neq 1$  and  $\delta < 2$ .

We presented some numerical values for the Rényi and Mathai-Houbold entropies in Table 2. We note a decrease in the entropy as the values of either  $\eta$  or  $\varphi$  increases for the Rényi. While for the M-H, we observed that the entropy increases with an increase in  $\eta$  and decreases with an increase in  $\varphi$ . It can further be seen that entropy can take negative values, which may be understood as a loss of information in physical systems [21].

**Table 2:** Rényi and Mathai-Houbold entropies for some values of the parameters

$\eta \downarrow$	$\varphi$	$\tau$	$\alpha$	Rényi	$\varphi \downarrow$	$\eta$	$\tau$	$\alpha$	Rényi
0.1				2.9245	0.7				1.1692
0.2				1.2632	0.9				0.9486
0.4				0.8462	1.2				0.7302
0.6	1.1	1.2	0.6	0.7562	1.4	0.5	1.2	0.6	0.6229
0.7				0.7243	1.6				0.5330
0.8				0.6940	1.8				0.4549
0.85				0.6790	2.0				0.3853
	$\varphi$	$\tau$	$\delta$	M-H	$\eta$	$\tau$	$\delta$		M-H
0.58				0.3928	0.68				2.0974
0.60				0.6530	0.70				1.4760
0.62				0.8740	0.72				0.9609
0.64	0.7	1.2	0.6	1.0635	0.74	0.5	1.2	1.6	0.5311
0.66				1.2278	0.76				0.1700
0.68				1.3716	0.78				-0.1352
0.70				1.4986	0.80				-0.3946

### 3. CHARACTERIZATION

Characterization guides an investigator in designing a stochastic model for a particular modelling problem to know if the model fits the conditions of a specific underlying probability distribution. The investigator will depend on the characterization of the chosen distribution. The technique characterizes distribution and its random variable when the distribution conditions are similar to those of the random variable. This section presents the IWW3 distribution characterization in two directions (i) in terms of the simple relationship for the ratio of two truncated moments and (ii) based on the hazard function. We employed a theorem due to [22], for the first characterization (see Theorem 1). Note that the first characterization can be utilized even when the CDF's closed-form does not exist.

#### 3.1. Characterizations based on two truncated moments

**Theorem 1.** Let  $(\Omega, \Sigma, P)$  be a given probability space and let  $H = [a_1, a_2]$ , - be an interval for some  $a_1 < a_2$  ( $a_1 = -\infty, a_2 = \infty$  might as well be allowed). Let  $X : \Omega \rightarrow H$  be a continuous random variable with the distribution function  $F$  and let  $q_1$  and  $q_2$  be two real functions defined on  $H$  such that

$$E[q_2(X)|X \geq x] = E[q_1(X)|X \geq x] \eta(x), x \in H$$

is defined with some real function  $\eta$ . Assume that  $q_1, q_2 \in C^1(H), \eta \in C^2(H)$  and  $F$  is twice continuously differentiable and strictly monotone function on the set . Finally, assume that the equation  $\eta q_1 = q_2$  has no real solution in the interior of  $H$ . Then  $F$  is uniquely determined by the functions  $q_1, q_2$  and  $\eta$ , particularly

$$F(x) = \int_a^x c \left| \frac{\eta'(u)}{\eta(u)q_1(u) - q_2(u)} \right| \exp[-s(u)] du$$

where the function  $s$  is a solution of the differential equation  $s' = \frac{\eta' q_1}{\eta q_1 - q_2}$  and  $C$  is a constant chosen to make  $\int_H dF = 1$ .

**Proposition 1:** Let  $X : \Omega \rightarrow (0, \infty)$  be a continuous random variable and let  $q_1(x) = 1$  and  $q_2(x) = \exp[-\kappa^\varphi]$ , where  $\kappa = e^{(x/\tau)^\eta} - 1$  and  $x > 0$ . The random variable  $X$  has PDF (3), if and only if the function  $\eta(x)$  defined in Theorem 1 has the form

$$\eta(x) = \frac{1}{2} \exp[-\kappa^\varphi], \quad x > 0.$$

**Proof.** Let the random variable  $X$  has the PDF (3), then

$$(1 - F(x))E[q_1(X)|X \geq x] = \exp[-\kappa^\varphi], \quad x > 0,$$

and

$$(1 - F(x))E[q_2(X)|X \geq x] = \frac{1}{2} \exp[-\kappa^\varphi], \quad x > 0.$$

Further,

$$\eta(x)q_1(x) - q_2(x) = -\frac{1}{2} \exp[-\kappa^\varphi] < 0, \quad \text{for } x > 0$$

Conversely, if  $\eta$  is given as above, then

$$s'(x) = \frac{\eta'(x)q_1(x)}{\eta(x)q_1(x) - q_2(x)} = \frac{\eta\varphi}{\tau}(x/\tau)^{\eta-1}(\kappa+1)\kappa^{\varphi-1} = \frac{f(x)}{\bar{F}(x)}, \quad \text{for } x > 0$$

Therefore, according to Theorem 1,  $X$  has PDF (3).

**Corollary 1.** Let  $X : \Omega \rightarrow (0, \infty)$  be a continuous random variable and let  $q_1(x)$  be as in Proposition 1. The PDF of  $X$  is (3) if and only if there exist functions  $q_2(x)$  and  $\eta(x)$  defined in Theorem 1 satisfying the differential equation

$$\frac{\eta'(x)q_1(x)}{\eta(x)q_1(x) - q_2(x)} = \frac{\eta\varphi}{\tau}(x/\tau)^{\eta-1}(\kappa+1)\kappa^{\varphi-1}, \quad \text{where } \kappa = e^{(x/\tau)^\eta} - 1 \text{ and } x > 0$$

**Remark 1.** The general solution of the differential equation in Corollary 1 is

$$\eta(x) = \exp[-\kappa^\varphi] \left[ - \int_0^\infty \frac{\eta\varphi}{\tau}(x/\tau)^{\eta-1}(\kappa+1)\kappa^{\varphi-1} \exp[-\kappa^\varphi] (q_1)^{-1} q_2 dx + D \right]$$

where  $D$  is a constant. Note that one set of functions satisfying the differential equation is given in Proposition 1 with  $D = 0$ .

### 3.2. Characterization based on hazard function

The hazard function,  $h(x)$  of a twice differentiable distribution function,  $F(x)$ , satisfy the following first order differential equation

$$\frac{f'(x)}{f(x)} = \frac{h'(x)}{h(x)} - h(x) \tag{8}$$

**Proposition 2:** Let  $X : \Omega \rightarrow (0, \infty)$  be a continuous random variable. The random variable  $X$  has PDF (3), if and only if its hazard function  $h(x)$  satisfy the following differential equation

$$h'(x) - \left\{ \frac{\eta}{\tau}(x/\tau)^{(\eta-1)} + \frac{(\eta-1)}{\tau}(x/\tau)^{(-1)} \right\} h(x) = \frac{\eta^2(\varphi-1)}{\tau^2}(x/\tau)^{2(\eta-1)}(\kappa+1)^2 \times \kappa^{\varphi-2}$$

under the boundary conditions  $h(0) \geq 0$  and  $\kappa = e^{(x/\tau)^\eta} - 1$ .

**Proof.** If random variable  $X$  has the hazard function given in (2), then

$$h'(x) = \frac{\eta^2(\varphi-1)}{\tau^2}(x/\tau)^{2(\eta-1)}(\kappa+1)^2\kappa^{\varphi-2} + \frac{\eta^2\varphi}{\tau^2}(x/\tau)^{2(\eta-1)}(\kappa+1)\kappa^{\varphi-1} + \frac{\eta(\eta-1)\varphi}{\tau^2}(x/\tau)^{(\eta-2)}(\kappa+1)\kappa^{\varphi-1}$$

and hence,

$$\frac{h'(x)}{h(x)} - h(x) = \frac{\eta(\varphi - 1)}{\tau} (x/\tau)^{(\eta-1)} (\kappa + 1)\kappa^{-1} + \frac{\eta}{\tau} (x/\tau)^{(\eta-1)} + \frac{(\eta - 1)}{\tau} (x/\tau)^{(-1)} - \frac{\eta}{\tau} (x/\tau)^{(\eta-1)} (\kappa + 1)\kappa^{\varphi-1} \tag{9}$$

Similarly ,

$$f'(x) = f(x) \left\{ \frac{\eta(\varphi - 1)}{\tau} (x/\tau)^{(\eta-1)} (\kappa + 1)\kappa^{-1} + \frac{\eta}{\tau} (x/\tau)^{(\eta-1)} + \frac{(\eta - 1)}{\tau} (x/\tau)^{(-1)} \right\} - \frac{\eta\varphi}{\tau} (x/\tau)^{(\eta-1)} (\kappa + 1)\kappa^{\varphi-1} f(x)$$

and thus,

$$\frac{f'(x)}{f(x)} = \frac{\eta(\varphi - 1)}{\tau} (x/\tau)^{(\eta-1)} (\kappa + 1)\kappa^{-1} + \frac{\eta}{\tau} (x/\tau)^{(\eta-1)} + \frac{(\eta - 1)}{\tau} (x/\tau)^{(-1)} - \frac{\eta\varphi}{\tau} (x/\tau)^{(\eta-1)} (\kappa + 1)\kappa^{\varphi-1} \tag{10}$$

Equations (9) and (10) satisfied the differential equation (8), and hence, the IWW3 random variable X has the hazard d function (2).

#### 4. PARAMETER ESTIMATION

In this section, we use the method of maximum likelihood to estimate the unknown parameters of the distribution for complete dataset. Let  $x_1, x_2, \dots, x_n$  be a random sample of size  $n$  from the IWW3 model with the vector of parameters  $\Psi = (\eta, \varphi, \tau)$ . Then the log-likelihood function of  $\Psi$  from the PDF (3) is

$$\ell(\Psi) = n \log \left( \frac{\eta\varphi}{\tau} \right) + (\eta - 1) \sum_{i=1}^n \log(z_i) + \varphi \sum_{i=1}^n z_i^\eta + (\varphi - 1) \sum_{i=1}^n \log(1 - e^{-z_i^\eta}) - \sum_{i=1}^n (e^{z_i^\eta} - 1)^\varphi$$

where  $z_i = \frac{x_i}{\tau}$ . The estimate  $\hat{\Psi} = (\hat{\eta}, \hat{\varphi}, \hat{\tau})^T$  of  $\Psi = (\eta, \varphi, \tau)^T$  is determined by maximizing the log-likelihood function  $\ell(\Psi)$  with respect to each of the IWW3 parameters. Thus, we have following score functions.

$$\frac{\partial \ell(\Psi)}{\partial \eta} = \frac{n}{\eta} + \sum_{i=1}^n \log z_i - \varphi \sum_{i=1}^n (e^{z_i^\eta} - 1)^{\varphi-1} e^{z_i^\eta} z_i^\eta \log(\tau z_i) + \varphi \sum_{i=1}^n z_i^\eta \log(\tau z_i) + (\varphi - 1) \sum_{i=1}^n z_i^\eta \log(\tau z_i) (e^{z_i^\eta} - 1)^{-1} \tag{11}$$

$$\frac{\partial \ell(\Psi)}{\partial \varphi} = \frac{n}{\varphi} - \sum_{i=1}^n (e^{z_i^\eta} - 1)^\varphi \log(e^{z_i^\eta} - 1) + \sum_{i=1}^n z_i^\eta + \sum_{i=1}^n \log(1 - e^{-z_i^\eta}) \tag{12}$$

and

$$\frac{\partial \ell(\Psi)}{\partial \tau} = -\frac{\eta}{\tau} \left[ n + \varphi \sum_{i=1}^n z_i^\eta (1 - e^{-z_i^\eta} (e^{z_i^\eta} - 1)^{\varphi-1}) - (\varphi - 1) \sum_{i=1}^n z_i^\eta (e^{z_i^\eta} - 1)^{-1} \right] \tag{13}$$

Solving equations (11)-(13) analytically may be intractable. Thus, a numerical approach is adopted to obtain the maximum likelihood estimates (MLEs) of the parameters  $\Psi = (\eta, \varphi, \tau)^T$  with a good set of initial values using R statistical package. To obtain the asymptotic interval estimation of

$\Psi = (\eta, \varphi, \tau)$ , we determine the observed Fisher information matrix. The  $3 \times 3$  Fisher information matrix is

$$J(\Psi) = - \begin{pmatrix} J_{\eta\eta}(\Psi) & J_{\eta\varphi}(\Psi) & J_{\eta\tau}(\Psi) \\ J_{\varphi\eta}(\Psi) & J_{\varphi\varphi}(\Psi) & J_{\varphi\tau}(\Psi) \\ J_{\tau\eta}(\Psi) & J_{\tau\varphi}(\Psi) & J_{\tau\tau}(\Psi) \end{pmatrix}$$

where the expressions of  $J_{\Psi_i, \Psi_j} = \frac{\partial^2 \ell(\Psi)}{\partial \Psi_i \partial \Psi_j}$ ,  $i, j = 1, 2, 3$ . Thus, the approximate variance of  $\Psi = (\eta, \varphi, \tau)$  can be obtained as

$$J^{-1}(\hat{\Psi}) = \begin{pmatrix} var(\hat{\eta}) & cov(\hat{\eta}, \hat{\varphi}) & cov(\hat{\eta}, \hat{\tau}) \\ cov(\hat{\varphi}, \hat{\eta}) & var(\hat{\varphi}) & cov(\hat{\varphi}, \hat{\tau}) \\ cov(\hat{\tau}, \hat{\eta}) & cov(\hat{\tau}, \hat{\varphi}) & var(\hat{\tau}) \end{pmatrix}$$

Hence, the  $100(1 - \alpha)\%$  asymptotic confidence intervals of  $\Psi = (\eta, \varphi, \tau)$  are given by  $\hat{\eta} \pm Z_{\frac{\alpha}{2}} \sqrt{var(\hat{\eta})}$ ,  $\hat{\varphi} \pm Z_{\frac{\alpha}{2}} \sqrt{var(\hat{\varphi})}$ , and  $\hat{\tau} \pm Z_{\frac{\alpha}{2}} \sqrt{var(\hat{\tau})}$ .

where  $Z_{\alpha}$  is the upper  $\alpha^{th}$  percentile of the standard normal distribution.

## 5. SIMULATION RESULTS

The main objective in this section is to evaluate the performance of the maximum likelihood method for estimating the IWW3 distribution parameters for a complete dataset via Monte Carlo simulation. For this purpose, we used six different combinations of the distribution parameters, including (1.8, 0.5, 0.5), (1.8, 0.7, 0.5), (1.8, 0.9, 0.5), (2, 0.5, 0.5), (2, 0.7, 0.5), and (2, 0.9, 0.5). The process is repeated 1000 times for four sample sizes  $n = 100, 150, 200,$  and  $300$ . Table 3 presents the MLEs, Biases, and Mean Square Errors (MSEs) of the parameters. Based on the results, we observe that the ML method performs well for estimating the distribution parameters. Also, as the sample size increases, the biases and the MSEs of the MLEs decrease as expected.

## 6. APPLICATIONS

In this section, we analyze two different datasets to assess the potentiality of the IWW3 distribution in practice. The datasets are Aarset data [23] and Meeka and Escoba data [19]. Both the two datasets have a bathtub-shaped hazard rate. We compared the results of the IWW3( $\Psi$ ) with Weibull and other Weibull extended models, including the exponentiated Weibull (EW) by [6], Weibull-Weibull(WW) by [10], Weibull-exponential (WE) by [11] and new Weibull-Weibull (NW-W) by [13]. To accomplish the purpose, we manage the maximum  $\ell(\hat{\Psi})$ , Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC) and Bayesian Information Criterion (BIC). The Kolmogorov-Smirnov (K-S) test is used to measure the closeness between the empirical and the fitted distribution. Generally, the smaller the value of these statistics, the better the model fit the dataset. All computations were done using RStudio 1.2.5042 software.

### 6.1. Aarset data

Here, we employed Aarset data [23], which is considered by many authors, such as [6], as standard data for assessing distributions with bathtub-shaped FR. It represents the failure times of fifty components placed on life test at time zero. The data revealed a bathtub-shaped FR, as shown by the TTT-plot in Figure 3. Table 4 presents the MLEs of the parameters of IWW3( $\Psi$ ) together with that of EW, WE, W, WW, and NW-W for the Aarset data. From Table 5, the IWW3( $\Psi$ ) has the smallest  $-\ell(\hat{\Psi})$ , AIC, CAIC, and BIC values, thus, the IWW3 model provides a best fit for the Aarset data. For the non-parametric goodness-of-fit statistics, the IWW3 model has the smallest K-S value with the highest p-value, which suggests that the IWW3 model has a better fit for the data set than the other competing models.

**Table 3:** MLEs, Biases, and MSEs of the distribution parameters.

<i>n</i>	$\eta$	$\varphi$	Estimates			Biases			MSEs		
			$\hat{\eta}$	$\hat{\varphi}$	$\hat{\tau}$	$\hat{\eta}$	$\hat{\varphi}$	$\hat{\tau}$	$\hat{\eta}$	$\hat{\varphi}$	$\hat{\tau}$
<b>100</b>	1.8	0.5	1.9714	0.5414	0.5263	0.1714	0.0414	0.0263	0.4260	0.1539	0.1529
	1.8	0.7	2.0089	0.9619	0.6985	0.2089	0.2619	0.1985	0.7313	2.2978	2.2686
	1.8	0.9	2.0485	1.5201	0.8309	0.2485	0.6201	0.3309	1.1138	3.2315	2.9566
	2.0	0.5	2.1910	0.5405	0.5189	0.1910	0.0405	0.0189	0.5269	0.0671	0.0658
	2.0	0.7	2.2326	0.9666	0.6548	0.2326	0.2666	0.1548	0.9076	1.3555	1.3084
	2.0	0.9	2.2756	1.5279	0.7601	0.2756	0.6279	0.2601	1.3719	2.1311	1.8045
<b>150</b>	1.8	0.5	1.9050	0.5146	0.5039	0.1050	0.0146	0.0039	0.2569	0.0013	0.0011
	1.8	0.7	1.9312	0.7964	0.5378	0.1312	0.0964	0.0378	0.4453	0.2475	0.2397
	1.8	0.9	1.9570	1.1864	0.6115	0.1570	0.2864	0.1115	0.6846	0.7578	0.6882
	2.0	0.5	2.1184	0.5147	0.5037	0.1184	0.0147	0.0036	0.3217	0.0012	0.0010
	2.0	0.7	2.1458	0.7968	0.5304	0.1458	0.0968	0.0304	0.5492	0.1516	0.1431
	2.0	0.9	2.1742	1.1878	0.5875	0.1742	0.2878	0.0875	0.8443	0.4836	0.4085
<b>200</b>	1.8	0.5	1.8735	0.5085	0.5027	0.0735	0.0085	0.0027	0.1610	0.0007	0.0006
	1.8	0.7	1.8896	0.7463	0.5112	0.0896	0.0463	0.0112	0.2781	0.0263	0.0243
	1.8	0.9	1.9039	1.0521	0.5388	0.1039	0.1521	0.0388	0.4310	0.1530	0.1314
	2.0	0.5	2.0825	0.5083	0.5024	0.0825	0.0083	0.0024	0.1995	0.0006	0.0005
	2.0	0.7	2.1005	0.7456	0.5090	0.1005	0.0456	0.0090	0.3430	0.0145	0.0125
	2.0	0.9	2.1154	1.0523	0.5315	0.1154	0.1523	0.0315	0.5321	0.1015	0.0793
<b>300</b>	1.8	0.5	1.8509	0.5050	0.5018	0.0509	0.0050	0.0018	0.1077	0.0004	0.0004
	1.8	0.7	1.8621	0.7222	0.5037	0.0621	0.0222	0.0037	0.1887	0.0011	0.0006
	1.8	0.9	1.8712	0.9718	0.5136	0.0712	0.0718	0.0136	0.2973	0.0362	0.0313
	2.0	0.5	2.0558	0.5053	0.5016	0.0558	0.0053	0.0016	0.1321	0.0003	0.0003
	2.0	0.7	2.0688	0.7222	0.5033	0.0688	0.0222	0.0033	0.2327	0.0010	0.0005
	2.0	0.9	2.0790	0.9729	0.5126	0.0790	0.0729	0.0126	0.3669	0.0366	0.0315

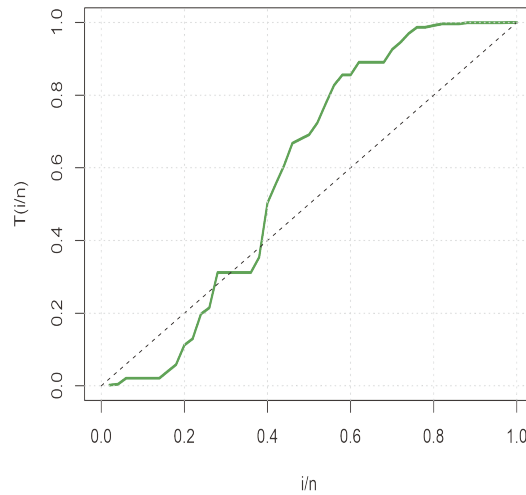
Figure 4 presents the plots of the fitted PDFs (see Figure 4, Fitted PDFs) and the estimated CDFs (see Figure 4, estimated CDFs), which equally illustrate that the IWW3 model has fitted the data well compared to the other competing models. Moreover, Figure 4 (estimated hazard rate function) has indicated that the hazard rate function is bathtub shaped, and hence, has ascertained the actual behavior of the data.

**Table 4:** MLEs and their standard errors (in parentheses) for the models fitted to the Aarset data.

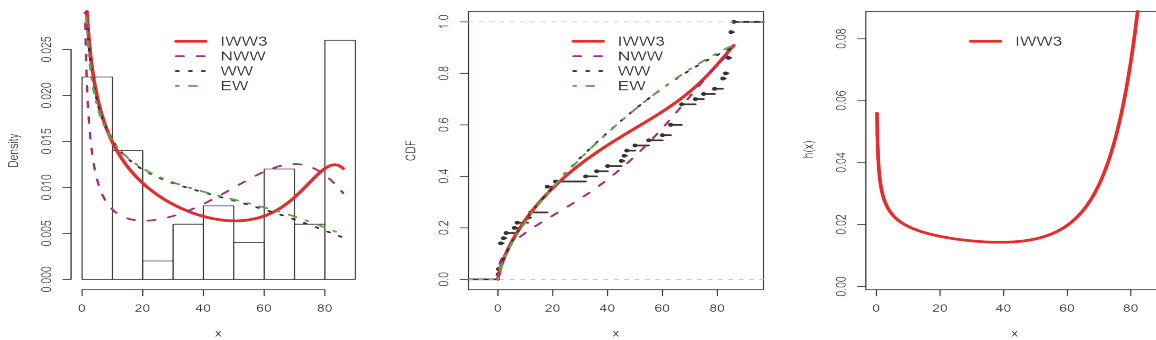
Models	$\hat{\eta}$	$\hat{\varphi}$	$\hat{\tau}$	$\hat{\theta}$	$\hat{a}$
EW	0.0109 (0.0009)	4.6713 (0.0246)	0.1450 (0.0217)		
WE	0.1742 (0.0621)	0.3851 (0.1029)	0.0778 (0.0257)		
W	0.9488 (0.1196)	44.847 (6.9313)			
WW	0.2705 (0.0700)	0.2558 (0.0584)	0.0066 (0.0059)	1.6026	
NW-W	0.4929 (0.0893)	0.0021 (0.0004)	1.4560 (0.0690)	(0.2263)	
IWW3	5.4238 (0.0040)	0.1363 (0.0230)	61.067 (3.6212)		

**Table 5:** The values of  $\ell(\hat{\Psi})$ , AIC, CAIC, BIC, K-S(with its p-value) statistics for the models fitted to the Aarset data.

Models	$-\ell(\hat{\Psi})$	AIC	CAIC	BIC	K-S	p-value
EW	229.136	464.272	464.7937	470.008	0.2057	0.0291
WE	225.6185	457.2371	457.7588	462.9731	0.1289	0.3774
W	241.0019	486.0037	486.259	489.8278	0.1933	0.0477
WW	220.902	449.804	450.6929	457.4521	0.1308	0.3598
NW-W	230.2974	466.5948	467.1166	472.3309	0.2040	0.0312
IWW3	218.3491	442.6982	443.2199	448.4342	0.1192	0.4762



**Figure 3:** TTT plot for Aarset data



**Figure 4:** Fitted PDFs (left panel), estimated CDFs (center panel) and estimated hazard rate function (right panel) for some of the fitted models to Aarset data.

## 6.2. Meeker and Escobar data

The second data used is [19] data, which represents the failure and running times of  $n = 30$  devices. It has been analyzed by many authors, including [24]. The data revealed a bathtub-shaped FR, as shown by the TTT-plot in Figure 5. Table 6 lists the MLEs of the parameters of IWW3( $\Psi$ ) together with that of EW, WE, W, WW, and NW-W for the data. From Table 7, it is noted that the IWW3( $\Psi$ ) has the smallest  $-\ell(\hat{\Psi})$ , AIC, CAIC, and BIC values, thus, the IWW3 model provides a better fit for the data. For the formal non-parametric goodness-of-fit statistic, the IWW3 model has the smallest value for K-S, with the highest p-value, which also ascertains the IWW3 model well fits the Meeker and Escobar data.

Figure 6 presents the plots of the fitted PDFs (see Figure 6, Fitted PDFs) and the estimated CDFs (see Figure 6, estimated CDFs), which illustrate that the IWW3 model has fitted the data well



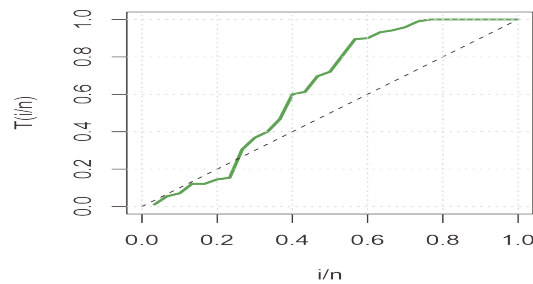
compare to the other competing models. Additionally, Figure 6 (estimated hazard rate function) has indicated that the hazard rate function is bathtub shaped, and hence, has ascertained the actual behavior of the data.

**Table 6:** MLEs and their standard errors (in parentheses) for the models fitted to the Meeker and Escobar data.

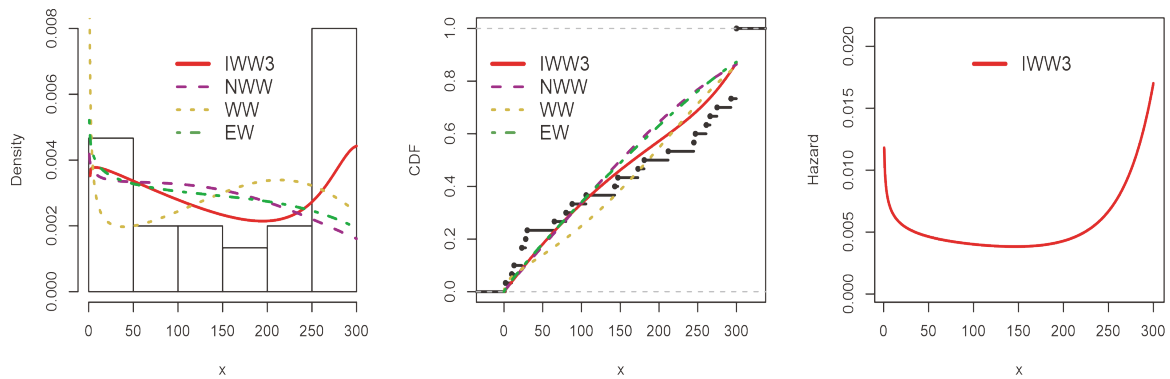
Models	$\hat{\eta}$	$\hat{\phi}$	$\hat{\tau}$	$\hat{\theta}$	$\hat{a}$
EW	0.0030 (0.0002)	5.5320 (0.2077)	0.1620 (0.0319)		0.5886 (0.1375)
WE	0.1258 (0.0669)	0.4976 (0.2022)	0.0187 (0.0093)		
W	1.2550 (0.2043)	180.652 (26.8740)			
WW	0.0860 (0.0409)	0.4936 (0.1760)	0.0539 (0.0304)	0.8357 (0.1147)	
NW-W	0.8785 (0.1653)	0.0020 (0.0003)	1.0669 (0.0371)		
IWW3	6.9601 (0.0085)	0.1479 (0.0297)	239.38 (14.149)		

**Table 7:** The values of  $\ell(\hat{\Psi})$ , AIC, CAIC, BIC, K-S(with its p-value) statistics for the models fitted to the Meeker and Escobar data.

Models	$-\ell(\hat{\Psi})$	AIC	CAIC	BIC	K-S	p-value
EW	177.9146	361.8293	362.7524	366.0329	0.2159	0.1219
WE	177.1573	360.3145	361.2376	364.5181	0.1724	0.3347
W	184.35	372.6999	373.1443	375.5023	0.2358	0.0712
WW	178.0157	364.0313	365.6313	369.6361	0.1650	0.3873
NW-W	180.3075	366.6149	367.538	370.8185	0.2234	0.1001
IWW3	170.804	347.608	348.5311	351.8116	0.1425	0.576



**Figure 5:** TTT plot for Meeker and Escobar data.



**Figure 6:** Fitted PDFs (left panel), estimated CDFs (center panel) and estimated hazard rate function (right panel) for some of the fitted models to Meeker and Escoba data.

## 7. CONCLUSION

This paper defined and studied a new generalized Weibull distribution called the Improved Weibull-Weibull (IWW3) distribution. It is a three-parameter flexible distribution with the ability to accommodate monotone and non-monotone failure rates lifetime data. We obtain explicit expressions for the moment generating function, moments, quantile function, Rényi entropy, and Mathai-Houbold entropy. Numerical results for median, Rényi entropy, Mathai-Houbold entropy and conduct a Monte Carlo simulation study to obtain some numerical results for the mean, variance, skewness, and kurtosis of the distribution. We also characterize the IWW3 model based on two truncated moments and in terms of the hazard function. Estimation of the distribution parameters is performed using the method of maximum likelihood, and the estimation method is assessed by Monte Carlo simulation experiments which yield consistent estimates in the samples considered. Two failure time data having non-monotone failure rate functions are analyzed to demonstrate the potentiality of the distribution.

### Disclosure statement

On behalf of The authors, I declare that no potential conflict of interest was reported.

### Funding

No funding was provided for the work.

## REFERENCES

- [1] Jiang, H., Xie, M., & Tang, L. C. (2008). On the odd Weibull distribution. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 222(4), 583-594.
- [2] Soulouknga, M. H., Doka, S. Y., Revanna, N., Djongy ang, N., & Kofane, T. C. (2018). Analysis of wind speed data and wind energy potential in Faya-Largeau, Chad, using Weibull distribution. *Renewable energy*, 121, 1-8.
- [3] Bayat, H., Rastgo, M., Zadeh, M. M., & Vereecken, H. (2015). Particle size distribution models, their characteristics and fitting capability. *Journal of hydrology*, 529, 872-889.
- [4] Wang, Y., & Peng, Z. (2020). Fatigue life prediction method of mechanical parts based on Weibull distribution. *In IOP Conference Series: Materials Science and Engineering*, 782 (2) p. 022068.
- [5] Dey, S., Sharma, V. K., & Mesfioui, M. (2017). A new extension of Weibull distribution with application to lifetime data. *Annals of Data Science*, 4, 31-61.
- [6] Mudholkar, G. S., & Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE transactions on reliability*, 42(2), 299-302.
- [7] Lai, C. D., Xie, M., & Murthy, D. N. P. (2003). A modified Weibull distribution. *IEEE Transactions on reliability*, 52(1), 33-37.

- [8] Cooray, K. (2006). Generalization of the Weibull distribution: the odd Weibull family. *Statistical Modelling*, 6(3), 265-277.
- [9] Silva, G. O., Ortega, E. M., & Cordeiro, G. M. (2010). The beta modified Weibull distribution. *Lifetime data analysis*, 16, 409-430.
- [10] Bourguignon, M., Silva, R. B., & Cordeiro, G. M. (2014). The Weibull-G family of probability distributions. *Journal of data science*, 12(1), 53-68.
- [11] Oguntunde, P. E., Balogun, O. S., Okagbue, H. I., & Bishop, S. A. (2015). The Weibull-exponential distribution: Its properties and applications. *Journal of Applied Sciences*, 15(11), 1305-1311.
- [12] Tahir, M. H., Cordeiro, G. M., Alizadeh, M., Mansoor, M., Zubair, M., & Hamedani, G. G. (2015). The odd generalized exponential family of distributions with applications. *Journal of Statistical Distributions and Applications*, 2, 1-28.
- [13] 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, 1-18.
- [14] Abd EL-Baset, A. A., & Ghazal, M. G. M. (2020). Exponentiated additive Weibull distribution. *Reliability Engineering & System Safety*, 193, 106663.
- [15] Mohammed, A. S., & Ugwuowo, F. I. (2021). On Transmuted Exponential-Topp Leon Distribution with Monotonic and Non-Monotonic Hazard Rates and its Applications. *Reliability: Theory & Applications*, 16(4 (65)), 197-209.
- [16] Mohammed, A. S., & Ugwuowo, F. I. (2020). A new Four-Parameter Weibull Distribution with Application to Failure Time Data. *FUDMA Journal of Sciences*, 4(3), 563-575.
- [17] Abba, B., Wang, H., & Bakouch, H. S. (2022). A reliability and survival model for one and two failure modes system with applications to complete and censored datasets. *Reliability Engineering & System Safety*, 223, 108460.
- [18] Wang, H., Abba, B., & Pan, J. (2022). Classical and Bayesian estimations of improved Weibull-Weibull distribution for complete and censored failure times data. *Applied Stochastic Models in Business and Industry*, 38(6), 997-1018.
- [19] Meeker, W. Q., Escobar, L. A., & Pascual, F. G. (2022). Statistical methods for reliability data. *John Wiley & Sons*.
- [20] Cordeiro, G. M., Ortega, E. M., & Lemonte, A. J. (2014). The exponential-Weibull lifetime distribution. *Journal of Statistical Computation and simulation*, 84(12), 2592-2606.
- [21] Bakouch, H. S., & Abd El-Bar, A. M. (2017). The exponential-Weibull lifetime distribution. *Journal of Statistical Computation and simulation*, 84(12), 2592-2606.
- [22] Glanzel, 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.
- [23] Aarset, M. V. (1987). How to identify a bathtub hazard rate. *IEEE transactions on reliability*, 36(1), 106-108.
- [24] Sarhan, A. M., & Apaloo, J. (2013). Exponentiated modified Weibull extension distribution. *Reliability Engineering & System Safety*, 112, 137-144.

# MARSHAL-OLKIN ALPHA POWER INVERSE RAYLEIGH DISTRIBUTION: PROPERTIES, ESTIMATION AND APPLICATIONS

Ismaila Olawale Adegbite

Department of Statistics, Osun State Polytechnic, Iree, Nigeria

[adegbiteonline@gmail.com](mailto:adegbiteonline@gmail.com)

•

Kayode Samuel Adekeye

Department of Mathematics and Statistics, Redeemer's University, Ede, Nigeria

[adekeyek@run.edu.ng](mailto:adekeyek@run.edu.ng)

•

Olubisi Lawrence Aako

Department of Mathematics and Statistics, Federal Polytechnic, Ilaro, Nigeria

[olubisi.aako@federalpolyilaro.edu.ng](mailto:olubisi.aako@federalpolyilaro.edu.ng)

## Abstract

*In this study, a new three-parameter distribution is introduced by extending the two-parameter Alpha Power Inverse Rayleigh distribution using Marshall-Olkin G approach. The proposed Marshall-Olkin Generalized Alpha Power Inverse Rayleigh (MOAPIR) distribution generalizes the Marshall-Olkin Inverse Rayleigh, Alpha Power Inverse Rayleigh, and Inverse Rayleigh distribution. The characterization and statistical properties of the proposed distribution such as hazard rate function, reversed hazard rate function, quantiles, moments, and order statistics were derived. The estimation of the MOAPIR distribution parameters is derived using the maximum likelihood estimation method. The performance of the proposed distribution was compared with other competing distribution using two real-life data. The goodness of fit criteria and the distribution function curve showed that the proposed distribution provides a better fit than other competing distributions of the same family of heavily positive skewed distribution.*

**Keywords:** Marshall-Olkin G family, Alpha Power Inverse Rayleigh distribution, Skewed distribution, distribution function, Statistical properties.

## I. Introduction

Marshal-Olkin G method of generalization (MO-G) proposed by Marshall and Olkin [1] is often used to generate a new family of distributions. Using the cumulative distribution function (CDF) of any distribution of a random variable  $X$ , the cumulative function of the new family of distributions is obtained by

$$G(x; \theta) = \frac{G(x)}{\theta + (1-\theta)G(x)} \quad \theta > 0, x \in \mathfrak{R} \quad (1)$$

where  $\theta$  is the location parameter.

Its corresponding probability density function (PDF) is

$$g(x; \theta) = \frac{\theta g(x)}{[\theta + (1-\theta)G(x)]^2} \quad \theta > 0, x \in \mathfrak{R} \quad (2)$$

Many authors such as; Ghitany [2], Ghitany et al.[3], Alice and Jose [4], Okasha and Kayid [5], Okasha et al. [6], Salah et al. [7], Gui [8], Krishna et al. [9], Al-Saiari et al. [10], Mahdavi and Kundu [11], Javed et al. [12], Maxwell et al. [13], Okasha et al. [14], Okasha et al. [15], Haj Ahmad and Almetwally [16], Abdul-Hadi et al. [17], Klakattawi et al. [18], and Aako et al. [19] have used MO-G to extend some base distributions by adding parameters to a well-established family of distribution to generate a new distribution.

This article proposes the generalization of APIR distribution proposed by Malik and Ahmad [20] based on the MO-G which we hereafter called the Marshall-Olkin Generalized Alpha Power Inverse Rayleigh (MOAPIR) distribution. The special cases and the statistical properties of MOAPIR were also presented. Furthermore, the method of maximum likelihood estimation was used to estimate the parameters of the proposed distribution and two data sets were used to demonstrate the performance of the proposed distribution in comparison with other competing distribution of the same family of distributions.

## 2. The Proposed Distribution

Let  $X_1, X_2, \dots$  be a sequence of independent and identically distributed random variables from the APIR distribution.

The cdf and pdf of the APIR distribution are presented in (3) and (4), respectively.

$$G_{APIR}(x; \alpha, \lambda) = \frac{\alpha e^{-\frac{\lambda}{x^2}} - 1}{\alpha - 1}, \quad x > 0, \alpha \neq 1, \lambda > 0, \quad (3)$$

$$g_{APIR}(x; \alpha, \lambda) = \frac{\log \alpha}{\alpha - 1} \frac{2\lambda}{x^3} e^{-\frac{\lambda}{x^2}} \alpha e^{-\frac{\lambda}{x^2}}, \quad x > 0, \alpha \neq 1, \lambda > 0 \quad (4)$$

where  $\alpha$  and  $\lambda$  are shape and scale parameters, respectively.

We applied the MO-G to the APIR distribution by inserting (3) into (1) and inserting (4) into (2) to have the CDF and PDF respectively, of a new generated distribution called the MOAPIR distribution.

If  $X$  is a random variable from MOAPIR distribution, we shall denote as  $X \sim MOAPIR(\alpha, \lambda, \theta)$ . The CDF of MOAPIR is

$$G_{MOAPIR}(x) = \begin{cases} \frac{\alpha e^{-\lambda x^{-2}} - 1}{\theta(\alpha - 1) + (1 - \theta)(\alpha e^{-\lambda x^{-2}} - 1)}, & x > 0, \alpha \neq 1, \lambda > 0, \theta > 0 \\ 0, & = 1 \end{cases} \quad (5)$$

and the corresponding PDF of MOAPIR distribution is

$$g_{MOAPIR}(x) = \begin{cases} \frac{(\alpha-1)2\lambda\theta \log(\alpha)x^{-3}e^{-\lambda x^{-2}}\alpha^{e^{-\lambda x^{-2}}}}{[(\alpha-1)\theta+(1-\theta)(\alpha^{e^{-\lambda x^{-2}}}-1)]^2} & x > 0, \alpha \neq 1, \lambda > 0, \theta > 0 \\ 0, & = 1 \end{cases} \quad (6)$$

To have a useful linear representation of the pdf of the proposed distribution, we used the generalized binomial expansion (GBE) in (7) and the power series in (8)

$$(1 - z)^2 = \sum_{k=0}^{\infty} (k + 1)z^k, |z| < 1, \quad (7)$$

$$\alpha^z = \sum_{m=0}^{\infty} (\log(\alpha))^m z^m \quad (8)$$

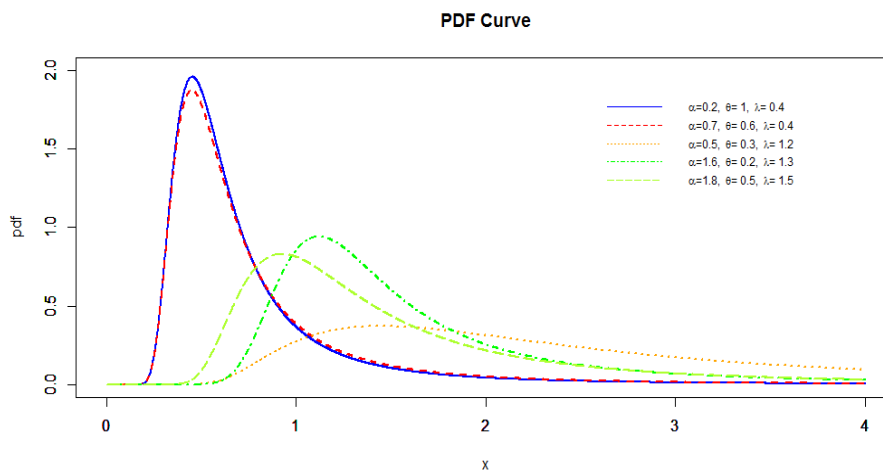
Applying the concept of GBE and power series in (7) and (8) into (6) if  $(\alpha > 0$  and  $\alpha \neq 1)$ , then we have

$$g_{MOAPIR}(x; \alpha, \lambda, \theta) = \sum_{k=0}^{\infty} \sum_{j=0}^k \sum_{m=0}^{\infty} W_{k,j,m} 2\lambda(m+1)x^{-3}e^{-(m+1)\lambda x^{-2}}, \quad (9)$$

where

$$W_{k,j,m} = \begin{cases} (-1)^j \binom{k}{j} (k+1) \frac{(\theta-1)^k (k-j+1)^m (\log(\alpha))^{m+1}}{\theta^{k+1} (\alpha-1)^{k+1} (m+1)!}, & \theta > 1 \\ (-1)^j \binom{k}{j} (k+1) \frac{(1-\theta)^k (j+1)^m (\log(\alpha))^{m+1}}{\theta^{k+1} (\alpha-1)^{k+1} (m+1)!}, & 0 < \theta < 1 \end{cases} \quad (10)$$

For some selected values of the parameters of MOAPIR, the cumulative distribution function and probability distribution function curves are presented in Figure 1. This is to show patterns of the behaviour of the parameters of the proposed distribution.



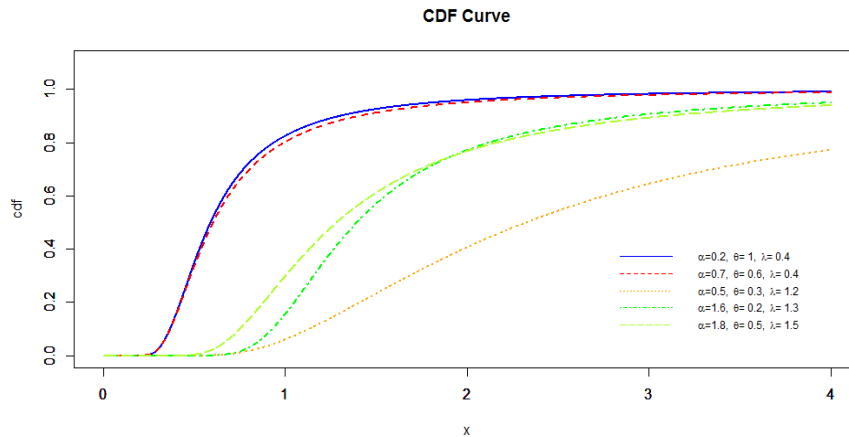


Figure 1: Plots of the PDF and CDF of MOPAIR distribution for selected values of the parameters

Figure 1 shows that MOAPIR is a skewed and unimodal distribution, in addition, the CDF values lies between 0 and 1 is an indication that MOAPIR has a true PDF.

### 2.1 Sub-models of MOAPIR Distribution

To show that the proposed MOAPIR distribution is a generalisation distribution of family of distributions, we varied the value of the parameters of the distribution.

If we substitute  $\alpha = 1$  in (6), then the expression will become

$$g_{MOIR}(x; \theta, \lambda) = \frac{2\lambda\theta x^{-3} e^{-\lambda x^{-2}}}{[\theta + (1 - \theta)e^{-\lambda x^{-2}}]^2} \quad x > 0, \lambda > 0, \theta > 0$$

which is the pdf of the Marshall-Olkin Inverse Rayleigh (MOIR).

Similarly, if  $\theta = 1$ , then the expression in (6) will become

$$g_{APIRD}(x; \alpha, \lambda) = \frac{\log \alpha \ 2\lambda}{\alpha - 1} x^3 e^{-\frac{\lambda}{x^2}} \alpha e^{-\frac{\lambda}{x^2}} \quad x > 0, \alpha \neq 1, \lambda > 0$$

which is the pdf of the Alpha Power Inverse Rayleigh (APIR) distribution proposed by Malik and Ahmad [20]. Also, when  $\alpha = \theta = 1$ , (6) will be reduced to the pdf of Inverse Rayleigh (IR) distribution proposed by Srinivasa, et al. [21] which is given by

$$g_{IR}(x; \lambda) = \frac{2\lambda}{x^3} e^{-\frac{\lambda^2}{x^2}} \quad x, \lambda > 0$$

Thus, the proposed MOAPIR has been proven to be a generalization distribution of the APIR family of distributions.

### 2.2 Reliability Analysis

#### 2.2.1 Survival Function

The survival function of MOAPIR distribution denoted by  $R_{MOAPIR}(x)$  is derived using the expression presented in (11)

$$R_{MOAPIR}(x) = \bar{G}(x) = 1 - G(x) \tag{11}$$

Substituting  $G(x)$  in (5) into (11), we have the Survival function of MOAPIR to be

$$R_{MOAPIR}(x) = \begin{cases} \frac{\alpha\theta(1-\alpha e^{-\lambda x^{-2}}-1)}{\theta(\alpha-1)+(1-\theta)(\alpha e^{-\lambda x^{-2}}-1)}, & x > 0, \alpha \neq 1, \lambda > 0, \theta > 0 \\ 0, & \alpha = 1 \end{cases} \tag{12}$$

2.2.2 Hazard Rate Function (HRF)

Let  $X$  be a random variable with pdf,  $g(x)$  and cdf,  $G(x)$ , then the hazard rate function (HRF) is derived by solving  $\frac{g(x)}{G(x)}$ .

Thus, if  $X$  is a MOAPIR random variable, then the HRF of the random variable  $X$  denoted by  $h_{MOAPIR}(x)$  is

$$h_{MOAPIR}(x) = \begin{cases} \frac{(\alpha-1)2\lambda \log(\alpha)x^{-3}e^{-\lambda x^{-2}}\alpha e^{-\lambda x^{-2}}-1}{\left[\theta(\alpha-1)+(1-\theta)(\alpha e^{-\lambda x^{-2}}-1)\right]\left(1-\alpha e^{-\lambda x^{-2}}-1\right)} & x > 0, \alpha \neq 1 \\ 0, & = 1 \end{cases} \tag{13}$$

where  $\lambda > 0, \theta > 0$ .

The pattern of the survival function and hazard rate function of the proposed distribution for various selected values of the distribution parameters are presented in Figure 2

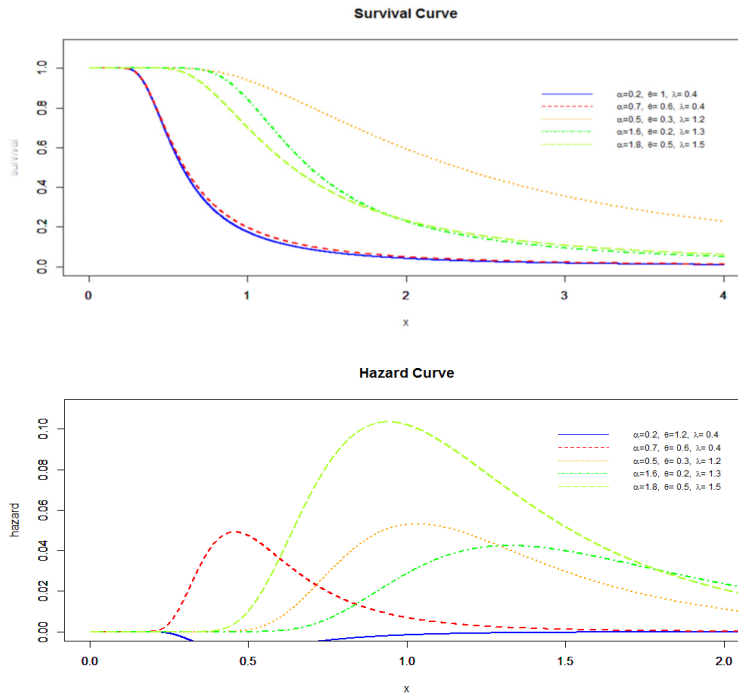


Figure 2: The Survival Function and Hazard Rate Function Curves of MOAPIR distribution



### 2.2.3 Reversed Hazard Rate Function

The reversed hazard rate (RHR) function of a random variable  $X$  from MOAPIR  $(\alpha, \theta, \lambda)$  distribution denoted as  $r_{MOAPIR}(x)$  is derived to be:

$$r_{MOAPIR}(x) = \begin{cases} \frac{(\alpha-1)2\lambda \log(\alpha)x^{-3}e^{-\lambda x^{-2}}\alpha e^{-\lambda x^{-2}}}{[\theta(\alpha-1)+(1-\theta)(\alpha e^{-\lambda x^{-2}}-1)](\alpha e^{-\lambda x^{-2}}-1)} & x > 0, \alpha \neq 1 \\ 0, & = 1 \end{cases} \quad (14)$$

where  $\lambda > 0, \theta > 0$

Figure 3 represents the RHRF curves for the MOAPIR  $(\alpha, \theta, \lambda)$  distribution for selected values of the distribution parameters

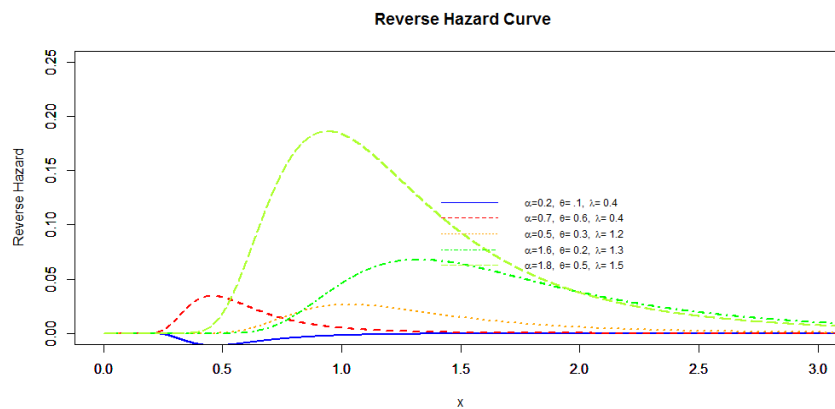


Figure 3: RHRF Curve of the MOAPIR distribution for some selected parameters values.

## 2.3 Statistical Properties

In this section, we derived the statistical properties of the MOAPIR distribution. The properties derived are quantiles, median, mean, variance, order statistics, and range.

### 2.3.1 Quantiles

Quantiles explain how many values in a distribution are above or below a certain limit and define special part of a data set. The quantile of any distribution of a random variable  $X$  is given by solving the expression in (15)

$$G(x_q) = q, \quad 0 < q < 1 \quad (15)$$

The  $q^{\text{th}}$  quantile function is obtained by solving (16)

$$q = \frac{\alpha e^{-\lambda x^{-2}} - 1}{\theta(\alpha-1) + (1-\theta)(\alpha e^{-\lambda x^{-2}} - 1)} \quad (16)$$

Hence,

$$x_q = G^{-1}(q) = \left[ \frac{1}{\lambda} \log \left( \frac{\log(\alpha)}{\log \left( \frac{1 + (\alpha\theta - 1)q}{1 + (\theta - 1)q} \right)} \right) \right]^{-\frac{1}{2}} \quad (17)$$

Using (17), we obtained the median, skewness, and kurtosis by determining the quantile of the MOAPIR distribution.

To obtain the First Quantile, substituting  $q = 0.25$  in (17), then we obtained

$$x_{0.25} = G^{-1}(0.25) = \left[ \frac{1}{\lambda} \log \left( \frac{\log(\alpha)}{\log\left(\frac{0.75+0.25\alpha\theta}{0.75+0.25\theta}\right)} \right) \right]^{-\frac{1}{2}} \quad (18)$$

For the Median =  $Q_2 = P_{50}$ , we use  $q = 0.50$  in (17) and obtained

$$\text{Median} = x_{0.5} = G^{-1}(0.5) = \left[ \frac{1}{\lambda} \log \left( \frac{\log(\alpha)}{\log\left(\frac{1+\alpha\theta}{1+\theta}\right)} \right) \right]^{-\frac{1}{2}} \quad (19)$$

For the Third Quantile =  $Q_3 = P_{75}$ , we use  $q = 0.75$  in (17) and obtained

$$Q_3 = x_{0.75} = G^{-1}(0.75) = \left[ \frac{1}{\lambda} \log \left( \frac{\log(\alpha)}{\log\left(\frac{0.25+0.75\alpha\theta}{0.25+0.75\theta}\right)} \right) \right]^{-\frac{1}{2}} \quad (20)$$

Using (18), (19) and (20), the Skewness ( $S_k$ ) and Kurtosis ( $K_{MOAPIR}$ ) of the MOAPIR distribution were obtained respectively as

$$S_{kMOAPIR} = \frac{G^{-1}(0.75) - 2G^{-1}(0.5) + G^{-1}(0.25)}{G^{-1}(0.75) - G^{-1}(0.25)} \quad (21)$$

and

$$K_{MOAPIR} = \frac{G^{-1}(0.875) - G^{-1}(0.625) - G^{-1}(0.375) + G^{-1}(0.125)}{G^{-1}(0.75) - G^{-1}(0.25)} \quad (22)$$

### 2.3.2 Moments

Let  $X$  be a random variable that has MOAPIR  $(\alpha, \lambda, \theta)$  distribution, the  $r^{\text{th}}$  moments of  $X$  is defined as

$$E[X^r] = \int_0^{\infty} x^r g(x) dx \quad (23)$$

$$= \int_0^{\infty} x^r \frac{(\alpha-1)2\lambda\theta \log(\alpha)x^{-3}e^{-\lambda x^{-2}}\alpha e^{-\lambda x^{-2}}}{[(\alpha-1)\theta + (1-\theta)(\alpha e^{-\lambda x^{-2}} - 1)]^2} dx \quad (24)$$

Using linear expressions of  $g_{MOAPIR}(x)$  in (9) and (10), we have

$$E[X^r] = \int_0^{\infty} x^r \sum_{k=0}^{\infty} \sum_{j=0}^k \sum_{m=0}^{\infty} W_{k,j,m} 2\lambda(m+1)x^{-3} dx \quad (25)$$

$$= \sum_{k=0}^{\infty} \sum_{j=0}^k \sum_{m=0}^{\infty} W_{k,j,m} ((m+1)\lambda)^{\frac{r}{2}} \Gamma\left(1 - \frac{r}{2}\right) \quad (26)$$

From (25) and (26), the mean and variance of a random variable  $X$  from MOAPIR distribution are:

$$E[X] = \sum_{k=0}^{\infty} \sum_{j=0}^k \sum_{m=0}^{\infty} W_{k,j,m} ((m+1)\lambda)^{\frac{1}{2}} \Gamma\left(\frac{1}{2}\right) \quad (27)$$

and

$$V(X) = \sum_{k=0}^{\infty} \sum_{j=0}^k \sum_{m=0}^{\infty} W_{k,j,m} ((m+1)\lambda) - \left( \Gamma\left(\frac{1}{2}\right) \sum_{k=0}^{\infty} \sum_{j=0}^k \sum_{m=0}^{\infty} W_{k,j,m} ((m+1)\lambda)^{\frac{1}{2}} \right)^2 \quad (28)$$

### 2.3.3 Order statistics

The pdf of the  $i^{\text{th}}$  order statistics  $X_{i:n}$  of a random sample  $X_1, X_2, \dots, X_n$  is

$$g_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!} g(x)[G(x)]^{i-1}[1 - G(x)]^{n-i}, \quad x > 0, \alpha > 0, \lambda > 0 \quad (29)$$

From (29), the pdf of the  $i^{\text{th}}$  order statistics  $X_{i:n}$  of MOAPIR distribution is obtained to be

$$g_{i:n}(x) = \frac{2\lambda}{(i-1)!(n-i)!} \sum_{k_1=0}^{\infty} \sum_{j_1=0}^{n-i} \sum_{j_2=0}^{i+k_1-j_1} W_{k_1, j_1, j_2} x^{-3} e^{\lambda x^{-2}} \alpha^{(n-i-j_1)+(i+k_1-j_2)} e^{\lambda x^{-2}} \quad (30)$$

where

$$W_{k_1, j_1, j_2} = (-1)^{j_1+j_2} \binom{n-i}{j_1} \binom{i+k_1-1}{j_2} \frac{\Gamma(n+k_1+1)(\theta-1)^{k_1}}{k_1! (\theta)^{i+k_1}}$$

### 2.3.4 Range of MOAPIR

Let  $X_{(1)}, X_{(2)}, X_{(3)}, \dots, X_{(n)}$  be the order statistics from the sample  $X_1, X_2, X_3, \dots, X_n$  of size  $n$  from a random variable that is of MOAPIR distribution, then the distribution of the range of the random variable  $X$ ,  $R(x)$  can be obtained by solving  $R(x) = g_{n:n}(x) - g_{1:n}(x)$

Using (30), the range of MOAPIR random variable is derived to be

$$R(x) = \frac{2\lambda}{(n-1)!} \left( \sum_{k_1=0}^{\infty} \sum_{j_2=0}^{n+k_1-1} W_{k_1, j_2} x^{-3} e^{\lambda x^{-2}} \alpha^{(n+k_1-j_2)} e^{\lambda x^{-2}} - \sum_{k_1=0}^{\infty} \sum_{j_1=0}^{n-1} \sum_{j_2=0}^{k_1} W_{k_1, j_1, j_2} x^{-3} e^{\lambda x^{-2}} \alpha^{(n-1-j_1)+(1+k_1-j_2)} e^{\lambda x^{-2}} \right) \quad (31)$$

## 2.4. Estimation of Parameters of MOAPIR Distribution

The parameters of the proposed distribution were derived using the maximum likelihood estimation approach. Let  $X_1, \dots, X_n$  be a random sample of size  $n$  from MOAPIR distribution, then the likelihood function of the MOAPIR distribution,  $L(x/\alpha, \lambda, \theta)$  is

$$L(x/\alpha, \lambda, \theta) = \prod_{i=1}^n g(x_i) = \frac{(\alpha-1)^n (\log(\alpha))^n \lambda^n 2^n \theta^n e^{-\lambda \sum_{i=1}^n x_i} \alpha^{\sum_{i=1}^n x_i} e^{-\lambda x_i^{-2}} \prod_{i=1}^n x_i^{-3}}{\prod_{i=1}^n \left[ (\alpha-1)\theta + (1-\theta) \left( \alpha^{e^{-\lambda x_i^{-2}}} - 1 \right) \right]^2} \quad (32)$$

By taking logarithm of the likelihood function, we have

$$\ell(x/\alpha, \lambda, \theta) = n \log((\alpha-1) \log(\alpha) 2\lambda\theta) - \lambda \sum_{i=1}^n x_i^{-2} + \log(\alpha) \sum_{i=1}^n e^{-\lambda x_i^{-2}} - 3 \sum_{i=1}^n \log(x_i) - 2 \sum_{i=1}^n \log \left[ (\alpha-1)\theta + (1-\theta) \left( \alpha^{e^{-\lambda x_i^{-2}}} - 1 \right) \right] \quad (33)$$

To obtain the MLEs of  $\alpha$ ,  $\lambda$  and  $\theta$ , we differentiate the expression in (33) with respect to  $\alpha$ ,  $\lambda$  and  $\theta$ . Thus, we have

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha-1} + \frac{n}{\alpha \log(\alpha)} + \frac{1}{\alpha} \sum_{i=1}^n e^{-\lambda x_i^{-2}} - 2 \sum_{i=1}^n \frac{\theta + (1-\theta) e^{-\lambda x_i^{-2}} \alpha^{e^{-\lambda x_i^{-2}}} - 1}{(\alpha-1)\theta + (1-\theta) \left( \alpha^{e^{-\lambda x_i^{-2}}} - 1 \right)} \quad (34)$$

$$\frac{\partial \ell}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^n x_i^{-2} - \log(\alpha) \sum_{i=1}^n x_i^{-2} e^{-\lambda x_i^{-2}} - 2 \sum_{i=1}^n \frac{(1-\theta) \log(\alpha) x_i^{-2} e^{-\lambda x_i^{-2}} \alpha^{e^{-\lambda x_i^{-2}}} - 1}{(\alpha-1)\theta + (1-\theta) \left( \alpha^{e^{-\lambda x_i^{-2}}} - 1 \right)} \quad (35)$$

and

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} - 2 \sum_{i=1}^n \frac{\alpha - \alpha^{e^{-\lambda x_i^{-2}}}}{(\alpha-1)\theta + (1-\theta) \left( \alpha^{e^{-\lambda x_i^{-2}}} - 1 \right)} \quad (36)$$

Solving (34), (35) and (36) by equating them to zero, we have

$$\frac{n}{\alpha-1} + \frac{n}{\alpha \log(\alpha)} + \frac{1}{\alpha} \sum_{i=1}^n e^{-\lambda x_i^{-2}} - 2 \sum_{i=1}^n \frac{\theta + (1-\theta)e^{-\lambda x_i^{-2}} \alpha e^{-\lambda x_i^{-2}} - 1}{(\alpha-1)\theta + (1-\theta)(\alpha e^{-\lambda x_i^{-2}} - 1)} = 0 \quad (37)$$

$$\frac{n}{\lambda} - \sum_{i=1}^n x_i^{-2} - \log(\alpha) \sum_{i=1}^n x_i^{-2} e^{-\lambda x_i^{-2}} - 2 \sum_{i=1}^n \frac{(1-\theta) \log(\alpha) x_i^{-2} e^{-\lambda x_i^{-2}} \alpha e^{-\lambda x_i^{-2}}}{(\alpha-1)\theta + (1-\theta)(\alpha e^{-\lambda x_i^{-2}} - 1)} = 0 \quad (38)$$

and

$$\frac{n}{\theta} - 2 \sum_{i=1}^n \frac{\alpha - \alpha e^{-\lambda x_i^{-2}}}{(\alpha-1)\theta + (1-\theta)(\alpha e^{-\lambda x_i^{-2}} - 1)} = 0 \quad (39)$$

The MLE of  $\alpha$ ,  $\lambda$  and  $\theta$  can not be obtained by solving (37), (38), and (39) analytically. Hence the Newton-Raphson iterative method would be used to accomplish the task of estimating the parameters.

### 3. Determination of Flexibility of the Proposed Distribution

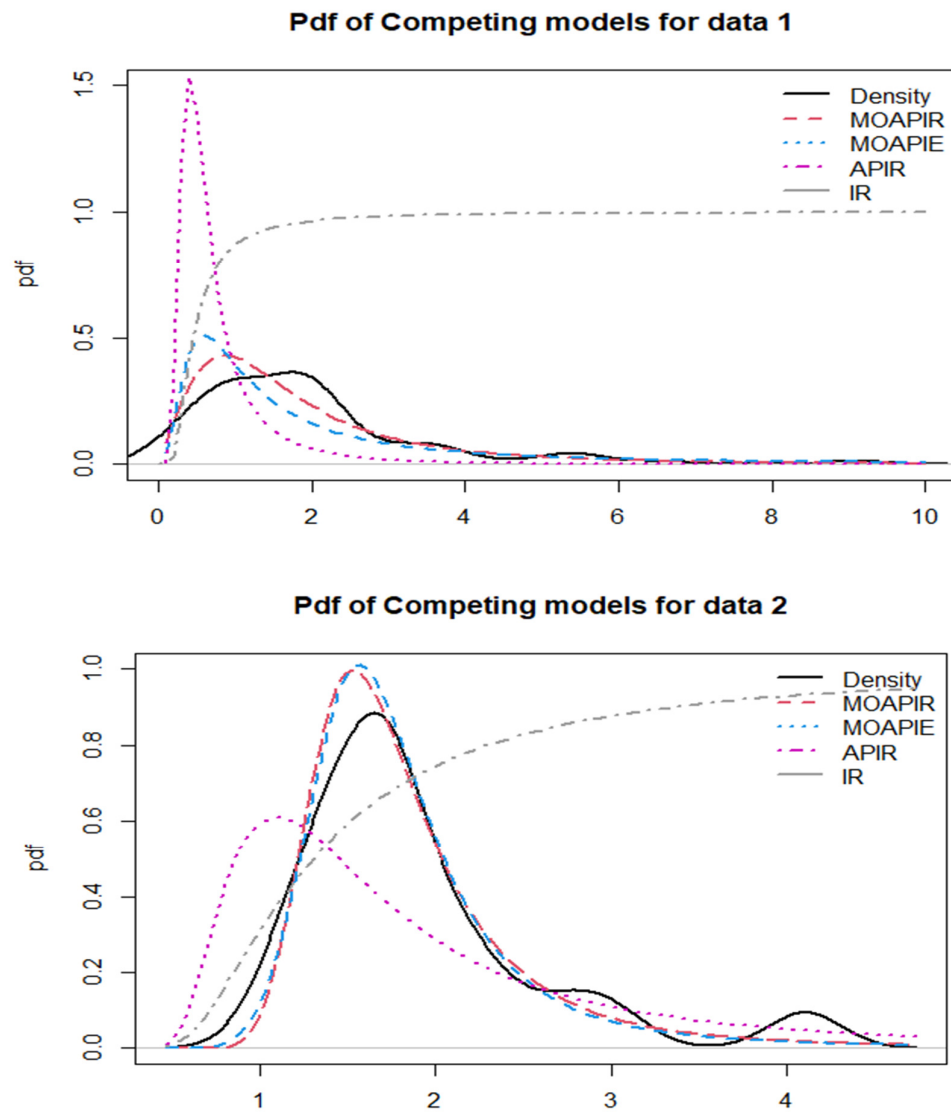
To access the flexibility of the proposed distribution, the MOAPIR distribution is compared with three competing distributions by using two real life data sets. The distributions considered in this study are the Marshall Olkin Alpha Power Inverse Exponential (MOAPIE), Alpha Power Inverse Rayleigh (APIR), and Inverse Rayleigh (IR) distributions.

**Data Set I** is on life of fatigue fracture of Kevlar 373/epoxy that are subjected to constant pressure at the 90% stress level until all had failed (Ogunde et al. [22]) and **Data Set II** is on the relief times of twenty patients receiving an analgesic as reported by Gross and Clark [23].

The summary statistics of the two datasets are presented in Table 1 and the density plot of the datasets along with the empirical density plots of the considered distributions are presented in Figures 4.

**Table 1: Summary Statistics of Datasets**

Data set	Min	Q1	Median	Q3	Mean	Variance	Max	Skewness	Kurtosis
I (n=76)	0.0251	0.0905	1.7361	2.2960	1.9590	2.4774	9.0960	1.9406	4.9474
II (n=20)	1.10	1.475	1.7	2.05	1.90	0.4958	4.10	1.5924	2.3465



**Figure 4:** Density Plot of Datasets I and II With MOAPIRD and other competing Distributions.

### 3.1 Parameter Estimation and Goodness of Fit Test

Four criteria, namely, the log-likelihood values (-LL), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Hannan-Quinn information criterion (HQIC) are used to select the best fitted model to the two data sets under consideration. The model with minimum value of each of the four criteria is adjudged as the best fit for the datasets under study. The estimated values of the parameters of the four distributions and the goodness of fit criteria are presented in Table 2.

**Table 2:** Estimated parameters and Criteria for goodness of fit

Data set	Distribution	Parameters			- LL	AIC	BIC	HQIC
		$\alpha$	$\lambda$	$\theta$				
I	MOAPIR	7.652	0.001136	853.4	<b>124.84</b>	<b>255.69</b>	<b>262.67</b>	<b>258.5</b>
	MOAPIE	1253	0.2375	0.5040	139.90	285.80	292.79	288.6
	APIR	6883	0.0293	-	6007.67	12019.33	12024.0	12021.2
	IR	-	0.1406	-	211.48	424.96	427.29	425.89
II	MOAPIR	51.84	7.5567	0.0071	<b>15.51</b>	<b>37.02</b>	<b>40.01</b>	<b>37.6</b>
	MOAPIE	1.114	10.80	0.0017	15.65	37.29	40.27	37.86
	APIR	1	1.801	-	694.52	1393.05	1395.03	1393.43
	IR	-	1.1749	-	28.27	58.54	59.54	58.73

\* the bold number represents the smallest value for each criterion.

#### 4. Discussion

The cumulative distribution function and probability density function of the proposed MOAPIR distribution were given in (5) and (6) respectively. Figure 1 illustrates the shape of the distribution when its parameters were varied and it was clear that the distribution is a positively heavily skewed, unimodal and a true distribution function. The survival function curve reflected that higher value of  $\lambda$ , ( $\lambda > 0.4$ ) will destruct the expected shape of the hazard function. Similarly, it is crystal clear that the value of  $\theta$  has no significant effect on the shape of the hazard function (see Figure 2). However, it was observed that variations in the values of the parameters significantly affect the pattern of the hazard rate function and the reversed hazard rate function (see Figure 2 and 3). Further analysis shows that  $\theta$  and  $\lambda$  have no significant effects on the skewness and kurtosis of the distribution but have influence on the mean, median and variance of the distribution. The summary statistics in Table 1 shows that the two data sets are heavily positively skewed data. Furthermore, the density plot in Figure 4 for both data sets indicated that the two data sets are heavily positively skewed. The fitted distributions as shown in Figure 4, reflected that the proposed MOAPIR is a more suitable distribution than all other competing distributions considered in this study. The results from the performance indices namely, -LL, AIC, BIC and HQIC confirmed that the proposed MOAPIR best fit the two data sets considered in this paper than the MOAPIE, APIR and IR distributions.

#### 5. Conclusion

In this paper, a new distribution called Marshall Olkin Alpha Power Inverse Rayleigh (MOAPIR) distribution was introduced. The pdf and cdf of the distribution were derived and some of its properties, such as hazard rate function, reversed hazard rate function, quantiles, moments, and order statistics were studied. The parameters of the proposed distribution were estimated using the Maximum likelihood estimation method. To access the flexibility of the proposed MOAPIR distribution with three competing distributions of the same family, namely the MOAPIE, APIR and IR distributions, two data sets were used. The results showed that the proposed MOAPIR distribution has minimum value of -LL, AIC, BIC and HQIC, and then, adjudged to be the best fit for the two data sets considered in this study. Therefore, the proposed distribution provides a better fit than other competing distributions of the same family of heavily positively skewed distribution.

Hence, for a heavily positive skewed data, the MOAPIR is a good distribution model to be used for further analysis.

## References

- [1] Marshall, A. W. & 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: 641 – 652.
- [2] Ghitany M. E. (2005). Marshall Olkin extended Pareto and its application. *International Journal of Applied Mathematics* 18:17 – 32.
- [3] Ghitany, M. E., Al-Awadhi, F. A. & Alkhalfan, L. A. (2007). Marshall–Olkin extended Lomax distribution and its application to censored data. *Communication Statistics- Theory and Methods* 36:1855 – 1866.
- [4] Alice T., Jose, K. K. (2005). Marshall–Olkin Logistic processes. *STARS International. Journal* 6: 1 – 11.
- [5] Okasha, H. M. & Kayid, M. (2016). A New Family of Marshall–Olkin Extended Generalized Linear Exponential Distribution. *Journal of Computational and Applied Mathematics* 296: 576 – 592.
- [6] Okasha, H. M., El-Baz, A. H., Tarabia, A. M. K. & Basheer, A. B. (2017). Extended inverse Weibull distribution with reliability application. *J Egypt Math Soc.* 25(3): 343 – 349.
- [7] Salah, M. M., Raqab, M. Z., & Ahsanullah, M. (2012). Marshall-Olkin exponential distribution: Moments of order statistics. *J. of Applied Statistical Science* 17(1): 81– 91.
- [8] Gui, W. (2013). Marshall-Olkin Extended Log-Logistic Distribution and Its Application in Minification Processes. *Applied Mathematical Sciences* 7(80): 3947 – 3961.
- [9] Krishna, E., Jose, K. K., Alice, T. & Ristic, M. M. (2013). The Marshall-Olkin Frechet distribution. *Communication Statistics Theory Methods* 42(1): 4091 – 4107.
- [10] Al-Saiari, A. Y., Mousa, S. A., & Baharith, L. A. (2016). Marshall-Olkin Extended Burr III Distribution. *International Mathematical Forum.* 11(13): 631 – 642.
- [11] Mahdavi, A. & Kundu, D. (2017) A new method for generating distributions with an application to exponential distribution. *Communication Statistics Theory Methods* 46(13): 6543 – 6557.
- [12] Javed, M., Nawaz, T., & Irfan, M. (2018). The Marshall-Olkin Kappa Distribution: Properties and Applications. *Journal of King Saud University- Science* 31(4): 684 - 691.  
<https://doi.org/10.1016/j.jksus.2018.01.001>.
- [13] Maxwell, O., Chukwu, A. U., Oyamakin, O. S., & Khaleel, M. A. (2019). The Marshall-Olkin Inverse Lomax Distribution (MO-ILD) with Application on Cancer Stem Cell. *Journal of Advances in Mathematics and Computer Science* 33(4): 1 – 12.
- [14] Okasha, H. M., El-Baz, A. H. & Basheer, A. M. (2020). On Marshall-Olkin Extended Inverse Weibull Distribution: Properties and Estimation using Type-II Censoring Data, *Journal of Statistics Applications & Probability Letters* 7(1): 9 – 21.
- [15] Okasha, H. M., Basheer, A.M. & El-Baz, A.H. (2021). Marshall-Olkin Extended Inverse Weibull Distribution: Different Method of Estimations, *Annals of Data Science* 8: 769 – 984.  
<https://doi.org/10.1007/s40745-020-00299-5>.
- [16] Haj Ahmad, H. A. & Almetwally, E. M. (2020). Marshall-Olkin Generalized Pareto Distribution: Bayesian and Non Bayesian Estimation. *Pak.j.stat.oper.res.* 16(1): 21 – 33.  
DOI: <http://dx.doi.org/10.18187/pjsor.v16i1.2935>.
- [17] Abdul-Hadi N. Ahmed, Z.M.N. & Hager N.A. (2021). New Compounded Family of Distributions, with Applications, *International Journal of Probability and Statistics*, 10 (3): 74 – 93.
- [18] Klakattawi, H.A.D, Elaal, M.A., Dey, S. & Baharith, L. (2022) A new generalized family of distributions based on combining Marshal-Olkin transformation with T-X family. *PLoSONE* 17(2):152 – 167.
- [19] Aako, O. L., Adewara, J. A. & Nkemnole, E. B. (2022). Marshall-Olkin Generalized Inverse Log-Logistic Distribution: Its properties and applications. *International Journal of Mathematical Sciences and Optimization: Theory and Applications.* 8(2): 15 – 31.

- [20] Malik, A. S. & Ahmad, S. P. (2018). A New Inverse Rayleigh Distribution: Properties and Application. *Inter. Journal of Scientific Research in Mathematical and Statistical Science* 5(5): 92 – 96.
- [21] Srinivasa, R.B., Kantam, R.R.L. & Reddy, J.P. (2013). Variable control charts based on percentiles of Inverse Rayleigh distribution. *Journal of Applied Probability and Statistics* 8(1): 46 – 57.
- [22] Ogunde, A. A., Fatoki, O. & Oseghale, O. I. (2017). On The Applications of Transmuted Inverted Weibull Distribution. *Global Journal of Science Frontier Research: F Mathematics and Decision Sciences* 17 (6): 24 – 38.
- [23] Gross, A.J. and Clark, V.A. (1975). *Survival Distributions: Reliability Applications in the Biometrical Sciences*. John Wiley, New York.



# ON DIFFERENT CLASSICAL ESTIMATION APPROACHES FOR TYPE I HALF LOGISTIC-TOPP- LEONE- EXPONENTIAL DISTRIBUTION

Akeem Ajibola Adepoju<sup>1</sup>, Sauta S. Abdulkadir<sup>2</sup>, Danjuma Jibasen<sup>3</sup>

<sup>1</sup>Department of Statistics, Faculty of Computing and Mathematical Sciences, Aliko Dangote  
University of Science and Technology, Wudil, Kano. Nigeria

<sup>2,3</sup>Department of Statistics, Faculty of Physical Sciences Modibbo Adama University, Yola,  
Adamawa. Nigeria

[akeebola@gmail.com](mailto:akeebola@gmail.com)<sup>1</sup>, [ssabdulkadir@mautech.edu.ng](mailto:ssabdulkadir@mautech.edu.ng)<sup>2</sup>, [djibasen2001@gmail.com](mailto:djibasen2001@gmail.com)<sup>3</sup>

## Abstract

*This paper aims to propose six methods of parameter estimation in order to examine the behavior of the new Type I Half Logistic Topp-leone Exponential distribution. The methods taking into consideration are Maximum Likelihood, Anderson\_Darling, Least Squares, Cramer\_von\_Mises, Maximum Product of Spacing, and Weighted Least Squares Methods. The results show that all the methods are consistent, since the estimates approach the true value of the parameters for all the methods. The bias, mean square error and mean relative estimates decay as the sample size is raised. The estimates of the six methods obtained for the model, indicated that MPS estimates is the closest to the true value of the parameters across the low, moderate and high sample sizes, invariable, the MPS produces the least biasness. Buttress more, the MPS produces the least MSE all through and remain the best estimator for low, moderate and high sample size of the model. Conclusively, MPS is the most consistent among the estimators for the model.*

**Keywords:** Type I Half Logistic Topp-leone Exponential distribution, maximum likelihood, Anderson\_Darling, least squares and weighted least squares Methods, Cramer\_von\_Mises, Maximum Product of Spacing

## I. Introduction

Exponential (Exp) distribution is an important and commonly explored probability distribution both in univariate, bivariate and multivariate cases. The Type I Half Logistic Topp-leone Exponential (TIHLTLExp) distribution was proposed by [1] as a generalized distribution. The distribution is characterized with two shape parameters and a scale parameter. The hazard rate shapes of the distribution are monotonically increasing, monotonically decreasing and bathtub in feature. It was revealed from the TIHLTLExp distribution analysis that the distribution potentiality is awesome in modeling a good number of life time data sets.

On the other hand, Exp distribution has witness different generalization where one or more shape parameter are introduced to extend it flexibility such can be found in the work of [2] where sine family was adopted to generalized the Exp distribution, in [3], the Type I half logistic exponentiated family was used to improve the Exp distribution. Other generalization of Exp

distribution includes Half logistic-truncated exponential distribution [4], A new extension named Lehmann type-II G class of distributions: Exp distribution [5] Lomax Exp distribution [6], Lehmann Type II-Lomax Distribution [7], Exponentiated Gamma Exp Distribution [8], Exponentiated Weibull Exp distribution [9], Topp-Leone generalized Exp power series distribution [10], new extension of Exp distribution [11], Type II Half Logistic Exp Distribution [12], Gamma-exponentiated Exp distribution [13], and the Type II half logistic exponentiated family [14] to mention but few.

Various methods have been developed and applied to estimate the some newly developed distribution. No particular estimation method is the best all round. However, some methods perform better than the other depending on the behavior of the distribution parameters. Six different classical methods are considered in this article. The classical approaches such as Maximum Likelihood Estimator (MLE), Anderson–Darling Estimator (ADE), Cramér–von Mises (CVM), Maximum Product Spacing (MPS), Least Square Estimator (LSE), and Weighted Least Square Estimator (WLSE) are explored. Articles that adopted some estimation methods includes, type II exponentiated half-logistic-PLo (TIIEtHL-PLo) distribution by [15], Parameter estimation methods adopted are MLE, LSE, WLSE, MPS, CVM, and ADE in the study. Inference on Kavya–Manoharan Kumaraswamy distribution by [16], estimation of polynomial Exp family of distributions by [17], estimation comparison for extreme value distribution by [18], Classical and Bayesian Approach Estimation of Weibull-Exp Distribution by [19], estimation preference inverse rayleigh frechet model by [20], estimation methods in Tasks of processing measurement results by [21], comparison of estimation methods for the (Three-Parameter) Lindley distribution by [22]. MLE, OLS, WLS, MPS, and CVM methods, different estimation approaches for Type I half-logistic topp–leone distribution by [23], comparative study of estimation for Pareto distribution by [24], some estimation methods for lindley distribution, estimation methods include MME, MLE, resulting identification of MLE to be the best estimator by [25], also, the weibull distribution parameters, three methods such as the MLE, MME and LSE regression method were considered and compared, from the result, the MME method was superior [26], LSE of distribution functions [27], MPS estimation with preference to the lognormal distribution [28] and parameters estimation for the (three-parameter) Reflected Weibull model. The MME, MLE, Location and Scale Parameters free ML estimator (LSPEE). The data transformation is the basis for LSPEE, Mont Carlo simulations show that the LSPEE outperform MME and MLE. The TIHLTLExp distribution was a newly distribution developed, however, only two methods MLE and MPS were used for parameter estimation.

This paper aims to investigate the behaviour of the TIHLTLExp model parameters using six estimation methods. The motivation for this study is the determination of the best model parameter’s estimator for low, moderate and high sample size of the TIHLTLExp distribution.

## II. Methods

### 2.0 Method of parameter estimation of TIHLTLExp distribution

In this section, we introduced the cumulative distribution function (cdf) and probability density function (pdf) of the Type Half Logistic Topp-leone Exponential Distribution.

$$F_{TIHLTLExp}(x; \beta, \theta, \lambda) = \frac{1 - \left[ 1 - \left[ 1 - (e^{-\lambda x})^2 \right]^\theta \right]^\beta}{1 + \left[ 1 - \left[ 1 - (e^{-\lambda x})^2 \right]^\theta \right]^\beta} \quad (1)$$

$$f_{TIHLTLExp}(x; \beta, \theta, \lambda) = \frac{4\beta\theta\lambda(e^{-\lambda x})^2 \left[ 1 - (e^{-\lambda x})^2 \right]^{\theta-1} \left[ 1 - \left[ 1 - (e^{-\lambda x})^2 \right]^\theta \right]^{\beta-1}}{\left[ 1 + \left[ 1 - \left[ 1 - (e^{-\lambda x})^2 \right]^\theta \right]^\beta \right]^2} \quad (2)$$

The method employed to be used to estimate the parameter include: MLE, ADE, CVM, MPS, LSE, and WLSE

### 2.1 Maximum Likelihood Estimation (MLE)

MLE is one of the widely explored estimation approaches. It is adopted in estimating the parameters of the TIHLTLExp model. if we randomly sampled  $X_i$  where  $i = 1, \dots, n$ , obtained from the TIHLTLExp distribution with parameter  $\Omega = \beta, \theta, \lambda$ . The log-likelihood function  $L(\Omega)$  of (1) is obtained as

$$L(\Omega) = n \log 4 + n \log \beta + n \log \theta + n \log \lambda + 2 \sum_{i=0}^n \log(e^{-\lambda x}) + (\theta - 1) \sum_{i=0}^n \log(1 - (e^{-\lambda x})^2) + (\beta - 1) \sum_{i=0}^n \log(1 - (1 - (e^{-\lambda x})^2)^\theta) - 2 \sum_{i=0}^n \log(1 + (1 - (1 - (e^{-\lambda x})^2)^\theta)^\beta) \quad (3)$$

By differentiating  $L(\Omega)$  in (3) with respect to  $\beta, \theta$  and  $\lambda$ , and the results set to zero will provide the estimators. Thus,

$$\frac{\delta L(\Omega)}{\delta \beta} = \frac{n}{\beta} + \sum_{i=0}^n \log(1 - (1 - (e^{-\lambda x})^2)^\theta) - \frac{\beta \log \sum_{i=0}^n (1 - (1 - (e^{-\lambda x})^2)^\theta)}{\sum_{i=0}^n (1 + (1 - (1 - (e^{-\lambda x})^2)^\theta)^\beta)} = 0 \quad (4)$$

$$\frac{\delta L(\Omega)}{\delta \theta} = \frac{n}{\theta} + \sum_{i=0}^n \log(1 - (e^{-\lambda x})^2) + \frac{\theta(\beta - 1) \sum_{i=0}^n \log(1 - (1 - (e^{-\lambda x})^2)^\theta)}{(1 - (1 - (e^{-\lambda x})^2)^\theta)} \quad (5)$$

$$-2\beta \sum_{i=0}^n \left( 1 + (1 - (1 - (e^{-\lambda x})^2)^\theta)^\beta \right) \frac{\theta \sum_{i=0}^n \log(1 - (1 - (e^{-\lambda x})^2)^\theta)}{(1 - (1 - (e^{-\lambda x})^2)^\theta)^\beta} = 0$$

$$\frac{\delta L(\Omega)}{\delta \lambda} = \frac{n}{\lambda} - \frac{2\lambda e^{-\lambda x} x}{(e^{-\lambda x})} + \frac{2\lambda(\theta-1)e^{-\lambda x} x}{(1-(e^{-\lambda x})^2)} + \frac{2\theta\lambda(\beta-1)(1-(e^{-\lambda x})^2)^{\theta-1} e^{-\lambda x} x}{(1-(1-(e^{-\lambda x})^2)^\theta)}$$

$$- \frac{2\beta \left(1 - (1 - (e^{-\lambda x})^2)^\theta\right)^{\beta-1} (1 - (e^{-\lambda x})^2)^{\theta-1} e^{-\lambda x} x}{\left(1 + \left(1 - (1 - (e^{-\lambda x})^2)^\theta\right)^\beta\right)}$$
(6)

## 2.2 Anderson–Darling Estimates (ADE)

The ADE was introduced by [30]. Applying ADE method for the TIHLTLExp distribution parameter  $\Omega = \beta, \theta, \lambda$

$$ADE_\Omega = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) \left\{ \log \left[ F_{TIHLTLExp} \left( x_{(i)}; \beta, \theta, \lambda \right) \right] + \log \left[ 1 - F_{TIHLTLExp} \left( x_{(n+1-i)}; \beta, \theta, \lambda \right) \right] \right\}$$
(7)

$$ADE_\Omega = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) \left[ \log \left[ \frac{1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^\theta \right]^\beta}{1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^\theta \right]^\beta} \right] + \log \left[ \frac{1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(n+1-i)}} \right)^2 \right]^\theta \right]^\beta}{1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(n+1-i)}} \right)^2 \right]^\theta \right]^\beta} \right] \right]$$
(8)

Thus, the estimates can be easily obtained by differentiating (8) with respect to.  $\beta, \theta$  and  $\lambda$  set the results to zero.

$$\frac{\delta ADE_\Omega}{\delta \beta} = -\frac{1}{n} \sum_{i=1}^n (2i-1) \left[ \frac{\varpi_i^{(p1)}(\beta, \theta, \lambda)}{\left[ F_{TIHLTLExp} \left( x_{(i)}; \beta, \theta, \lambda \right) \right]} - \frac{\varpi_{n+1-i}^{(p1)}(\beta, \theta, \lambda)}{\left[ 1 - F_{TIHLTLExp} \left( x_{(n+1-i)}; \beta, \theta, \lambda \right) \right]} \right] = 0$$
(9)

$$\frac{\delta ADE_\Omega}{\delta \theta} = -\frac{1}{n} \sum_{i=1}^n (2i-1) \left[ \frac{\varpi_i^{(p2)}(\beta, \theta, \lambda)}{\left[ F_{TIHLTLExp} \left( x_{(i)}; \beta, \theta, \lambda \right) \right]} - \frac{\varpi_{n+1-i}^{(p2)}(\beta, \theta, \lambda)}{\left[ 1 - F_{TIHLTLExp} \left( x_{(n+1-i)}; \beta, \theta, \lambda \right) \right]} \right] = 0$$
(10)

$$\frac{\delta ADE_\Omega}{\delta \lambda} = -\frac{1}{n} \sum_{i=1}^n (2i-1) \left[ \frac{\varpi_i^{(p3)}(\beta, \theta, \lambda)}{\left[ F_{TIHLTLExp} \left( x_{(i)}; \beta, \theta, \lambda \right) \right]} - \frac{\varpi_{n+1-i}^{(p3)}(\beta, \theta, \lambda)}{\left[ 1 - F_{TIHLTLExp} \left( x_{(n+1-i)}; \beta, \theta, \lambda \right) \right]} \right] = 0$$
(11)

where

$$\varpi_i^{(p1)}(\beta, \theta, \lambda) = \frac{2\beta \log \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^\theta \right] \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^\theta \right]}{\left[ 1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^\theta \right]^\beta \right] \left[ 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^\theta \right]^\beta \right]^3}$$
(12)

$$\varpi_i^{(p2)}(\beta, \theta, \lambda) = \frac{2\beta\theta \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^\theta \right]^\beta \log \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]}{\left[ 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^\theta \right]^\beta \right]^2} \quad (13)$$

$$\varpi_i^{(p3)}(\beta, \theta, \lambda) = \frac{4\beta\theta \left[ 1 - \left[ 1 - \left( e^{-\lambda x} \right)^2 \right]^\theta \right]^{2\beta-1} \left[ 1 - \left( e^{-\lambda x} \right)^2 \right]^{\theta-1} x e^{-\lambda x}}{\left[ 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x} \right)^2 \right]^\theta \right]^\beta \right]^2} \quad (14)$$

$$\varpi_{n+1-i}^{(p1)}(\beta, \theta, \lambda) = \frac{2\beta \log \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(n+1-i)}} \right)^2 \right]^\theta \right] \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(n+1-i)}} \right)^2 \right]^\theta \right]}{\left[ 1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(n+1-i)}} \right)^2 \right]^\theta \right]^\beta \right] \left[ 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(n+1-i)}} \right)^2 \right]^\theta \right]^\beta \right]^3} \quad (15)$$

$$\varpi_{n+1-i}^{(p2)}(\beta, \theta, \lambda) = \varpi_i^{(p2)}(\beta, \theta, \lambda) = \frac{2\beta\theta \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(n+1-i)}} \right)^2 \right]^\theta \right]^\beta \log \left[ 1 - \left( e^{-\lambda x_{(n+1-i)}} \right)^2 \right]}{\left[ 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(n+1-i)}} \right)^2 \right]^\theta \right]^\beta \right]^2} \quad (16)$$

$$\varpi_{n+1-i}^{(p3)}(\beta, \theta, \lambda) = \frac{4\beta\theta \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(n+1-i)}} \right)^2 \right]^\theta \right]^{2\beta-1} \left[ 1 - \left( e^{-\lambda x_{(n+1-i)}} \right)^2 \right]^{\theta-1} x_{(n+1-i)} e^{-\lambda x_{(n+1-i)}}}{\left[ 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(n+1-i)}} \right)^2 \right]^\theta \right]^\beta \right]^2}, \quad (17)$$

### 2.3 Cramér-von Mises Estimators. (CVM)

CVM was proposed [31]. The concept of this approach is to minimize the following function with respect to parameter  $\Omega = \beta, \theta, \lambda$ . The CVM distance function for TIHLTLExp distribution is defined by

$$CVM_\Omega = \frac{1}{12n} + \sum_{i=1}^n \left[ F_{TIHLTLExp} \left( x_{(i)}; \beta, \theta, \lambda \right) - \frac{2i-1}{2n} \right]^2 = \frac{1}{12n} + \sum_{i=1}^n \left[ \frac{\left( 1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x} \right)^2 \right]^\theta \right]^\beta \right)}{\left( 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x} \right)^2 \right]^\theta \right]^\beta \right)} - \frac{2i-1}{2n} \right]^2 \quad (18)$$

Thus, the estimates of the TIHLTLExp distribution parameter under CVM method is obtained by differentiating the (18) with respect to  $\beta, \theta$  and  $\lambda$  and set it to zero.

$$\frac{\delta CVM_{\Omega}}{\delta \beta} = 2 \sum_{i=1}^n \varpi_i^{(p1)}(\beta, \theta, \lambda) \left[ \frac{\left( \frac{1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}}{1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}} \right)}{\left( \frac{1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}}{1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}} \right)} \right] - \frac{2i-1}{2n} = 0 \quad (19)$$

$$\frac{\delta CVM_{\Omega}}{\delta \theta} = 2 \sum_{i=1}^n \varpi_i^{(p2)}(\beta, \theta, \lambda) \left[ \frac{\left( \frac{1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}}{1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}} \right)}{\left( \frac{1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}}{1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}} \right)} \right] - \frac{2i-1}{2n} = 0 \quad (20)$$

$$\frac{\delta CVM_{\Omega}}{\delta \lambda} = 2 \sum_{i=1}^n \varpi_i^{(p3)}(\beta, \theta, \lambda) \left[ \frac{\left( \frac{1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}}{1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}} \right)}{\left( \frac{1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}}{1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}} \right)} \right] - \frac{2i-1}{2n} = 0 \quad (21)$$

$\varpi_i^{(p1)}(\beta, \theta, \lambda)$ ,  $\varpi_i^{(p2)}(\beta, \theta, \lambda)$  and  $\varpi_i^{(p3)}(\beta, \theta, \lambda)$  are defined in (12), (13) and (14) respectively. The (31) provides more details

### 2.3 Maximum Product of Spacing (MPS)

The MPS approach of estimating the TIHLTLExp distribution parameters  $\Omega = \beta, \theta, \lambda$  are produced by maximizing the equations below with respect to the parameters:

$$MPS_{\Omega} = \frac{1}{n+1} \sum_{i=1}^{n+1} \log(G_i) \quad (22)$$

Where

$$G_i = F_{TIHLTLExp}(x_{(i)}; \beta, \theta, \lambda) - F_{TIHLTLExp}(x_{(i-1)}; \beta, \theta, \lambda)$$

$$F_{TIHLTLExp}(x_{(0)}; \beta, \theta, \lambda) = 0, \quad F_{TIHLTLExp}(x_{(n+1)}; \beta, \theta, \lambda) = 1$$

and

$$\sum_{i=1}^{n+1} G_i = 1$$

Thus the  $MPS_{\Omega}$  estimates are obtained by differentiating the equation (22) with respect to the parameters

where  $F_{TIHLTLExp}(x_{(i)}; \beta, \theta, \lambda)$  is the cdf of the TIHLTLExp distribution defined in (1)

## 2.4 Least Square Estimates (LSE)

LSE was introduced by [32]. The LSE of the TIHLTLExp distribution parameters  $\Omega = \beta, \theta, \lambda$  are obtained by minimizing, the equation below. The LSE function is defined by

$$LSE_{\Omega} = \sum_{i=1}^n \left[ F_{TIHLTLExp}(x_{(i)}; \beta, \theta, \lambda) - \frac{i}{n+1} \right]^2 = \sum_{i=1}^n \left[ \frac{\left( 1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta} \right)}{\left( 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta} \right)} - \frac{i}{n+1} \right]^2 \quad (23)$$

Thus, the LSE can be obtained by differentiating equation [23] with respect to the  $\beta, \theta$  and  $\lambda$ , and set it to zero

$$\frac{\delta LSE_{\Omega}}{\delta \beta} = 2 \sum_{i=1}^n \varpi_i^{(p1)}(\beta, \theta, \lambda) \left[ \frac{\left( 1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta} \right)}{\left( 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta} \right)} - \frac{i}{n+1} \right] = 0 \quad (24)$$

$$\frac{\delta LSE_{\Omega}}{\delta \theta} = 2 \sum_{i=1}^n \varpi_i^{(p2)}(\beta, \theta, \lambda) \left[ \frac{\left( 1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta} \right)}{\left( 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta} \right)} - \frac{i}{n+1} \right] = 0 \quad (25)$$

$$\frac{\delta LSE_{\Omega}}{\delta \lambda} = 2 \sum_{i=1}^n \varpi_i^{(p3)}(\beta, \theta, \lambda) \left[ \frac{\left( 1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta} \right)}{\left( 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta} \right)} - \frac{i}{n+1} \right] = 0 \quad (26)$$

$\varpi_i^{(p1)}(\beta, \theta, \lambda)$ ,  $\varpi_i^{(p2)}(\beta, \theta, \lambda)$  and  $\varpi_i^{(p3)}(\beta, \theta, \lambda)$  are defined in (12), (13) and (14) respectively.

## 2.5 Weighted Least Square Estimates (WLSE)

Similarly, the WLSE was introduced by [32]. The WLSE of the TIHLTLExp distribution parameters  $\Omega = \beta, \theta, \lambda$  are produced by minimizing the equation below with respect to the  $\beta, \theta$  and  $\lambda$ . The WLSE function is defined by

$$WLSE_{\Omega} = \sum_{i=1}^n \left( \frac{(n+1)^2 (n+1)}{i(n+1-i)} \right) \left[ F_{TIHLTLExp}(x_{(i)}; \beta, \theta, \lambda) - \frac{i}{n+1} \right]^2 \quad (27)$$

$$\frac{\delta WLSE_{\Omega}}{\delta \theta} = 2 \sum_{i=1}^n \varpi_i^{(p2)}(\beta, \theta, \lambda) \left[ \frac{\left( \frac{1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}}{1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}} \right)}{\left[ 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}} \right)} - \frac{i}{n+1} \right] = 0 \quad (28)$$

$$\frac{\delta WLSE_{\Omega}}{\delta \beta} = 2 \sum_{i=1}^n \varpi_i^{(p1)}(\beta, \theta, \lambda) \left[ \frac{\left( \frac{1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}}{1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}} \right)}{\left[ 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}} \right)} - \frac{i}{n+1} \right] = 0 \quad (29)$$

$$\frac{\delta WLSE_{\Omega}}{\delta \lambda} = 2 \sum_{i=1}^n \varpi_i^{(p3)}(\beta, \theta, \lambda) \left[ \frac{\left( \frac{1 - \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}}{1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}} \right)}{\left[ 1 + \left[ 1 - \left[ 1 - \left( e^{-\lambda x_{(i)}} \right)^2 \right]^{\theta} \right]^{\beta}} \right)} - \frac{i}{n+1} \right] = 0 \quad (30)$$

$\varpi_i^{(p1)}(\beta, \theta, \lambda)$ ,  $\varpi_i^{(p2)}(\beta, \theta, \lambda)$  and  $\varpi_i^{(p3)}(\beta, \theta, \lambda)$  are defined in (12), (13) and (14) respectively.

### III. Results

#### 3.1 Simulation study

Now, the performance of the MLE, ADE, CVM, MPS, LSE and WLSE method is investigated for TIHLTLExp parameters through Monte Carlo simulation study while considering 10,000 replications. Data were generated with different sample sizes (10,30,50,100,200). The estimates, Bias, Mean square error (MSE) and Mean relative estimate were obtained by R software. Thus, obtained as follows

**Table 1:** Estimates of different estimation methods for parameter lambda=1.5, theta=1 and beta=1

n	Estimation methods																	
	MLE			AD			CVM			MPS			LS			WLS		
10	1.65	1.71	1.45	1.63	1.39	1.33	1.80	2.28	1.48	1.42	1.18	1.14	1.57	1.55	1.29	1.55	1.38	1.34
30	1.46	1.19	1.22	1.46	1.17	1.21	1.50	1.21	1.23	1.43	1.07	1.07	1.48	1.18	1.19	1.36	1.15	1.28
50	1.42	1.13	1.20	1.42	1.11	1.18	1.49	1.15	1.17	1.45	1.06	1.05	1.46	1.12	1.16	1.29	1.10	1.33
100	1.40	1.08	1.17	1.40	1.07	1.17	1.45	1.09	1.15	1.47	1.04	1.04	1.44	1.08	1.14	1.27	1.08	1.31
200	1.40	1.06	1.14	1.41	1.05	1.14	1.43	1.06	1.14	1.48	1.03	1.03	1.43	1.05	1.11	1.24	1.05	1.32

**Table 2:** Bias of different estimation methods for parameter lambda=1.5, theta=1 and beta=1

n	Estimation methods																	
	MLE			AD			CVM			MPS			LS			WLS		
10	0.61	0.71	0.45	0.47	0.39	0.33	0.61	1.28	0.48	0.29	0.19	0.14	0.38	0.55	0.29	0.44	0.38	0.34
30	0.29	0.19	0.22	0.26	0.18	0.21	0.27	0.21	0.23	0.15	0.07	0.07	0.23	0.18	0.19	0.31	0.15	0.28
50	0.24	0.13	0.20	0.19	0.10	0.18	0.21	0.15	0.17	0.11	0.06	0.05	0.19	0.12	0.16	0.33	0.11	0.33
100	0.19	0.08	0.17	0.18	0.07	0.17	0.16	0.09	0.15	0.07	0.04	0.04	0.15	0.08	0.14	0.31	0.08	0.31
200	0.16	0.06	0.14	0.15	0.05	0.14	0.14	0.06	0.14	0.05	0.03	0.03	0.12	0.05	0.11	0.32	0.05	0.32



**Table 3:** Mean square error of different estimation methods for parameter  $\lambda=1.5$ ,  $\theta=1$  and  $\beta=1$

n	Estimation methods																				
	MLE			AD			CVM			MPS			LS			WLS					
10	2.22	2.57	0.81	1.02	0.81	0.45	3.26	2.18	0.92	0.21	0.26	0.10	0.55	0.59	0.39	1.01	1.18	0.43			
30	0.16	0.12	0.18	0.15	0.12	0.17	0.18	0.17	0.21	0.05	0.03	0.03	0.12	0.14	0.16	0.16	0.10	0.19			
50	0.11	0.05	0.14	0.08	0.04	0.11	0.10	0.08	0.13	0.02	0.02	0.01	0.09	0.06	0.11	0.17	0.04	0.23			
100	0.08	0.01	0.10	0.07	0.02	0.10	0.06	0.03	0.09	0.01	0.01	0.01	0.06	0.02	0.09	0.14	0.02	0.20			
200	0.06	0.01	0.08	0.06	0.01	0.07	0.05	0.01	0.08	0.01	0.00	0.00	0.04	0.01	0.06	0.15	0.01	0.19			

**Table 4:** Mean relative estimates of different estimation methods for parameter  $\lambda=1.5$ ,  $\theta=1$  and  $\beta=1$

n	Estimation methods																				
	MLE			AD			CVM			MPS			LS			WLS					
10	0.41	0.71	0.45	0.31	0.39	0.33	0.40	1.28	0.48	0.19	0.19	0.14	0.26	0.60	0.29	0.29	0.38	0.34			
30	0.20	0.19	0.22	0.18	0.18	0.21	0.18	0.21	0.23	0.10	0.07	0.07	0.16	0.18	0.19	0.21	0.15	0.28			
50	0.16	0.13	0.20	0.13	0.10	0.18	0.14	0.15	0.17	0.07	0.06	0.05	0.13	0.12	0.16	0.22	0.11	0.33			
100	0.13	0.08	0.17	0.12	0.07	0.17	0.10	0.09	0.15	0.05	0.04	0.04	0.10	0.08	0.14	0.21	0.08	0.31			
200	0.11	0.06	0.14	0.10	0.05	0.14	0.09	0.06	0.14	0.03	0.03	0.03	0.08	0.05	0.11	0.21	0.05	0.32			

**Table 5:** Mean square error ranking for different estimation methods for parameter  $\lambda=1.5$ ,  $\theta=1$  and  $\beta=1$

n	Estimation methods																				
	MLE			AD			CVM			MPS			LS			WLS					
10	5+5+6=16 <sup>5</sup>			4+4+3=11 <sup>4</sup>			6+6+5=17 <sup>6</sup>			1+1+1=3 <sup>1</sup>			2+2+2=6 <sup>2</sup>			3+3+4=10 <sup>3</sup>					
30	4.5+3.5+4=12 <sup>5</sup>			3+3.5+3=9.5 <sup>3</sup>			6+6+6=18 <sup>6</sup>			1+1+1=3 <sup>1</sup>			2+5+2=9 <sup>2</sup>			4.5+2+5=11.5 <sup>4</sup>					
50	5+3+4=12 <sup>4</sup>			2+1.5+2.5=6 <sup>2</sup>			4+5+4=13 <sup>5</sup>			1+1+1=3 <sup>1</sup>			3+4+2.5=9.5 <sup>3</sup>			6+1.5+6=13.5 <sup>6</sup>					
100	5+1.5+4.5=11 <sup>3.5</sup>			4+4+4.5=12.5 <sup>5</sup>			2.5+6+2.5=11 <sup>3.5</sup>			1+1.5+1=3.5 <sup>1</sup>			2.5+4+2.5=9 <sup>2</sup>			6+4+6=16 <sup>6</sup>					
200	4.5+4+4.5=13 <sup>5</sup>			4.5+4+3=11.5 <sup>3.5</sup>			3+4+4.5=11.5 <sup>3.5</sup>			1+1+1=3 <sup>1</sup>			2+4+2=8 <sup>2</sup>			6+4+6=16 <sup>6</sup>					

**Table 6:** Best estimation methods based on the Monte Carlo simulation study

Rank/n	Estimation methods				
	10	30	50	100	200
1 <sup>st</sup>	MPS	MPS	MPS	MPS	MPS
2 <sup>nd</sup>	LS	LS	AD	LS	LS
3 <sup>rd</sup>	WLS	AD	LS	MLE/CVM	AD/CVM
4 <sup>th</sup>	AD	WLS	MLE	MLE/CVM	AD/CVM
5 <sup>th</sup>	MLE	MLE	CVM	AD	MLE
6 <sup>th</sup>	CVM	CVM	WLS	WLS	WLS

#### IV. Discussion

Table 1-6 is the illustration of simulation study conducted. The six methods (MLE, ADE, CVM, MPS, LS, WLS) explored in this article. The Table 1 reveals various estimates for the TIHLTExp parameters across the six methods explored. The estimates of the estimation methods approach the true value of the parameters as the sample sizes increases. Table 2 illustrate the biases of the different methods explored, one can deduced that the biases reduces as the sample sizes increases. Table 3 illustrates the mean square error MSE, the MSE values decay as the sample sizes increases. It is evidenced that the Mean relative estimates of different estimation methods decay as the sample sizes increasing, this is illustrated in Table 4. It is evidence from the results that the six estimators possess consistency property.

The ranking of the performance of methods explored in this article is achieved in the Table 4. In Table 5, summation of the rank is done across the three parameters of the distribution. The preference of estimation methods is summarized in table 6 and the sample size are categorized as low (10,30), moderate (50) and finally high (100, 200). For the low, moderate and high sample sizes, the MPS is the best. The second best estimator for low and high sample sizes is LS and the second

best estimator for moderate sample size is AD. However, the worst estimator for low sample size is CVM, while the worst estimator for moderate and high sample sizes is WLS. Conclusively, since MPS outperform other estimation methods at low, moderate and high sample sizes, it is suggested that MPS should be adopted for analyzing the TIHLTLExp model. Alternatively, LS could be consider for estimating low and high sample size while AD for moderate sample size.

## References

- [1] Adepoju, A. A., Abdulkadir, S. S., & Jibasen, D. (2023). The Type I Half Logistics-Topp-Leone-G Distribution Family: Model, its Properties and Applications. *UMYU Scientifica*, 2(4), 09-22.
- [2] Isa A. M., Sule O. B., Ali B. A., Akeem A. A., and Ibrahim I. I. (2022). Sine-Exponential Distribution: Its Mathematical Properties and Application to Real Dataset. *UMYU Scientifica*, 1(1), 127 – 131.
- [3] Bello O., A., Doguwa S., I., Yahaya A., Jibril H., M. (2020). A Type I Half Logistic Exponentiated-G Family of Distributions: Properties and Application, *Communication in Physical Sciences*, 7(3):147-163.
- [4] Gul A, Sandhu A.J., Farooq M., Adil M., Hassan Y., Khan F. (2023) Half logistic-truncated exponential distribution: Characteristics and applications. *PLoS ONE* 18(11): e0285992.
- [5] Balogun, O. S., Arshad, M. Z., Iqbal, M. Z., & Ghamkhar, M. (2021). A new modified Lehmann type-II G class of distributions: exponential distribution with theory, simulation, and applications to the engineering sector. *F1000Research*, 10, 483
- [6] Ijaz, M., Asim, S. M. & Alamgir (2019). Lomax exponential distribution with an application to real life data. *Plos One* 14, 12, pp. e0225827,
- [7] Isa A. M., Kaigama A., Adepoju A. A., Bashiru S. O., Lehmann Type II-Lomax Distribution: Properties and Application to Real Data Set. (2023). *Communication in Physical Sciences*, 9(1):63 – 72
- [8] Jabeen, S. & Para, B. (2018). Exponentiated Gamma Exponential Distribution. *Sohag Journal of Mathematics: An International Journal*, 5, 3, pp. 79-84.
- [9] Elgarhy, M. Shakil, M & Kibria, B. M. G. (2017). Exponentiated Weibull Exponential distribution with application. *Application and Applied Mathematics: An International Journal*, 12, 2, pp. 710-725.
- [10] Kunjiratanachot, N.; Bodhisuwan, W.; Volodin, A. The Topp-Leone generalized exponential power series distribution with applications. (2018). *J. Probab. Stat. Sci.* 16, 197–208.
- [11] Almarashi, A. M., Elgarhy, M., Elsehetry, M. M., Kibria, B. M., & Algarni, A. (2019). A new extension of exponential distribution with statistical properties and applications. *Journal of Nonlinear Sciences & Applications (JNSA)*, 12(3).
- [12] Elgarhy M., Muhammad, A., & Ismat, P. (2018). Type II Half Logistic Exponential Distribution with Applications, *Annals of Data Science*.
- [13] Ristic, M. M., & Balakrishnan, N. (2011). The gamma-exponentiated exponential distribution. *Journal of Statistical Computation and Simulation*, 82, 1191-1206.
- [14] Bello O., A., Doguwa S., I., Yahaya A., Jibril H., M. (2021). A Type II Half Logistic Exponentiated-G Family of Distributions with Applications in Survival Analysis, *FUDMA Journal of Science*, 5(3):177-190.
- [15] Hassan E. A. A., Elgarhy M., Eldessouky E. A., Hassan O. H. M. Amin E. A. Almetwally, E. M. (2023). Different Estimation Methods for New Probability Distribution Approach Based on Environmental and Medical Data. *Axioms*. 12, 220.
- [16] Alotaibi B. N., Elbatal I., Shrahili M., Al-Moisheer A., Elgarhy M., and Almetwally E. M.

(2023). Statistical inference for the Kavya–Manoharan Kumaraswamy model under ranked set sampling with applications. *Symmetry*, 15:3, 587.

[17] Mukherjee I., Maiti S. S., Singh V. V. (2021). On estimation of the PDF and CDF of the one-parameter polynomial exponential family of distributions. *Communications in statistics- Theory and Methods*. 52:1104-120.

[18] Yilmaz A., Kara M., Ozdemir O. (2021). Comparison of different estimation methods for extreme value distribution. *Journal of applied statistics*. 48, 13-15.

[19] Adepoju A. A., Usman M., Alkassim R. S., Sani S. S. Adamu K.. Parameter (shape) Estimation of Weibull-Exponential Distribution Using Classical and Bayesian Approach Under Different Loss Functions. (2021). Royal Statistical Society Nigeria Local Group Conference Proceedings. 182-190

[20] Adepoju A. A., Issa A. A., Magaji A. A., Nasir M. S., Aliyu A. M., Preference of bayesian methods over classical method in estimating the scale parameter of inverse rayleigh frechet distribution. (2021). Royal Statistical Society Nigeria Local Group Conference Proceedings. 158-167.

[21] Artyushenko V. M., Volovach V. I. ((2021)). Methods of Estimating the Form of the Probability Distribution Density in Tasks of Processing Measurement Results. International Conference on Automatics and Energy (ICAE 2021) Journal of Physics: Conference Series. 2096 012136

[22] Thamer M. K., Zine R. (2021). Comparison of Five Methods to Estimate the Parameters for the Three-Parameter Lindley Distribution with Application to Life Data. *Hindawi Computational and Mathematical Methods in Medicine* Volume, Article ID 2689000, 14 pages

[23] ZeinEldin R. A., Chesneau C, Jamal F., Elgarhy M. (2019). Different estimation methods for Type I Half-Logistic Topp–Leone Distribution. *Mathematics*, 7, 985.

[24] Warsono, Gustavia E., Kurniasari D., Amanto, Antonio Y. (2019). On the Comparison of the Methods of Parameter Estimation for Pareto Distribution. *IOP Conf. Series: Journal of Physics: Conf. Series* 1338, 012042 e.

[25] Al-bayati R. S. S. (2018) Some Estimation Methods for Lindley Distribution M, [M.S. Thesis], College of Science, Mustansiriyah University, Baghdad, Iraq.

[26] Lei Y. (2008). Evaluation of three methods for estimating the Weibull distribution parameters of Chinese pine (*Pinus tabulaeformis*). *Journal of forest science*, 54, (12): 566–571

[27] Venkatraman J. J. S, S., and Wilson J. R., (1988) "Least-squares estimation of distribution functions in Johnson's translation system," *Journal of Statistical Computation and Simulation*, vol. 29, no. 4, pp. 271–297.

[28] Cheng R., Amin N. (1979). Maximum product of spacings estimation with application to the lognormal distribution, *Communications in Statistics-Theory and Methods*, Math. Report, p. 791.

[29] Macdonald P. (1971). An estimation procedure for mixtures of distributions. *Journal of the Royal Statistical Society*, vol. 20, pp. 102–107.

[30] Anderson, T.W.; Darling, D.A. (1952). Asymptotic theory of certain 'goodness-of-fit' criteria based on stochastic processes. *Ann. Math. Stat.* 23, 193–212.

[31] Macdonald, P.D.M. (1971). Comment on 'An estimation procedure for mixtures of distributions' by Choi and Bulgren. *J. R. Stat. Soc. B*, 33, 326–329.

[32] Swain, J.J.; Venkatraman, S.; Wilson, J.R. (1988). Least-squares estimation of distribution functions in johnson's translation system. *J. Stat. Comput. Simul.* 29, 271–297.

# ON THE $Q$ -RAYLEIGH DISTRIBUTION AND ITS APPLICATIONS

IBRAHIM SADOK



Department of Mathematics and Computer Science, Faculty of Exact Sciences,  
University of Bechar, Algeria  
ibrahim.sadok@univ-bechar.dz

## Abstract

*This paper introduces the two-parameter  $q$ -Rayleigh distribution, a powerful extension of the classical Rayleigh model for analysing real-world data. Compared to the Rayleigh, the  $q$ -Rayleigh incorporates a novel pathway parameter  $q$ , offering greater flexibility in capturing diverse data patterns. We delve into the mathematical properties of the  $q$ -Rayleigh, including its hazard rate function and quantile function, and explore parameter estimation through maximum likelihood methods. We demonstrate its superior fit compared to the widely-used Rayleigh distribution for real-world data. Moreover, we explore its application in reliability analysis. This comprehensive study makes the  $q$ -Rayleigh a compelling choice for modelling data exhibiting gradual transitions and enhanced flexibility.*

**Keywords:**  $q$ -Rayleigh distribution, Statistical properties, Parameter estimation, Modelling data

## 1. INTRODUCTION

The Rayleigh distribution, originally introduced by Rayleigh [13], is a notable probability distribution that serves as a specialized model and a modified variant of the Weibull distribution. Widely applicable across diverse disciplines, including medicine, engineering, finance, astronomy, and physics, the Rayleigh distribution has garnered significance due to its versatile utility in modelling various phenomena. Its pivotal role has led to extensive research, resulting in the proposal of several extensions by numerous scholars. Noteworthy examples include the truncated Rayleigh distribution, explored by Khalaf and Al-Kadim [8], and the Rayleigh Gamma-Gompertz distribution, investigated by Al-Noor and Asri [4]. Additionally, Rahman [12] introduced the Cubic Transformed Inverse Rayleigh distribution, and Adnan et al. [1] developed the Weibull Lindley Rayleigh distribution. These extensions and modifications reflect the adaptability and applicability of the Rayleigh distribution in different contexts. The probability density function (pdf) and cumulative distribution function (cdf) of the Rayleigh distribution are given respectively, by

$$f(x; \sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}; \quad \sigma > 0, x \geq 0 \quad (1)$$

$$F(x; \sigma) = 1 - e^{-\frac{x^2}{2\sigma^2}}; \quad \sigma > 0, x \geq 0 \quad (2)$$

The  $q$ -distribution, a concept integral to mathematical physics and probability theory, exhibits a broader generality compared to classical distributions. Originating from the pioneering work of Tsallis [19], the landscape of probability distributions has expanded significantly through the introduction of  $q$ -type distributions. This extension involves incorporating the  $q$  Tsallis

parameter, setting the stage for an extensive body of research on this topic. A notable array of  $q$ -type distributions has emerged as a result, showcasing the versatility of this concept. Notable examples include the  $q$ -exponential distributions proposed by Amari and Ohara [3],  $q$ -Gaussian distributions elucidated by Sato [16], and the  $q$ -Gamma distribution investigated by Zhang et al. [20]. Additionally,  $q$ -Weibull distributions have been introduced by researchers such as Picoli et al. [10]. These  $q$ -type distributions represent a rich and diverse set of mathematical formulations, contributing to the enhanced understanding and modelling of complex phenomena in various scientific disciplines. The cornerstone of  $q$ -type distributions is the  $q$ -exponential function:

$$\exp_q(x) = \begin{cases} [1 + (1 - q)x]^{\frac{1}{1-q}}, & 1 + (1 - q)x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

This function introduces a parameter  $q$ , that bestows a remarkable degree of adaptability in shaping the distribution, empowering it to effectively model non-trivial data patterns that often elude conventional approaches. Building upon this foundation, we introduce the  $q$ -Rayleigh distribution, a  $q$ -analogue poised to potentially expand the scope of modelling possibilities for intricate data relationships.

Recently, Gül [6] introduced the  $q$ -Rayleigh distribution for the case of  $q < 1$  and discussed the estimation of unknown parameters through maximum likelihood and least squares methods. In this paper, we extend the exploration of mathematical properties to two cases:  $q < 1$  and  $1 < q < 2$ . The analysis encompasses the survival function, hazard rate function, quantile function, limiting behaviour, and moments of the distribution. Furthermore, we delve into intriguing results concerning extreme value properties associated with the  $q$ -Rayleigh distribution. We employ the maximum likelihood estimator for parameter estimation in this new distribution. To assess its performance, we compare the  $q$ -Rayleigh distribution with the standard Rayleigh distribution using diverse real-life time data sets.

The rest of the paper is organised as follows. Section 2 introduces the novel  $q$ -Rayleigh distribution, providing a comprehensive exploration of its specific cases. Section 3 delves into the mathematical and statistical properties of this distribution, elucidating its asymptotic behaviours. Section 4, meticulously elucidates the method of maximum likelihood estimation. In Section 5, we employ the newly proposed model on two distinct datasets concerning the treatment of head and neck cancer patients with radiation plus chemotherapy, as well as COVID-19 mortality rates data from Italy. A comparative analysis with the  $q$ -Rayleigh and Rayleigh models is conducted, affirming the superior fit of the  $q$ -Rayleigh model. The conclusive Section brings together the findings, summarizing the key insights and implications derived from the exploration of the innovative  $q$ -Rayleigh distribution.

## 2. THE $q$ -RAYLEIGH DISTRIBUTION

### 2.1. Distributional characteristics

The pdf of the  $q$ -Rayleigh distribution is defined as

$$f_q(x) = (2 - q) \frac{x}{\sigma^2} \exp_q \left[ -\frac{x^2}{2\sigma^2} \right], \quad x > 0 \quad (4)$$

where  $\sigma > 0$  and  $q < 2$  are shape parameters, and  $\eta > 0$  is a scale parameter.

By introducing  $\beta = \sigma^{-2}$  and using  $\exp_q(x)$  in equation (3), the pdf of the  $q$ -Rayleigh distribution, for  $x > 0$  and for  $q < 1$ , can be rewritten as

$$f_q(x) = (2 - q)\beta x \left[ 1 - (1 - q) \frac{\beta x^2}{2} \right]^{\frac{1}{1-q}}, \quad q < 1 \text{ and } x \in \left[ 0, \left( \frac{\beta}{2} (1 - q) \right)^{-1/2} \right] \quad (5)$$

For  $x > 0$  and  $q > 1$ , the pdf of the  $q$ -Rayleigh distribution is expressed as:

$$f_q(x) = (2 - q)\beta x \left[ 1 + (q - 1)\frac{\beta x^2}{2} \right]^{-\frac{1}{q-1}}, \quad 1 < q < 2 \text{ and } x \in [0, +\infty) \quad (6)$$

The cumulative distribution function (cdf) of the  $q$ -Rayleigh distribution, when  $q < 1$  is defined as

$$F_q(x) = 1 - \left[ 1 - (1 - q)\frac{\beta x^2}{2} \right]^{\frac{2-q}{1-q}} \quad (7)$$

If  $1 < q < 2$ , the cdf function of the  $q$ -Rayleigh distribution, formulated as follows:

$$F_q(x) = 1 - \left[ 1 + (q - 1)\frac{\beta x^2}{2} \right]^{\frac{q-2}{q-1}} \quad (8)$$

## 2.2. Survival function

In the context of the  $q$ -Rayleigh distribution, the survival function (sf), denoted by  $S(x)$ , represents the probability that an individual or entity survives beyond time  $t$ . Its mathematical expression is as follows

$$S(x) = P(X > t) = 1 - F(x)$$

$$S_q(x) = \left[ 1 - (1 - q)\frac{\beta x^2}{2} \right]^{\frac{2-q}{1-q}}, \quad \text{for } q < 1,$$

$$S_q(x) = \left[ 1 + (q - 1)\frac{\beta x^2}{2} \right]^{\frac{2-q}{1-q}}, \quad \text{for } 1 < q < 2$$

## 2.3. Hazard function

The concept of risk within the context of survival analysis is characterized by the hazard rate function (hrf),  $h(x)$ . This function measures the immediate risk of an event (e.g., death) for an individual who has survived until that time. Its formal representation is as follows

$$h(x) = P(X > t) = \frac{f(x)}{S(x)}$$

The hrf of  $q$ -Rayleigh distribution for  $q < 1$  is defined as

$$h_q(x) = \frac{(2 - q)\beta x}{1 - (1 - q)\frac{\beta x^2}{2}}$$

In the case of  $1 < q < 2$ , the hrf of  $q$ -Rayleigh distribution is characterized by

$$h_q(x) = \frac{(2 - q)\beta x}{1 + (q - 1)\frac{\beta x^2}{2}}$$

## 2.4. Cumulative hazard function

The probability of an event occurring before a given time is quantified by the cumulative hazard function (chf), presented below

$$H(x) = -\ln(1 - F(x))$$

The chf for the  $q$ -Rayleigh distribution, with  $q < 1$ , is expressed as follows

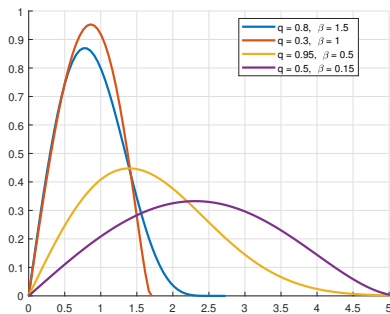
$$H_q(x) = \frac{2 - q}{q - 1} \ln \left[ 1 - (1 - q)\frac{\beta x^2}{2} \right].$$

For the case where  $1 < q < 2$ , the chf of the  $q$ -Rayleigh distribution is given by

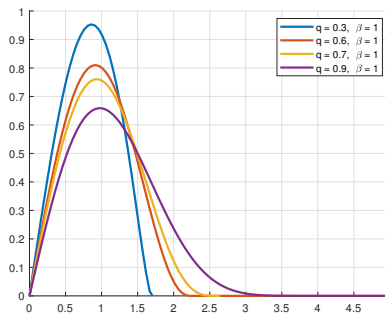
$$H_q(x) = \frac{2-q}{q-1} \ln \left[ 1 + (q-1) \frac{\beta x^2}{2} \right]$$

### 2.5. Graphical Study of $q$ -Rayleigh distribution under various functions

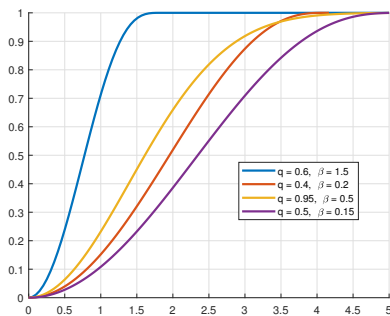
Driven by a desire to understand the nuanced behaviour of the  $q$ -Rayleigh distribution, we embark on a detailed exploration of its key functions (pdf, cdf, sf, and hrf) across a range of parameter values. By meticulously analysing the illustrative figures presented below, we uncover fascinating insights into how varying parameters sculpt the behaviour of this versatile distribution. Complementing our theoretical exploration, we presented illustrative figures to visually depict the distribution's characteristics, enhancing accessibility and understanding.



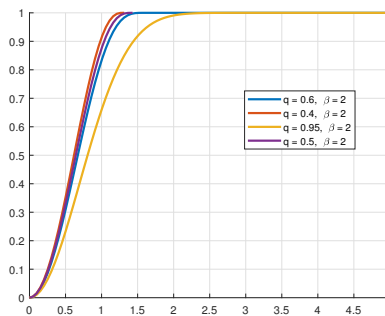
**(a)** Graph of the pdf of the  $q$ -Rayleigh distribution when all the parameters are changed



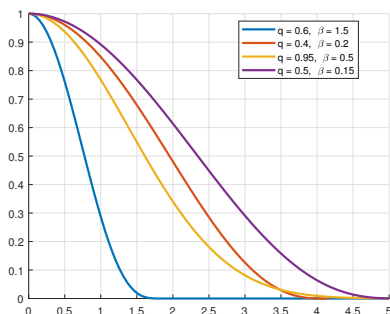
**(b)** Graph of the pdf of the  $q$ -Rayleigh distribution when changing the  $q$  values and  $\beta$  is fixed



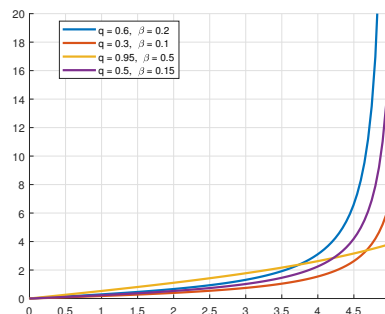
**(c)** Graph of the cdf of the  $q$ -Rayleigh distribution when all the parameters are changed



**(d)** Graph of the cdf of the  $q$ -Rayleigh distribution when changing the  $q$  values and  $\beta$  is fixed

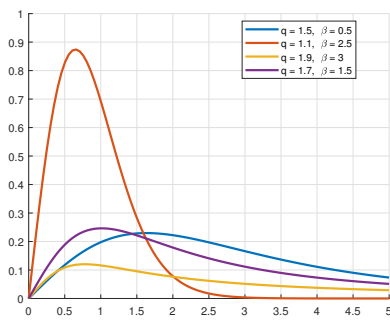


**(e)** Graph of the sf of the  $q$ -Rayleigh distribution with different parameter values

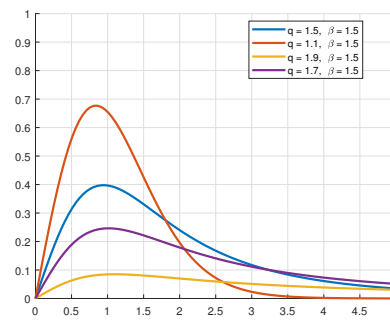


**(f)** Graph of the hrf of the  $q$ -Rayleigh distribution with different parameter values

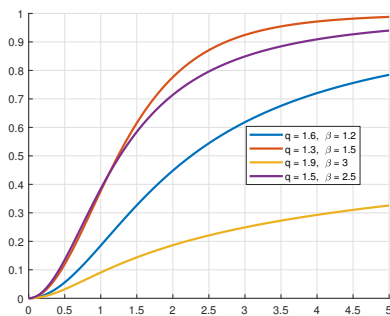
**Figure 1:** Graphical representation of the key functions of the  $q$ -Rayleigh distribution:  $q < 1$



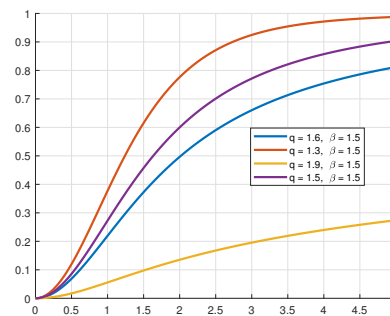
**(a)** Graph of the pdf of the  $q$ -Rayleigh distribution when all the parameters are changed



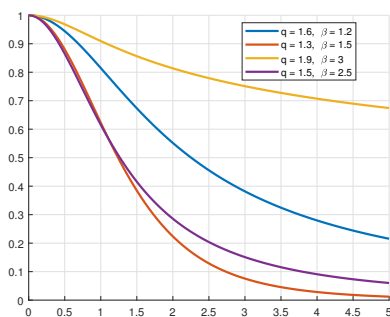
**(b)** Graph of the pdf of the  $q$ -Rayleigh distribution when changing the  $q$  values and  $\beta$  is fixed



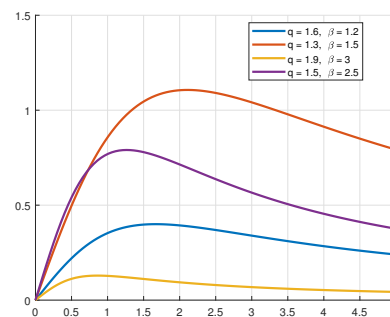
**(c)** Graph of the cdf of the  $q$ -Rayleigh distribution when all the parameters are changed



**(d)** Graph of the cdf of the  $q$ -Rayleigh distribution when changing the  $q$  values and  $\beta$  is fixed



**(e)** Graph of the sf of the  $q$ -Rayleigh distribution with different parameter values



**(f)** Graph of the hrf of the  $q$ -Rayleigh distribution with different parameter values

**Figure 2:** Graphical representation of the key functions of the  $q$ -Rayleigh distribution:  $1 < q < 2$

Figures 1 and 2 showcase the graphical representation of the key functions of the  $q$ -Rayleigh distribution for cases where  $q < 1$  and  $1 < q < 2$ , respectively. Examining the probability density function graphs (1(a), 1(b), 7(a), and 7(b)), it becomes evident that the distribution exhibits skewness and a high degree of adaptability to diverse parameter values.

In Figures 1(c), 1(d), 2(c) and 2(d), we observe cumulative density plots that serve to validate the distribution's suitability as a probability distribution. Additionally, Figures 1(e) and 2(e) portray the survival function, revealing distinct patterns of fast and slow decreases. The hazard rate function graphs (1(f), 2(f)) further contribute to the distribution's versatility, showcasing a range of shapes including increasing, decreasing, and constant. This variability allows for the effective fitting of datasets with diverse forms, a characteristic that the  $q$ -Rayleigh distribution adeptly demonstrates. In essence, our exploration underscores the distribution's capability to accommodate different data sets, making it a valuable tool in statistical analysis.



### 3. PROPERTIES

This section delves into the mathematical and statistical characteristics of the  $q$ -Rayleigh distribution.

#### 3.1. Limiting Behaviour

**Lemma 1.** As the parameter  $q$  approaches 1, the pdf of the  $q$ -Rayleigh distribution, (denoted as  $f_q(x)$ ), converges to the standard Rayleigh distribution.

**Proof.** For  $q < 1$ , the limiting pdf for  $q = 1$  is

$$\begin{aligned} \lim_{q \rightarrow 1} f_q(x) &= \beta x \lim_{q \rightarrow 1} \left\{ \left[ 1 - (1-q) \frac{\beta x^2}{2} \right]^{\frac{-1}{(1-q)\beta x^2/2}} \right\}^{-\beta x^2/2} \\ &= \beta x \exp\left(-\frac{\beta x^2}{2}\right) \end{aligned}$$

a Rayleigh pdf.

The established proof methodology can be directly applied to the range  $1 < q < 2$ , yielding an analogous conclusion. ■

#### 3.2. Quantile Function

The quantile function of  $X$ , denoted as  $Q(u)$  and defined as  $Q(u) = F^{-1}(u)$ , can be derived by inversely solving equations (7) and (8) as follows

$$\begin{aligned} Q_q(x) &= \left[ \frac{2}{\beta(1-q)} \left( 1 - (1-u)^{\frac{1-q}{2-q}} \right) \right]^{\frac{1}{2}}, \text{ for } q < 1, \\ Q_q(x) &= \left[ \frac{2}{\beta(q-1)} \left( -1 + (u-1)^{\frac{1-q}{2-q}} \right) \right]^{\frac{1}{2}}, \text{ for } 1 < q < 2 \end{aligned}$$

#### 3.3. Moments

This section presents the moment function for the  $q$ -Rayleigh distribution, where moments serve as quantitative indicators associated with the function's shape. The moments of the  $q$ -Rayleigh distribution can be derived as follows:

$$E(X^s) = \int_0^{+\infty} x^s f_q(x) dx$$

If  $q < 1$ ,

$$\begin{aligned} E(X^s) &= \int_0^{\left(\frac{\beta}{2}(1-q)\right)^{-1/2}} x^s (2-q)\beta x \left[ 1 - (1-q) \frac{\beta x^2}{2} \right]^{\frac{1}{1-q}} dx \\ &= \frac{2-q}{(1-q)^{1+s/2}} \left(\frac{2}{\beta}\right)^{s/2} B\left(\frac{1+s}{1-q}, 1 + \frac{s}{2}\right) \end{aligned}$$

where,

$$B(p, q) = \int_0^1 t^{p-1} (1-t)^{q-1} dt = \int_0^{+\infty} \frac{t^{p-1}}{(1+t)^{p+q}} dt$$

denotes the beta function. It follows that the mean and variance of the  $q$ -Rayleigh random variable when  $q < 1$  are

$$E(X) = \frac{2-q}{(1-q)^{1+1/2}} \left(\frac{2}{\beta}\right)^{1/2} B\left(\frac{2}{1-q}, \frac{3}{2}\right)$$

$$Var(X) = \frac{2(2-q)}{(1-q)\beta} \left[ B\left(\frac{3}{1-q}, 2\right) - \frac{2-q}{(1-q)^{3/2}} B^2\left(\frac{2}{1-q}, \frac{3}{2}\right) \right]$$

If  $1 < q < 2$ ,

$$E(X^s) = \int_0^{+\infty} x^s (2-q)\beta x \left[ 1 + (q-1) \frac{\beta x^2}{2} \right]^{-\frac{1}{q-1}} dx$$

$$= \frac{2-q}{(q-1)^{1+s/2}} \left(\frac{2}{\beta}\right)^{s/2} B\left(\frac{1}{q-1} - \frac{s}{2} - 1, \frac{s}{2} + 1\right)$$

provided  $\frac{1}{q-1} - \frac{s}{2} > 1$ . Consequently, the mean and variance of the  $q$ -Rayleigh random variable can be expressed as follows

$$E(X) = \frac{2-q}{(q-1)^{1+1/2}} \left(\frac{2}{\beta}\right)^{1/2} B\left(\frac{1}{q-1} - \frac{3}{2}, \frac{3}{2}\right)$$

$$Var(X) = \frac{2(2-q)}{(q-1)\beta} \left[ B\left(\frac{1}{q-1} - 2, 2\right) - \frac{2-q}{(q-1)^{3/2}} B^2\left(\frac{1}{q-1} - \frac{3}{2}, \frac{3}{2}\right) \right]$$

### 3.4. Extreme value properties

**Theorem 1.** Let  $\{X_i, i = 1, \dots, n\}$  be independent and identically distributed random variables (r.v.) following the  $q$ -Rayleigh distribution, then  $U = \min_{1 \leq i \leq n} X_i$  has also the same distributional form.

**Proof.** For  $q < 1$  the survival function is  $S_q(x) = \left[ 1 - (1-q) \frac{\beta x^2}{2} \right]^{\frac{2-q}{1-q}}$ . Then,

$$S_q(x) = P \left[ \min_{1 \leq i \leq n} X_i > x \right]$$

$$= \prod_{i=1}^n P [X_i > x]$$

$$= \prod_{i=1}^n \left[ 1 - (1-q) \frac{\beta x^2}{2} \right]^{\frac{2-q}{1-q}}$$

$$= \left[ 1 - (1-q) \frac{\beta x^2}{2} \right]^{n \frac{2-q}{1-q}} \rightarrow e^{-n \frac{\beta x^2}{2}} \text{ as } q \rightarrow 1$$

For  $1 < q < 2$  the survival function is  $S_q(x) = \left[ 1 + (q-1) \frac{\beta x^2}{2} \right]^{\frac{2-q}{1-q}}$ . Then,

$$S_q(x) = \left[ 1 + (q-1) \frac{\beta x^2}{2} \right]^{-n \frac{2-q}{q-1}} \rightarrow e^{-n \frac{\beta x^2}{2}} \text{ as } q \rightarrow 1$$

■

**Theorem 2.** Let  $\{X_i, i = 1, \dots, n\}$  be independent and identically distributed random variables (r.v.) following the  $q$ -Rayleigh distribution, then  $V = \max_{1 \leq i \leq n} X_i$  has also the same distributional form.

**Proof.** For  $q < 1$  the cdf is  $F_q(x) = 1 - \left[1 - (1 - q)\frac{\beta x^2}{2}\right]^{\frac{2-q}{1-q}}$ . Then,

$$\begin{aligned} F_q(x) &= P\left[\max_{1 \leq i \leq n} X_i \leq x\right] \\ &= \prod_{i=1}^n P[X_i \leq x] \\ &= \prod_{i=1}^n \left[1 - \left[1 - (1 - q)\frac{\beta x^2}{2}\right]^{\frac{2-q}{1-q}}\right] \\ &= \left[1 - \left[1 - (1 - q)\frac{\beta x^2}{2}\right]^{\frac{2-q}{1-q}}\right]^n \rightarrow \left[1 - e^{-\frac{\beta x^2}{2}}\right]^n \text{ as } q \rightarrow 1 \end{aligned}$$

Similarly for  $1 < q < 2$  the cdf of  $V$  is

$$F_q(x) = \left[1 - \left[1 + (q - 1)\frac{\beta x^2}{2}\right]^{\frac{q-2}{q-1}}\right]^n \rightarrow \left[1 - e^{-\frac{\beta x^2}{2}}\right]^n \text{ as } q \rightarrow 1$$

■

#### 4. ESTIMATION OF PARAMETERS

This section explores the estimation of the unknown parameters in the  $q$ -Rayleigh distribution through the application of the maximum likelihood estimation method (MLE).

Let  $x_1, x_2, \dots, x_n$  represent a random sample obtained from the  $q$ -Rayleigh distribution. The subsequent expression outlines the logarithm of the likelihood function corresponding to the pdf represented in equation (5) for  $q < 1$  is

$$\ln L = n \ln(2 - q) + n \ln \beta + \sum_{i=1}^n \ln(x_i) + \frac{1}{1 - q} \sum_{i=1}^n \ln\left(1 - (1 - q)\frac{\beta x_i^2}{2}\right) \quad (9)$$

The maximum likelihood estimates of the parameters  $(q, \beta)$  are found by taking a partial derivative of  $\ln L$  with respect to  $q$  and  $\beta$ , equating the derivatives to zero, and evaluating them at  $\hat{q}, \hat{\beta}$

$$\begin{aligned} \frac{\partial \ln L}{\partial q} &= -\frac{n}{2 - q} + \frac{1}{(1 - q)^2} \sum_{i=1}^n \ln\left(1 - (1 - q)\frac{\beta x_i^2}{2}\right) + \frac{1}{1 - q} \sum_{i=1}^n \frac{\beta x_i^2}{2 - (1 - q)\beta x_i^2} \\ \frac{\partial \ln L}{\partial \beta} &= \frac{n}{\beta} - \sum_{i=1}^n \frac{x_i^2}{2 - (1 - q)\beta x_i^2} \end{aligned}$$

In the range where  $1 < q < 2$ , the log-likelihood corresponding to the pdf in equation (6) takes the form

$$\ln L = n \ln(2 - q) + n \ln \beta + \sum_{i=1}^n \ln(x_i) - \frac{1}{q - 1} \sum_{i=1}^n \ln\left(1 + (q - 1)\frac{\beta x_i^2}{2}\right) \quad (10)$$

Upon differentiating the log-likelihood function in terms of the parameters  $q$  and  $\beta$ , one obtains the following expressions:

$$\begin{aligned} \frac{\partial \ln L}{\partial q} &= -\frac{n}{2 - q} + \frac{1}{(q - 1)^2} \sum_{i=1}^n \ln\left(1 + (q - 1)\frac{\beta x_i^2}{2}\right) - \frac{1}{q - 1} \sum_{i=1}^n \frac{\beta x_i^2}{2 + (q - 1)\beta x_i^2} \\ \frac{\partial \ln L}{\partial \beta} &= \frac{n}{\beta} - \sum_{i=1}^n \frac{x_i^2}{2 + (q - 1)\beta x_i^2} \end{aligned}$$

The partial derivatives of the log-likelihood function with respect to  $q$  and  $\beta$  are non-linear in both cases ( $q < 1$  and  $1 < q < 2$ ). This non-linearity poses a challenge for directly finding closed-form solutions for the MLEs of  $q$  and  $\beta$ . While closed-form solutions involve expressing the estimates as explicit mathematical expressions in terms of the data, numerical optimization methods often involve iterative algorithms to find approximate solutions.

$$\begin{aligned} \max \quad & \ln L \\ \text{s.t.} \quad & q < 2, \\ & \beta > 0, \end{aligned} \tag{11}$$

Despite theoretical challenges in rigorously proving the uniqueness of the solution to optimization problem (11), empirical evidence suggests a strong case for its singularity. Employing a specific optimization algorithm across a wide range of initial parameter values consistently yielded convergence to the same solution, demonstrating remarkable robustness and providing compelling support for uniqueness in practical applications. While a formal proof remains elusive, this robust empirical evidence bolsters the validity of the solution for practical applications within this domain.

### 5. APPLICATION TO REAL LIFE DATA

In this section, we have employed various sets of real-life failure time data to demonstrate the appropriateness of the  $q$ -Rayleigh distribution. Additionally, we have conducted a comparative analysis with the conventional Rayleigh distribution, highlighting the advantages and nuances of our proposed model. This exploration not only showcases the versatility of the  $q$ -Rayleigh distribution but also provides valuable insights into its performance in comparison to the widely accepted standard Rayleigh distribution.

To assess the flexibility of the proposed distribution, we utilized several model selection criteria, such as -log-likelihood (-LL), Kolmogorov–Smirnov (KS) statistics, and associated  $p$ -values. The analyses were carried out using Matlab software. It is important to note that a superior distribution is identified by smaller values of -LL and KS statistics. Additionally, a more favourable distribution, particularly in terms of  $p$ -values, is characterized by a significance level that aligns with the chosen threshold ( $<0.005$ ), further contributing to the comprehensive evaluation of the proposed distribution’s fit to the data.

**Dataset 1:** In medical research, the assessment and comparison of treatment regimens are commonplace. A deeper comprehension of cancer genetics has broadened the spectrum of treatment options for various cancers falling under the umbrella of head and neck cancers, including those affecting the oral cavity, throat, larynx, para-nasal sinuses, and salivary glands. The three primary types of cancer treatments encompass primary, adjuvant, and palliative approaches. Within these categories, diverse treatment regimens such as surgery, radiation, chemotherapy, hormone therapy, immune therapy, and targeted drug therapy are employed.

Efron [5] conducted a randomized clinical trial comparing two treatment arms for head and neck cancer patients: radiation therapy alone (Arm A) and radiation plus chemotherapy (Arm B). The study recorded survival times (in days) for 51 patients in Arm A and 44 patients in Arm B. In this investigation, we specifically focus on the data from Arm B, examining the appropriateness of fitting the data to the  $q$ -Rayleigh distribution. The results are subsequently juxtaposed with those obtained using the standard Rayleigh distribution for a comprehensive evaluation.

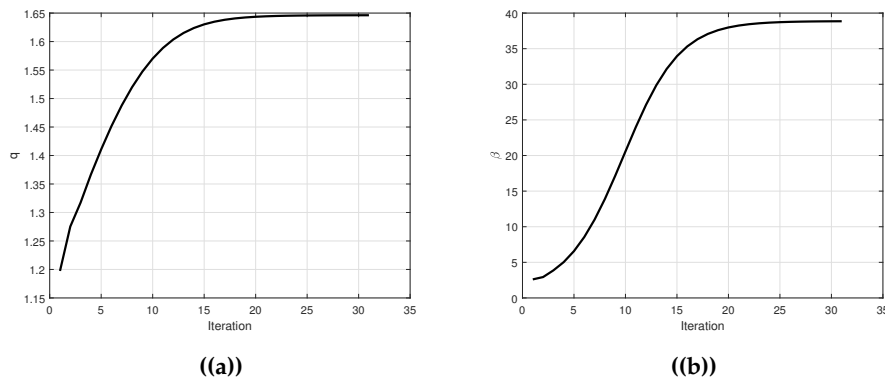
37	84	92	94	112	119	127	130	133	140	146
155	159	169	173	179	194	195	209	249	281	319
339	432	469	519	528	547	613	633	725	759	817
1092	1245	1331	1557	1771	1776	1897	2023	2146	2297	

**Table 1:** Database of Arm B (Sample size 44).

Model	Estimated Parameters		Model Selection		
	$\hat{q}$	$\hat{\beta}$	-LL	KS	$p$ -value
$q$ -Rayleigh	1.6462	38.8526	37.1631	0.10769	0.033834
Rayleigh	-	1.9505	69.9683	0.50856	$1.8308 \times 10^{-6}$

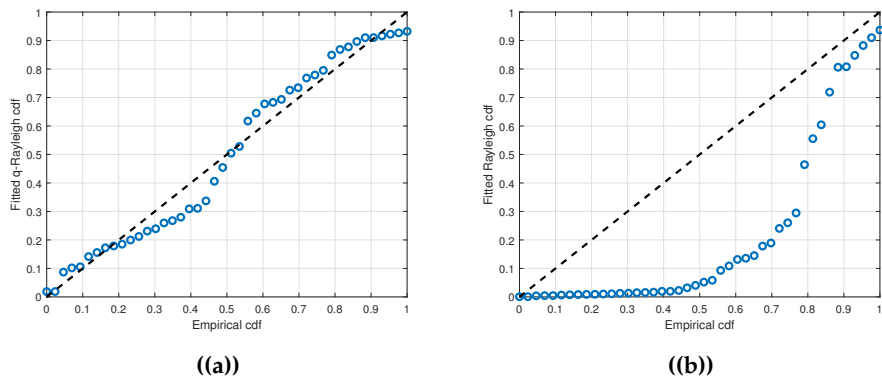
**Table 2:** Estimates of fitted distribution for Arm B data.

Table 2 outlines estimates for fitted distributions of Arm B data, comparing the  $q$ -Rayleigh and Rayleigh models. The negative log-likelihood values, are substantially lower for the  $q$ -Rayleigh model than for the Rayleigh model, suggesting superior fit for the former. Additionally, the KS statistic is smaller for the  $q$ -Rayleigh model compared to the Rayleigh model, reinforcing the notion that the former provides a more accurate representation of the data. The associated  $p$ -value for the KS statistic is also notably smaller for the  $q$ -Rayleigh model, underscoring its statistical significance in capturing the observed data distribution.



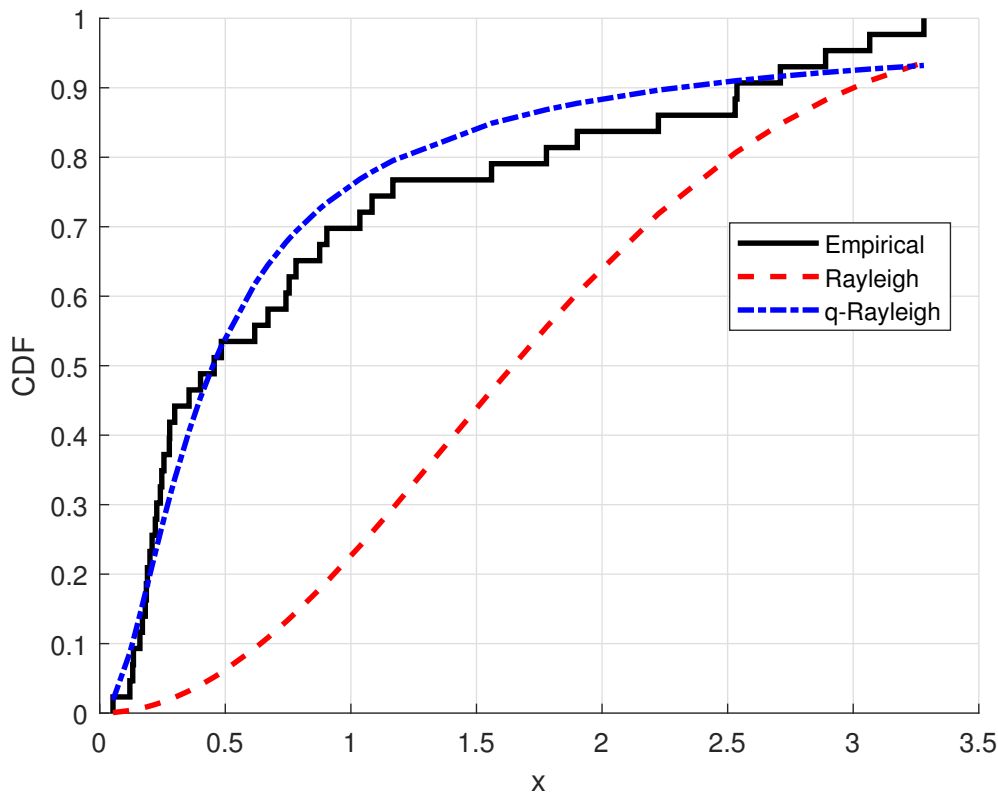
**Figure 3:** Convergence of Newton-Raphson Method for Parameters Estimations  $q$  (a) and  $\beta$  (b) for Arm B data.

Figure 3 illustrates the convergence of the Newton-Raphson method for parameter estimations of  $q$  and  $\beta$ . The convergence is achieved within 31 iterations.



**Figure 4:** PP plot for fitted  $q$ -Rayleigh (a) and Rayleigh (b) for Arm B data.

Figure 4 represents the Probability-Probability (PP) plot for fitted  $q$ -Rayleigh (a) and Rayleigh distribution. the PP plot for the  $q$ -Rayleigh model provides a more accurate representation of the data which implies that the former is considered better than that of the Rayleigh model. A visually superior alignment of points along the line in the PP plot for the  $q$ -Rayleigh model compared to the Rayleigh model indicates that the former better captures the distributional characteristics of the data, reinforcing the notion that the  $q$ -Rayleigh model is a more suitable fit for the observed dataset.



**Figure 5:** Empirical, Rayleigh, and  $q$ -Rayleigh cdf's for Arm B data.

In Figure 5, the cdf's of the Empirical, Rayleigh, and  $q$ -Rayleigh models are presented. A superior fit for the  $q$ -Rayleigh model is suggested when examining these cdf's, signifying its enhanced capability to accurately represent the observed data in comparison to the conventional Rayleigh model. This might be evidenced by a closer alignment of the  $q$ -Rayleigh cdf to the empirical cdf, suggesting that the additional parameter  $q$  improves the model's ability to capture the nuances in the data distribution.

**Dataset 2:** Authentic data pertaining to COVID-19 mortality rates in Italy is utilized to assess the goodness of fit of the  $q$ -Rayleigh distribution. The dataset spans a period of 59 days, commencing from February 27 to April 27, 2020, capturing the temporal evolution of mortality rates during this critical period. The detailed information, including date-specific mortality rates, is organized and presented in Table 4, forming the basis for conducting a rigorous statistical analysis to evaluate the appropriateness of the  $q$ -Rayleigh distribution in modelling the observed COVID-19 mortality trends in Italy.

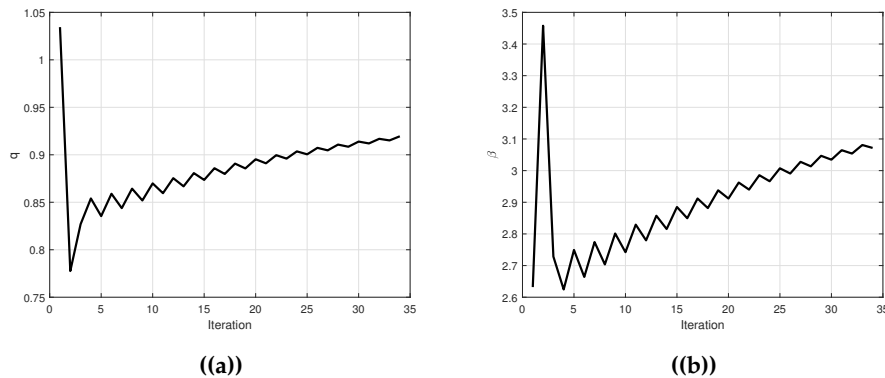
4.571	7.201	3.606	8.479	11.410	8.961	10.919	10.908	6.503	18.474	11.010	17.337
16.561	13.226	15.137	8.697	15.787	13.333	11.822	14.242	11.273	14.330	16.046	11.950
10.282	11.775	10.138	9.037	12.396	10.644	8.646	8.905	8.906	7.407	7.445	7.214
6.194	4.640	5.452	5.073	4.416	4.859	4.408	4.639	3.148	4.040	4.253	4.011
3.564	3.827	3.134	2.780	2.881	3.341	2.686	2.814	2.508	2.450	1.518	

**Table 3:** COVID-19 Data in Italy from February 27 to April 27, 2020.

Model	Estimated Parameters		Model Selection		
	$\hat{q}$	$\hat{\beta}$	-LL	KS	$p$ -value
$q$ -Rayleigh	0.91949	3.0717	18.7225	0.14996	0.0068288
Rayleigh	-	3.6895	18.7643	0.7544	$1.2615 \times 10^{-22}$

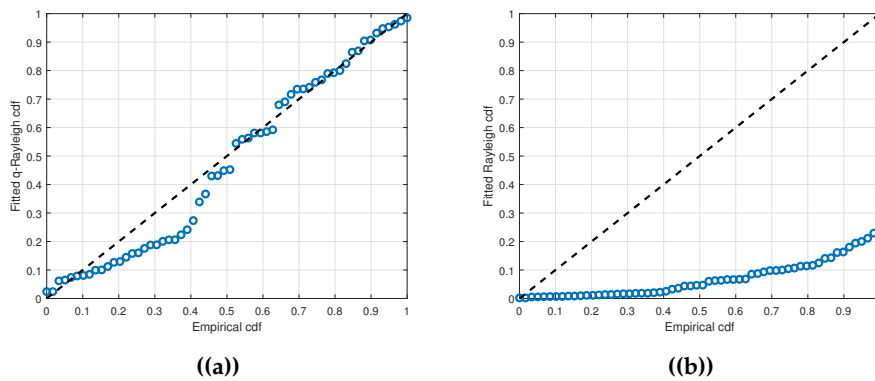
**Table 4:** Estimates of fitted distribution for COVID-19 data.

Table 4 compares two distribution models applied to COVID-19 data: the  $q$ -Rayleigh and Rayleigh distributions. The  $q$ -Rayleigh model exhibits a lower negative log-likelihood value and a smaller KS statistic compared to the Rayleigh model. Additionally, the  $q$ -Rayleigh model has a notably lower  $p$ -value, indicating a better fit to the observed COVID-19 data. These collective indicators of model performance suggest the superiority of the  $q$ -Rayleigh distribution in capturing the underlying distribution of the COVID-19 dataset during the specified period.



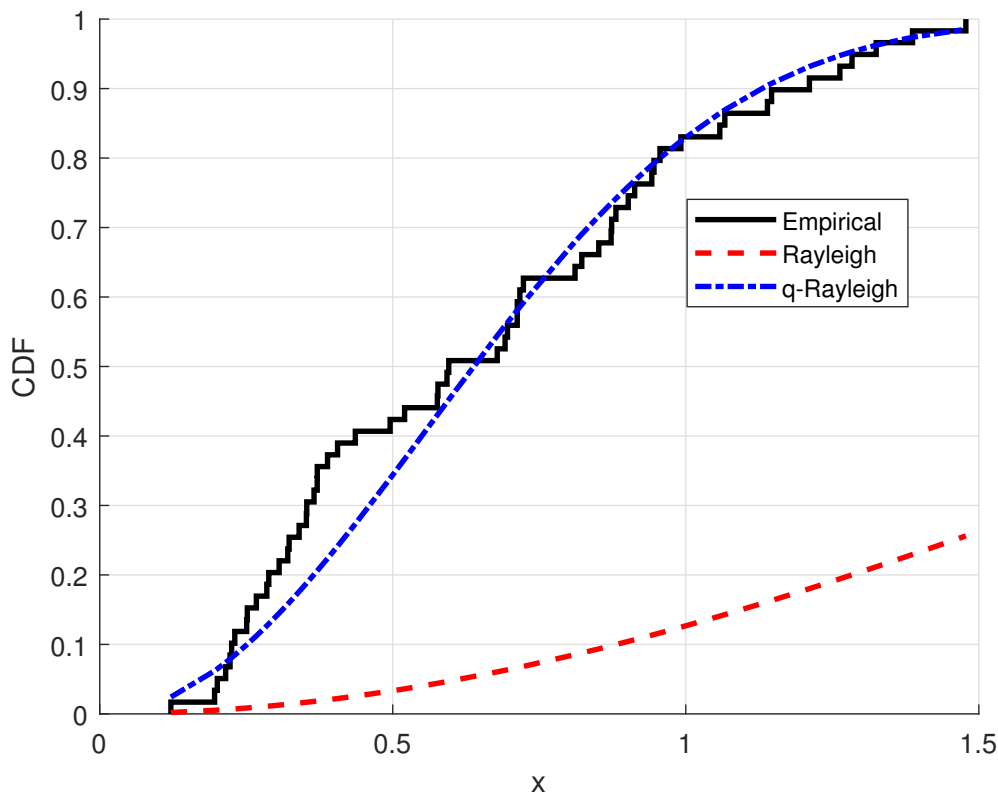
**Figure 6:** Convergence of Newton-Raphson Method for Parameters Estimations  $q$  (a) and  $\beta$  (b) for COVID-19 data.

Figure 6 demonstrates the convergence of the Newton-Raphson method in estimating the parameters  $q$  and  $\beta$ . The convergence is successfully attained after 34 iterations.



**Figure 7:** PP plot for fitted  $q$ -Rayleigh (a) and Rayleigh (b) for COVID-19 data.

Figure 7 displays the PP plots illustrating the fitted  $q$ -Rayleigh and Rayleigh distributions concerning COVID-19 data. The PP plots provide a compelling visual diagnosis. The  $q$ -Rayleigh's points align closely with the diagonal, indicating a superior fit and capturing the nuances of the observed distribution. Conversely, the Rayleigh's deviations highlight potential inaccuracies in its representation. This comparative analysis, therefore, underscores the  $q$ -Rayleigh's superior efficacy in describing the intricacies of COVID-19 data.



**Figure 8:** Empirical, Rayleigh, and  $q$ -Rayleigh cdf's for COVID-19 data.

Figure 8 offers a compelling insight into the process of selecting models for COVID-19 data analysis. The empirical cdf serves as the reference, with the  $q$ -Rayleigh model demonstrating a remarkable level of fidelity. Its curve closely follows the trajectory of the observed data,



contrasting with the Rayleigh model's comparatively less precise fit. Consequently, the  $q$ -Rayleigh model emerges as the preferred choice, providing a more accurate and insightful representation of the pandemic's patterns.

## 6. CONCLUSION

In this research paper, we have introduced a novel category of two-parameter distributions termed as the " $q$ -Rayleigh distribution". This distribution is formulated by utilizing the Rayleigh distribution as the foundational distribution and incorporating the  $q$ -exponential function as the generator function. To evaluate the characteristics of the model, we derived survival, hazard, and cumulative hazard functions for the  $q$ -Rayleigh distribution, analysing them graphically. Additionally, we explored extreme value properties.

The graphical examination of the  $q$ -Rayleigh distribution, employing various functions with diverse parameter values, demonstrated that the proposed distribution exhibits favourable properties in terms of its density function. We applied mathematical and statistical properties to assess the  $q$ -Rayleigh distribution, confirming its adherence to the aforementioned characteristics. The parameters of the  $q$ -Rayleigh distribution were estimated through the maximum likelihood estimation method.

To validate the goodness of fit, we employed the KS test,  $p$ -value and PP plot. Additionally, we conducted a comparison by examining the empirical cdf against those of the  $q$ -Rayleigh and Rayleigh distributions. Furthermore, we applied the  $q$ -Rayleigh distribution to cancer mortalities and COVID-19 data. The proposed distribution outperformed other distributions based on model selection criteria. In light of these findings, the  $q$ -Rayleigh distribution emerges as more adaptable and flexible in fitting real-life failure time data. We anticipate that this proposed distribution will find broader applications across diverse research domains, including reliability analysis, medical engineering, economics, and beyond.

## REFERENCES

- [1] Adnan, H., Hassan, N. J., & Jassim, H. K. (2023, February). The Weibull Lindley Rayleigh distribution. In AIP Conference Proceedings (Vol. 2414, No. 1, p. 040064). AIP Publishing LLC.
- [2] Almongy, H. M., Almetwally, E. M., Aljohani, H. M., Alghamdi, A. S., & Hafez, E. H. (2021). A new extended Rayleigh distribution with applications of COVID-19 data. *Results in Physics*, 23, 104012.
- [3] Amari, S. I., & Ohara, A. (2011). Geometry of  $q$ -exponential family of probability distributions. *Entropy*, 13(6), 1170-1185.
- [4] Al-Noor, N. H., & Assi, N. K. (2021, March). Rayleigh Gamma gompertz distribution: Properties and applications. In AIP Conference Proceedings (Vol. 2334, No. 1). AIP Publishing.
- [5] Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American statistical Association*, 83(402), 414-425.
- [6] G?l, H. H. (2023, August).  $q$ -Rayleigh Distributions: Properties, Estimation and Applications. In 2023 5th International Conference on Problems of Cybernetics and Informatics (PCI) (pp. 1-3). IEEE.
- [7] Jose, K. K., & Naik, S. R. (2009). On the  $q$ -Weibull distribution and its applications. *Communications in Statistics—Theory and Methods*, 38(6), 912-926.
- [8] Khalaf, R. Z., & Al-Kadim, K. A. (2020, July). Truncated Rayleigh Pareto Distribution. In *Journal of Physics: Conference Series* (Vol. 1591, No. 1, p. 012106). IOP Publishing.
- [9] Merovci, F., & Elbatal, I. (2015). Weibull Rayleigh distribution: Theory and applications. *Appl. Math. Inf. Sci*, 9(5), 1-11.
- [10] Picoli Jr, S., Mendes, R. S., & Malacarne, L. C. (2003).  $q$ -exponential, Weibull, and  $q$ -Weibull distributions: an empirical analysis. *Physica A: Statistical Mechanics and its Applications*, 324(3-4), 678-688.

- [11] Provost, S. B., Saboor, A., Cordeiro, G. M., & Mansoor, M. (2018). On the q-generalized extreme value distribution. *REVSTAT-Statistical Journal*, 16(1), 45-70.
- [12] Rahman, M. M. (2022). Cubic Transmuted Rayleigh Distribution: Theory and Application. *Austrian Journal of Statistics*, 51(3), 164-177.
- [13] Rayleigh, L. (1880). XII. On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 10(60), 73-78.
- [14] Sadok, I., & Masmoudi, A. (2022). New parametrization of stochastic volatility models. *Communications in Statistics-Theory and Methods*, 51(7), 1936-1953.
- [15] Sadok, I., Zribi, M., & Masmoudi, A. (2023). Non-informative Bayesian estimation in dispersion models. *Hacettepe Journal of Mathematics and Statistics*, 1-18.
- [16] Sato, A. H. (2010, December). q-Gaussian distributions and multiplicative stochastic processes for analysis of multiple financial time series. In *Journal of Physics: Conference Series* (Vol. 201, No. 1, p. 012008). IOP Publishing.
- [17] Saritha, K. N., Rao, G. S., & Rosaiah, K. Survival analysis of cancer patients using a new Lomax Rayleigh distribution. *Journal of Applied Mathematics, Statistics and Informatics*, 19(1), 19-45.
- [18] Sundaram, N., & Jayakodi, G. (2023). A study on statistical properties of a new class of q-exponential-Weibull distribution with application to real-life failure time data. *Reliability: Theory & Applications*, 18(3 (74)), 582-595.
- [19] Tsallis, C. (2009). Introduction to nonextensive statistical mechanics: approaching a complex world (Vol. 1, No. 1, pp. 1-2). New York: Springer.
- [20] Zhang, F., Shi, Y., Keung Tony Ng, H., & Wang, R. (2016). Tsallis statistics in reliability analysis: Theory and methods. *The European Physical Journal Plus*, 131, 1-20.

# MI-K-MEAN ALGORITHM: A NEW APPROACH FOR FINANCIAL RISK ANALYSIS WITH MISSING DATA IMPUTATION IN BIG DATA

Ravindra Kumar<sup>1</sup>, Diwakar Shukla<sup>2</sup>, Kamlesh Kumar Pandey<sup>3</sup>

<sup>1</sup>Department of Computer Science and Application, Dr Harisingh Gour Vishwavidyalaya  
Sagar, M.P., India

<sup>2</sup>Department of Mathematics and Statistics, Dr Harisingh Gour Vishwavidyalaya  
Sagar, M.P., India

<sup>3</sup>Department of Vocational Education, Indira Gandhi National Tribal University  
Amarkantak, M.P., India

<sup>1</sup>chakravarti.ravindra@gmail.com, <sup>2</sup>diwakarshukla@rediffmail.com, <sup>3</sup>kamleshamk@gmail.com

## Abstract

*The data mining is a tool of searching information from the data warehouse. Several mining algorithms exist in literature, one of the most common is the usual K-mean procedure. This generates centroids after every round of iteration. It is assumed that sample data is completely cleaned and noise free before the start of execution of the usual K-mean algorithm. If  $\alpha\%$  values are missing in sample data then after cleaning only  $(100-\alpha)\%$  values are available for the execution of the usual K-mean algorithm. Such bears a loss of information that affects the decision. This paper considers this problem and resolves such issue by replacing the missing data through imputed values calculated by the available values, called Mean Imputation (MI). It helps in financial risk analysis quite a lot because of risk prediction being taken on a larger sample (cleaned and imputed both). Several imputation procedures are available in literature. This paper considers the financial risk data as sample where the missing values of sample are imputed by the usual Mean-Imputation (MI) method and then on complete sample. Proposed MI-K-mean strategy is compared with no imputation usual procedure and found more efficient over the four-evaluation criterion of cluster formation while applying on risk data analysis.*

**Keywords:** Missing Data, Mean Imputation (MI), Credit card risk, K-mean clustering, Big data

## 1. Introduction

Financial risk calculation is used to bifurcate the customer as per the account information in a bank. It is the possibility of potential losses in direct investments caused by the effects of corporate credit, tax financing, other economic factors and corresponding economic shocks. Risk computation is an important method to provide the general description about a customer as per the credit score, which helps to the bank manager for taking the decision about distributing the loan. Financial risk is a measure to identify and analyze the existing financial risk factors, determine the likelihood and severity of probable new risks, and it provide scientific basis for risk for evaluation prevention and control [18]. Loan risk analysis plays a vital role among banking system where bank can identify the customer those who are exposed with good and bad risk. For the decision-making process the human analysis is more complex for large amount of financial data. Financial risk [23] includes risk identification, risk assessment and risk treatment. Risk identification and assessment is a part relating to evaluate the account of a customer for the financial risks and their sources through account details. Moreover, qualitative and quantitative methods are important to measure the size of the risk and generating the risk warning.

Data mining provides the general description of the data for the analysis to predict the values and to forecast the solution for making the decisions on it [31]. Data pre-processing is an important step for data analysis used to clean the data for removing the noise. This is time-consuming task to perform some calculation over data.

Missing data calculation is very tedious task for finding the location and manage them. For missing data handling, the imputation is the process for substituting the value in place of missing value [9]. Imputation methods can be included with various statistical methods like mean, median and mode. Imputation [10] is commonly used in computer science and related fields for several reasons such as handling non responded data, preprocessing data, maintaining dataset integrity, improving model performance and data analysis for visualization. Maintaining Dataset Integrity, in many applications, maintaining the integrity of the dataset is vital. Removing rows or columns with missing values [14] might result in a significant loss of data, reducing the representativeness and potential insights that can be obtained. Imputation allows for retaining the maximum amount of information available in the dataset. Imputation can also lead to improved model performance by reducing the potential bias and noise introduced by missing data. By imputing missing values [26], models can utilize the complete dataset to learn patterns, relationships, and make more accurate predictions.

Clustering is a technique to find the homogeneity of the particular group of objects. Cluster analysis is very useful in big data to category of studied object is not known in advance, to group the similarities into a particular category based on the degree of affinity therefore the same category can achieve the maximum similarity and minimize the dissimilarities. Moreover, the different categories achieve the maximum homogeneity and minimum heterogeneity. Cluster model selection is the process involved as per objective of the problem [25]. Objective of the problem define as per domain and may be vary as per the model selection so that the model selection is important concern for the prediction. Clustering analysis method can be categorized into three different types: trying to calculate an optimal data partition [15] to divide the given data into a specific number of clusters; trying to find out a method for the cluster structure; and trying to find a method based on statistical model for potential cluster modelling.

The K-means [12] cluster analysis technique effectively ignore the subjective negative impact caused by the artificial threshold value and ignores the missing data aspect, therefore it can more accurately and objectively describe the state intervals of different financial risks. On the basis of previous summary and analysis, this paper provide the current research status and significance of financial risk using the imputation and k-means [17] algorithm, elaborated the development background, current status and future challenges of the K-means clustering algorithm using imputation method, introduced the related works of similarity measure and item clustering with imputation[13], proposed a financial risk indicator system based on the K-means[20] clustering algorithm, performed evaluation parameter and data processing, constructed a financial risk based model based on the K-means [28] clustering algorithm with imputation [19], the dataset stored the values in credit card [33]. Study results of this paper provide a reference for further researches on financial risk based on K-means [22] clustering algorithm with imputation and the removal of the data in big data mining.

This paper is organized in nine sections. Second contains technical part of background of big data, imputation, clustering and evaluation methodologies. The third section is based on main problem undertaken in the paper while fourth section is based on motivation and hypothesis creation of research. The solution as in the form of proposed procedure is in section 5 whereas section 6 is with a new MI-K mean algorithm which is crux of this study. Section 7 reveals the flow of execution of this new algorithm and Section 8 supports the outcomes with numerical data. The last section 9 contains conclusion of all findings in a nutshell.

## 2. Background Technical Aspect

### 2.1 Big Data

Big Data cluster is a term for a collection of datasets who are so large and complex that it becomes difficult to process using on-demand database management tools or using traditional data processing applications [21]. Big data can be classified mainly in three basic categories like Volume, Variety and Velocity. This large volume of data continuously increases day by day by using electronic gadgets and through web-based platforms. The social media is a major source of big data [24] and others sources are like medical, insurance, marketing, weather forecasting etc. The main source are social media like Facebook, WhatsApp, Twitter etc. where at every second the volume of data increases drastically. In a day data generate with different forms of text messages, audio-recording, images, videos, log files etc. Big data parameter is important to discuss along with challenges [13] of this technology.

### 2.2 Missing Data types, Techniques and Classification

Data collection and analysis is the major part in for research and development. This step be performed very carefully however due to the large data there is chance to miss any value for the entry or missing due to any other reason [7]. Missing observation has mainly three patterns MAR, MCAR, MNAR [32]. For handling the missing values in datasets various strategies exist like try to find out the missing data [9], leave out the incomplete data and go-ahead for the next step, replace the missing data [27] as per the mean value etc.

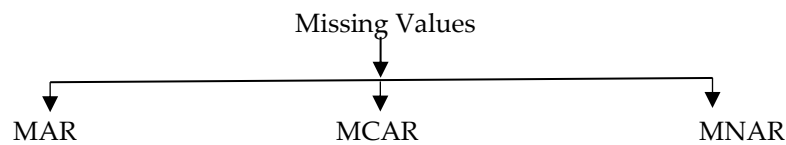


Figure 1. Missing Data Types

(a) MAR (Missing at Random): MAR [30] values give the same value in the particular group which it is belongs to the observed data.

(b) MCAR (Missing Completely at Random): In MCAR [32] finding the missing values in the same for all cases where the values are not available in the observation.

(c) MNAR (Missing Not at Random): MNAR is the missing value unknown to us, it is very difficult to finding in the observation.

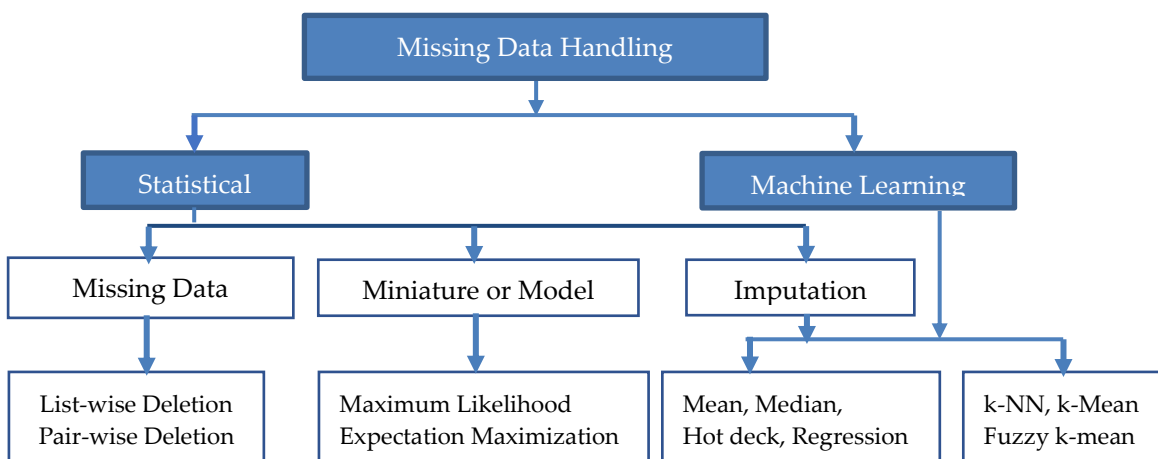


Figure 2: Classification of Missing data handling techniques

### 2.3 Data Clustering

Clustering is the method of a given data points, partition them into a set of groups which are as similar as possible. Data Clustering technique to find the complete or incomplete data clustering [8], is a data exploration technique used in various fields, including data science and machine learning. It involves grouping a set of data points into clusters, objects are similar to each other and performed the Big data clustering [11] in the dataset. The K-means [14] algorithm is the most frequently used clustering method. Moreover K-means [16] clustering algorithm is to select K data as the initial centroid of each category and divide them into K categories according to the principle of one category with the smallest distance, and then further the divided mean values are judged according to the square error criterion function.

Thakur and Shukla [1] proposed the missing data estimation based on the chaining technique in survey sampling. The method section included estimation, missing data, chaining, imputation, bias, mean squared error (MSE), factor type (F-T), chain type estimator, double sampling.

Thakur et al. [5] presented some new concept on mean estimation with imputation using two-phase sampling design. In this paper a imputation using in a sample survey in presence of missing data and one of the substitution techniques of missing observations is applied.

Shukla et al. [2] proposed some new aspects on imputation using sampling. Methods included estimation, missing data, imputation, bias, mean squared error (m.s.e.), compromised estimator, factor-type compromised imputation (FTCI). The number of causes that affect the quality of survey and missing data is one of such that keeps sample incomplete.

Pandey and Shukla [3] deployed a new approach on stratified linear systematic sampling-based clustering approach for detection of financial risk group by mining of big data. Risk analysis is beneficial for taking the business decision for finding the unknown risks such as credit risk, debit risk, operational risk and financial risk.

Jager et al. [6] integrated a benchmark for data imputation methods. This paper provides the detailed information about the missing data and its categories such that MAR, MCAR and MNAR. The method section included data quality, data cleaning, imputation, missing data, benchmark, MCAR, MNAR, MAR. The data preprocessing method selection for automated data quality improvement.

Pandey et al. [4] employed max-min distance sort heuristic-based initial centroid method of partitional clustering for big data mining. The methods included big data clustering, Initial centroid algorithm, convergence speed, stratified sampling, K-means, K-means++, MDSHK-means.

### 2.4 Cluster Evaluation Parameters

(a) Silhouette Score: The silhouette score is a measurement of how similar an object is to its own cluster (cohesion) compared to other clusters (separation). For the calculation using the mean intra-cluster distance (a) and the mean nearest-cluster distance (b) for each sample.

$$\text{Silhouette Coefficient} = (b - a) / \max(a, b)$$
$$\text{Percentage Gain} = \frac{|\text{Strategy B score} - \text{Strategy A Score}|}{\text{Strategy A}} \times 100$$

(b) Davies Bouldon Score (DBS): This Davies Bouldon Score is calculated as the average similarity measure of each cluster with its most similar cluster, where similarity is the ratio of within-cluster distances to between-cluster distances, which is simply the average of the similarity measures

of each cluster with a cluster most similar.

$$\bar{R} = \frac{1}{N} \sum_{i=1}^N R_i$$

$$\text{Percentage Gain} = \frac{|\text{Strategy A score} - \text{Strategy B Score}|}{\text{Strategy A}} \times 100$$

(c) Mutual Information Score (MIS): Mutual Information is a measure of the similarity between two labels of the same data. Where  $|U_i|$  is the number of the samples in cluster  $U_i$  and  $|V_j|$  is the number of the samples in cluster  $V_j$ , the Mutual Information between clustering  $U$  and  $V$  is given below.

$$MI(U, V) = \sum_{i=1}^{|U|} \sum_{j=1}^{|V|} \frac{|U_i \cap V_j|}{N} \log \frac{N|U_i \cap V_j|}{|U_i||V_j|}$$

$$\text{Percentage Gain} = \frac{|\text{Strategy B score} - \text{Strategy A Score}|}{\text{Strategy A}} \times 100$$

(d) Rand Index Score: Rand Index is calculating a similarity between two cluster results by taking all points identified within the same cluster. This value is equal to 0 when points are assigned into clusters randomly and it equals to 1 when the two cluster results are same.

$$RI = \frac{\text{Number of pairs in same cluster(actual)} \times \text{Number of pairs in same cluster(predicted)}}{\text{Total number of possible pairs}}$$

$$\text{Percentage Gain} = \frac{|\text{Strategy B score} - \text{Strategy A Score}|}{\text{Strategy A}} \times 100$$

(e) Adjusted Rand Index (ARI): ARI is used to measure the similarity between two clustering by considering all the pairs of the  $n$  samples and calculating the counting pairs of the assigned in the same or different clusters in the actual and predicted.  $E$  is indicating Expected.

$$ARI = \frac{\text{Number of pairwise true positive prediction} - E[RI]}{\text{Average number of pairs in same cluster for actual and predicted} - E[RI]}$$

$$\text{Percentage Gain} = \frac{|\text{Strategy B score} - \text{Strategy A Score}|}{\text{Strategy A}} \times 100$$

### 2.5 Mean Imputation (MI)

- Step I: Take sample of  $n$  observations.
- Step II: Find missing values in dataset (out of  $n$ ).
- Step III: Let  $k$  ( $k < n$ ) values of dataset are found missing.
- Step IV: Find mean of  $(n-k)$  values in sample data. Let it is denoted as  $\bar{x}^k$ .
- Step V: Replace all missing values in the dataset by  $\bar{x}^k$ .

## 3. Problem Undertaken

This paper aims to explore about the application of imputation techniques over clustering methods applicable to financial risk calculation in the big data environment. Cluster evaluation parameters evaluates the cluster accuracy and provide the efficient result for creating the clusters. This paper aspires to contribute the existing literature by providing efficient evidence and theoretical insights for data cluster calculation in the financial risk data setup when missing data is replaced by the imputed values. In view of combination of clustering and imputation need new algorithm which is a problem considered herein what follows.

### 4. Motivation

The data cleaning procedure reduces the sample size of financial risk data by eliminating noise presence therein. Noise may be in the term of missing values. One can think of that if such are replaced using the known values by an appropriate imputation method then larger sample size will be available for applying the usual K-mean algorithm which may produce efficient result of financial risk clustering. The financial risk is dangerous and require large data size for prediction.

#### 4.1. Hypothesis

- (a) Is there significant effect of imputed data against missing observations on the cluster evaluation parameters?
- (b) Comparing risk reduction for imputed sample with the cleaned sample.
- (c) Is the risk a decreasing function of imputed values in sample?

### 5. Proposed Procedure

Two strategies are given below.

Strategy A: A new algorithm is proposed named after “MI-K-mean algorithm” which considers entire sample data n (using imputation).

Strategy B: Usual K-mean algorithm applicable over only cleaned data which is less in sample size due to cleaning.

This paper presents a comparison between Strategy A (Proposed) and Strategy B (usual method) for data mining. The step-wise execution of algorithm is as under:

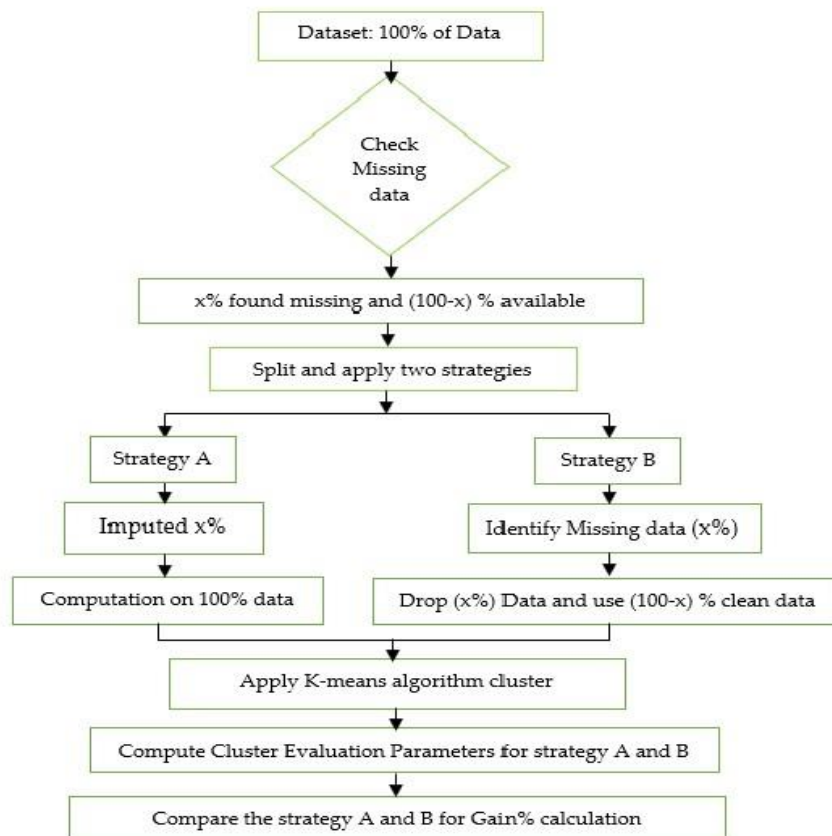


Figure 3: Basic Model for workflow of proposed method with imputation and usual method



The figure 3 contains two strategies A and B for the cluster formation by K-means algorithm on financial risk data. Strategy A uses replacement of missing values by an appropriate imputation method (MI-imputation) while strategy B contains cleaned data after eliminating the missing. The sample size for B is smaller that strategy A.

## 6. Proposed MI-K-mean Algorithm (Step-wise)

Imputation and clustering based proposed strategy A in order to find cluster is as under

(i) Input

1.  $N = \{a_1, a_2, a_3, \dots, a_n\}$  is the data points to the financial risk-based D dataset
2. K= Required number of clusters.

(ii) Output

$C = \{c_1, c_2, c_3, \dots, c_n\}$

(iii) Dataset Description

- 1 Dataset Head: `data.head()` [35]
- 2 Dataset Shape: `data.shape()` [35]
- 3 Dataset Statistical description: `data.describe()` [35]

(iv) Missing values Identification

Method for finding missing values (null values) in dataset: `data.isnull.sum()` [35]

(v) Dropping the Missing Values (Removal of the data)

Methods for deleting the missing data

- (a) Row wise deletion: `data.dropna(axis=0)` [35]
- (b) Column wise deletion: `data.dropna(axis=1)` [35]

(vi) Imputation: Mean Imputation

`data['Column_name'] = data['Column_name'].fillna(data['Column_name'].mean())` [35]

(vii) Clustering K-mean algorithm

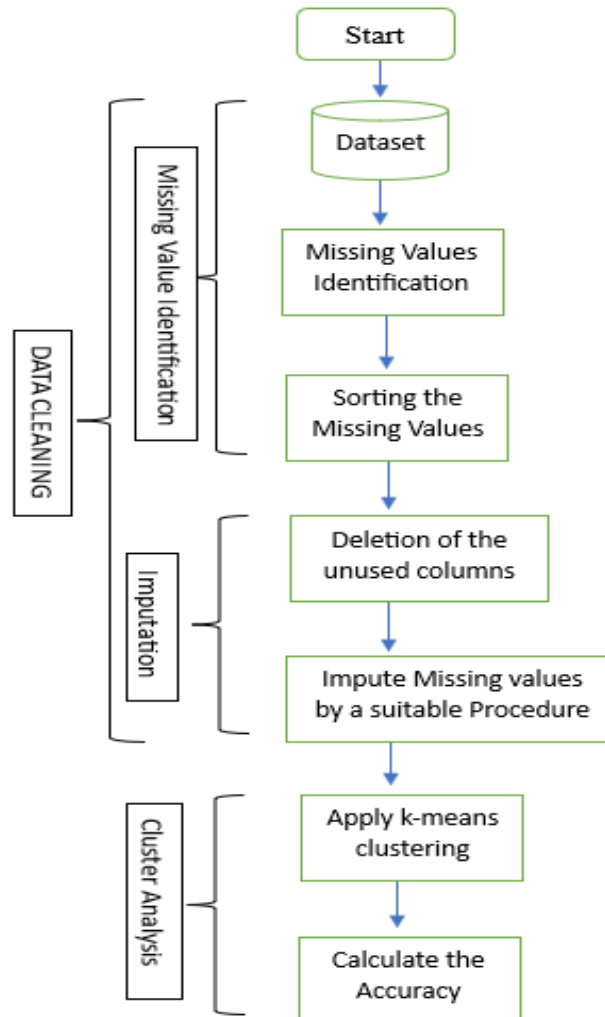
- 1 K-Means Clustering (Un-supervise Clustering Method)
- 2 Select k random values
- 3 Find out the optimality of clusters using Elbow or Silhouette method
- 4 Calculate the cluster centers and centroid
- 5 Find out the clusters

(viii) Cluster Evaluation

- 1 Calculation of Silhouette Score (SC) for the cluster evaluation
- 2 Calculation of Davies Bouldon Score (DBS) for the cluster evaluation
- 3 Calculation of Mutual Information Score (MIS) for the cluster evaluation
- 4 Calculation of Rand Index (RI) for the cluster evaluation
- 5 Calculation of Adjusted Rand Index (ARI) for the cluster evaluation

### 7. Implementation Procedure of MI-K-algorithm

Model for implementation of the proposed Strategy A needs several steps in data cleaning process such as missing value identification and performed imputation methods to obtain clusters on financial risk data.



**Figure 4:** Implementation model for proposed method (Strategy A)

### 8. Empirical Analysis

An Empirical study has been performed for applying efforts on the computing environment, datasets, existing algorithm, evaluation criteria and results.

(A) Experiment Environment and Credit Card General Loan Risk Dataset [34]

**Table 1:** Description of the Credit Card General Dataset

ID	Dataset	Objects	Attributes	Class	Data source
data	CC General	8950	18	2	www.kaggle.com

(B) Computing Environment

The computing environment for the proposed clustering approach using mean imputation technique is developed in Anaconda Navigator (anaconda 3) Jupyter and Google Collab notebook. The experimental environment is configured with an Intel(R) Core (TM) i5-2430M CPU @ 2.40GHz, 256 GB SSD, 4GB DDR3 RAM, Windows 10 Pro, Python 3.10.11, Microsoft Edge browser.

(C) Results and Discussion

**Table 2:** Original Dataset Credit Card General Data Analysis

ind	CUST_ID	BALANCE	BALANCE	PURCHASE	ONECINS	CASH_A	PURCHASE	ONECINS	CASH_A	PURCHASE	CREDIT	PAYMENTS	MINIMUM	PRC	TENURE			
0	10001	40.90	0.82	95.40	0	95	0.00	0.17	0.00	0.08	0.00	0	2	1000	201.80	139.51	0.00	12
1	10002	3202.47	0.91	0.00	0	0	6442.95	0.00	0.00	0.00	0.25	4	0	7000	4103.03	1072.34	0.22	12
2	10003	2495.15	1.00	773.17	773.2	0	0.00	1.00	1.00	0.00	0.00	0	12	7500	622.07	627.28	0.00	12
3	10004	1666.67	0.64	1499.00	1499	0	205.79	0.08	0.08	0.00	0.08	1	1	7500	0.00	NaN	0.00	12
4	10005	817.71	1.00	16.00	16	0	0.00	0.08	0.08	0.00	0.00	0	1	1200	678.33	244.79	0.00	12

Table 2 shows the description of the sample dataset [ Total 19 columns and five rows result using head() method]. Actual analysis performed over 8950 rows.

**Table 3:** Data Size

data.shape	Rows
Before Removal Shape Size	8950
After Removal Shape Size	8636

Table 3 shows that there are total 8950 row and after noise removal (cleaning) 8636 rows remained.

**Table 4:** Reduced data for analysis

index	CUST_ID	BALANCE	CREDIT_LIMIT	PAYMENTS	MINIMUM_PAYMENTS	TENURE
0	10001	40.90	1000	201.80	139.50	12
1	10002	3202.46	7000	4103.03	1072.34	12
2	10003	2495.14	7500	622.06	627.28	12
3	10004	1666.67	7500	0	NaN (Missing)	12
4	10005	817.71	1200	678.33	244.791	12

Table 4 shows that only six columns have been taken for analysis besides that all area available.

**Table 5:** Descriptive analysis of dataset

index	CUST_ID	BALANCE	CREDIT_LIMIT	PAYMENTS	MINIMUM_PAYMENTS	TENURE
count	8636	8636	8636	8636	8636	8636
mean	14477.9188	1601.225	4522.091	1784.478	864.3049	11.5343
std	2565.75979	2095.571	3659.24	2909.81	2372.566	1.3109
min	10001	0	50	0.0495	0.0191	6
25%	12267.75	148.0952	1600	418.5592	169.1635	12
50%	14469.5	916.8555	3000	896.6757	312.4523	12
75%	16698.25	2105.196	6500	1951.142	825.4965	12
max	18950	19043.14	30000	50721.48	76406.21	12

Table 5 shows descriptive analysis of data after removal of missing data.

**Table 6:** Count of missing values in sample data

CUST_ID	0
BALANCE	0
CREDIT_LIMIT	1 (Missing Values)
PAYMENTS	0
MINIMUM_PAYMENTS	313 (Missing Values)
TENURE	0

Table 6 provides information about the total number of missing fields(values) among six columns.

**Table 7:** Statistic description of the data

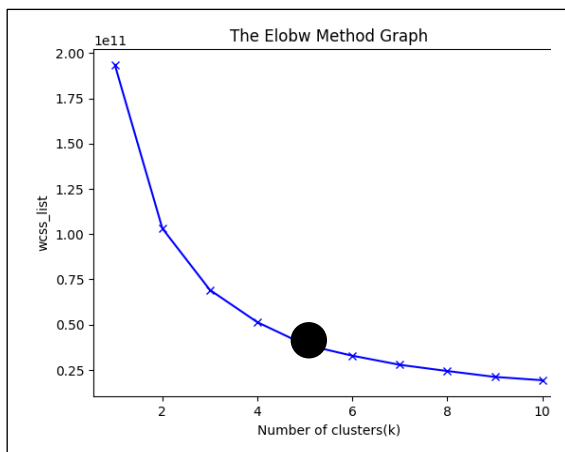
index	CUST_ID	BALANCE	CREDIT_LIMIT	PAYMENTS	MINIMUM_PAYMENTS	TENURE
count	8950	8950	8949	8950	8637	8950
mean	14475.5	1564.475	4494.449	1733.144	864.2065	11.517
std	2583.787	2081.532	3638.816	2895.064	2372.447	1.338
min	10001	0	50	0	0.019163	6
25%	12238.25	128.2819	1600	383.2762	169.1237	12
50%	14475.5	873.3852	3000	856.9015	312.3439	12
75%	16712.75	2054.14	6500	1901.134	825.4855	12
max	18950	19043.14	30000	50721.48	76406.21	12

Table 7 provides descriptive analysis statistics after the imputation of missing data (using mean imputation)

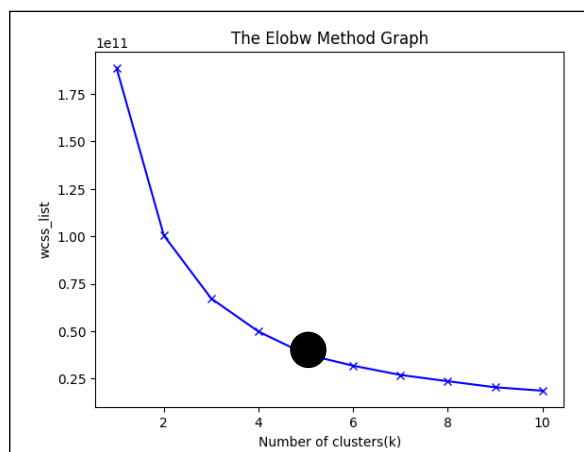
**Table 8:** Silhouette Method for finding the optimal cluster

k(clusters)	(Strategy A Proposed)	(Strategy B)(Usual)
k=2	0.4155	0.4000
k=3	0.3681	0.3691
k=4	0.2788	0.2810
<b>k=5</b>	<b>0.2792</b>	<b>0.2812</b>
k=6	0.1937	0.1919
k=7	0.1875	0.18771
k=8	0.1311	0.1089
k=9	0.1358	0.1234
k=10	0.1147	0.1245

Table 8 shows the optimal cluster is obtained at k=5 using Silhouette Method



**Figure 1.** Strategy A (Elbow method)



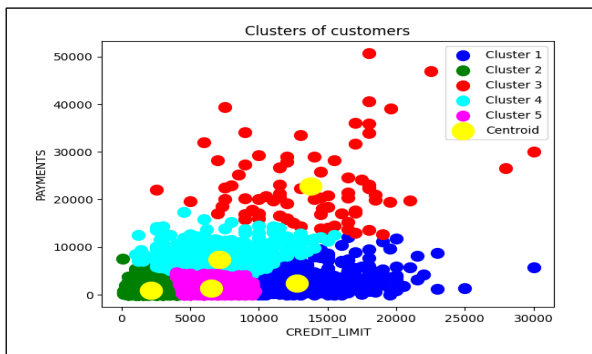
**Figure 2.** Strategy B (Elbow method)

Fig.1 and 2 reveals that graphical representation of the cluster optimality using strategy. Using Elbow method at k=5 the proposed strategy A is better than strategy B.

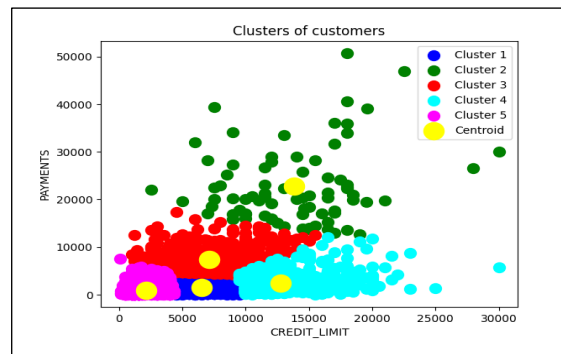
**Table 9:** Centroid or cluster center calculation

S.No	Strategy A	Strategy B
	Cluster Centers or Centroids	Cluster Centers or Centroids
1	[12751.36, 2347.60]	[ 6487.49, 1460.97]
2	[ 2113.72, 892.23]	[13822.78, 22720.13]
3	[13775.00 , 22802.04]	[ 7121.49, 7406.53]
4	[13775.00 , 22802.04],	[12731.04, 2403.48]
5	[ 6482.10, 1394.93]	[ 2103.22 , 923.39]

Table 9 reveals that Centroid obtained for five clusters (k=5) where k denotes the optimal number of clusters

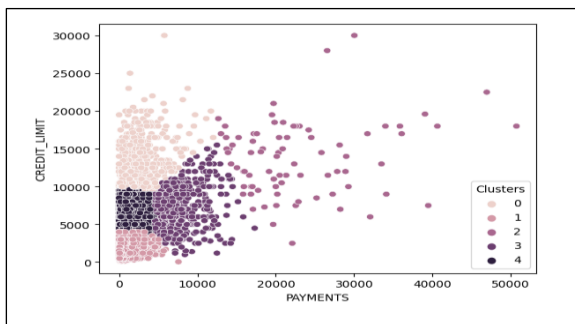


**Figure 3:** Strategy A (Cluster representation)

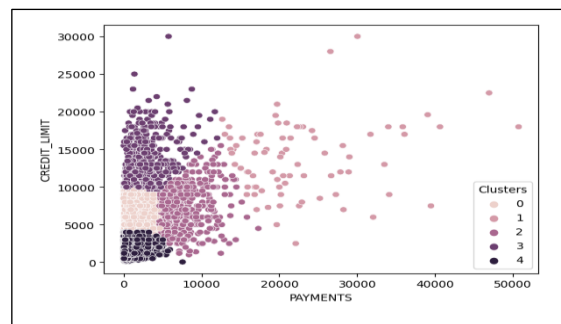


**Figure 4:** Strategy B (Cluster representation)

Comparing figure 3 and 4 one can observe that figure 3 is showing betterment (strategy A) in comparison to figure 4(strategy B) plotted using Matplotlib library.



**Figure 5:** Strategy A



**Figure 6:** Strategy B

Considering figure 5 and figure 6, the strategy A is better than B plotted using Seaborn library.

**Table 10:** Cluster evaluation parameters Strategy A and Strategy B

Cluster Evaluation	Strategy A	Strategy B	Percentage Gain (%)
Silhouette Score	0.287495857	0.303354945	5.51%
Devies Bouldon Score	1.283730623	1.013587896	21.04%
Mutual Information Score	1.075392591	0.381224234	64.55%
Rand Index Score	1	0.655464085	34.45%
Adjusted Rand Index	1	0.276573174	72.34%

Table 10 shows percentage gain due to five evaluation parameters of clusters by Silhouette, Davies Bouldon, Mutual Information, Rand Index and Adjusted Rand Index Score. There is significant percentage gain in four cluster evaluation criterions.

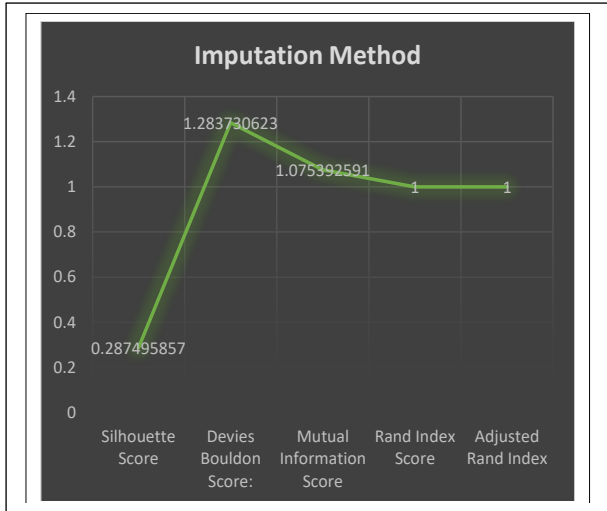


Figure 7: Strategy A

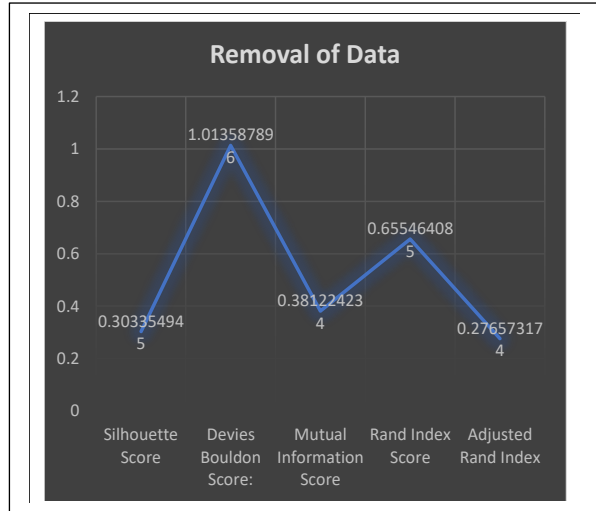


Figure 8: Strategy B

Figure 7 and 8 have graphical representation of the evaluation criterion on five parameters. Only first criteria (Silhouette score) bearing the low value but all other four evaluations showing gain, so strategy A is better than strategy B.

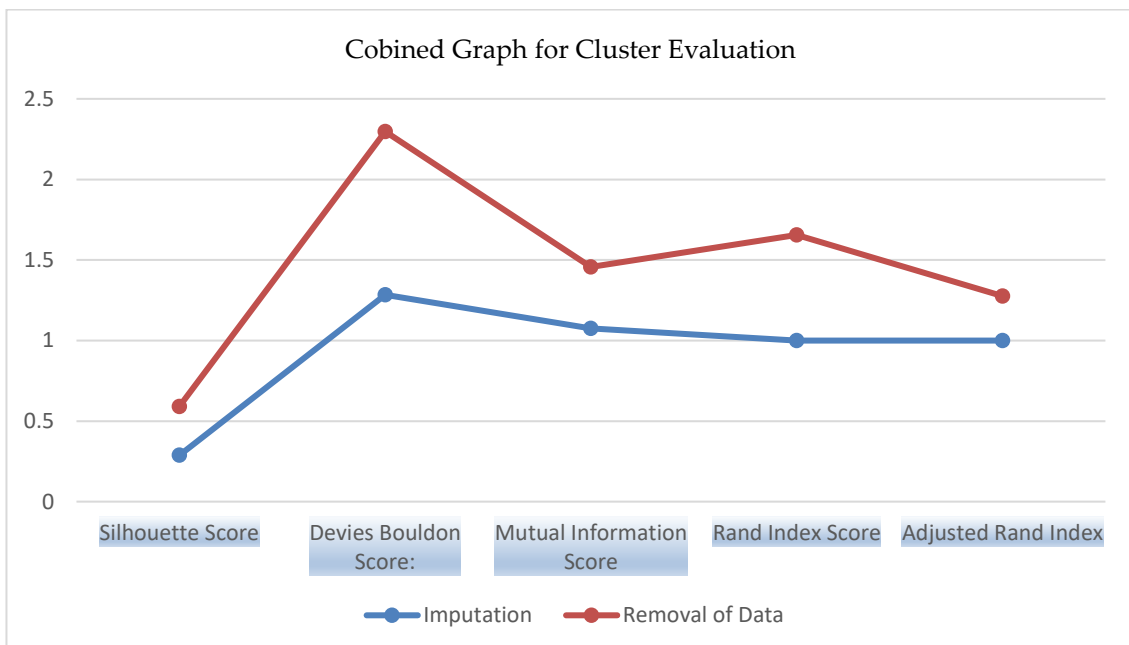
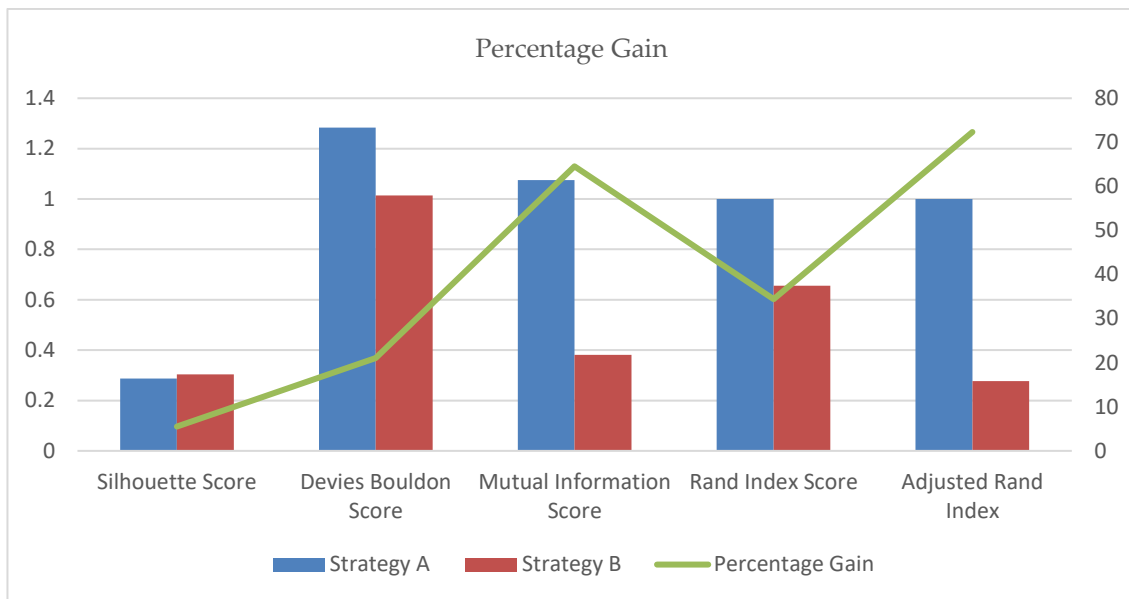


Figure 9: Combined Graph for cluster evaluation parameters

Figure 9 have a combined graphical representation of the evaluation criterion showing betterment for the proposed strategy A.



**Figure 10:** Percentage Gain between the Strategy A and Strategy B

Figure 10 reveals the percentage gain due to strategy A over the strategy B. So using the mean imputation (MI) of missing data one can find better result by using proposed algorithm (Strategy A)

## 9. Conclusion

This paper has presented a new algorithm named after MI-K-mean algorithm for the analysis of financial risk data. When risk factor exists then more input data values are required to reach the better decision. So, for such situation, the usual K-mean algorithm fail to form creating the efficient clusters. The proposed algorithm MI-K-mean (Strategy A) found efficient over the four cluster evaluation parameters while applying over the financial risk data. Table 10, figure 7, figure 8, figure 9 and figure 10 are supporting this fact. The MI-K-mean algorithm contain a new imputation-based approach which is unique feature. It opens bright avenues for further research while dealing with the risk data.

## References

- [1] Thakur, N. S., & Shukla, D. (2022). Missing data estimation based on the chaining technique in survey sampling. *Statistics in Transition new series*, 23(4):91-111.
- [2] D. Shukla, N. S. Thakur, and S. Pathak. (2013). Some new aspects on imputation in sampling. *African Journal of Mathematics and Computer Science Research*, 6(1):5-15.
- [3] Pandey, K.K., Shukla, D. (2022). Stratified linear systematic sampling-based clustering approach for detection of financial risk group by mining of big data. *Int J Syst Assur Eng Manag*, 13:1239-1253.
- [4] Pandey, K. K., & Pradhan, N. (2014). An analytical and comparative study of various data preprocessing method in data mining. *International Journal of Emerging Technology and Advanced Engineering*, 4(10):174-180.
- [5] Shukla, D., Thakur, N. S. (2008). Estimation of mean with imputation of missing data using Factor Type Estimator. *Statistics in Transition*, 9(1):33-48.

- [6] Jäger, S., Allhorn, A., & Biebmann, F. (2021). A Benchmark for Data Imputation Methods. *Frontiers in Big Data*, 4:693674- 88.
- [7] Ali, A., Emran, N. A., & Asmai, S. A. (2021). Missing values compensation in duplicates detection using hot deck method. *Journal of Big Data*, 8(1):1-19.
- [8] Jung Wun Lee & Ofer Harel (2023). Incomplete clustering analysis via multiple imputation. *Journal of Applied Statistics*, 50(9):1962-1979.
- [9] Doretti, M., Geneletti, S. & Stanghellini, E. (2018). Missing data: A unified taxonomy guided by conditional independence." *International Statistical Review*, 86(2):189–204.
- [10] Singh, S. & Horn, S. (2000). Compromised imputation in survey sampling, *Metrika*, 51: 266–276.
- [11] Khondoker M.R. (2018). Big data clustering, Wiley StatsRef: Statistics Reference Online. John Wiley & Sons Ltd, Chichester, 1:1–10.
- [12] Cao F, Liang J, Jiang G. (2009). An initialization method for the k-means algorithm using neighborhood model, *Comput Math with Appl*, 58:474–483.
- [13] Madhu, G. & Nagachandrika G. (2016). A New Paradigm for Development of Data Imputation Approach for Missing Value Estimation. *International Journal of Electrical and Computer Engineering*, 6:3222-3228.
- [14] Biessmann & Felix (2019). DataWig: Missing Value Imputation for Tables, *J. Mach. Learn. Res.* 20:1751-1756.
- [15] Mahmud M.S., Huang J.Z. and Salloum S. (2020). A survey of data partitioning and sampling methods to support big data analysis, *Big Data Mining Analytics*, 3:85–101.
- [16] Zahra S., Ghazanfar M.A., Khalid A. (2015). Novel centroid selection approaches for k-means-clustering based recommender systems. *Inf Sci (Ny)*, 320:156–189.
- [17] Jain A.K. (2010). Data clustering: 50 years beyond k-means. *Pattern Recognition Letter*, 31:651–666.
- [18] Zhu Z. & Liu N. (2021). Early Warning of Financial Risk Based on K-Means Clustering. Algorithm. *Complex*, 55:16831-12.
- [19] Woźnica, K., & Biecek, P. (2020). Does imputation matter? Benchmark for predictive models. ArXiv. /abs/pp. 2007.02837.
- [20] I.D. Borlea, R.E. Precup, F. Dragan & A.B. Borlea (2017). Centroid update approach to K-means clustering. *Advances in Electrical and Computer Engineering*, 17(4):3–10.
- [21] Aggarwal, C.C., Reddy, C.K. (eds.): Data Clustering: Algorithms and Applications (2013).
- [22] Xu J., Xu B. & Zhang W. (2009). Stable initialization scheme for k-means clustering. *Wuhan Univ J Nat Sci*, 14:24–28.
- [23] C. Diks, C. Hommes & J. Wang (2019). Critical slowing down as an early warning signal for financial crises?. *Empirical Economics*, 57(4):1201–1228.
- [24] Fahad A, Alshatri N, Tari Z. (2014). A survey of clustering algorithms for big data: taxonomy and empirical analysis. *IEEE Trans Emerg Top Computer*, 2(3):267–279.
- [25] Fränti P., Sieranoja S. (2019). How much can k-means be improved by using better initialization and repeats?. *Pattern Recognition*, 93:95– 112.
- [26] Schafer, J. L. & J. W. Graham (2002). Missing data: our view of the state of the art. *Psychol Methods* 7 (2):147-177.
- [27] Molenberghs G., C. Beunckens, C. Sotito, & M. G. Kenward (2008). Every missingness not at random model has a missingness at random counterpart with equal fit. *J. R. Stat. Soc. Ser. B. Stat. Methodology*, 70 (2):371–388.



- [28] M. Z. Hossain, M. N. Akhtar, R. B. Ahmad and M. Rahman (2019). A dynamic K-means clustering for data mining. *Indonesian Journal of Electrical Engineering and Computer Science*, 13(2): 521–526.
- [29] Mohan, K. & J. Pearl (2014b). On the testability of models with missing data. *In AISTATS*, 1:643–650.
- [30] Pearl, J. & K. Mohan (2013). Recoverability and testability of missing data: Introduction and summary of results, Technical Report R-417. *University of California, Los Angeles*.
- [31] Aggarwal C.C., Reddy C.K. (2014), Data clustering algorithms and applications. *CRC Press, United States*, 1:589–601.
- [32] Bhaskaran, K., Smeeth, L., (2014), What is the difference between missing completely at random and missing at random? *International Journal of Epidemiology*, 43(4):1336–1339.
- [33] UCI Repository for dataset: <https://archive.ics.uci.edu>
- [34] Kaggle for dataset: <https://www.kaggle.com/datasets>
- [35] Scikit Learn library: <https://scikit-learn.org>

# A TYPE I HALF LOGISTIC TOPP-LEONE INVERSE LOMAX DISTRIBUTION WITH APPLICATIONS IN SKINFOLDS ANALYSIS

Akeem Ajibola Adepoju<sup>1</sup>, Sauta S. Abdulkadir<sup>2</sup>, Danjuma Jibasen<sup>3</sup>, Jamiu S. Olumoh<sup>4</sup>

<sup>1</sup>Department of Statistics, Faculty of Computing and Mathematical Sciences, Aliko Dangote University of Science and Technology, Wudil, Kano. Nigeria

<sup>2,3</sup>Department of Statistics, Faculty of Physical Sciences Modibbo Adama University, Adamawa. Nigeria

<sup>4</sup>Department of Mathematics and Statistics, American University of Nigeria, Adamawa. Nigeria  
[akeebola@gmail.com](mailto:akeebola@gmail.com)<sup>1</sup>, [ssabdulkadir@mautech.edu.ng](mailto:ssabdulkadir@mautech.edu.ng)<sup>2</sup>, [djibasen2001@gmail.com](mailto:djibasen2001@gmail.com)<sup>3</sup>,  
[jamiu.olumoh@aun.edu.ng](mailto:jamiu.olumoh@aun.edu.ng)<sup>4</sup>

## Abstract

*This paper proposed a novel distribution parameterized by four parameters. This is achieved by compounding the potentials properties of the Type I half logistic topp-leone generalized distribution family with the properties of the inverse lomax distribution to form the novel Type I half logistic topp-leone inverse lomax distribution. The novel distribution is potentially capable of extending classical inverse lomax distribution. The potentiality of the shape of the probability density function of the novel distribution is worth recognizing since it produces right skewed, approximately normal, left skewed and a reverted J-shaped. Decreasing life failure shape is also observed. Distinctive features of the novel distribution such as moments, entropy, moment generating function, reliability and hazard function were derived. The estimation method explored in this study is maximum likelihood estimation. It is adopted to estimate the novel distribution unknown parameters. Real life data set was adopted to investigate the potentiality and applicability of the novel model. The type I half logistic topp-leone inverse lomax distribution outperform the recent models.*

**Keywords:** Type I Half Logistic Topp-leone-G family, Inverse Lomax, Maximum Likelihood Estimation.

## I. Introduction

Extension of the classical models have received tremendous attention, and the new extension is applicable to real life problems ranging from medical science, environmental, economics, demography, engineering, industrial statistics, biological sciences, and actuary science. There are several approaches to improve the classical distribution. However, the recent approaches provide the parents distributions with more shapes capacities and model flexibility through the generalized distribution families.

The type I half logistic Topp-leone –G (TIHLTL-G) distribution family was proposed by [1]. The family is characterized with two shapes – and the hazard rate shape which includes increasing, decreasing and bathtub shapes. The family is seen with potentiality capable of improve the classical model such as exponential model. On the other hand, the Lomax (L) distribution [2] sometimes refers to as Pareto type II distribution, coined from the second kind of generalized beta distribution. The L distribution is purposely applied to solve problems in insurance, biological sciences,

economics, reliability modeling, lifetime and engineering and other areas [3]. According to [4], the L distribution is an excellent distribution with potential of modeling survival complexity, and life-experimentation (engineering) and survival analysis.

The Inverse Lomax (IL) distribution is an excellent replacement for some closely related distributions like Inverse Weibull, Lomax, Gamma, Weibull distribution. Reason being that IL distribution possesses decreasing and upside-down bathtub hazard rate shape. Researchers, analysts and statistician found the IL distribution has a viable model useful in modeling diverse data sets. The [5] illustrated that IL distribution is among the inverted distribution family with noticeable flexibility in modeling various data sets, especially the non-monotonic failure rate. The IL distribution has also witness diverse extension, as it can be seen in [5], [6], [7], [8], [9], [10], [11], [12] and [13].

The author, [14] study reliability data using generalized IL distribution. The breaking stress of carbon fibres data was investigated by [15], the statistical methods for reliability data was study by [16], the [17] analyzed competing risks survival data, the reliability assessment under extended Chen distribution by [18]

In a scenario where a random variable  $X$  emanated from IL distribution, with cumulative distribution function (cdf) and probability density function (pdf) is expressed as:

$$H(x) = \left[1 + \frac{a}{x}\right]^{-b} \tag{1}$$

$$h(x) = abx^{-2} \left[1 + \frac{a}{x}\right]^{-b-1} \tag{2}$$

The cdf likewise the pdf of the TIHLTL-G by [1] are distinctly stated below

$$F_{TIHLTL-G}(x; \theta, \beta, \hat{\lambda}) = \frac{1 - \left[1 - \left[1 - \left(1 - (H(x; \hat{\lambda}))\right)^2\right]^\beta\right]^\theta}{1 + \left[1 - \left[1 - \left(1 - (H(x; \hat{\lambda}))\right)^2\right]^\beta\right]^\theta} \tag{3}$$

$$f_{TIHLTL-G}(x; \theta, \beta, \hat{\lambda}) = \frac{4\theta\beta h(x; \hat{\lambda}) [1 - H(x; \hat{\lambda})] \left[1 - \left[1 - H(x; \hat{\lambda})\right]^2\right]^{\beta-1} \left[1 - \left[1 - \left[1 - H(x; \hat{\lambda})\right]^2\right]^\beta\right]^{\theta-1}}{\left[1 + \left[1 - \left[1 - \left(1 - H(x; \hat{\lambda})\right)^2\right]^\beta\right]^\theta\right]^2} \tag{4}$$

The justification for this study lies in the fact that the IL distribution is noticed to have suffered from lack of pliability in the tail and peak features. This call for extensions of the IL distribution, diverse extensions has been witnessed. However, some of the extensions lack good flexibility. This motivates us to introduce a new attractive extension with TIHLTL-G with two shape parameters which can offer additional flexibility and improve the goodness of fit of the IL distribution.

## II. Methods

### 2.1 Type I Half Logistic Topp Leone Inverse Lomax (TIHLTL-IL) Distribution

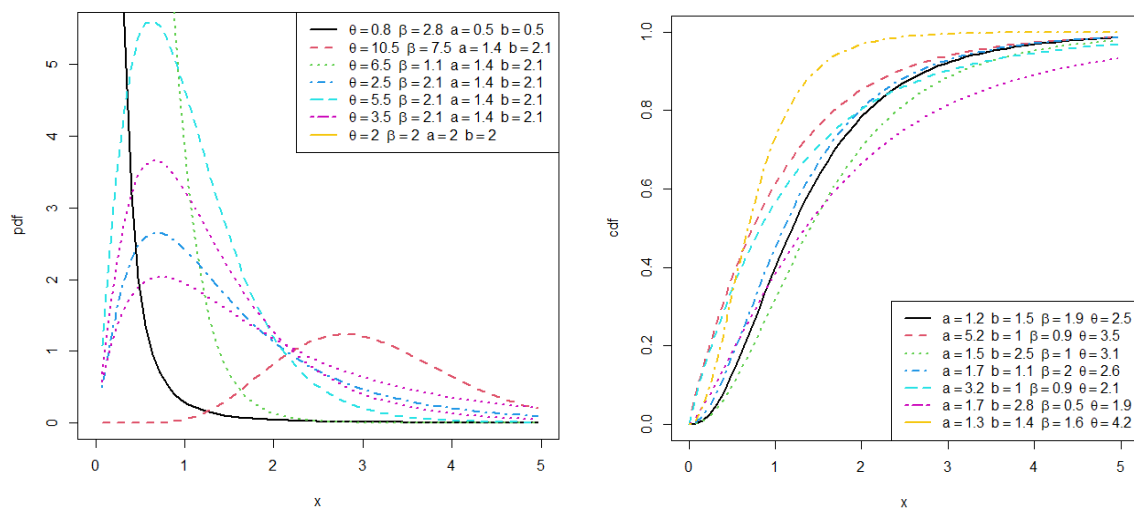
This section introduces the novel TIHLTL-IL distribution cdf and pdf. The pdf plots, densities expansion, and statistical features of the novel distribution. This characterization of this new distribution will be revealed and evaluated. The cdf and pdf of the novel TIHLTL-IL distribution is

obtained as:

$$F_{TIHLTL-IL}(x; \theta, \beta, a, b) = \frac{1 - \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right)^2 \right]^\beta \right]^\theta}{1 + \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right)^2 \right]^\beta \right]^\theta} \quad (5)$$

$$f_{TIHLTL-IL}(x; \theta, \beta, a, b) = \frac{4\theta\beta abx^{-2} \left[ 1 + \frac{a}{x} \right]^{-b-1} \left[ 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right] \left[ 1 - \left[ 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right]^2 \right]^{\beta-1} \left[ 1 - \left[ 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right]^2 \right]^\beta \left[ 1 - \left[ 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right]^2 \right]^{\theta-1}}{\left[ 1 + \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right)^2 \right]^\beta \right]^\theta \right]^2} \quad (6)$$

where  $x \geq 0$ ,  $a > 0$  scale and  $b, \beta, \theta > 0$  are shape parameters.



**Figure 1:** Pdf and cdf plots of TIHLTL-IL distribution with various parameters' choices.

From Figure 1, it is noticeable that the TIHLTL-IL distribution pdf can be seen as reversed J, approximately symmetric, right-skewed shaped. The additional plus observed in the new model is that, it revealed different forms of shapes with certain versatility in skewness, kurtosis and mode. The pdf is capable of modeling a heavy tailed and approximately symmetric data. The plus observed from the new model cannot be attributed to the IL distribution. The cdf plot of the novel TIHLTL-IL distribution converges to one. Its probability values range from zero to one. This implies that the novel TIHLTL-IL is a valid distribution.

### 2.2.1 Density expansion of TIHLTL-IL distribution

Consider the generalized binomial expansion expressed below

$$(1+z)^{-p} = \sum_{j=0}^{\infty} (-1)^j \binom{p+j-1}{j} z^j \tag{7}$$

$$(1-z)^r = \sum_{k=0}^{\infty} (-1)^k \binom{r}{k} z^k \tag{8}$$

Now consider the pdf given in (6) for expansion.

$$f_{THLTL-IL}(x; \beta, \theta, a, b) = 4\theta\beta abx^{-2} \left[1 + \frac{a}{x}\right]^{-b-1} \left[1 - \left[1 + \frac{a}{x}\right]^{-b}\right] \left[1 - \left[1 - \left[1 + \frac{a}{x}\right]^{-b}\right]^2\right]^{\beta-1} \left[1 - \left[1 - \left[1 - \left[1 + \frac{a}{x}\right]^{-b}\right]^2\right]^{\beta}\right]^{\theta-1} \left[1 + \left[1 - \left[1 - \left(1 - \left[1 + \frac{a}{x}\right]^{-b}\right)^2\right]^{\beta}\right]^{\theta}\right]^{-2}$$

now, consider this term for expansion using the generalized binomial expansion in (7) and (8)

$$\left[1 + \left[1 - \left[1 - \left(1 - \left[1 + \frac{a}{x}\right]^{-b}\right)^2\right]^{\beta}\right]^{\theta}\right]^{-2} = \sum_{j=0}^{\infty} (-1)^j \binom{1+j}{j} \left[1 - \left[1 - \left(1 - \left[1 + \frac{a}{x}\right]^{-b}\right)^2\right]^{\beta}\right]^{\theta j}$$

$$\left[1 - \left[1 - \left(1 - \left[1 + \frac{a}{x}\right]^{-b}\right)^2\right]^{\beta}\right]^{\theta j + \theta - 1} = \sum_{k=0}^{\infty} (-1)^k \binom{j\theta + \theta - 1}{k} \left[1 - \left(1 - \left[1 + \frac{a}{x}\right]^{-b}\right)^2\right]^{\beta k}$$

Following similar approach of expansion, we obtained the simplified version of the pdf.

Hence the pdf is rewritten as

$$f_{THLTL-IL}(x; \beta, \theta, a, b) = \sum_{j=0}^{\infty} \zeta_{\psi} x^{-2} \left[1 + \frac{a}{x}\right]^{-b(1+j)-1} \tag{9}$$

where  $\zeta_{\psi} = 4\theta\beta ab \sum_{k,l,m=0}^{\infty} (-1)^{(k+l+m)} \binom{1+j}{j} \binom{j\theta + \theta - 1}{k} \binom{\beta k + \beta - 1}{l} \binom{2l+1}{m}$

Similarly, cdf expansion goes same way.

Consider the  $[F_{THLTL-IL}(x; \theta, \beta, a, b)]^k$

$$F[F_{THLTL-IL}(x; \theta, \beta, a, b)]^k = \left[1 - \left[1 - \left[1 - \left(1 - \left[1 + \frac{a}{x}\right]^{-b}\right)^2\right]^{\beta}\right]^{\theta}\right]^k \left[1 + \left[1 - \left[1 - \left(1 - \left[1 + \frac{a}{x}\right]^{-b}\right)^2\right]^{\beta}\right]^{\theta}\right]^{-k}$$

Now, we expand this using the generalized binomial expansion in (8) and (9)

$$\left[ 1 + \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{X} \right]^{-b} \right)^2 \right]^\beta \right]^\theta \right]^{-w} = \sum_{j=0}^{\infty} (-1)^j \binom{w+j-1}{j} \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{X} \right]^{-b} \right)^2 \right]^\beta \right]^{\theta j}$$

$$\left[ 1 - \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{X} \right]^{-b} \right)^2 \right]^\beta \right]^\theta \right]^w = \sum_{k=0}^{\infty} (-1)^k \binom{w}{k} \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{X} \right]^{-b} \right)^2 \right]^\beta \right]^{\theta k}$$

Following similar approach of expansion, we obtained the simplified version of the cdf.

Hence the  $[F_{TIHLTL-IL}(x; \theta, \beta, a, b)]^k = \sum_{j=0}^{\infty} \alpha_{\zeta} \left[ 1 + \frac{a}{X} \right]^{-bn}$  (10)

where  $\zeta_{\psi} = \sum_{k, l, m, n=0}^{\infty} (-1)^{(k+l+m)} \binom{w+j-1}{j} \binom{w}{k} \binom{\theta(k+j)}{l} \binom{\beta l}{m} \binom{2m}{n}$

### 2.3 Properties of TIHLTL-IL distribution

In this section, derivation of the TIHLTL-IL distribution statistical properties is done. Properties explored are moments, probability weighted moment, entropy, reliability function, hazard function and quantile function.

#### 2.3.1 Moments

Moments of any distributions avails researcher the chance to investigate and reveal some important properties such as kurtosis, skewness, dispersion and central tendency. Assuming  $z$  is a random variable.

$$E(Z^r) = \int_0^{\infty} z^r f(z) dz$$
 (11)

To obtain the Ms of the TIHLTL-IL distribution, we substitute (9) and (10) in (11). Then we have

$$E(X^r) = \zeta_{\psi} \int_0^{\infty} \left[ 1 + \frac{a}{X} \right]^{-b(1+m)-1} X^{r-2} dX$$

where  $\int_0^{\infty} \left[ 1 + \frac{a}{X} \right]^{-b(1+m)-1} X^{r-2} dX = a^r B[(1-r), (b(1+m)+r)]$

$$E(X^r) = \sum_{\psi} \zeta_{\psi} a^r B[(1-r), (b(1+m)+r)]$$
 (12)

### 2.3.2 Probability Weighted Moments (PWMs)

The PWMs generally represented mathematically as:

$$\omega_{r,s} = E \left[ z^r F(z)^s \right] = \int_0^\infty z^r f(z) F(z)^s dz \tag{13}$$

In order to obtain the PWMs of the TIHLTL-IL distribution, we substitute (9) and (10) in (13) and make  $k = s$ . Then we have,

$$\omega_{r,s} = \sum_{j=0}^{\infty} \zeta_{\psi} \alpha_{\zeta} a^r B \left[ (1-r), (b(1+m+n)+r) \right]$$

now,

$$\int_0^\infty \left[ 1 + \frac{a}{x} \right]^{-b(1+m+n)-1} x^{r-2} dx = a^r B \left[ (1-r), (b(1+m+n)+r) \right]$$

$$\omega_{r,s} = \sum_{j=0}^{\infty} \zeta_{\psi} \alpha_{\zeta} a^r B \left[ (1-r), (b(1+m+n)+r) \right] \tag{14}$$

### 2.3.3 Entropy

Entropy is applied as a metric of uncertainty or randomness, which exists in a random observation of its real population composition. A larger value of entropy signifies greater uncertainty in the data. It follows that continuous random variable  $X$  under the Shannon entropy is expressed as:

$$H_{\alpha}(x) = \frac{1}{1-\alpha} \log \int_{-\infty}^{\infty} f_{TIHLTL-IL}(x; \beta, \theta, a, b)^{\alpha} dx \tag{15}$$

$$f_{TIHLTL-IL}(x; \beta, \theta, a, b)^{\alpha} = \left[ \sum_{j=0}^{\infty} \zeta_{\psi} x^{-2} \left[ 1 + \frac{a}{x} \right]^{-b(1+m)-1} \right]^{\alpha}$$

$$= \left[ \sum_{j=0}^{\infty} \zeta_{\psi} \right]^{\alpha} \left[ \left[ 1 + \frac{a}{x} \right]^{-b(1+m)-1} x^{-2} \right]^{\alpha}$$

$$\text{let } \left[ \left[ 1 + \frac{a}{x} \right]^{-b(1+m)-1} x^{-2} \right]^{\alpha} = \Phi^{\alpha}$$

$$\text{then, } H_{\alpha}(x) = \frac{1}{1-\alpha} \left[ \sum_{j=0}^{\infty} \zeta_{\psi} \right]^{\alpha} \log \int_{-\infty}^{\infty} \Phi^{\alpha} dx \tag{16}$$

### 2.3.4 Reliability Function

The reliability function generally represented mathematically as:

$$R(z) = 1 - F(z) \tag{17}$$

Now, the reliability function for TIHLTL-IL distribution can be obtained from (17) as

$$\begin{aligned}
 R(x) &= 1 - \frac{\left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right)^2 \right]^\beta \right]^\theta}{\left[ 1 + \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right)^2 \right]^\beta \right]^\theta} \\
 &= \frac{2 \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right)^2 \right]^\beta \right]^\theta}{\left[ 1 + \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right)^2 \right]^\beta \right]^\theta}
 \end{aligned} \tag{18}$$

### 2.3.5 Hazard function

The hazard function is generally represented mathematically as:

$$H(z) = \frac{f(z)}{R(z)}$$

Then, the hazard function for TIHLTL-IL distribution can be obtained as,

$$H(x) = \frac{4\theta\beta abx^{-2} \left[ 1 + \frac{a}{x} \right]^{-b-1} \left[ 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right] \left[ 1 - \left[ 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right]^2 \right]^{\beta-1} \left[ 1 - \left[ 1 - \left[ 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right]^2 \right]^\beta \right]^{\theta-1}}{\left[ 1 + \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right)^2 \right]^\beta \right]^\theta \right] \left[ 1 - \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right)^2 \right]^\beta \right]^\theta \right]} \tag{19}$$

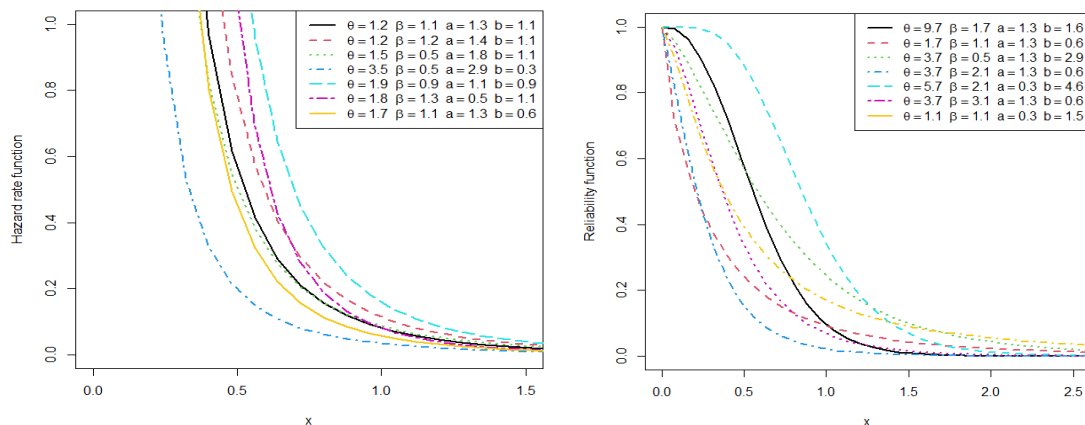


Figure 2: Plots of hazard rate and reliability function of TIHLTL-IL distribution with various parameter choices



### 2.3.6 Quantile Function

The quantile function of the TIHLTL-IL distribution is expressed below

$$F_{TIHLTL-IL}(x; \theta, \beta, a, b) = u$$

$$\frac{1 - \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right)^2 \right]^{\beta} \right]^{\theta}}{1 + \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{x} \right]^{-b} \right)^2 \right]^{\beta} \right]^{\theta}} = u \tag{20}$$

$$x = \frac{a}{\left[ 1 + \left[ 1 - \left[ 1 - \left[ 1 - \left[ \frac{1-u}{u+1} \right]^{\frac{1}{\theta}} \right]^{\frac{1}{\beta}} \right]^2 \right]^{\frac{1}{-b}} \right]^{\theta}} \tag{21}$$

### 2.3.7 Maximum Likelihood Estimation (MLE)

MLE is an approach channeled towards parameter estimation which has gained spread in terms of usage. This article adopted this method to estimate the parameters of the TIHLTL-IL model. Consider a randomly sampled  $X_i$  from the TIHLTL-IL distribution with parameter  $\Psi = (\theta, \beta, a, b)$   $\beta$ , where  $i = 1, \dots, n$ . The log-likelihood function for TIHLTL-IL model  $L(\Psi)$  is obtained as

$$L(\Psi) = n \log 4 + n \log \theta + n \log \beta + n \log a + n \log b - 2 \sum_{i=0}^n \log x_i - (b+1) \sum_{i=0}^n \log \left[ 1 + \frac{a}{x_i} \right]$$

$$+ \sum_{i=0}^n \log \left[ 1 - \left[ 1 + \frac{a}{x_i} \right]^{-b} \right] + (\beta - 1) \sum_{i=0}^n \log \left[ 1 - \left[ 1 - \left[ 1 + \frac{a}{x_i} \right]^{-b} \right]^2 \right]$$

$$+ (\theta - 1) \sum_{i=0}^n \log \left[ 1 - \left[ 1 - \left[ 1 - \left[ 1 + \frac{a}{x_i} \right]^{-b} \right]^2 \right]^{\beta} \right]$$

$$- 2 \sum_{i=0}^n \log \left[ 1 + \left[ 1 - \left[ 1 - \left( 1 - \left[ 1 + \frac{a}{x_i} \right]^{-b} \right)^2 \right]^{\beta} \right]^{\theta} \right] \tag{22}$$

By differentiating  $L(\Psi)$  in (22) with respect to  $\theta, \beta, a$  and  $b$ , the resulting equation is set to zero will produce the MLE estimates.

$$\text{Let } K = \left[ 1 - \left[ 1 - \left[ 1 - \left[ 1 + \frac{a}{x_i} \right]^{-b} \right]^2 \right]^\beta \right], \quad P = \left[ 1 - \left[ 1 - \left[ 1 + \frac{a}{x_i} \right]^{-b} \right]^2 \right], \quad Q = \left[ 1 - \left[ 1 + \frac{a}{x_i} \right]^{-b} \right] \text{ and } L = \left[ 1 + \frac{a}{x_i} \right]$$

Thus,

$$\frac{\delta L(\Psi)}{\delta \theta} = \frac{n}{\theta} + \sum_{i=0}^n \log K - 2 \sum_{i=0}^n [1 + K^\theta] \times \frac{\log K}{[1 + K^\theta]} = 0 \quad (23)$$

$$\frac{\delta L(\Psi)}{\delta \beta} = \frac{n}{\beta} - \sum_{i=0}^n \log P + (\theta - 1) \sum_{i=0}^n P^\beta \frac{\log P}{K} - 2\theta \sum_{i=0}^n \frac{K^{\theta-1}}{[1 + K^\theta]} [K]^\theta \log K = 0 \quad (24)$$

$$\begin{aligned} \frac{\delta L(\Psi)}{\delta b} &= \frac{n}{b} - \sum_{i=0}^n \log M + \sum_{i=0}^n M^{-b} \frac{\log M}{Q} + 2(\beta - 1) \sum_{i=0}^n \frac{QM^{-b} \log M}{Q} \\ &+ 2\beta(\theta - 1) \sum_{i=0}^n \frac{[P]^{\beta-1} QM^{-b} \log M}{K} - 4\theta\beta \sum_{i=0}^n \frac{K^{\theta-1} P^{\beta-1} Q}{[1 + K^\theta]} = 0 \end{aligned} \quad (25)$$

$$\begin{aligned} \frac{\delta L(\Psi)}{\delta a} &= \frac{n}{a} - (b+1) \sum_{i=0}^n \frac{x_i^{-1}}{M} + b \sum_{i=0}^n \frac{x_i^{-1} M^{-b-1}}{Q} - 2b(\beta - 1) \sum_{i=0}^n \frac{x_i^{-1} QM^{-b-1}}{P} \\ &- 2\beta(\theta - 1) \sum_{i=0}^n \frac{x_i^{-1} P^{\beta-1} QM^{-b-1}}{K} + 4\theta\beta \sum_{i=0}^n \frac{x_i^{-1} K^{\theta-1} P^{\beta-1} QM^{-b-1}}{[1 + K^\theta]} = 0 \end{aligned} \quad (26)$$

### 2.3.8 Information Criterion

The information criteria considered in this study include Bayesian (BIC), Akaike's (AIC), Hannan-Quinn (HQIC) and lastly, Consistent Akaike's (CAIC) Information Criterion. Their statistics are expressed mathematically as follow;

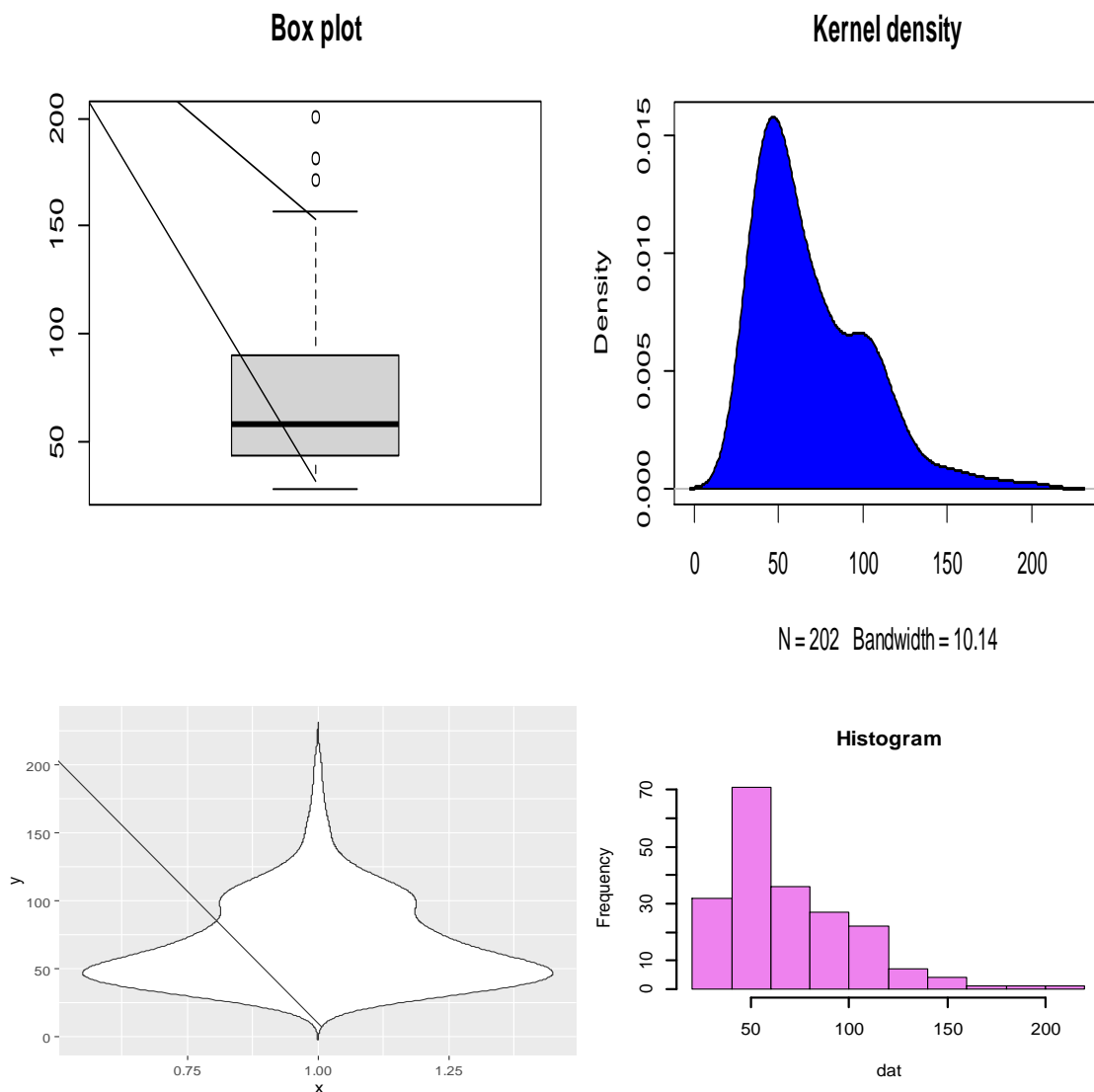
$$\left\{ \begin{array}{l} AIC = -2\ell + 2q \\ CAIC = -2\ell + \frac{2qn}{n-1-q} \\ BIC = -2\ell + q \log(n) \\ HQIC = -2\ell + 2q \log(\log(n)) \end{array} \right. \quad (27)$$

Where  $q$  and  $n$  are the number of distribution's estimated parameters, and the observations size, while  $\ell$  represents the log-likelihood (maximized) of the parameter vector  $\Psi = (\theta, \beta, a, b)$ . The preferred model according to this criterion is the one with least values estimated from the model

### III. Results

#### 3.1 Application

This section provides application to real-life data sets, demonstrating the applicability and flexibility of the TIHLTL-IL distribution against its comparators such as exponentiated generalized inverse lomax (EGIL) distribution [14] and half logistic inverse lomax (HLIL) distribution [9]. The choice of the distribution with most applicability and flexibility is determined by the distribution with the large likelihood's values and the lowest information criteria's values.



**Figure 3:** Boxplot, kernel density, violin and histogram of the data set.

The boxplot reveals information on the data set given below. It provides us with the necessary overview of the dispersion and the location of the data set. First of all, three outliers (171.1 181.7 200.8) are revealed from the data set, same is seen from the histogram, the minimum and maximum (28.0, 156.6) without outliers), first and third quartiles (43.80, 88.95) and the median (57.9) of the distribution. The kernel density and the histogram revealed that the data set is positively skewed, meaning that, bulk number of the observations is concentrated in left side of the distribution.

The data set represents the sum of skin folds in 202 athletes collected at the Australian Institute of Sports, it was previously studied by [19].

The data set is:

28.0, 98, 89.0, 68.9, 69.9, 109.0, 52.3, 52.8, 46.7, 82.7, 42.3, 109.1, 96.8, 98.3, 103.6, 110.2, 98.1, 57.0, 43.1, 71.1, 29.7, 96.3, 102.8, 80.3, 122.1, 71.3, 200.8, 80.6, 65.3, 78.0, 65.9, 38.9, 56.5, 104.6, 74.9, 90.4, 54.6, 131.9, 68.3, 52.0, 40.8, 34.3, 44.8, 105.7, 126.4, 83.0, 106.9, 88.2, 33.8, 47.6, 42.7, 41.5, 34.6, 30.9, 100.7, 80.3, 91.0, 156.6, 95.4, 43.5, 61.9, 35.2, 50.9, 31.8, 44.0, 56.8, 75.2, 76.2, 101.1, 47.5, 46.2, 38.2, 49.2, 49.6, 34.5, 37.5, 75.9, 87.2, 52.6, 126.4, 55.6, 73.9, 43.5, 61.8, 88.9, 31.0, 37.6, 52.8, 97.9, 111.1, 114.0, 62.9, 36.8, 56.8, 46.5, 48.3, 32.6, 31.7, 47.8, 75.1, 110.7, 70.0, 52.5, 67, 41.6, 34.8, 61.8, 31.5, 36.6, 76.0, 65.1, 74.7, 77.0, 62.6, 41.1, 58.9, 60.2, 43.0, 32.6, 48, 61.2, 171.1, 113.5, 148.9, 49.9, 59.4, 44.5, 48.1, 61.1, 31.0, 41.9, 75.6, 76.8, 99.8, 80.1, 57.9, 48.4, 41.8, 44.5, 43.8, 33.7, 30.9, 43.3, 117.8, 80.3, 156.6, 109.6, 50.0, 33.7, 54.0, 54.2, 30.3, 52.8, 49.5, 90.2, 109.5, 115.9, 98.5, 54.6, 50.9, 44.7, 41.8, 38.0, 43.2, 70.0, 97.2, 123.6, 181.7, 136.3, 42.3, 40.5, 64.9, 34.1, 55.7, 113.5, 75.7, 99.9, 91.2, 71.6, 103.6, 46.1, 51.2, 43.8, 30.5, 37.5, 96.9, 57.7, 125.9, 49.0, 143.5, 102.8, 46.3, 54.4, 58.3, 34.0, 112.5, 49.3, 67.2, 56.5, 47.6, 60.4, 34.9.

Table 1: The descriptive statistics of the data set

N	Min	Max	median	Mean	Var	Skewness	Kurtosis
202	28	200.8	58.6	69.0218	106.0501	1.1659	1.3220

Table 2: The Estimates of the MLE based on data set

Models	$\theta$	$\beta$	$a$	$b$
TIHLTL-IL	4.7115	26.0005	1.1422	2.7533
EG-IL	12.1005	17.9792	1.2394	2.6083
TIHL-IL	-	0.0401	7.5925	12.5563

Table 3: The Performance evaluation based on data set

Models	$\ell$	AIC	AICC	BIC	HQIC
TIHLTL-IL	-954.0944	1916.1819	1916.392	1929.422	1921.543
EG-IL	-1094.129	2196.258	2196.461	2209.491	2201.612
TIHL-IL	-963.0500	1932.100	1932.251	1942.025	1936.116

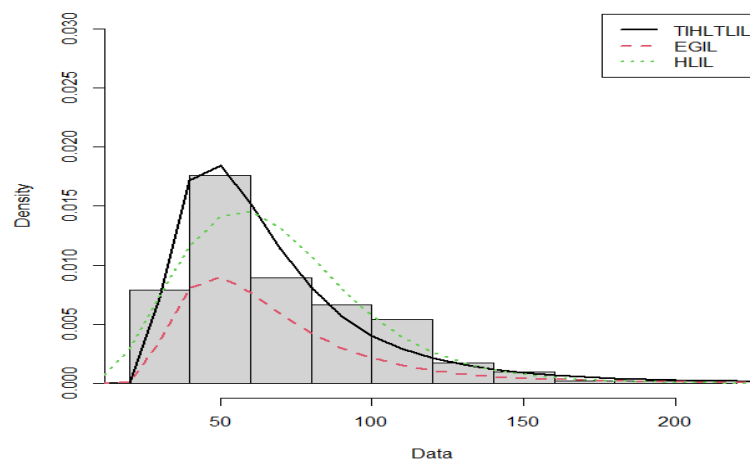


Figure 4: Fitted pdfs plot for the TIHLTL-IL, EGIL and HLIL distribution to the data set

#### IV. Discussion

This paper introduced a novel model called the Type I Half Logistic Topp-leone Inverse Lomax TIHLTL-IL model. The model's properties was defined and studied. We explored some useful statistical features of this novel model, including probability weighted moment, moments, moment generating functions, entropy, reliability functions, hazard function and the quantile functions. In order to have an insight of the model and to buttress our study, different plots were constructed such as the Pdf and cdf plots of TIHLTL-IL model, one can deduced that the model displays capability of handling data set with left and right skewed shape, reverse J-shape and approximately symmetric shape while the cdf plot confirmed the validity of the model. Investigation was conducted to visualize the data set using boxplot kernel density, violin and histogram, meanwhile, the boxplot suggest that there are three outliers in the data set, violin and histogram show a direction of the extreme values, indicating a positively skewed data set. The hazard shape reveals that the model can handle data set with monotonic decreasing life failure. We present MLE method, in order to estimate the unknown model's parameters. We delve in the applicability and flexibility of the novel model using a real life data set. The information criterion reveals that the proposed TIHLTL-IL model outclassed the related models. This claim is also supported by the fitted pdf plot.

#### References

- [1] Adepoju, A. A., Abdulkadir, S. S., & Jibasen, D. (2023). The Type I Half Logistics-Topp-Leone-G Distribution Family: Model, its Properties and Applications. *UMYU Scientifica*, 2(4), 09-22.
- [2] Kleiber, C. and Kotz, S. (2003). Statistical size distributions in economics and actuarial sciences. *John Wiley, Sons Inc*, Hoboken, New Jersey.
- [3] Al-Zaharani, B. and Al-Sobhi, M. (2013). On parameters estimation of Lomax distribution under general progressive censoring. *Journal of Quality and Survival Engineering* 1-7.
- [4] Hassan, A.S. and Al-Ghamdi, A.S. (2009). Optimum step stress accelerated life testing for Lomax distribution. *Journal of Applied Sciences Research*, 5: 2153-2164.
- [5] Ogunde, A.A., Chukwu A.U. and Oseghale I.O. (2023). The Kumaraswamy Generalized Inverse Lomax distribution and applications to reliability and survival data. *Scientific African*, 19.
- [6] Hassan, A.S., Al-Omar, A.I., Ismail, D.M. and Al-Anzi, A. (2021). A new generalization of the inverse Lomax distribution with statistical properties and applications. *International Journal of Advanced and Applied Sciences*, 8(4): 89-97.
- [7] Isa A. M., Kaigama A., Adepoju A. A., Bashiru S. O., Lehmann Type II-Lomax Distribution: Properties and Application to Real Data Set. (2023). *Communication in Physical Sciences*, 9(1):63 – 72
- [8] Kumar, D., Yadav, A.S., Kumar, P., Kumar, P., Singh, S.K. and Singh, U. (2021). Transmuted Inverse Lomax Distribution and its Properties. *International Journal of Agricultural and Statistical Sciences*, 17(1):1-8.
- [9] Al-Marzouki, S., Jamal, F., Chesneau, C., Elgarhy, M. (2021). Half Logistic Inverse Lomax Distribution with Applications. *Symmetry*. 13, 309.
- [10] Almarashi, A.M. (2021). A New Modified Inverse Lomax Distribution: Properties, Estimation and Applications to Engineering and Medical Data. *Computer Modeling in Engineering & Sciences*, 127(2):621-643.
- [11] ZeinEldin, R.A., Haq, M. A., Hashmi, S. and Elsehety, M. (2020). Alpha power transformed inverse lomax distribution with different methods of estimation and applications. *Complexity*, 2020(1), 1-15.
- [12] Hassan, A. S., and Mohamed, R. E. (2019). Weibull inverse lomax distribution. *Pakistan Journal of Statistics and Operation Research*, 15(3), 587-603.
- [13] Maxwell, O., Kayode, A.A., Onyedikachi, I.P., Obi-Okpala, I. and Victor, E.U. (2019). Useful generalization of the inverse Lomax distribution: Statistical Properties and Application to Lifetime

---

Data. *America Journal of Biomedical Science & Research*, 6(3): 258-265.

[14] Sule Omeiza Bashiru and Ibrahim Ismaila Itopa (2023). Modeling of reliability and survival data with exponentiated generalized inverse lomax distribution. *RT&A*, No 4(76), volume 18. 493-501

[15] Isa A. M., Sule O. B., Ali B. A., Akeem A. A., and Ibrahim I. I. (2022). Sine-Exponential Distribution: Its Mathematical Properties and Application to Real Dataset. *UMYU Scientifica*, 1(1), 127 – 131.

[16] Lee, E.T. and Wang, J. W. (2003). Statistical methods for survival data analysis (3rd Edition), *John Wiley and Sons*, New York, USA, 535 Pages, ISBN 0-471-36997-7.

[17] Rehman H., Chandra N., Abuzaid A. H. (2023). Analysis and modelling of competing risks survival data using modified Weibull additive hazards regression approach. *Journal of Mathematics & Statistics*. . *Hacettepe* Volume 52 (5) 1263 – 1281.

[18] Ibrahim, M., Aidi, K., Ali, M.M. (2023). A Novel Test Statistic for Right Censored Validity under a new Chen extension with Applications in Reliability and Medicine. *Ann. Data. Sci.* 10, 1285–1299.

[19] Hosseini B., Afshari M., Alizadeh M. (2018). The generalized odd gamma-G family of distributions: Properties and applications, *Austrian J. Stat.* 47(2), 69–89.

# A COMPRESSIVE STUDY ON FAULT DETECTION AND DIAGNOSIS FOR RELIABLE OPERATION OF HVAC, ENERGY BUILDINGS AND MACHINERIES

M. S. PATIL, G.M. MALWATKAR

Government College of Engineering, Jalgaon  
mahesh.patil@gcoej.ac.in, gajanan.malwatkar@gcoej.ac.in

## Abstract

*In Heating, Ventilation, and Air Conditioning (HVAC) systems, faults can be occurred due to various reasons such as drift deviation, valve/fan failure, water clogging, air filter obstruction, temperature sensor failure and so on. Similarly in electrical machineries faults can be occurred due to multiple causes such as phase reversal, over or under voltage, starter open/short circuit, bearing problems, insulation breakdown, overloading, thermal unbalance, environmental as well as other technical issues. The faults analysis at various stages of electrical systems are critically important for reliable operation of the system. In view of reliability and safety operations of modern sophisticated electrical systems, faults analysis and its diagnosis are necessary to avoid unaccountable losses. The faults at various stages, its causes, methods of detection and diagnosis, fault classifications are included in this work. The comment on effectiveness methods of detection of fault and diagnosis are included for electrical systems. In the industries, systems are incorporated with monitoring capacity for detection of faults at easy and early stage. This paper mainly focused on advancements in fault detection and diagnosis (FDD) methods with short review of various recent methods. This includes system information representation, methods of FDD, description of faults, fault classification, and decision actions related to maintenance, providing a systematic overview of the current state of FDD. Furthermore, the paper underscores the pivotal roles of FDD in electrical systems, emphasizing its effectiveness in identifying faulty states and taking pre-emptive actions against potential failures or accidents. The discussion extends to developments of current research in FDD approaches for electrical machineries with system monitoring, accompanied by short review of diverse and valuable FDD methodologies. The study concludes by addressing comments on recent trends, future directions, challenges, and prospective solutions in the hybrid and dynamic landscape of FDD.*

**Keywords:** Fault types and classification, Fault detection and diagnosis (FDD), HVAC, Electrical machines, Energy buildings, Reliability

## 1. INTRODUCTION

In the era of Industry 4.0, processes are evolving smart systems these are well equipped with advanced sensing devices to collect process related data for fault detection and process monitoring. As industries embrace full automation, meticulous supervision, involving process maintenance, control and corrective actions, becomes imperative to ensure operational efficiency [1]-[2]. Maintaining reliable and optimal performance in industrial processes is a challenge, often susceptible to various faults, is a key challenge. Among the array of FDD in process supervision techniques, is important issue of control methodology. Industries seek to enhance their process performance by leveraging advanced FDD capabilities, which primarily involve monitoring process behavior and uncovering faults, their characteristics, and root causes [3]- [4]. Efficient and accurate detection and diagnosis tools are crucial for sustaining high process yield and throughput. FDD

has garnered substantial attention across diverse industrial sectors and academia over several decades, offering benefits such as cost reduction, improved quality, and enhanced productivity [5]. This is particularly evident in safety-critical applications like robotics, autonomous vehicles, surveillance, and manufacturing systems, where FDD plays a pivotal role in ensuring human safety and preventing infrastructure loss. Modern systems and equipment demand the integration of FDD not only for safety but also for increased production and reliable operation. A robust FDD system, as highlighted in recent research, encompasses overall system health monitoring, diverse malfunction handling, and precise fault identification and localization for safe component removal. Over the past three decades, extensive work has been conducted on FDD, resulting in various techniques. These range from approaches of model based such as structural graphs and observer-based to approaches of data-driven employing classification, pattern recognition, and neural networks. Model-based FDD relies on accurate mathematical models, making it suitable for smaller systems with explicit models but susceptible to disturbances and uncertainties. In contrast, method of data-driven extract information from predicted signals to predict faults, with approaches of signal-based divided into statistical methods. The advent of technology has brought intelligent systems to the forefront, posing challenges in developing knowledge bases from raw historic information. The representation of information based includes knowledge explicit through production rules or expert systems and knowledge implicit in machine learning (ML) classifiers. Earlier reviews, have focused on model-based or data-driven FDD techniques, spectral approaches, and deep learning. This review provides a comprehensive overview, encompassing both traditional and signal processing based FDD approaches, with a specific emphasis on artificial intelligence-based methods. Covering the fundamental elements of FDD systems and prevalent techniques, this article contributes valuable insights to the FDD field for HVAC and electrical machineries. Challenges in real-time datasets includes the presence of outliers, which are often detected using unsupervised methods. In the approach of semi supervised learning which leverages both unlabelled and labelled data, providing a better choice. Data-driven FDD methods have gained significant attention across diverse industries, playing a pivotal role in monitoring of complex industrial process. The effectiveness of these approaches relies on the quality of historical data and the analytical models employed [6]. While various data-driven FDD methods exist, PCA-based and PLS-based approaches stand out for their simplicity and efficiency in detecting and diagnosing process faults. In literature of data-driven methodologies, it has been focused on PLS-based and PCA-based monitoring of process schemes. Many academicians have addressed modifications necessary for successful implementation and proposed an integrated adaptive residual generation technique to address uncertainty issues. The control techniques of fault-tolerant and data driven based FDD methods have been developed by Wang et al. [7], discussing their advances and general developments. In the work, researcher presented application example and outlined direction of research work, highlighting issues of FDD [8]. It is details by the Yin et. al [9] that data driven process was fundamental monitoring and diagnosis of faults including PLS, PCA, ICA and FDA. The study covered characteristics, computational complexities, design, and algorithms of these data-driven methods. In another work of Qin [11] provided data driven approaches and applications. In the study of it has been discussed the modelling on the basis of latent variable and fault detection work which are approaches for diagnosis and identification. Sensors are often limited to data transmission and sensing capabilities. Periodically, they send sensed data to a remote node that houses FDD blocks. They then wait for that node to make a determination on the presence of faults. With the aforementioned limitations in mind, we suggest a distributed sensor-fault detection and diagnosis system such that, immediately following data sensing, the sensor's fault detection block starts to function. This will conserve the energy used for periodic data transfer to and from a remote node in addition to providing a speedy determination regarding the existence of a malfunction. Additionally, this plan will offer chances to alert users or automatically halt system operations prior to a monetary loss or harm to human life. On the other hand, a central node implements the fault diagnosis block. Due to the fact that diagnosis is not time-sensitive, data exchanges between nodes may cause delays in the system. The use of this strategy minimizes the sensor's computing burden. In actuality, diagnosis is computationally



expensive because, in contrast to detection models that just distinguish between normal and abnormal conditions, the model must learn higher-level representations in order to distinguish among the fault types. A central node hosts problem diagnostics in the suggested distributed system design, whereas sensor nodes handle fault detection. The fault detection and reliability or model operation are active area not only in the industrial systems but also in the multidisciplinary fields[10]-[11].

This paper organised into the following sections: Section 1, included the introduction about the fault classification, FDD methods with overview of general faults category. Section 2, where a short discussion made on the work of fault categorization and various methods if FDD for general applications which includes categories from the general industrial processes. The survey on the major techniques for HVAC and energy buildings is included in Section 3 while Section 4 contains survey of electrical machineries. The paper ends with some remarks and conclusions which is part of Section 5.

## 2. FAULT CATEGORISATION AND DETECTION METHODS

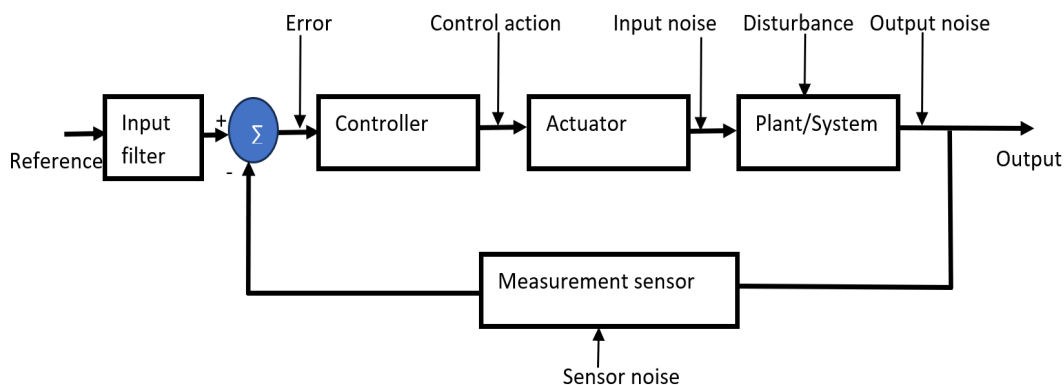
Beyond system representation and redundancy, the selection of an appropriate FDD system is heavily influenced by the nature of faults. A primary classification categorizes faults into software, hardware, communication and networking faults. Hardware faults encompass sensor, actuator, plant/process, and structural faults. Software faults include bit-flips, subroutine execution failures, runtime issues, and other software malfunctions. Networking and communication faults involve protocol incompatibility, packet transmission failures, and non-recoverable data packets. The general block diagram of industrial process control system is shown in Fig. 1. In the control theory, various academicians had been contributed to control the process using modified controller [12, 13, 14], with advances in control strategies [15, 16]. There are many types of faults and malfunctions which are based on industrial systems and described below in short

- **Sensor Category**

In sensor category, there are faulty components due to current/voltage sensor, speed, position sensor, absolute encoder type sensors or in general the fault components because of sensors which are used in industrial systems. The fault descriptions may include one or more reasons such as additive and/or multiplicative fault, abrupt voltage, power failure dropout, encoder fault, open circuit fault, multiple hard and soft failures and so on.

- **Actuator or final control element category**

This category can be based on electrical, mechanical or hydraulic or pneumatic elements. Faults can be because of armature and field winding, fault in stator winding, defect in insulation, rotor and bearing faults, rotor axis misalignment, gear box defects, fault in electrohydrostatic aerospace actuators. The description of actuators can be found by means



**Figure 1:** General block diagram of industrial process with faults and/or disturbance points

of drift, open and short circuit faults, magnetic fields degradation associated with windings. There can be loss of effectiveness of actuators, defects or faults in inner-race, outer-race and ball problem/ damage, excessive wear of bearing, less lubrication and problem in axis of rotor. Also, mechanical imbalance of rotor and broken motor bar are causes of faults. Friction and leakage losses are also said to be cause of regular faults in case of mechanical systems.

- **Controller or Control action category**  
Transient faults arise from sudden changes within the system and can be disappear after some time. Permanent faults cause lasting damage, requiring repair or replacement. Intermittent faults cycle between active and inactive states, while incipient faults exhibit gradual or slowly changes in the state variables of faulty components. This comprehensive categorization aids in understanding and addressing various types of problems in faults and malfunctions that may impact systems.
- **Plant/Process category**  
The faulty components under the category of plant/ process includes engine, part of plant, intelligent in automatic wind turbine, robotic manipulator, centrifugal pump malfunction, chemical/petrochemical plant or any specific units etc. The fault description includes various reasons such as misconduct of diagnosis in engine, leakage in tank, disengagement in DC motor, mechanical and electrical motor faults, transducers, final control elements and faults in torque converter additive magnitude joint faults, bearing defects, open and short circuit faults, duct and damper leaking fan and sensor failures.
- **Software and hardware category**  
The faults may be bit-flips, execution failure in routine functioning, faults of structural functioning in network, faults in communication network. The bit-flips can cause detection as well as correction of mainly leading faults and dependently faults, error probe and fault prone attributes. The network part is due to network faults on chip switches. The faulty nodes in source to destination transmission leads to communication faults which can be main reason to loss of signal or information. Any faulty controller output, electronic throttle controller and electric power steering controller included in this category. Fault description of controller response is partial loss of control effectiveness, degradation of throttle damping and return spring and friction loss prognosis. Another classification focuses on the dynamics and nature of faults, distinguishing between permanent, transient, incipient and intermittent faults.

In view of system representation, information and redundancy considerations, the next crucial factor in selecting appropriate FDD systems is the categorization of faults. A common classification divides faults into software, hardware communication and networking faults. Hardware faults encompass sensor, actuator, plant/process, and structural faults. Software faults include bit-flips, sub-routine execution failures, runtime issues, and other software malfunctions. Networking and communication faults involve protocol incompatibility, packet transmission failures, and non-recoverable data packets. The recent work of FDD methods are summarized in Table 1 with short remark and applicable domain. The FD methods available in the literature are shown in Fig.2.

Intelligent manufacturing has garnered substantial interest from both academia and industry in recent times. Intelligence is essential to the chemical and petroleum industries for both productivity and safety. This explains the recent decades' rapid development of FDD. A vast amount of measurement data is accessible to extract the useful information for process monitoring and optimization schemes because of the use of advanced computer and information technologies. Numerous applications involve the installation of sensors in hard-to-reach locations, which makes tasks like battery replacement or recharging challenging. In actuality, certain locations such as deep within woods to track weather patterns and identify fires or other possible calamities are more frequently equipped with limited-resource sensors than easily accessible ones. Furthermore,

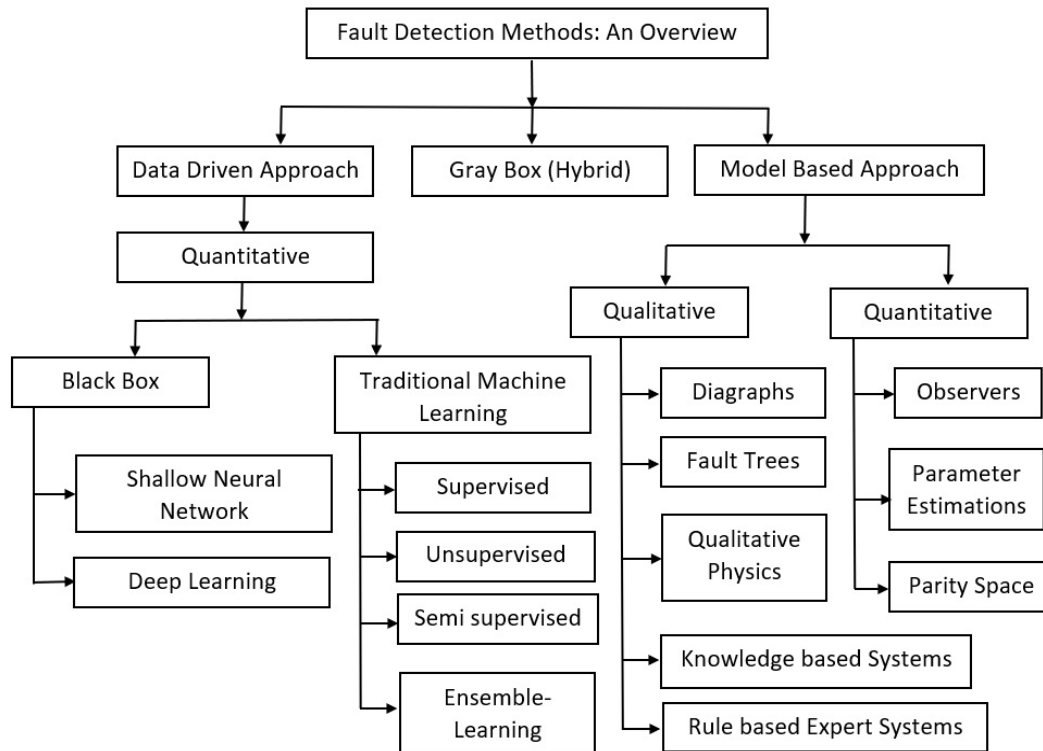


Figure 2: Fault detection methods

their battery life, memory, and processing power are all constrained. Therefore, while building FDD systems for such sensors, the following issues should be met.

- Smart Fault Detection: It is important to find errors as soon as they arise and before they cause significant losses. If these errors could be predicted, that would be much better for reliable operation of the system.
- Low Computation Cost: Because sensors have limited energy and computational capacity, FDD systems must operate effectively using the available resources.

### 3. DOMESTIC AND HVAC APPLICATIONS

For more than 20 years, there has been active study into FDD for applications involving air conditioning (AC) systems. Still, the vast majority of techniques were created for commercial structures. Although a lot of this work is applicable to the domestic marketplace, there are certain possibilities and problems specific to this industry that should be thought to be apart of the industrial refrigeration and commercial HVAC systems. A fault in the measurement of mixed air temperature, for instance, could affect the data gathered by fans, coils, dampers and elements of airflow loop, thereby affecting the condition of the indoor environment. The anomalous airflow supply to building zones caused by the problems related to the chilled water supply could have an adverse effect on the chiller water supply by means of the control loop. Additionally, it's feasible that other subsystems servicing the same zone make up for the effects of one system defect (such as an inadequate air supply to the zone). With the exception of fault propagation, identification and compensation, there are some fault symptoms become more difficult to identify and isolate than others because they are less evident than others. One instance is the faults of the Variable air volume (VAV) terminals, which could be hard to find because there aren't enough sensors [51], [52]. Finally, the possibility of many faults occurring simultaneously during HVAC

system operation makes FDD more difficult to understand because of contradicting or mutually worsening fault symptoms.

### 3.1. Domestic Applications

This subsection reviews various methods of FDD for AC systems with fault types and evaluates it with appropriate methods. Opportunities for advancement exist in the field of applying these techniques to the domestic and residential market, such as: (a) Taking into account the degree of fault diagnosis (FD) that is most economical in the residential market. (b) Reducing the number of sensors needed for FDD. The two most widely acknowledged advantages of accurately identifying and treating issues with AC systems are: (1) lower energy usage and (2) lower maintenance expenses. But even with these advantages, the expense of applying FDD techniques in the domestic AC industry has not been considered worthwhile. This study is primarily focused on developing cutting edge FDD techniques for domestic air conditioning systems and exploring ways to lower costs of these methods. But the goal of this subsection is to lay out a more thorough grasp of the advantages that efficient FDD offers. The person living in the house stands to gain the most from lower energy usage, and the owner benefits most from lower maintenance expenses. Nonetheless, FDD offers the owner and occupant additional significant advantages. The home-owner may gain more from the contented tenants than from lower maintenance expenses, and the tenant may profit more from the dependable comfort that FDD may offer than from lower electrical bills. Moreover, there is a lengthy value chain for ACs, and the home-owner and occupant are just two of the participants. FDD does, however, also help a great deal of other organizations. A straightforward illustration of the AC value chain and the advantages

**Table 1:** Short summary of data driven and knowledge/model based approaches for FDD

Approach	Method	Literature work	Remark / Applicability
Data driven	PCA/ICA	[17, 18, 19, 20, 21]	Complex processes Nonlinear fault diagnosis Minimization of false alarms
-	CVA	[22], [23],[24, 25], [25]	Time domain approach Consistent performance
-	Data type	[26], [27],[28],[29]	Biochemical/nuclear ANN& Wavelet transform Very complex systems
Model based	Observer type	[30],[31],[32]	Sliding mode observer Non linear processes Parameter varying processes
-	Parity equation	[33],[34],[35]	Parity relation of input-output Use of optimisation Time varying system
-	Supervised learning	[36],[37],[38],[39]	Use of support vector machine Grid search,genetic algorithm Efficient algorithms
-	Unsupervised learning	[40],[41],[42]	Online fault detection Use of CNN Hybrid fault detection
-	AI based	[43],[44],[45],[46]	Real time systems Smart NN based approach Predicted fault detection
-	Knowledge based	[47],[48],[49],[50]	Neuro-fuzzy and BN approach Use of fault isolation Advanced algorithm for FDD



**Figure 3:** Benefits of automated FDD methods in HVAC applications[53]

that FDD offers to its different organizations can be found in Fig. 3. Peak demand will drop as a result of lower AC loads, which will substantially reduce the demands on the companies that generate, transmit, and distribute power. Efficient FDD techniques may also enhance the commissioning of AC systems and give personnel a way to confirm the efficacy of their work. Home-owners may reduce the burden on the AC service sector during the hot/summer by identifying problems before they become apparent and taking appropriate action during the shoulder seasons. In addition, the home-owner’s repair expenses would go down. Effective FDD techniques could offer input on system design and sales to the dealer and manufacturer, allowing them to determine which systems have a track record of dependability and where improvements can be made. Lastly, by lowering carbon emissions from power plants and refrigerant leakage, better AC operations have a major positive environmental impact. In order for FDD to be widely adopted in the domestic cum residential sector, it is necessary to comprehend the advantages of this diagnosis at every stage of the value chain. The increased costs of the FDD system must be covered by someone, and if these benefits are attained, several parties may have to split the cost. Electric grid operators can, for instance, offer consumers who install the FDD system a financial rebate as a sort of incentive. Manufacturers may also provide the dealer with a cheaper FDD-enabled system. In order to get input, the dealer, installer, and service provider may also provided access to the available FDD data. The research described below shows that there is scope for considerable progress in the areas considering following points.

- **Reduced maintenance costs**  
 In Downey and Proctor’s work [72], almost 13,000 air conditioners, both home and commercial, were examined. The study’s took into account a number of variables pertaining to the state of air conditioners, including performance and operational parameters, indoor environmental conditions, and interior building circumstances where cooling constraints were important considerations among other things. The authors concluded that whereas 57% of the systems did not meet refrigerator level specifications, 65% of domestic/residential and 71% of light commercial systems needed maintenance as well as repairing. Breuker and Braun’s work examined frequent rooftop AC defects and their effects, and it assessed the relative cost of servicing for each fault through analysis of record. The influence of performance indicators in terms of simple issue detection and timely, affordable repair was the author’s main concern. It was determined that the average impact of the faults on

cooling capacity and coefficient of performance (COP) indices were significant because they raise energy costs and/or cause building occupants to feel less comfortable and because they offer a standard measurement for comparing the effects of various faults. Additionally, according to a database available in the literature of 6000 separate fault cases, it was noted that 24% of total repair expenses were attributable to compressor faults.

- Reduced electricity costs

According to Proctor and Downey’s analysis in [73], normal HVAC servicing practices fail to address two crucial parameters that affect equipment performance, which is why air conditioners and heat pumps perform below their designed capacity and efficiency. Two factors that the authors examined were inadequate airflow and an inaccurate refrigerant charge. Researchers have noted that domestic air conditioners work at least 17% less efficiently than what is stated on their efficiency ratings. According to the authors of [74], duct leakage, poor indoor air flow, and inaccurate refrigerant charges were among the most common errors. Furthermore, the authors found that only fixing issues with charges and ventilation might result in an average 16% boost in efficiency. According to [73], residential air conditioners function at a minimum of 17% less efficiently than their rated capacity. According to the authors of [74], duct leakage, poor indoor air flow, and inaccurate refrigerant charges were among the most common errors. Furthermore, the authors found that only fixing issues with charges and ventilation might result in an average 16% boost in

**Table 2:** Short summary of recent work (2020-2024) on FD methods for AC systems

Fault	Method	Ref.	System
Drift deviation	Kernel PCA and double layer	[54]	HVAC
Coil valve dampers	long-short term memory ANN,GA & multilinear regression	[55]	AHU
Leakage and fouling	On field measurement	[56]	Heat pumps
Gas & liquid line restrictions	CPA	[57]	AC with microtube condenser
Failures in valve	PCA & hybrid data mining	[58]	VRF AC
Liquid floodback(compressor)			
Compressor liquid & refrigerant charge	SVM,shallow NN deep learning	[59]	VRF AC
Fan failure, damper stuck	COP-deep learning	[60]	AC
water clogging, air duck leakage	SVM,multilayer perception		
Several faults	IoT & cyber physical system	[61]	HVAC
Refrigerant charge faults & condenser fouling	Virtual sensors & fault impact	[62]	HVAC
Valve,fan, temp sensors	Grey box	[63]	HVAC
Air temperature sensors	Hybrid approach	[64]	HVAC
Air filter obstruction	Physical based	[65]	HVAC
Chiller faults	AI-twin architecture	[66]	HVAC
Valve & temperature chiller faults	learning based	[67]	chilled beam
Condenser/evaporator fouling	Convolutional network	[68]	HVAC chiller
Fouling of condenser	ML	[69]	Roof top units
reduced water flow,refrigerant	Adaptive NN	[70]	Chiller
Condenser fouling	Feature recognition	[71]	Chiller
reduced water flow,refrigerant	Spectral regression		

efficiency.

- Improved commissioning  
Commercial and industrial buildings were analyzed by Rogers and Rasmussen [75] for power usage and 15-minute peak demands. The writers noted that a specific and effective side The refrigerant charge is wrong in over 60% of domestic ACs [73]. Furthermore, 47% of home systems are excessive, in comparison to the suggested sizing computation [74].
- Reduced peak demand  
Reduced efficiency results in higher levels of peak demand and total energy consumption. The summertime peak demand levels in Texas are approximately 25 % greater than the wintertime peak levels [75], mostly because of chiller loads and air conditioning.

The specifics of FDD techniques based on quantitative models are provided in [76]. Process fault diagnosis has a plethora of literature covering anything from statistical techniques to artificial intelligence (AI) and analytical procedures. From a modeling standpoint, certain techniques necessitate precise process models, semi-quantitative models, or qualitative models. On the other side of the spectrum, some methods just use historical process data and do not require any kind of model information. Furthermore, based on process information, various search strategies can be used to carry out diagnostics. Any candidate who is not an expert in these tactics will frequently find it challenging to navigate such a confusing array of alternatives and methodologies. The fault diagnosis techniques are categorized into three main types and are covered in three sections. According to [76], there are three types of model-based approaches: process history, quantitative model, and qualitative model. The researchers also provided a general mathematical framework that included a multi-step, complete FDD algorithm in addition to this categorization. In addition, it examined unprocessed measurements to produce helpful characteristics that were applied to identify certain issues. In general, the three-part review is a useful tool for comprehending the whole FDD methodology. Nevertheless, the review lacked relevant application-related information. According to [77], equipment that is not adequately maintained, deteriorated, or managed wastes between 15 and 30 percent of the energy utilized in commercial buildings. A large portion of this waste might be avoided if automated condition-based maintenance were widely used. The foundation for condition-based maintenance of engineered systems is provided by prognostics and automated FDD. Applications for energy building systems, such as HVAC and refrigeration have been researched and showcased. However, a plenty of of research and development has been done in the past ten years with the goal of creating FDD techniques for HVAC and refrigeration systems. In the work of [78] provides an overview of automated FDD research conducted since 2004 that is pertinent to the commercial building industry. The evaluation divides automated FDD techniques into three categories and updates an earlier review that was carried out in 2004. A selection of automated FDD examples from the major category are examined in order to determine which approaches are best suited for system construction and to comprehend the advantages and disadvantages of each approach. Additionally described in the dispersion of studies based on HVAC systems and automated FDD techniques. The current article can be used as a reference by industries and researchers to choose an acceptable automated FDD approach.

### 3.2. Applications in Energy Buildings

There must be Recognize the difficulties and complexity of FDD. There are three levels of complexity associated with this FDD problem: (1) building complexity, which arises from the existence of different building types and characteristics; (2) HVAC complexity, which arises from the intricate coupling of components of HVAC to meet various building needs; and (3) fault complexity of HVAC system, which arises from complex and variable fault symptoms. More than 40% of a building's energy is used by the HVAC system, one of the most significant mechanical systems. Problems with HVAC system operation can lead to interior environmental problems, such as low indoor air quality and thermal comfort, which can have an impact on

occupant health and productivity. [79], [80], and [[81]]. The study conducted by Jasmin et al. [79] examined the impact of a flexible space layout design on energy demand and thermal comfort within a contemporary open-plan office setting. The scholars evaluated the suitability of four control zoning methodologies in conjunction with three distinct HVAC systems, radiant ceiling, mechanical ventilation and a thermally active building system using dynamic thermal modelling. According to their findings, mechanical ventilation systems required a more intricate control plan to maintain thermal comfort, whereas thermally active and radiant ceilings building systems offer potential options for flexible office spaces for a typical location.

In the meantime, broken or malfunctioning HVAC systems waste a lot of energy and reduce the energy efficiency of buildings. According to estimates, problems with the HVAC system and lighting systems together might raise the use of these sectors by 4 % which is  $\approx 18\%$ , or 0.35 and 1.7 quads of US yearly consumption, respectively [82]. Given the complexity in HVAC systems with several coupling components and the intricate interactions between HVAC systems, buildings, and inhabitants, maintaining fault-free functioning of HVAC systems is difficult. The development of computer methods, such as the emergence of deep learning methods and building management systems which are utilizing more affordable sensors to support building operations allowing for the potential application of FDD. A promising method for guaranteeing HVAC system faultlessness. FDD techniques are typically categorized into three types : process history based, quantitative model-based, as well as qualitative model-based [83]. Simplified or detailed physics based models are typically useful in quantitative model based approaches to monitor variations between measured system status and anticipated system operation conditions. Qualitative model based techniques typically follow expert guidelines or fundamental FDD concepts. Process history based techniques rely on data, as they examine system sensing data directly to identify and diagnose HVAC system operating conditions. More broadly, knowledge-based approaches can be broadly defined as quantitative and qualitative model based methods that draw from engineering or physics knowledge in FDD. Data-driven techniques can be defined as FDD methods that solely rely on system sensing data. Zhao et al.s work [84] includes a thorough literature review of AI based fault detection and diagnosis (FDD) methods for energy systems built in the 20 years between 1998 and 2018, summarizing the benefits and drawbacks of the available AI-based techniques and outlining the most crucial areas for future research.

There are numerous types of structures for both residential and commercial purpose, including multi-family or single-family and cottages type (e.g., office, school, shopping center). These buildings serve a variety of purposes, which contributes to different building operation patterns throughout the day. Additionally, the physical characteristics of buildings vary greatly between designs and vintages, including window-to-wall ratios, zone configurations, and insulation levels. Ultimately, the behaviour of building occupants varies and is stochastic, resulting in a variety of characteristics for the load profile. Each of these results in distinct patterns and demands for heating and cooling; hence, building with HVAC interactions change, which in turn adds to the inherent FDD complexity. Examples of applications of FDD include campus buildings [85], manufacturing buildings [86], commercial buildings in hot [87], [88], mild [89], uncertain climates [[90]], etc. The intricacy of HVAC systems needs to be examined, as they have different types, capacities, and modes of operation that are driven by the growth of HVAC techniques and the various requirements for preserving the indoor environment [91]. Variable air volume, variable refrigerant flow, and direct expansion systems are common system types found in existing buildings; Table 2 provides examples of these systems to which FDD has been applied. These days, HVAC systems are typically made up of a broad range of closely coupled sub-systems (such as air handling units chillers, cooling towers, air distribution systems and so on.) to effectively maintain the typical indoor environment of buildings [92], [93]. Therefore, in order to achieve FDD, the established technique must take into account the interdependencies, or mutual influence caused by controlling feedback loops which couple HVAC sub-systems, in addition to properly handling probable software and hardware errors within the sub-systems. The inconsistent design of the HVAC system, its real functioning, and the FDD mechanism further exacerbate the issue. HVAC systems frequently operate under unanticipated circumstances while they are in use (such



as an oversized or undersized system). Moreover, rather of being used for FDD, the sensor elements in HVAC systems are made for feedback control of HVAC during regular operation. Each of them adds to the FDDs HVAC system's complexity. The symptom complexity of system problems is a direct outcome of the buildings and HVAC systems. In particular, the process of detecting and diagnosing faults with symptom propagation and compensation is complicated by the interdependencies between sub-systems and within sub-system components [85], [93].

The significance of FDD in HVAC systems for ensuring building energy performance and occupant service has garnered attention from the building HVAC research community on a regular basis. The developed FDD techniques were summarized and categorized in a number of previous assessments. For instance, in their FDD application, Katipamula and Brambley had divided FDD techniques into the three categories previously described and have included a brief discussion of the advantages and disadvantages of each kind of method [77]. There was also discussion on how these methods could be applied to particular HVAC & R(Refrigeration) areas. In a further investigation, Yu et. al. and colleagues have integrated and synthesized FDD research from 2005 to 2017 with the identical classification [94]. Yu et al [94] have examined analytical-based, knowledge based, and data driven approaches specifically for FDD of Air Handling Units (AHUs). Frank and colleagues [95] evaluated the obstacles and difficulties facing FDD in small commercial buildings. In their analysis of the progress made in each stage of the FDD process data sources, feature creation, fault detection, and fault diagnosis Shi and Brien [96] have identified a number of issues that need to be resolved in subsequent FDD studies. The applications of AI-based methods in FDD have garnered a lot of attention lately, and numerous writers have talked about the crucial next research projects in the subject of FDD. Mirnaghi and Haghghat [97] have examined data-driven methods that combine supervised, unsupervised, and hybrid learning for large-scale HVAC system fault diagnosis and repair. Li and neill [98] have concentrated on examining FDDs fault modeling for HVAC systems.

#### 4. APPLICATIONS IN ELECTRICAL MACHINERY

In industrial processes, dependability and safety are essential components. Numerous industries depend heavily on rotating machinery, which is prone to malfunction because of its lengthy operating lifespan and difficult working circumstances [99]. The various faults occurs in electrical machineries is enlisted in Fig. 4 and the detailed information is enlisted in the work of Asad et. al. [100] and [101]. The vibration signals of rolling element bearings always appear as low signal noise ratio, nonstationary statistical parameters while operating under demanding conditions (such as time-varying speed and load, high shocks), which makes diagnostic techniques challenging. To ensure smooth functioning under erratic situations, faults in various electrical machinery components must be found. A few instances of rotating machinery parts are motors, engines, shafts, bearings, gears, pumps, and blades. Qu et al. [102] have developed and tested AE-based methodologies and acoustic emission (AE) sensors for gearbox failure diagnosis. For the diagnosis of gearbox faults, AE-based methods demand far larger sampling rates than vibration analysis-based methods. It is therefore debatable whether, at the same sampling rate, an AE-based technique would perform better or at least as well as vibration analysis-based techniques. The first known attempt to compare the gearbox fault detection performance of AE and vibration analysis based methodologies using the same sampling rate was made by the authors in their comparative study for gearbox tooth damage level diagnostics using AE and vibration measurements. The study also mentioned that the lab experiments are conducted using a gearbox test rig to seed and test partial tooth cut faults. After conducting a comparative analysis, the authors concluded that, as compared to the vibration-based technique, the AE-based approach has the ability to distinguish between different levels of gear tooth damage. Mechanical resonance can easily impair vibration signals, but AE signals operate more steadily. The researcher concludes that vibration signal condition indicators are inconsistent with the extent of gear tooth damage because vibration is less sensitive than AE to minute tooth damage in the low speed range, making it challenging to identify gear faults.

Sakthivel et al.'s work [103] focuses on vibration-based problem diagnostics for single-block centrifugal pumps. Experiments have been conducted on the pump under various fault scenarios as well as in good working order. Drawn pump characteristic graphs show discharge vs. efficiency under both ideal and unfavorable situations. Authors noted that the pump's efficiency is high when everything is working properly, and for any malfunction, it falls into a range of values that is significantly lower than when everything is working properly. It is clear that if any of the study's studied pump flaws were present, the pump's efficiency would drop precipitously. Therefore, it is imperative that this fault identification investigation be completed. Additionally, in the same work, a mono-block monoblock centrifugal pump is used to model six classical states: normal, bearing fault, impeller fault, seal fault, impeller and bearing fault together, and cavitation. Using the C4.5 decision tree approach, a set of features has been retrieved and classified in the simulation. It is noted that based on the discussion and findings, it is safe to conclude that the C4.5 algorithm and vibration signals are suitable options for real-world defect diagnosis of monoblock centrifugal pumps. By Haidong et al. [104], a novel technique for rolling bearing fault diagnosis termed deep wavelet auto-encoder (DWAE) with extreme learning machine (ELM) was presented for intelligent rolling bearing fault detection. The study utilized wavelet function as a nonlinear activation function to create a wavelet auto-encoder (WAE) that is capable of efficiently capturing signal properties. To improve the capacity for unsupervised feature learning, a DWAE including several WAEs was built, and ELM was chosen as the classifier to precisely identify various bearing problems. The technique was used to examine the experimental vibration signals from the bearings. According to the authors' data, the created method is more effective than both standard deep learning methods and traditional methods in eliminating the need for human feature extraction. Haidong et al. [104] note that combining the wavelet function with deep learning and extreme learning machine improves fault diagnosis of rotating bearings greatly. Vibration signature analysis has historically been used to identify shafting system misalignments. The temperature increase at the source is also caused by these misalignments, couplings and bearings. Mohanty et al.'s contribution [105] describes an experimental investigation that used a thermal imaging camera to measure the shaft couplings' temperature in order to discover misalignment in systems early on. In order to identify flaws, the effects of load, speed, and misalignment on the different types of couplings and their temperature rise have been investigated. Before the temperature of the coupling achieves its steady state value, it is utilized to measure the misalignment in the system. In order to correlate with the thermal imaging, vibration

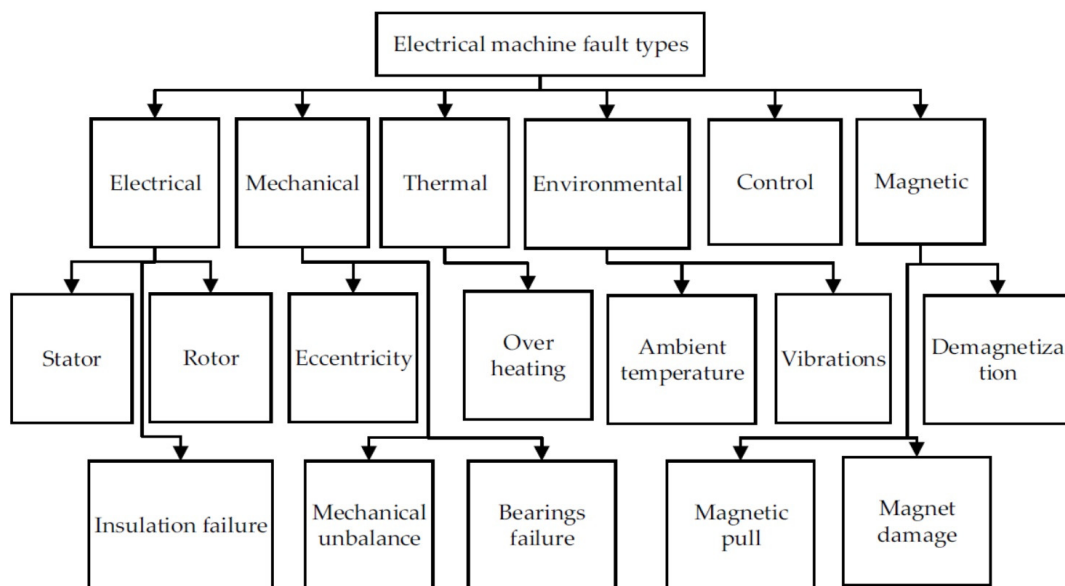


Figure 4: Fault types in electrical machines [100]

measurements at the bearing locations under various load and speed circumstances were also made using accelerometers and single point laser vibrometers. The recorded transient spatial temperature distribution on the couplings, the researchers discovered, also points to the shafting system's misalignment. The approach can be applied to automated thermography-based detection systems to identify misalignments from far-off places where traditional vibration monitoring would be challenging. Notably, the method can be applied to windmill gearbox problem detection at elevated positions where traditional contact type instrumentation would be very laborious.

Kubiak and colleagues conducted the failure analysis of the 150 MW gas turbine blades [106]. The 150 MW gas turbine used as the study's basis experienced a forced breakdown due to abnormally strong vibrations, which reduced output power to nearly nil. The analyzers' diagnostic task is to identify the primary reason behind the blades' failure. Additional research revealed that low cycle fatigue was the initial cause of the blade failure, which resulted in a crack that spread throughout the securing pin hole (stress raiser) at the blade's root. A few suggestions are made in light of the study to prevent gas turbine blades faults and failures. In light of the need for and impact of condition monitoring and fault diagnosis in induction motors (IMs) as well as the need for further study, Choudhary et al. [107] provided a state-of-the-art review that details various IM fault types and their corresponding diagnostic approaches. Numerous surveillance The methods that are available for diagnosing IM faults have been noted and shown. The researchers announced that there is a lot of potential for the use of non-invasive data collection methods in autonomous, timely maintenance scheduling and failure aspect prediction of dynamic machinery. The shortcomings of traditional sensors and monitoring schemes will be addressed by the use of non-invasive type instruments, which will remove the requirement to attach the sensor on the machine and provide speedy measurement, non-intrusiveness, and high accuracy. The thermal imaging approach is thought to be an effective tool for online instant messaging monitoring without human intervention when compared to other non-invasive techniques. For practical applications, combining infrared thermal imaging techniques with artificial intelligence-based methods can speed up decision-making even more.

An oil monitoring approach for engine wear evaluation was examined in the work of Bin Fan et al. [108]. The oil samples underwent quick on-site analysis using online visual ferrograph (OLVF). The wear debris concentration for the abnormal engines was discovered to have a low index of particle coverage area (IPCA) by the authors. Large debris was also infrequently seen on OLVF ferrograms, which was congruent with the findings of analytical ferrography. The cause of this was examined and addressed. In order to reduce the number of manual confirmations that require disassembling the oil pans, the researchers looked into an oil monitoring technique of wear evaluation for the 9-min engine hot tests. Oil samples from engines are quickly analyzed on-site using OLVF. However, the authors also mentioned that it is challenging to collect the larger wear debris by sampling at half of the oil level because of the short operating duration and the elimination of wear debris. Low concentration of the small wear debris is also a sign of abnormal wear during the 9-min hot test. The amount of small debris in the oil samples from the normal engines is greater than that of the abnormal engines.

## 5. CONCLUSIONS

In this study, FDD approaches for typical applications of electrical applications is reviewed to study the impact of early diagnosis of the faults. The various approaches of FDD for domestic applications, HVAC and electrical machineries applications are studied with the contributions of academicians and researchers from the literature. A disturbance or fault rejection strategy is main focus of the this paper for the reliable and maintaince free operation of the HVAC and electrical machines. In this study, recent techniques has been incorporated in view to design and implement the FDD methods. The approaches of FDD included are admittance as per earlier techniques due to fast FD algorithm with use of computational facilities on the ground of soft faults. Recent algorithms have improved fault detection strategy which will not affected due to parameter uncertainties and model mismatch in operational reliability of the system or modelling.

The study uses to generates minimum fouling in the system performance by considering the sensor, actuator and control signal.

The presented review includes

- The work which have been contributed for domestic applications in recent years
- The work which focuses on FDD of HVAC systems and its major failures due to non maintenance/repairing of the systems.
- The contribution which clearly mention that the routine maintenance can reduce the cost of expenses to avoid replacing the equipment.
- The study includes typical faults and its mitigation by advanced FDD methods with

It has been mentioned that the performances of the any system obtained through FDDs have been interest of reliability operation of the systems. The detection of fault can be made by parallel simulations or by means of dynamic identifications either through data driven or knowledge based approach. The FDD implied dynamically for the HVAC or any other devices to avoid any malfunction would be interest of the researcher through the use of advanced techniques such as deep learning, machine learning or AI approaches. In future work, the hybrid concepts can be useful for the FDD of sensors and other components of the industrial complex systems. There are the opportunities and challenges for the real time applications at micro level in the focused domains including defence, sensitive chemical and petrochemical as well as other industries.

## REFERENCES

- [1] Liang Ming and Jinsong Zhao. Review on chemical process fault detection and diagnosis. In *2017 6th international symposium on advanced control of industrial processes (AdCONIP)*, pages 457–462. IEEE, 2017.
- [2] Xian Du. Fault detection using bispectral features and one-class classifiers. *Journal of Process Control*, 83:1–10, 2019.
- [3] Inseok Hwang, Sungwan Kim, Youdan Kim, and Chze Eng Seah. A survey of fault detection, isolation, and reconfiguration methods. *IEEE transactions on control systems technology*, 18(3):636–653, 2009.
- [4] Rolf Isermann. *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*. Springer Science & Business Media, 2005.
- [5] Karim Salahshoor, Mojtaba Kordestani, and Majid S Khoshro. Fault detection and diagnosis of an industrial steam turbine using fusion of svm (support vector machine) and anfis (adaptive neuro-fuzzy inference system) classifiers. *Energy*, 35(12):5472–5482, 2010.
- [6] Nassim Laouti, Nida Sheibat-Othman, and Sami Othman. Support vector machines for fault detection in wind turbines. *IFAC Proceedings Volumes*, 44(1):7067–7072, 2011.
- [7] Wang Hongm, Chai Tian-You, Ding Jin-Liang, and Martin Brown. Data driven fault diagnosis and fault tolerant control: some advances and possible new directions. *Acta Automatica Sinica*, 35(6):739–747, 2009.
- [8] Shen Yin, Steven X Ding, Adel Haghani, Haiyang Hao, and Ping Zhang. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark tennessee eastman process. *Journal of process control*, 22(9):1567–1581, 2012.
- [9] S Joe Qin. Survey on data-driven industrial process monitoring and diagnosis. *Annual reviews in control*, 36(2):220–234, 2012.
- [10] Ramgopal Dhaka, Bhoopendra Pachauri, and Anamika Jain. Two-dimensional srgm with delay in debugging by considering the uncertainty factor and predictive analysis. *Reliability: Theory & Applications*, 16(SI 2 (64)):82–94, 2021.
- [11] Amitkumar Patil, Gunjan Soni, Anuj Prakash, and Mangey Ram. Intelligent valve fault diagnosis approach for reciprocating compressor based on acoustic signals. *Reliability: Theory & Applications*, 16(SI 2 (64)):35–47, 2021.

- [12] GM Malwatkar, SH Sonawane, and LM Waghmare. Tuning pid controllers for higher-order oscillatory systems with improved performance. *ISA transactions*, 48(3):347–353, 2009.
- [13] SR Shiledar, GM Malwatkar, IS Jadhav, and GV Lakhekar. Design of discrete sliding mode controller for higher order system. *Reliability: Theory & Applications*, 16(SI 1 (60)):90–97, 2021.
- [14] GM Malwatkar, AA Khandekar, VG Asutkar, and LM Waghmare. Design of centralized pi/pid controller: interaction measure approach. In *2008 IEEE Region 10 and the Third international Conference on Industrial and Information Systems*, pages 1–6. IEEE, 2008.
- [15] VS Biradar and GM Malwatkar. Second order sliding mode control for robust performance of the systems. *Reliability: Theory & Applications*, 17(4 (71)):358–370, 2022.
- [16] Shaktikumar R Shiledar and Gajanan M Malwatkar. A compressive study on discrete time sliding mode control with advances and applications. *International Journal of Systems, Control and Communications*, 13(4):283–310, 2022.
- [17] Jinane Harmouche, Claude Delpha, and Demba Diallo. Incipient fault detection and diagnosis based on kullback–leibler divergence using principal component analysis: Part i. *Signal processing*, 94:278–287, 2014.
- [18] Evan L Russell, Leo H Chiang, and Richard D Braatz. Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis. *Chemometrics and intelligent laboratory systems*, 51(1):81–93, 2000.
- [19] Gang Li, S Joe Qin, and Tao Yuan. Data-driven root cause diagnosis of faults in process industries. *Chemometrics and Intelligent Laboratory Systems*, 159:1–11, 2016.
- [20] Sang Wook Choi, Changkyu Lee, Jong-Min Lee, Jin Hyun Park, and In-Beum Lee. Fault detection and identification of nonlinear processes based on kernel pca. *Chemometrics and intelligent laboratory systems*, 75(1):55–67, 2005.
- [21] Xi Sun, Horacio J Marquez, Tongwen Chen, and Muhammad Riaz. An improved pca method with application to boiler leak detection. *ISA transactions*, 44(3):379–397, 2005.
- [22] Shallon Stubbs, Jie Zhang, and Julian Morris. Fault detection in dynamic processes using a simplified monitoring-specific cva state space modelling approach. *Computers & Chemical Engineering*, 41:77–87, 2012.
- [23] Benben Jiang, Dexian Huang, Xiaoxiang Zhu, Fan Yang, and Richard D Braatz. Canonical variate analysis-based contributions for fault identification. *Journal of Process Control*, 26:17–25, 2015.
- [24] Benben Jiang, Xiaoxiang Zhu, Dexian Huang, and Richard D Braatz. Canonical variate analysis-based monitoring of process correlation structure using causal feature representation. *Journal of Process Control*, 32:109–116, 2015.
- [25] Cristobal Ruiz-Cárcel, Yi Cao, D Mba, Liyun Lao, and RT Samuel. Statistical process monitoring of a multiphase flow facility. *Control Engineering Practice*, 42:74–88, 2015.
- [26] Xi Zhang and Karlene A Hoo. Effective fault detection and isolation using bond graph-based domain decomposition. *Computers & chemical engineering*, 35(1):132–148, 2011.
- [27] Kamal Hadad, Meisam Pourahmadi, and Hosein Majidi-Maraghi. Fault diagnosis and classification based on wavelet transform and neural network. *Progress in nuclear energy*, 53(1):41–47, 2011.
- [28] P Konar and P Chattopadhyay. Bearing fault detection of induction motor using wavelet and support vector machines (svms). *Applied Soft Computing*, 11(6):4203–4211, 2011.
- [29] Kris Villez, Babji Srinivasan, Raghunathan Rengaswamy, Shankar Narasimhan, and Venkat Venkatasubramanian. Kalman-based strategies for fault detection and identification (fdi): Extensions and critical evaluation for a buffer tank system. *Computers & chemical engineering*, 35(5):806–816, 2011.
- [30] Yunkai Wu, Bin Jiang, Ningyun Lu, Hao Yang, and Yang Zhou. Multiple incipient sensor faults diagnosis with application to high-speed railway traction devices. *ISA transactions*, 67:183–192, 2017.
- [31] Farzin Piltan and Jong-Myon Kim. Bearing fault diagnosis using an extended variable structure feedback linearization observer. *Sensors*, 18(12):4359, 2018.

- [32] Emanuel Bernardi and Eduardo J Adam. Observer-based fault detection and diagnosis strategy for industrial processes. *Journal of the Franklin Institute*, 357(14):10054–10081, 2020.
- [33] Hendrik M Odendaal and Thomas Jones. Actuator fault detection and isolation: An optimised parity space approach. *Control Engineering Practice*, 26:222–232, 2014.
- [34] Maiying Zhong, Yang Song, and Steven X Ding. Parity space-based fault detection for linear discrete time-varying systems with unknown input. *Automatica*, 59:120–126, 2015.
- [35] Achmad Widodo and Bo-Suk Yang. Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical systems and signal processing*, 21(6):2560–2574, 2007.
- [36] Jonghyuck Park, Ick-Hyun Kwon, Sung-Shick Kim, and Jun-Geol Baek. Spline regression based feature extraction for semiconductor process fault detection using support vector machine. *Expert Systems with Applications*, 38(5):5711–5718, 2011.
- [37] Tribeni Prasad Banerjee and Swagatam Das. Multi-sensor data fusion using support vector machine for motor fault detection. *Information Sciences*, 217:96–107, 2012.
- [38] Yingchao Xiao, Huangang Wang, Lin Zhang, and Wenli Xu. Two methods of selecting gaussian kernel parameters for one-class svm and their application to fault detection. *Knowledge-Based Systems*, 59:75–84, 2014.
- [39] Chen Jing and Jian Hou. Svm and pca based fault classification approaches for complicated industrial process. *Neurocomputing*, 167:636–642, 2015.
- [40] Shiyong Zhong, Qiaojun Wen, and Zhiqiang Ge. Semi-supervised fisher discriminant analysis model for fault classification in industrial processes. *Chemometrics and Intelligent Laboratory Systems*, 138:203–211, 2014.
- [41] Jueun Kwak, Taehyung Lee, and Chang Ouk Kim. An incremental clustering-based fault detection algorithm for class-imbalanced process data. *IEEE Transactions on Semiconductor Manufacturing*, 28(3):318–328, 2015.
- [42] Ki Bum Lee, Sejune Cheon, and Chang Ouk Kim. A convolutional neural network for fault classification and diagnosis in semiconductor manufacturing processes. *IEEE Transactions on Semiconductor Manufacturing*, 30(2):135–142, 2017.
- [43] Joe-Air Jiang, Cheng-Long Chuang, Yung-Chung Wang, Chih-Hung Hung, Jiing-Yi Wang, Chien-Hsing Lee, and Ying-Tung Hsiao. A hybrid framework for fault detection, classification, and location-part ii: implementation and test results. *IEEE transactions on power delivery*, 26(3):1999–2008, 2011.
- [44] Ihab Samy, Ian Postlethwaite, and Da-Wei Gu. Survey and application of sensor fault detection and isolation schemes. *Control Engineering Practice*, 19(7):658–674, 2011.
- [45] S Sina Tayarani-Bathaie and K Khorasani. Fault detection and isolation of gas turbine engines using a bank of neural networks. *Journal of Process Control*, 36:22–41, 2015.
- [46] Likun Ren and Weimin Lv. Fault detection via sparse representation for semiconductor manufacturing processes. *IEEE Transactions on Semiconductor Manufacturing*, 27(2):252–259, 2014.
- [47] David Leung and Jose Romagnoli. An integration mechanism for multivariate knowledge-based fault diagnosis. *Journal of Process Control*, 12(1):15–26, 2002.
- [48] Roozbeh Razavi-Far, Hadi Davilu, Vasile Palade, and Caro Lucas. Model-based fault detection and isolation of a steam generator using neuro-fuzzy networks. *Neurocomputing*, 72(13-15):2939–2951, 2009.
- [49] Kim Verbert, R Babuška, and Bart De Schutter. Combining knowledge and historical data for system-level fault diagnosis of hvac systems. *Engineering Applications of Artificial Intelligence*, 59:260–273, 2017.
- [50] Mihiran Galagedarage Don and Faisal Khan. Dynamic process fault detection and diagnosis based on a combined approach of hidden markov and bayesian network model. *Chemical Engineering Science*, 201:82–96, 2019.
- [51] Fu Xiao, Yang Zhao, Jin Wen, and Shengwei Wang. Bayesian network based fdd strategy for variable air volume terminals. *Automation in Construction*, 41:106–118, 2014.

- [52] Haitao Wang, Youming Chen, Cary WH Chan, and Jianying Qin. An online fault diagnosis tool of vav terminals for building management and control systems. *Automation in Construction*, 22:203–211, 2012.
- [53] Amir Rafati, Hamid Reza Shaker, and Saman Ghahghahzadeh. Fault detection and efficiency assessment for hvac systems using non-intrusive load monitoring: A review. *Energies*, 15(1):341, 2022.
- [54] Xiuying Yan, Ting Guan, Kaixing Fan, and Qing Sun. Novel double layer bilstm minor soft fault detection for sensors in air-conditioning system with kpca reducing dimensions. *Journal of Building Engineering*, 44:102950, 2021.
- [55] Narges Torabi, H Burak Gunay, William O'Brien, and Ricardo Moromisato. Inverse model-based virtual sensors for detection of hard faults in air handling units. *Energy and Buildings*, 253:111493, 2021.
- [56] F Pelella, L Viscito, and AW Mauro. Soft faults in residential heat pumps: Possibility of evaluation via on-field measurements and related degradation of performance. *Energy Conversion and Management*, 260:115646, 2022.
- [57] Yifeng Hu and David P Yuill. Impacts of common faults on an air conditioner with a microtube condenser and analysis of fault characteristic features. *Energy and Buildings*, 254:111630, 2022.
- [58] Yuzhou Wang, Zhengfei Li, Huanxin Chen, Jianxin Zhang, Qian Liu, Junfeng Wu, and Limei Shen. Research on diagnostic strategy for faults in vrf air conditioning system using hybrid data mining methods. *Energy and Buildings*, 247:111144, 2021.
- [59] Zhenxin Zhou, Guannan Li, Jiangyu Wang, Huanxin Chen, Hanlu Zhong, and Zihan Cao. A comparison study of basic data-driven fault diagnosis methods for variable refrigerant flow system. *Energy and Buildings*, 224:110232, 2020.
- [60] Noor Asyikin Sulaiman, Md Pauzi Abdullah, Hayati Abdullah, Muhammad Noorazlan Shah Zainudin, and Azdiana Md Yusop. Fault detection for air conditioning system using machine learning. *IAES International Journal of Artificial Intelligence*, 9(1):109, 2020.
- [61] Sana Ullah Jan, Young Doo Lee, and In Soo Koo. A distributed sensor-fault detection and diagnosis framework using machine learning. *Information Sciences*, 547:777–796, 2021.
- [62] Woohyun Kim and Je-Hyeon Lee. Fault detection and diagnostics analysis of air conditioners using virtual sensors. *Applied Thermal Engineering*, 191:116848, 2021.
- [63] Akshay Ranade, Gregory Provan, Alie El-Din Mady, and Dominic O'Sullivan. A computationally efficient method for fault diagnosis of fan-coil unit terminals in building heating ventilation and air conditioning systems. *Journal of Building Engineering*, 27:100955, 2020.
- [64] Hesam Hassanpour, Prashant Mhaskar, John M House, and Timothy I Salisbury. A hybrid modeling approach integrating first-principles knowledge with statistical methods for fault detection in hvac systems. *Computers & Chemical Engineering*, 142:107022, 2020.
- [65] Antonio Gálvez, Alberto Diez-Oliván, Dammika Seneviratne, and Diego Galar. Fault detection and rul estimation for railway hvac systems using a hybrid model-based approach. *Sustainability*, 13(12):6828, 2021.
- [66] Xiang Xie, Jorge Merino, Nicola Moretti, Pieter Pauwels, Janet Yoon Chang, and Ajith Parlikad. Digital twin enabled fault detection and diagnosis process for building hvac systems. *Automation in Construction*, 146:104695, 2023.
- [67] Liping Wang, James Braun, and Sujit Dahal. An evolving learning-based fault detection and diagnosis method: Case study for a passive chilled beam system. *Energy*, 265:126337, 2023.
- [68] Cunxiao Shen, Hanyuan Zhang, Songping Meng, and Chengdong Li. Augmented data driven self-attention deep learning method for imbalanced fault diagnosis of the hvac chiller. *Engineering Applications of Artificial Intelligence*, 117:105540, 2023.
- [69] Mohammed G Albayati, Jalal Faraj, Amy Thompson, Prathamesh Patil, Ravi Gorthala, and Sanguthevar Rajasekaran. Semi-supervised machine learning for fault detection and diagnosis of a rooftop unit. *Big Data Mining and Analytics*, 6(2):170–184, 2023.
- [70] Ke Yan and Xiaokang Zhou. Chiller faults detection and diagnosis with sensor network and adaptive 1d cnn. *Digital Communications and Networks*, 8(4):531–539, 2022.

- [71] Xi Bai, Muxing Zhang, Zhenghao Jin, Yilin You, and Caihua Liang. Fault detection and diagnosis for chiller based on feature-recognition model and kernel discriminant analysis. *Sustainable Cities and Society*, 79:103708, 2022.
- [72] Mark S Breuker and James E Braun. Common faults and their impacts for rooftop air conditioners. *Hvac & R Research*, 4(3):303–318, 1998.
- [73] John Proctor and Tom Downey. Transforming routine air conditioner maintenance practices to improve equipment efficiency and performance. In *Proceedings of the 1999 international energy program evaluation conference*, pages 1–12, 1999.
- [74] Chris Neme, Steven Nadel, and John Proctor. Energy savings potential from addressing residential air conditioner and heat pump installation problems. American Council for an Energy-Efficient Economy Washington, DC, 1999.
- [75] Austin P Rogers and Bryan P Rasmussen. Opportunities for consumer-driven load shifting in commercial and industrial buildings. *Sustainable Energy, Grids and Networks*, 16:243–258, 2018.
- [76] Venkat Venkatasubramanian, Raghunathan Rengaswamy, Kewen Yin, and Surya N Kavuri. A review of process fault detection and diagnosis part i quantitative model based methods. *Computers & chemical engineering*, 27(3):293–311, 2003.
- [77] Srinivas Katipamula and Michael R Brambley. Methods for fault detection, diagnostics, and prognostics for building systems a review, part i. *Hvac& R Research*, 11(1):3–25, 2005.
- [78] Woohyun Kim and Srinivas Katipamula. A review of fault detection and diagnostics methods for building systems. *Science and Technology for the Built Environment*, 24(1):3–21, 2018.
- [79] Jasmin Anika Gärtner, Francesco Massa Gray, and Thomas Auer. Assessment of the impact of hvac system configuration and control zoning on thermal comfort and energy efficiency in flexible office spaces. *Energy and Buildings*, 212:109785, 2020.
- [80] Pawel Wargocki, Jan Sundell, W Bischof, G Brundrett, Povl Ole Fanger, F Gyntelberg, SO Hanssen, P Harrison, A Pickering, Olli Seppänen, et al. The role of ventilation and hvac systems for human health in nonindustrial indoor environments. a supplementary review by euroven group. In *9th International Conference on Indoor Air Quality and Climate*, 2002.
- [81] Olli Seppanen, William J Fisk, and QH Lei. Effect of temperature on task performance in office environment. 2006.
- [82] Kurt W Roth, Detlef Westphalen, Patricia Llana, and Michael Feng. The energy impact of faults in us commercial buildings. 2004.
- [83] Jianli Chen, Liang Zhang, Yanfei Li, Yifu Shi, Xinghua Gao, and Yuqing Hu. A review of computing-based automated fault detection and diagnosis of heating, ventilation and air conditioning systems. *Renewable and Sustainable Energy Reviews*, 161:112395, 2022.
- [84] Yang Zhao, Tingting Li, Xuejun Zhang, and Chaobo Zhang. Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future. *Renewable and Sustainable Energy Reviews*, 109:85–101, 2019.
- [85] Debashis Dey and Bing Dong. A probabilistic approach to diagnose faults of air handling units in buildings. *Energy and Buildings*, 130:177–187, 2016.
- [86] Qianjun Mao, Xi Fang, Yunpeng Hu, and Guannan Li. Chiller sensor fault detection based on empirical mode decomposition threshold denoising and principal component analysis. *Applied Thermal Engineering*, 144:21–30, 2018.
- [87] Shengwei Wang and Jianying Qin. Sensor fault detection and validation of vav terminals in air conditioning systems. *Energy conversion and management*, 46(15-16):2482–2500, 2005.
- [88] Cheng Fan, Fu Xiao, and Chengchu Yan. A framework for knowledge discovery in massive building automation data and its application in building diagnostics. *Automation in Construction*, 50:81–90, 2015.
- [89] Guannan Li, Yunpeng Hu, Huanxin Chen, Haorong Li, Min Hu, Yabin Guo, Shubiao Shi, and Wenju Hu. A sensor fault detection and diagnosis strategy for screw chiller system using support vector data description-based d-statistic and dv-contribution plots. *Energy and Buildings*, 133:230–245, 2016.



- [90] Ding Li, Donghui Li, Chengdong Li, Lin Li, and Long Gao. A novel data-temporal attention network based strategy for fault diagnosis of chiller sensors. *Energy and Buildings*, 198:377–394, 2019.
- [91] Zhimin Du, Ling Chen, and Xinqiao Jin. Data-driven based reliability evaluation for measurements of sensors in a vapor compression system. *Energy*, 122:237–248, 2017.
- [92]
- [93] Fu Xiao, Shengwei Wang, Xinhua Xu, and Gaoming Ge. An isolation enhanced pca method with expert-based multivariate decoupling for sensor fdd in air-conditioning systems. *Applied Thermal Engineering*, 29(4):712–722, 2009.
- [94] Yuebin Yu, Denchai Woradechjumroen, and Daihong Yu. A review of fault detection and diagnosis methodologies on air-handling units. *Energy and Buildings*, 82:550–562, 2014.
- [95] Stephen Frank, Xin Jin, Daniel Studer, and Amanda Farthing. Assessing barriers and research challenges for automated fault detection and diagnosis technology for small commercial buildings in the united states. *Renewable and Sustainable Energy Reviews*, 98:489–499, 2018.
- [96] Zixiao Shi and William O'Brien. Development and implementation of automated fault detection and diagnostics for building systems: A review. *Automation in Construction*, 104:215–229, 2019.
- [97] Maryam Sadat Mirnaghi and Fariborz Haghighat. Fault detection and diagnosis of large-scale hvac systems in buildings using data-driven methods: A comprehensive review. *Energy and Buildings*, 229:110492, 2020.
- [98] Yanfei Li and Zheng Oneill. A critical review of fault modeling of hvac systems in buildings. In *Building Simulation*, volume 11, pages 953–975. Springer, 2018.
- [99] Ming Zhao, Jing Lin, Xiaoqiang Xu, and Xuejun Li. Multi fault detection of rolling element bearings under harsh working condition using imf based adaptive envelope order analysis. *Sensors*, 14(11):20320–20346, 2014.
- [100] Bilal Asad, Toomas Vaimann, Anton Rassolkin, Ants Kallaste, and Anouar Belahcen. A survey of broken rotor bar fault diagnostic methods of induction motor. *Electrical, Control and Communication Engineering*, 14(2):117–124, 2018.
- [101] Siddique Akbar, Toomas Vaimann, Bilal Asad, Ants Kallaste, Muhammad Usman Sardar, and Karolina Kudelina. State-of-the-art techniques for fault diagnosis in electrical machines: Advancements and future directions. *Energies*, 16(17):6345, 2023.
- [102] Yongzhi Qu, David He, Jae Yoon, Brandon Van Hecke, Eric Bechhoefer, and Junda Zhu. Gearbox tooth cut fault diagnostics using acoustic emission and vibration sensors a comparative study. *Sensors*, 14(1):1372–1393, 2014.
- [103] NR Sakthivel, V Sugumaran, and S Babudevasenapati. Vibration based fault diagnosis of monoblock centrifugal pump using decision tree. *Expert Systems with Applications*, 37(6):4040–4049, 2010.
- [104] Shao Haidong, Jiang Hongkai, Li Xingqiu, and Wu Shuaipeng. Intelligent fault diagnosis of rolling bearing using deep wavelet auto-encoder with extreme learning machine. *Knowledge-Based Systems*, 140:1–14, 2018.
- [105] AR Mohanty and S Fatima. Shaft misalignment detection by thermal imaging of support bearings. *IFAC-PapersOnLine*, 48(21):554–559, 2015.
- [106] J Kubiak, G Urquiza, JA Rodriguez, G González, I Rosales, G Castillo, and J Nebradt. Failure analysis of the 150 mw gas turbine blades. *Engineering Failure Analysis*, 16(6):1794–1804, 2009.
- [107] Anurag Choudhary, Deepam Goyal, Sudha Letha Shimi, and Aparna Akula. Condition monitoring and fault diagnosis of induction motors: A review. *Archives of Computational Methods in Engineering*, 26:1221–1238, 2019.
- [108] Bin Fan, Song Feng, Yitong Che, Junhong Mao, and Youbai Xie. An oil monitoring method of wear evaluation for engine hot tests. *The International Journal of Advanced Manufacturing Technology*, 94:3199–3207, 2018.

# M/M/C QUEUE WITH MULTIPLE WORKING VACATIONS AND SINGLE WORKING VACATION UNDER ENCOURAGED ARRIVAL WITH IMPATIENT CUSTOMERS

PRAKATI P, JULIA ROSE MARY K



*Department of Mathematics, Nirmala College for Women, India*

prakatidhanam@gmail.com

juliakulandaisamy@gmail.com

## Abstract

*This paper demonstrates an M/M/C queuing model with Multiple working vacations and also single working vacation under encouraged arrival with impatient customers. The queuing model with the servers adopting multiple working vacation policy and single working vacation are determined separately and it is observed that the servers during working vacation(s) will be serving the customers at a slower service rate when compared during regular busy period. In addition to the above conditions, if there is a rapid increase in the customers' arrival i.e, if encouraged arrival occurs and due to this sudden growth of the queue, there may be a impatience in the behaviour of the customer. With these considerations, an M/M/C Queuing model is analysed with two vacation policies separately by applying PGF method and thus the performance measures for an M/M/C Queue with Multiple Working Vacations and Single Working Vacation under Encouraged arrival with impatient customers are evaluated.*

**Keywords:** Multiple Working Vacations(MWV), Single Working Vacation (SWV), Encouraged Arrival, Impatient behaviour, Performance Measures

## 1. INTRODUCTION

In our daily life, we meet up with the scenario of waiting in queues to get our work done, for example - to make bank deposit, mail a package, obtain food in cafeteria etc. Waiting in queue is a matter of personal annoyance and it also costs the amount of time that we waste by waiting in queues. It may affect the efficiency of the service provided and is a major factor in both the quality of life and also affecting the efficiency of a nation's economy. Great inefficiencies also occur because of waiting.

For example, making machines wait to be repaired may result in less production, delay in telecommunication transmission due to saturated lines may cause data glitches etc. In fact, we have become accustomed to considerable amounts of waiting. Origin of Queuing theory in research was contributed by Agner Krarup Erlang, who created models to describe the system of incoming calls at the Copenhagen Telephone Exchange Company.

An M/M/s queuing system in which the servers under going vacation was analysed in [7]. In Queuing vacation policy, an overview of some general decomposition results were attained and the methodology used to obtain those results for two vacation models were analysed in [3]. Moreover, the literature on statistical analysis of queuing systems were briefly discussed in [2].

It can be observed that in numerous industrial sector, the concept of Queuing with servers' vacation is implemented. An M/G/1 Queue with vacation policy used in the scenarios like

maintenance of production systems, where machines or equipment mainly degrade while being operated were evaluated and for such queuing model, an explicit expression for the distribution of the time it takes until the specific amount of work has been served were derived in [1].

In General, Systems with vacations are usually modeled and analyzed by queuing theory. An approach for modeling and analyzing finite-source multi-server systems with single and multiple vacations of servers or all stations were presented using the Generalized Stochastic Petri nets model in [11]. During any service, the servers may undergo breakdown simultaneously both in regular busy period and working vacation period due to the failure of a main control unit. This scenario was discussed by modeling and analysing a Markovian multiserver finite buffer queue under synchronous working vacation policy in [5].

A multiserver queuing system with customers' impatience until the end of service under single and multiple vacation policies were examined in [6]. Situations like arrival of the customers following Poisson distribution but the general distribution followed by the administration rendering service with various vacations were detailedly discussed in [10].

The concept of impatient behaviours like balking and reneging with the availability of heterogeneous servers in an M/M/c queue was analysed in [16]. Moreover, the time-dependent system size probabilities were derived explicitly using generating function and also the time-dependent mean, variance, busy period distribution and steady-state probabilities were also obtained. In addition to this, performance of an M/M/c/K Queuing Models applied in Healthcare Things for Medical Monitoring were evaluated in [14].

The impatient nature of the customer during any service may be expressed if there is a delay in the service and the delay may be due to lack of servers or slow service provided. Queues with slow servers and impatient customers were considered and the mean queue size were derived. Also, Several extreme cases were investigated and numerical results are presented in [12].

An M/M/1 queue with single and multiple working vacations with impatient customers were studied and Closed-form solutions and various performance measures like, the mean queue lengths and the mean waiting times were derived and the stochastic decomposition properties were verified for both multiple and single working vacation cases in [13]. Likely, the impatient behaviour of the customers with single and multiple synchronous working vacations in an M/M/C queue was analysed in [9]. Performance nature of a Markovian Queue with Impatient Customers and Working Vacation were derived in [8].

It is obvious that in the case of any discounts or offers provided during any sale or if any sudden demand is created for a product or a service, then there will be a rapid growth in the arrival of the customers, which is termed as encouraged arrival. The concept of encouraged arrival in an M/M/c/N queuing systems with reneging, retention and Feedback customers were discussed in [15]. The stationary system size probabilities were obtained recursively for the above model, while the steady state behavior of the M/M/1/N queuing model with encouraged or discouraged arrivals and impatient customers are obtained in [4].

With the aid of the above discussed concepts, an M/M/C Queuing model during encouraged arrival under going single working vacation and multiple working vacations with impatient behaviour of the customers are analysed separately.

In this paper, between the two vacation policies analysed, multiple working vacation is considered first in which if a server returns to an empty queue, then he goes for another vacation immediately, thus working vacation occurs multiple times. Whereas, in the later vacation policy, the server takes only a single vacation each time. Thus for an M/M/C Queue during encouraged arrival with impatient behaviour under going multiple working vacation is derived with explicit formulations followed by the same queuing model with single working vacation.

## 2. METHODS

An M/M/c queuing model with encouraged arrival following multiple working vacations with impatient customers is considered. Customers arriving to be served follow Poisson process and the arrival rate is denoted by the parameter  $\lambda_w$ . If there is a sudden increase in the arrival of

the customers,i.e., encouraged arrival occurring in the system follows poisson process with the encouraged arrival rate  $\lambda_w(1 + \delta)$ .

Since the considered model denotes 'c' servers, there may be maximum of 'c' servers available, to serve the customers according to FCFS rule. When a customer arrives and find all the servers in the system are busy, then he needs to wait until he gets served and thus the waiting line or the queue begins.

The time taken for each server to complete the work during regular busy period follows exponential distribution and denoted with the service rate  $\mu_w$ . Thus the traffic intensity or the stability of the system during regular busy period is considered as  $\rho = \frac{\lambda_w(1+\delta)}{c\mu_w} < 1$

After completion of a service, if there is no customer in the system, then all the 'c' servers will take vacation promptly and the duration of working vacation for each servers is exponentially distributed with parameter  $\eta'$ . As all the servers in the system undergo vacation, even if a single customer arrives, then any one of the server will return from his vacation and start serving the arrived customer. Thus the concept of working even during vacation for the arrival of customers is termed as working vacation period, and the service rate following exponential process during working vacation period is  $\mu_{wv}$  and it is observed that the service rate during working vacation is slower than the regular busy period i.e.,  $\mu_{wv} < \mu_w$

It is obvious that if the servers return from their vacation and when the system is non empty, the service rate of the servers changes from  $\mu_{wv}$  to  $\mu_w$  indicating that the regular busy period begins. Suppose, if the servers find no customer waiting in the queue after returning from their vacation, they immediately leave for another vacation. In such cases, if a customer waits in the queue for a longer time, as all the 'c' servers are in working vacation period, he may become impatient in waiting and the impatient behaviour of the customer at the time T is exponentially distributed with parameter  $\gamma_w$  which is considered to be independent of the customers in that moment.

The customer waiting in the queue may exit the queue and never returns if its service has not been completed before the time T expires. The inter arrival times, service times, vacation duration times and impatient time are all taken to be mutually independent. To construct this system, we define a two dimensional continuous time discrete state Markov chain as  $\{(M(t), N(t)), t \geq 0\}$  with state space  $s = \{(0, 0) \cup \{(n, j)\}, n \geq 1, j = 0, 1\}$

Where  $M(t)$  denotes the total number of customers in the system at time t and  $N(t)$  denotes the state of the system at time t with

$N(t) = \{1 \text{ when the servers are in non-vacation period at time } t\}$  and

$N(t) = \{0 \text{ when the servers are in working vacation period at time } t\}$ .

### 2.1. Steady State Equations and its Solutions for Multiple Working Vacations Model:

The steady state transition probabilities are defined by

$$P_{nj} = P\{M(t) = n, N(t) = j\}, n \geq 0, j = 0, 1$$

Now, the set of balance equations as

$$\lambda_w(1 + \delta)P_{00} = (\mu_{wv} + \gamma_w)P_{1,0} + \mu_w P_{1,1}, \tag{1}$$

$$[\lambda_w(1 + \delta) + \eta' + n(\mu_{wv} + \gamma_w)]P_{n,0} = \lambda_w(1 + \delta)P_{n-1,0} + (n + 1)((\mu_{wv} + \gamma_w)P_{n+1,0}, \text{ if } n \geq 1, \tag{2}$$

$$(\lambda_w(1 + \delta) + \mu_w)P_{1,1} = \eta' P_{1,0} + 2\mu_w P_{2,1}, \tag{3}$$

$$(\lambda_w(1 + \delta) + n\mu_w)P_{n,1} = \lambda_w(1 + \delta)P_{n-1,1} + (n + 1)\mu_w P_{n+1,1} + \eta' P_{n,0}, \text{ if } 2 \leq n \leq c - 1, \tag{4}$$

$$(\lambda_w(1 + \delta) + c\mu_w)P_{n,1} = \lambda_w(1 + \delta)P_{n-1,1} + c\mu_w P_{n+1,1} + \eta' P_{n,0} \text{ if } n \geq c. \tag{5}$$

By letting the probability generating functions as

$$P_0(z) = \sum_{n=0}^{\infty} z^n P_{n,0},$$

$$P_1(z) = \sum_{n=1}^{\infty} z^n P_{n,1}.$$

with  $P_0(1) + P_1(1) = 1$  and  $P_0'(z) = \sum_{n=1}^{\infty} n z^{n-1} P_{n,0}$ .

Now, By Multiplying Eq(2) with  $z^n$  and adding over 'n' and rearranging the terms, the differential equation is attained as :

$$(\mu_{wv} + \gamma_w)(1 - z)P_0'(z) = [\lambda_w(1 + \delta)(1 - z) + \eta']P_0(z) - (\eta'P_{0,0} + \mu_w P_{1,1}). \quad (6)$$

Likely multiplying Eq(4) and Eq(5) by  $z^n$  and adding over 'n', the following equation is obtained,

$$(1 - z)(\lambda_w(1 + \delta)z - c\mu_w)P_1(z) = \eta'zP_0(z) - (\eta'P_{0,0} + \mu_w P_{1,1})z + \mu_w(1 - z) \sum_{n=1}^c (n - c)z^n P_{n,1}. \quad (7)$$

Let us consider ,

$$A = \eta'P_{0,0} + \mu_w P_{1,1}. \quad (8)$$

Then, for  $z \neq 1$ ,

$$P_0'(z) - \left[ \frac{\lambda_w(1 + \delta)}{(\mu_{wv} + \gamma_w)} + \frac{\eta'}{(\mu_{wv} + \gamma_w)(1 - z)} \right] P_0(z) = - \frac{A}{(\mu_{wv} + \gamma_w)(1 - z)}. \quad (9)$$

Eq(9) is an ordinary linear differential equation with constant coefficients To solve the equation, an integrating factor can be considered as

$$I.F = e^{-\int \left[ \frac{\lambda_w(1 + \delta)}{(\mu_{wv} + \gamma_w)} + \frac{\eta'}{(\mu_{wv} + \gamma_w)(1 - z)} \right] dz} = e^{-\frac{\lambda_w(1 + \delta)z}{(\mu_{wv} + \gamma_w)} (1 - z) \frac{\eta'}{(\mu_{wv} + \gamma_w)}}$$

The General solution to Eq(9) is given by:

$$\frac{d}{dz} \left[ e^{-\frac{\lambda_w(1 + \delta)z}{(\mu_{wv} + \gamma_w)} (1 - z) \frac{\eta'}{(\mu_{wv} + \gamma_w)}} P_0(z) \right] = \left[ \frac{-A}{(\mu_{wv} + \gamma_w)(1 - z)} \right] e^{-\frac{\lambda_w(1 + \delta)z}{(\mu_{wv} + \gamma_w)} (1 - z) \frac{\eta'}{(\mu_{wv} + \gamma_w)}}. \quad (10)$$

Now, integrating from 0 to z, following equation is attained,

$$P_0(z) = \left[ e^{\frac{\lambda_w(1 + \delta)z}{(\mu_{wv} + \gamma_w)} (1 - z) \frac{\eta'}{(\mu_{wv} + \gamma_w)}} P_0(0) - \frac{A}{(\mu_{wv} + \gamma_w)} \int_0^z e^{-\frac{\lambda_w(1 + \delta)z}{(\mu_{wv} + \gamma_w)} (1 - x) \frac{\eta'}{(\mu_{wv} + \gamma_w)}} dx \right]. \quad (11)$$

then,

$$P_0(1) = e^{\frac{\lambda_w(1 + \delta)}{(\mu_{wv} + \gamma_w)}} \left[ P_0(0) - \frac{A}{(\mu_{wv} + \gamma_w)} \int_0^1 e^{-\frac{\lambda_w(1 + \delta)z}{(\mu_{wv} + \gamma_w)} (1 - x) \frac{\eta'}{(\mu_{wv} + \gamma_w)}} dx \right] \lim_{z \rightarrow 1} (1 - z)^{-\frac{\eta'}{(\mu_{wv} + \gamma_w)}}. \quad (12)$$

Since  $0 \leq P_0(1) = \sum_{n=0}^{\infty} z^n P_{n,0} \leq 1$  and  $\lim_{z \rightarrow 1} (1 - z)^{-\frac{\eta'}{(\mu_{wv} + \gamma_w)}} = \infty$ , and thus the existing term is as follows

$$P_{0,0} = P_0(0) = \frac{A}{(\mu_{wv} + \gamma_w)} L \tag{13}$$

Where  $L = \int_0^1 e^{-\frac{\lambda_w(1+\delta)z}{(\mu_{wv}+\gamma_w)}} (1-x)^{\frac{\eta'}{(\mu_{wv}+\gamma_w)}-1} dx.$  (14)

Defin  $Z(\lambda_w(1 + \delta), \eta') = -\lambda_w(1 + \delta)^{-\eta'} e^{-\lambda_w(1+\delta)} (-\Gamma(\eta', -\lambda_w(1 + \delta)) + \Gamma(\eta'))$  (15)

where  $\Gamma(z)$  is the  $\Gamma$  function which is represented as

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt \tag{16}$$

and  $\Gamma(y, z) = \int_z^\infty e^{-t} t^{y-1} dt.$  (17)

some calculations give

$$L = Z\left(\frac{\lambda_w(1 + \delta)}{(\mu_{wv} + \gamma_w)}, \frac{\eta'}{(\mu_{wv} + \gamma_w)}\right). \tag{18}$$

By Eq(8) and Eq (13), it is observed that

$$P_{0,0} = \frac{\eta' P_{0,0} + \mu_w P_{1,1}}{(\mu_{wv} + \gamma_w)} L = \frac{L \mu_w}{\mu_{wv} + \gamma_w - \eta' L} P_{1,1}. \tag{19}$$

Now, using the value of A from Eq(13) in Eq(11),  $P_0(z)$  is obtained as

$$P_0(z) = \frac{e^{\frac{\lambda_w(1+\delta)z}{(\mu_{wv}+\gamma_w)}}}{(1-z)^{\frac{\eta'}{(\mu_{wv}+\gamma_w)}}} \left[1 - \frac{1}{L} \int_0^z e^{-\frac{\lambda_w(1+\delta)x}{(\mu_{wv}+\gamma_w)}} (1-x)^{\frac{\eta'}{(\mu_{wv}+\gamma_w)}-1} dx\right] P_{0,0}. \tag{20}$$

By applying L'Hospital's rule to Eq(20), we get

$$P_0(1) = \frac{(\mu_{wv} + \gamma_w)}{\eta' L} P_{0,0} \tag{21}$$

and now substituting the value of  $P_{0,0}$  from Eq(19), the following relation is obtained

$$\eta' P_0(1) = \eta' P_{0,0} + \mu_w P_{1,1}. \tag{22}$$

From Eq(7),  $P_1(z)$  is attained as,

$$P_1(z) = \frac{[\eta' P_0(z) - A]z}{(\lambda_w(1 + \delta)z - c\mu_w)(1 - z)} - \frac{\mu_w}{(\lambda_w(1 + \delta)z - c\mu_w)} F(z), \tag{23}$$

where,

$$F(z) = \sum_{n=1}^c (n - c) z^n P_{n,1}. \tag{24}$$

It is clear from Eq(20) that  $P_0(z)$  is a function of  $P_{0,0}$  and the ratio between the time of the servers on working vacation and the system is empty. Similarly from Eq(23),  $P_1(z)$  is a function of  $P_0(z)$ , A and F(z). Hence, if  $P_{0,0}$  and  $P_{j,1}(j=1,2,...c)$  are obtained,  $P_0(z)$  and  $P_1(z)$  can be determined completely.

### 2.2. Performance Measures

By using L'Hospital's rule in Eq(23), we get

$$P_1(1) = \frac{[\eta'P_0(1) - A] + \eta'P_0'(1)}{c\mu_w - \lambda_w(1 + \delta)} + \frac{\mu_w}{c\mu_w - \lambda_w(1 + \delta)}F(1), \tag{25}$$

where

$$F(1) = \sum_{n=1}^c (c - n)P_{n,1}. \tag{26}$$

Using Eq(22) and Eq(8) in Eq(25), we get,

$$P_1(1) = \frac{\eta'}{c\mu_w - \lambda_w(1 + \delta)}E(L_0) + \frac{\mu_w}{c\mu_w - \lambda_w(1 + \delta)}F(1). \tag{27}$$

Now, by applying L'hospital's rule to Eq(6), we have

$$E(L_0) = \lim_{z \rightarrow 1} P_0'(z) = \frac{-\lambda_w(1 + \delta)P_0(1) + \eta'P_0'(1)}{-(\mu_{wv} + \gamma_w)} = \frac{-\lambda_w(1 + \delta)P_0(1) - E(L_0)}{(\mu_{wv} + \gamma_w)} \text{ which implies} \tag{28}$$

$$P_0(1) = \frac{\eta' + \mu_{wv} + \gamma_w}{\lambda_w(1 + \delta)}E(L_0). \tag{29}$$

As  $P_0(0) + P_0(1) = 1$ , from Eq(27) and Eq(29), the expected number of customers during working vacation period is obtained as

$$E(L_0) = \frac{\lambda_w(1 + \delta)(1 - \rho)}{\eta' + \mu_{wv}(1 - \rho) + \gamma_w(1 - \rho)} - \frac{\frac{\lambda_w(1 + \delta)}{c}}{\eta' + \mu_{wv}(1 - \rho) + \gamma_w(1 - \rho)}F(1). \tag{30}$$

On substituting Eq(30) in Eq(29), the probability that the system in working vacation period is as

$$P(J = 0) = P_0(1) = \frac{(1 - \rho)(\eta' + \mu_{wv} + \gamma_w)}{\eta' + \mu_{wv}(1 - \rho) + \gamma_w(1 - \rho)} - \frac{\frac{\eta' + \mu_{wv} + \gamma_w}{c}}{\eta' + \mu_{wv}(1 - \rho) + \gamma_w(1 - \rho)}F(1) \tag{31}$$

and the probability that the system is in busy period is found as

$$P(J = 1) = P_1(1) = 1 - P_0(1) = \frac{(\eta'\rho)}{\eta' + \mu_{wv}(1 - \rho) + \gamma_w(1 - \rho)} + \frac{\frac{\eta' + \mu_{wv} + \gamma_w}{c}}{\eta' + \mu_{wv}(1 - \rho) + \gamma_w(1 - \rho)}F(1). \tag{32}$$

$E(L_1)$  can be obtained by differentiating Eq(23) and using L'Hospital's rule,

$$\begin{aligned} i.e., E(L_1) &= \lim_{z \rightarrow 1} P_1'(z) \\ &= \lim_{z \rightarrow 1} \left\{ \frac{-\lambda_w(1 + \delta)[z(-A + \eta'P_0(z))]}{(1 - z)(\lambda_w(1 + \delta)z - c\mu_w)^2} + \frac{-A + \eta'P_0(z) + z\eta'P_0'(z)}{(1 - z)(\lambda_w(1 + \delta)z - c\mu_w)} \right. \\ &\quad \left. + \frac{z(-A + \eta'P_0(z))}{(1 - z)^2(\lambda_w(1 + \delta)z - c\mu_w)} + \mu_w \frac{[(c\mu_w - \lambda_w(1 + \delta)z)F'(z) + \lambda_w(1 + \delta)F(z)]}{(c\mu_w - \lambda_w(1 + \delta)z)^2} \right\} \end{aligned} \tag{33}$$

$$= \frac{\eta'(c\mu_w - \lambda_w(1 + \delta))E(L_0(L_0 - 1)) + 2c\mu_w\eta'E(L_0)}{2(c\mu_w - \lambda_w(1 + \delta)z)^2} + \frac{F'(1)}{c(1 - \rho)} + \frac{\rho F(1)}{(c(1 - \rho))^2} \tag{34}$$

where

$$F'(1) = \frac{dF(z)}{dz} \text{ at } z=1$$

$$= \sum_{j=1}^c (c-j)P_{j,1} \tag{35}$$

Now, the value of  $P_0''(1)$  is obtained on differentiating Eq(6) twice on both sides as

$$(\eta' + \gamma_w)(1-z)P_0''(z) + 2\lambda_w(1+\delta)P_0'(z) = [\lambda_w(1+\delta)(1-z) + \eta' + 2(\mu_{wv} + \gamma_w)]P_0''(z) \tag{36}$$

where

$$P_0'''(z) = \frac{d^3 P_0(z)}{dz^3}$$

By letting  $z=1$  in Eq(36), we get  $P_0''(1) = \frac{2\lambda_w(1+\delta)}{\eta' + 2(\mu_{wv} + \gamma_w)} P_0'(1)$  (37)

or it can also be denoted as,  $E(L_0(L_0 - 1)) = \frac{2\lambda_w(1+\delta)EL_0}{\eta' + 2(\mu_{wv} + \gamma_w)}$ . (38)

Now, substituting, Eq(38) into Eq(34), the Mean number of customers, when the system in regular busy period is obtained as

$$E[L_1] = \frac{\rho\eta'}{(1-\rho)} \left[ \frac{1}{\eta' + 2(\mu_{wv} + \gamma_w)} + \frac{1}{\lambda_w(1+\delta)(1-\rho)} \right] E[L_0] + \frac{1}{c(1-\rho)} F'(1) + \frac{\rho}{c(1-\rho)^2} F(1) \tag{39}$$

Hence,  $E[L] = E[L_0] + E[L_1]$

$$= 1 + \frac{\rho\eta'}{(1-\rho)} \left[ \frac{1}{\eta' + 2(\mu_{wv} + \gamma_w)} + \frac{1}{\lambda_w(1+\delta)(1-\rho)} \right] \left[ \frac{\lambda_w(1+\delta)(1-\rho) - \frac{\lambda_w(1+\delta)}{c} F(1)}{\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho)} \right] + \frac{1}{c(1-\rho)} F'(1) + \frac{\rho}{c(1-\rho)^2} F(1) \tag{40}$$

Substituting Eq(31) in Eq(21) results in  $P_{(0,0)} = \frac{\eta'k}{(\mu_{wv} + \gamma_w)} P_0(1)$

$$= \frac{\eta'k}{(\mu_{wv} + \gamma_w)} \left[ \frac{(1-\rho)(\eta' + \mu_{wv} + \gamma_w)}{\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho)} - \frac{(\eta' + \mu_{wv} + \gamma_w)}{c} \right] F(1). \tag{41}$$

Suppose, the state of the system is  $(n,1)$ , then the service rates of the servers are  $n\mu_w$  for  $n \leq c$  and  $c\mu_w$  for  $n > c$  respectively.

In this manner, the expected number of customers served per unit of time is given by

$$N_s = \sum_{n=1}^c n\mu_w P_{n,1} + \sum_{n=c+1}^{\infty} c\mu_w P_{n,1} = \mu_w [cP_1(1) - F(1)] \tag{42}$$

and the proportion of customers served per unit of time is given by

$$P_s = \frac{N_s}{\lambda_w(1+\delta)} = \frac{1}{c\rho} [cP_1(1) - F(1)] \tag{43}$$

where  $P_1(1)$  is given by Eq(32).

If the state of the system is  $(n,1)$ ,  $n \geq 1$ , the rate of customer abandonment of a customer due to impatience is  $n\gamma_w$ . Thus the mean rate of the customer abandonment due to impatience is given by

$$R_a = \sum_{n=1}^{\infty} n\gamma_w P_{n,0} = \gamma_w E[L_0]. \tag{44}$$

Thus an M/M/c Queuing model with Multiple Working Vacations under encouraged arrival with impatient behaviour is evaluated.



### 2.3. Single Working Vacation Model:

A Single working Vacation policy define that the server(s) in the queuing system takes vacation immediately , when he find no customers waiting in the queue. At the end of the working vacation, if the server find the system non empty, then he starts his regular busy period by shifting his service rate from  $\mu_{wv}$  to  $\mu_w$ . If not,the server will remain idle in the system itself than going for vacation and waits until the customer arrives for the new busy period. To construct this system, we defin a Markov chain as  $\{(M(t), N(t)), t \geq 0\}$  with state space as in Multiple Working Vacations for Single Working Vacation also.  $s = \{(n, j)\}, n \geq 0, j = 0, 1\}$

Where  $M(t)$  denotes the total number of customers in the system at time  $t$  and  $N(t)$  denotes the state of the system at time  $t$  with

$$N(t) = \{1 \text{ when the servers are in non-vacation period at time } t\} \text{ and}$$

$$N(t) = \{0 \text{ when the servers are in working vacation period at time } t\}.$$

### 2.4. Steady State Equations and its Solutions for Single Working Vacation Model:

Now, the set of balance equations as

$$(\lambda_w(1 + \delta) + \eta')P_{00} = (\mu_{wv} + \gamma_w)P_{1,0} + \mu_w P_{1,1}, \tag{45}$$

$$[\lambda_w(1 + \delta) + \eta' + n(\mu_{wv} + \gamma_w)]P_{n,0} = \lambda_w(1 + \delta)P_{n-1,0} + (n + 1)((\mu_{wv} + \gamma_w)P_{n+1,0}, \quad \text{if } n \geq 1, \tag{46}$$

$$(\lambda_w(1 + \delta))P_{0,1} = \eta'P_{0,0}, \tag{47}$$

$$(\lambda_w(1 + \delta) + n\mu_w)P_{n,1} = \lambda_w(1 + \delta)P_{n-1,1} + (n + 1)\mu_w P_{n+1,1} + \eta'P_{n,0}, \quad \text{if } 1 \leq n \leq c - 1, \tag{48}$$

$$(\lambda_w(1 + \delta) + c\mu_w)P_{n,1} = \lambda_w(1 + \delta)P_{n-1,1} + c\mu_w P_{n+1,1} + \eta'P_{n,0} \quad \text{if } n \geq c. \tag{49}$$

By letting the probability generating functions as

$$R_0(z) = \sum_{n=0}^{\infty} z^n P_{n,0},$$

$$R_1(z) = \sum_{n=1}^{\infty} z^n P_{n,1}.$$

with  $R_0(1) + R_1(1) = 1$  and  $R'_0(z) = \sum_{n=1}^{\infty} n z^{n-1} P_{n,0}$ .

Now, By Multiplying Eq(46) with  $z^n$  and adding over 'n' and rearranging the terms, the differential equation is attained as :

$$(\mu_{wv} + \gamma_w)(1 - z)R'_0(z) = [\lambda_w(1 + \delta)(1 - z) + \eta']R_0(z) - (\mu_w P_{1,1}). \tag{50}$$

Likely multiplying Eq(48) and Eq(49) by  $z^n$  and adding over 'n', the following equation is obtained,

$$(1 - z)(\lambda_w(1 + \delta)z - c\mu_w)R_1(z) = \eta'zR_0(z) - (\eta'P_{0,0} + \mu_w P_{1,1})z + z^2\eta'P_{0,0} + \mu_w(1 - z) \sum_{n=1}^c (n - c)z^n P_{n,1}. \tag{51}$$

Then, for  $z \neq 1$ ,

$$R'_0(z) - \left[ \frac{\lambda_w(1 + \delta)}{(\mu_{wv} + \gamma_w)} + \frac{\eta'}{(\mu_{wv} + \gamma_w)(1 - z)} \right]R_0(z) = - \frac{\mu_w P_{1,1}}{(\mu_{wv} + \gamma_w)(1 - z)}. \tag{52}$$

Solving the differential equation, as in Multiple Working Vacations Model we get,

$$R_0(z) = \frac{e^{\frac{\lambda_w(1+\delta)z}{(\mu_{wv}+\gamma_w)}}}{(1-z)^{\frac{\eta'}{(\mu_{wv}+\gamma_w)}}} \left[ 1 - \frac{1}{L} \int_0^z e^{-\frac{\lambda_w(1+\delta)z}{(\mu_{wv}+\gamma_w)}} (1-x)^{\frac{\eta'}{(\mu_{wv}+\gamma_w)}-1} dx \right] P_{0,0}. \tag{53}$$

Thus a similar expression for  $R_0(z)$  as in Multiple Working Vacations Model and here we arrive at,

$$R_0(0) = P_{0,0} = \frac{(\mu_w P_{1,1})}{(\mu_{wv} + \gamma_w)} L \tag{54}$$

$$R_0(1) = \frac{(\mu_{wv} + \gamma_w)}{\eta' L} P_{0,0} \tag{55}$$

and from Eq(54) and Eq(55), the following relation is obtained

$$\eta' R_0(1) = \mu_w P_{1,1}. \tag{56}$$

From Eq(51),  $R_1(z)$  is attained as,

$$R_1(z) = \frac{[\eta' R_0(z) - A]z + z^2 \eta' P_{0,0}}{(\lambda_w(1+\delta)z - c\mu_w)(1-z)} - \frac{\mu_w}{(\lambda_w(1+\delta)z - c\mu_w)} F(z), \tag{57}$$

where,

$$F(z) = \sum_{n=1}^c (n-c) z^n P_{n,1}. \tag{58}$$

It is clear from Eq(53) that  $R_0(z)$  is a function of  $P_{0,0}$  and the ratio between the time of the servers on working vacation and the system is empty. Similarly from Eq(57),  $R_1(z)$  is a function of  $R_0(z)$ , A and F(z). Hence, if  $P_{0,0}$  and  $P_{j,1}(j=1,2,\dots,c)$  are obtained,  $P_0(z)$  and  $P_1(z)$  can be determined completely.

### 2.5. Performance Measures

By using L'Hospital's rule in Eq(57), we get

$$R_1(1) = \frac{[\eta' E(L)_0] + B}{c\mu_w - \lambda_w(1+\delta)} + \frac{\mu_w}{c\mu_w - \lambda_w(1+\delta)} F(1) \tag{59}$$

where

$$B = \eta'(2-c)P_{0,0} \quad \text{and} \quad F(1) = \sum_{n=1}^c (n-c)P_{n,1}. \tag{60}$$

Using Eq(22) and Eq(8) in Eq(25), the following equation is obtained,

$$P_1(1) = \frac{\eta'}{c\mu_w - \lambda_w(1+\delta)} E(L_0) + \frac{\mu_w}{c\mu_w - \lambda_w(1+\delta)} F(1). \tag{61}$$

Now, applying L'hospital's rule to Eq(6), we have

$$E(L_0) = \lim_{z \rightarrow 1} P'_0(z) = \frac{-\lambda_w(1+\delta)P_0(1) + \eta' P'_0(1)}{-(\mu_{wv} + \gamma_w)} = \frac{-\lambda_w(1+\delta)P_0(1) - E(L_0)}{\mu_{wv} + \gamma_w} \text{ which implies} \tag{62}$$

$$P_0(1) = \frac{\eta' + \mu_{wv} + \gamma_w}{\lambda_w(1+\delta)} E(L_0). \tag{63}$$

As  $P_0(0) + P_0(1) = 1$ , from Eq(27) and Eq(29), is the expected number of customers during working vacation period is obtained as

$$E(L_0) = \frac{\lambda_w(1+\delta)(1-\rho)}{\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho)} - \frac{-\rho\eta'(2-c)P_{0,0}}{\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho)} - \frac{\lambda_w(1+\delta)}{c} F(1). \quad (64)$$

On substituting Eq(30) in Eq(29), the probability that the system in working vacation period is as

$$P(J=0) = R_0(1) = \frac{(1-\rho)(\eta' + \mu_{wv} + \gamma_w)}{\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho)} - \frac{\rho\eta'(\eta' + \mu_{wv} + \gamma_w)(2-c)P_{0,0}}{\lambda_w(1+\delta)[\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho)]} - \frac{\eta' + \mu_{wv} + \gamma_w}{c} F(1)$$

where

$$X = \frac{\rho\eta'(\eta' + \mu_{wv} + \gamma_w)(2-c)P_{0,0}}{\lambda_w(1+\delta)[\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho)]}, \quad \text{then}$$

$$P(J=0) = R_0(1) = \frac{(1-\rho)(\eta' + \mu_{wv} + \gamma_w)}{\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho)} - X - \frac{\eta' + \mu_{wv} + \gamma_w}{c} F(1) \quad (65)$$

and the probability that the system is in busy period is as follows

$$P(J=1) = R_1(1) = 1 - R_0(1) = \frac{(\eta'\rho)}{\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho)} + \frac{\rho\eta'(\eta' + \mu_{wv} + \gamma_w)(2-c)P_{0,0}}{\lambda_w(1+\delta)[\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho)]} + \frac{\eta' + \mu_{wv} + \gamma_w}{c} F(1).$$

since we know that,

$$X = \frac{\rho\eta'(\eta' + \mu_{wv} + \gamma_w)(2-c)P_{0,0}}{\lambda_w(1+\delta)[\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho)]},$$

we get

$$P(J=1) = R_1(1) = 1 - R_0(1) = \frac{(\eta'\rho)}{\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho)} + X + \frac{\eta' + \mu_{wv} + \gamma_w}{c} F(1). \quad (66)$$

Now,  $E(L_1)$  can be obtained by differentiating Eq(58) and using L'Hospital's rule,

$$E(L_1) = \lim_{z \rightarrow 1} R_1'(z) = \lim_{z \rightarrow 1} \left\{ \frac{-\lambda_w(1+\delta)[z(-A + \eta'R_0(z)) + z^2\eta'P_{0,0}]}{(1-z)(\lambda_w(1+\delta)z - c\mu_w)^2} + \frac{-A + \eta'R_0(z) + 2z\eta'P_0'(z) + z\eta'R_0'(z)}{(1-z)(\lambda_w(1+\delta)z - c\mu_w)} + \frac{z(-A + \eta'R_0(z) + z^2\eta'P_{0,0})}{(1-z)^2(\lambda_w(1+\delta)z - c\mu_w)} + \mu_w \frac{[(c\mu_w - \lambda_w(1+\delta)z)F'(z) + \lambda_w(1+\delta)F(z)]}{(c\mu_w - \lambda_w(1+\delta)z)^2} \right\} \quad (67)$$

$$= \frac{\eta'(c\mu_w - \lambda_w(1+\delta))E(L_0(L_0 - 1)) + 2c\mu_w\eta'E(L_0) + 2\eta'[(2(c\mu_w - \lambda_w(1+\delta)) - c\lambda_w(1+\delta))P_{0,0}]}{2(c\mu_w - \lambda_w(1+\delta))^2} + \frac{F'(1)}{c(1-\rho)} + \frac{\rho F(1)}{(c(1-\rho))^2} \quad (68)$$

where

$$F'(1) = \frac{dF(z)}{dz} \text{ at } z = 1$$

$$= \sum_{j=1}^c (c-j)P_{j,1} \tag{69}$$

Now, the value of  $R_0''(1)$  is obtained on differentiating Eq(50) twice on both sides and proceeding similarly as in Multiple Working Vacations, We get

$$E(L_0(L_0 - 1)) = \frac{2\lambda_w(1 + \delta)}{\eta' + 2(\mu_{wv} + \gamma_w)} E(L_0). \tag{70}$$

Now, substituting, Eq(70) into Eq(69),the Mean number of customers, when the system in regular busy period is obtained as

$$E[L_1] = \frac{\rho\eta'}{(1-\rho)} \left\{ \left[ \frac{1}{\eta' + 2(\mu_{wv} + \gamma_w)} + \frac{1}{\lambda_w(1 + \delta)(1-\rho)} \right] E[L_0] + \left[ \frac{1}{\lambda_w(1 + \delta)} - \frac{1}{\mu_w(1-\rho)} \right] P_{0,0} \right\} + \frac{1}{c(1-\rho)} F^1 + \frac{\rho}{c(1-\rho)^2} F(1). \tag{71}$$

$$E[L] = E[L_0] + E[L_1] = \left\{ 1 + \frac{\rho\eta'}{(1-\rho)} \left[ \frac{1}{\eta' + 2(\mu_{wv} + \gamma_w)} + \frac{1}{\lambda_w(1 + \delta)(1-\rho)} \right] \left[ \frac{\lambda_w(1 + \delta)(1-\rho) - \rho B - \frac{\lambda_w(1+\delta)}{c} F(1)}{\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho)} \right] \right\} + Y + \frac{1}{c(1-\rho)} F'(1) + \frac{\rho}{c(1-\rho)^2} F(1) \tag{72}$$

$$\text{wher } eY = \frac{\rho\eta'}{(1-\rho)} \left[ \frac{1}{\lambda_w(1 + \delta)} - \frac{1}{\mu_w(1-\rho)} \right] P_{0,0}$$

Substituting Eq(65) in Eq(55) results in  $P_{(0,0)} = \frac{\eta'^k}{(\mu_{wv} + \gamma_w)} R_0 1$

$$= \frac{\eta'^k}{(\mu_{wv} + \gamma_w)} \left[ \frac{\lambda_w(1 + \delta)(1-\rho)((\eta' + \mu_{wv} + \gamma_w) - \frac{(\lambda_w(1+\delta)\eta' + \mu_{wv} + \gamma_w)}{c})}{\eta' + \mu_{wv}(1-\rho) + \gamma_w(1-\rho) + \frac{k\eta'^2\rho(2-c)(\eta' + \mu_{wv} + \gamma_w)}{(\mu_{wv} + \gamma_w)}} \right]. \tag{73}$$

Suppose, the state of the system is (n,1),then the service rates of the servers are  $n\mu_w$  for  $n \leq c$  and  $c\mu_w$  for  $n > c$  respectively.

Thus, the expected number of customers served per unit of time is given by

$$N_s = \sum_{n=1}^c n\mu_w P_{n,1} + \sum_{n=c+1}^{\infty} c\mu_w P_{n,1} = \mu_w [cP_1(1) - F(1)] \tag{74}$$

and the proportion of customers served per unit of time is given by

$$P_s = \frac{N_s}{\lambda_w(1 + \delta)} = \frac{1}{c\rho} [cP_1(1) - F(1)] \tag{75}$$

wher  $eP_1(1)$  is given by Eq(66).

If the state of the system is (n,1),  $n \geq 1$ , the rate of customer abandonment of a customer due to impatience is  $n\gamma_w$ . Thus the mean rate of the customer abandonment due to impatience is given by

$$R_a = \sum_{n=1}^{\infty} n\gamma_w P_{n,0} = \gamma_w E[L_0]. \tag{76}$$

Hence, an M/M/c Queuing model with Multiple Working Vacations under encouraged arrival with impatient behaviour is evaluated.

### 3. RESULTS

In this paper, an M/M/C Queuing model under Multiple working vacations and single working vacation with impatient behaviour of the customer during encouraged arrival are analysed. It is observed that for the system of steady state equations, performance measures like Mean Queue length ( $E[L]$ ), Probability that the system is in working vacation period ( $P[J=0]$ ), Probability that the system is in regular busy period ( $P[J=1]$ ) are evaluated for the two different vacation policies separately.

### 4. DISCUSSION

On comparing the performance measures between the two vacation policies, from Eq (65) and Eq(31), it is observed that the difference between the probability of the system ( $P[J=0]$ ) in single working vacation and that during multiple working vacations, we notice that by reducing the term "X" from the probability of the system in multiple working vacations, we attain the probability of the system in single working vacation. Likely, from Eq (66) and Eq(32), it is clear that the probability of the system in regular busy period during single working vacation is obtained by adding the term "X" to the probability of the system in regular busy period during multiple working vacation. Moreover, while comparing the mean queue length during the two different vacation policies, we observe that from Eq (72) and Eq(40),  $E(L)$  in single working vacation is the addition of the term "Y" and the term  $\rho B$  to the existing mean queue length of multiple working vacations.

### 5. CONCLUSION

As the M/M/c Queuing model with Multiple and single working vacation with impatient behaviour of the customers during encouraged arrival is analysed, apart from deriving the explicit formulations, some of the characteristic measures are also discussed. It can be concluded that, with the impact of the terms "X", "Y" and "B" in multiple working vacations an M/M/C Queuing model with impatient behaviour of the customer during encouraged arrival can be shifted to Single working vacation. However, for an efficient functioning of the queue a single working vacation can be suggested. In future work, numerical examples may be evaluated to evident the obtained result.

### REFERENCES

- [1] Adan, I. Boxma, O. Claeys, D. and Kella, O. (2018). A Queuing System with Vacations after a Random Amount of Work. *SIAM Journal on Applied Mathematics*, 78:1697–1711.
- [2] Bhat, U. N. and Rao, S. S. (1987). Statistical analysis of queuing systems. *Queuing Syst*, 1:217–247.
- [3] Doshi, B. T. (1986). Queuing systems with vacations A survey. *Queuing Syst*, 1:29–66
- [4] Hanumantha Rao, S. Vasanta Kumar, V. Sathish Kumar, K. (2020). Encouraged Or Discouraged Arrivals Of An M/M/1/N Queuing System With Modified Reneging. *Advances in Mathematics:Scientific Journal*, 9:6641–6647.
- [5] Jain, M. and Upadhyaya, S. (2011). Synchronous working vacation policy for finite-buffer multiserver queuing system. *Appl. Math. Comput*, 217:9912–9916.
- [6] Kadi, M. Bouchentouf, A. A. and Yahiaoui, L. (2020). On a multi-server queuing system with customers' impatience until the end of service under single and multiple vacation policies. *Applications and Applied Mathematics*, 15:740–763.
- [7] Levy, Y. and Yechiali, U. (1976). M/M/S queues with server vacations. *INFO*, 14:153–163.
- [8] Majid, S. (2023). Performance Analysis of a Markovian Queue with Impatient Customers and Working Vacation. *J. Oper. Res. Soc. China*, 11:133–156.

- [9] Majid, S. and Manoharan, P. (2018). Impatient customers in an M/M/c queue with single and multiple synchronous working vacations. *Pak. J. Stat. Oper. Res.*, 14:571–594.
- [10] Maragathasundari, S. Eswar, S. and Somasundaram, R. S. (2022). A study on phases of service and multi-vacation policy in a non-Markovian queuing system. *International Journal of Mathematics in Operational Research*, 21:444–465.
- [11] Nawel, G. and Malika, I. (2010). Numerical investigation of finite-source multi-server systems with different vacation policies. *Journal of Computational and Applied Mathematics*, 234:625–635.
- [12] Perel, N. and Yechiali, U. (2010). Queues with slow servers and impatient customers. *Eur. J. Oper. Res.*, 201:247–258.
- [13] Selvaraju, N. and Cosmika, G. (2013). Impatient customers in an M/M/1 queue with single and multiple working vacations. *Computers and Industrial Engineering*, 65:207–215.
- [14] Silva, F. A. Nguyen, T. A. Brito, C. Min, D. and Lee, J. W. (2021). Performance Evaluation of an Internet of Healthcare Things for Medical Monitoring Using M/M/c/K Queuing Models. *IEEE Access*, 9:55271–55283.
- [15] Som, B. K. and Sunny Seth. (2018). M/M/c/N queuing systems with encouraged arrivals, reneging, retention and Feedback customers. *Yugoslav Journal of Operations Research*, 28:6–6.
- [16] Sudhesh, R. and Azhagappan, A. (2019). Analysis of an M/M/c queue with heterogeneous servers, balking and reneging. *International Journal of Operational Research*, 36:293–309

# BEHAVIORAL ANALYSIS AND MAINTENANCE DECISIONS OF WOOD INDUSTRIAL SUBSYSTEM USING STOCHASTIC PETRI NETS SIMULATION MODELING

URVASHI<sup>1</sup>, SHIKHA BANSAL<sup>2\*</sup>

•

1. Research Scholar, Department of Mathematics, SRM Institute of Science  
and Technology, Delhi-NCR Campus, Ghaziabad, 201204, India  
urvashigodara8@gmail.com

2\*. Assistant Professor, Department of Mathematics, SRM Institute of Science  
and Technology, Delhi-NCR Campus, Ghaziabad, 201204, India  
srbansal2008@gmail.com

\* Corresponding author

## Abstract

*This study aims to optimize the productivity of the plywood manufacturing system within the wood industry. A Petri nets simulation-based technique has been used to evaluate the availability analysis of the plywood manufacturing system. A Petri nets model is created to represent the modeling of the plywood system. The model is subsequently simulated using the licensed program Petri Nets (PN) GRIF 2023.7. This simulation is used to evaluate the performance of the system. In the PN simulation model, timed transitions are fired based on the failure and repair rate of the system. Immediate transitions, on the other hand, have their own guard function for firing which is coded using a logical AND-OR gate. This study also assesses the impact of the repairman on the system's availability. The system's availability is optimized by increasing the number of repairmen. However, once a specific number of repairmen is reached, the system's availability remains constant. This research is highly valuable for determining the optimal number of maintenance staff needed for the wood industrial system.*

**Keywords:** Availability, Maintenance, Performance, Petri-Nets, Repairman, Simulation.

## 1. INTRODUCTION

In the context of engineering, reliability is the average time between failures (MTBF), which indicates how consistently a system performs without malfunctioning. To attain a high level of reliability and availability, it is necessary to implement a strong design, utilize high-quality components, and employ effective fault detection methods to reduce the amount of time that the system is not operational. These indicators are essential in sectors where system failures can have major effects on the economy and public safety, such as manufacturing and energy. Prior studies have extensively investigated diverse facets of designing industrial systems in the realm of reliability, applying varied methodologies to enhance system performance. The subsequent part presents a concise overview of the literature that encompasses the study undertaken by several scholars in the domain of reliability.

Tan and Kramer [1] provided an approximation for the financial impact of an unexpected plant shutdown, stating that it results in a revenue loss ranging from 500 to 100,000 USD per hour. Angela and David [2] have presented a method utilizing the Petri Net approach to analyze the dependability and safety of an industrial-scale production system. Regattieri and Bellom [3] implemented the Innovative lay-up system in the plywood production process leading to a substantial boost in productivity (about 19%) and a significant reduction in the number of personnel (-54%). Bansal and Tyagi [4] assessed the reliability of the screws mill production system by employing a combination of standby and parallel arrangement, and using the orthogonal matrix strategy. Kumar Amit et al. [5] maximized the efficiency of the ethanol manufacturing system and conducted a comparative analysis between the genetic algorithm (GA) and particle swarm optimization algorithm (PSO). The findings indicate that the PSO method outperforms the genetic algorithm in optimizing the system. Kumar Narendra et al. [6] enhanced the efficiency and assessed the reliability of milk pasteurization by the utilization of a probabilistic Petri net methodology. Narendra et al. [7] assessed the efficiency of the veneer gluing system using a stochastic Petri net methodology and determined that the thermal press is the most important component in this system. Malik and Tewari [8] utilized the particle swarm optimization technique to enhance the efficiency of the coal handling system. They achieved a performance improvement of 99.33% with an average population of 40 and 93.31% with an average generation size of 70. Tyagi et al. [9, 10] assessed the availability of every component in the leaf spring production facility by utilizing the matrix method and Markov birth-death methodology. They employ the C programming language to solve mathematical problems. Kalaivani and Kannan [11] evaluating reliability properties of a linear consecutive k-out-of-n: The F system in this uses asymptotic confidence intervals, mean time to failure, and reliability function for different sample sizes and parameter combinations using likelihood estimation, Monte Carlo (MC) training, and real data visualization. Chaudhary and Bansal [12] evaluated the reliability of the hydroelectric power station using the Laplace transformation method. Kumar Sudhir and Tewari [13] utilized the Petri module of GRIF to assess the performance of the coal handling system by manipulating the failure and repair rates. Subsequently, the performance was optimized using the particle swarm optimization strategy. Godara and Bansal [14] assessed the availability of a multi-state machine using an artificial neural network methodology, where neural weights are determined based on the system's failure and repair rate. Tyagi and Bansal [15] enhanced the efficiency of the wastewater treatment process by employing the Runge-Kutta numerical technique. They develop a mathematical model utilizing a probabilistic strategy and a Markovian technique. Rathi et al. [16] assessed the dependability of the parallel and cold standby unit in the system. Godara and Bansal [17] assessed the dependability as well as the availability of the steam turbine generating facility using a neural network methodology and boolean function technology. Urvashi and Shikha Bansal [18] assessed the dependability factor and availability of the threshing machine plant system using both a general and copula distribution. They found that the copula distribution yielded superior results compared to the general distribution.

The research provides a comprehensive examination of the behavior and performance of the standby plywood manufacturing system. In this work, the performance of the system is enhanced through the utilization of the stochastic Petri nets simulation method. After enhancing the system's performance, an analysis is conducted to determine the impact of the number of repairmen on the system's availability. The goal is to identify the optimal number of repairmen required to efficiently repair the system and maximize its availability.

The subsequent sections of this paper are structured in the following way: Section 2 provides a comprehensive overview of the plywood manufacturing system, including a detailed explanation of each subsystem within the system. Section 3 explains how the system is modeled using Petri nets. This section specifically details the process of formulating the Petri nets model for the plywood system. Section 4 focuses on optimizing the performance of the system. This is achieved by varying the failure repair rate and increasing the number of repairmen to enhance system performance. The conclusion of this research is provided in Section 5.



## 2. SYSTEM DESCRIPTION

The plywood system plays a crucial role in the wood industry. Plywood is a flexible type of manufactured wood that is created by bonding together small pieces or pieces of wood, referred to as layers or sheets. The layers are often arranged with their grain horizontal to neighboring layers, so augmenting the reliability and stability of the eventual product. The manufacture of the plywood system involves multiple processes and steps.

Figure 1 displays the schematic representation of this. The subsystems of this system are organized in a hybrid structure, with each subsystem described as follows.

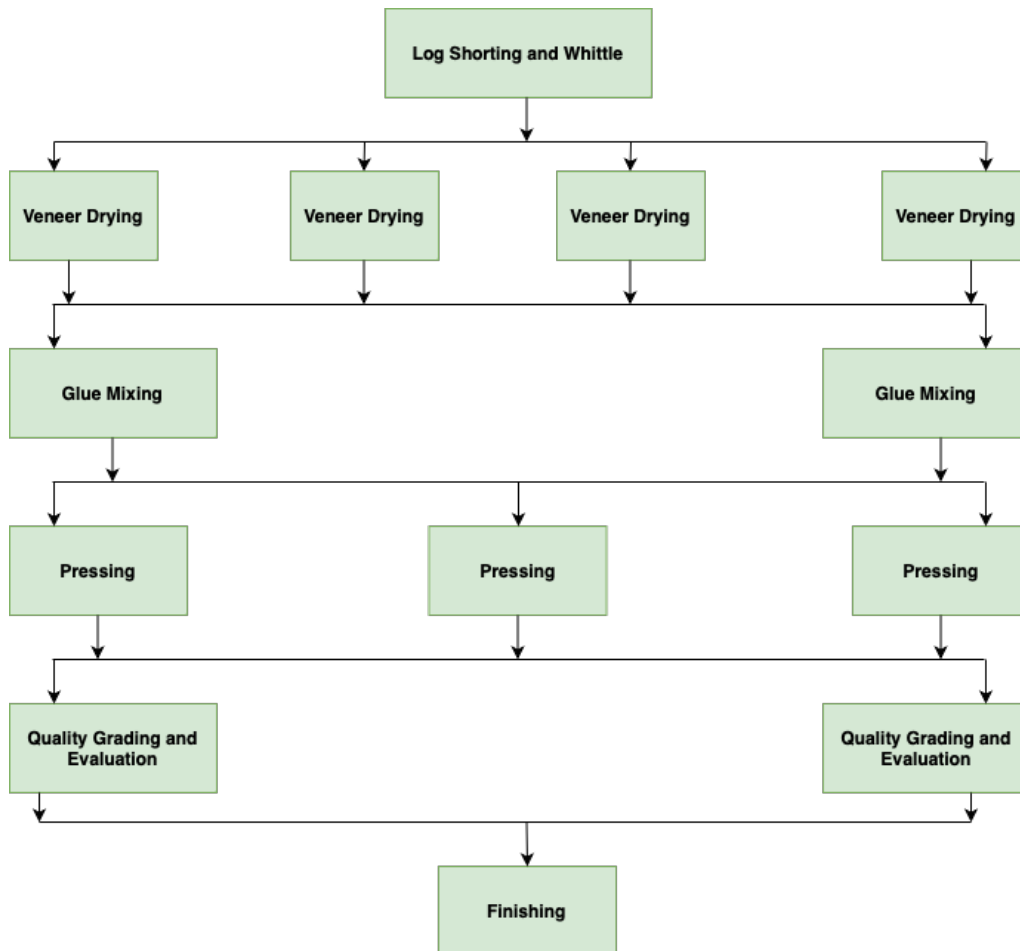


Figure 1: Block Diagram of Plywood Manufacturing System

- **Log Shorting and Whittle (LSW) :-** This is the first phase of the plywood production process, wherein a carefully chosen log of high-quality wood is picked based on the desired type of plywood to be produced. This piece of wood is commonly referred to as whittle. The whittle is straight and has a substantial size, making it ideal for creating an enormous amount of layers.
- **Veneer Drying (VD) :-** This subsystem is essential in the plywood fabrication process as its main objective is to decrease the wetness of the wood layers. Effectively dry veneers are crucial for guaranteeing the excellence, durability, and constancy of the ultimate hardwood goods. This subsystem consists of two types of drier equipment: a spinning dryer and a continuous veneer dryer. These dryers are designed to achieve consistent drying by controlling the heat and airflow. This subsystem consists of four units that are arranged in a

standby configuration.

- **Glue Mixing (GM) :-** The veneer layers in this subsystem are bonded together through the use of adhesive or glue. The subsystem consists of two machines, namely the glue distributors and the blending machine, which is arranged in standby mode.
- **Pressing :-** Once the veneer layer is bonded using glue, it is subjected to pressure and heat in a machine. Specifically, three machines are arranged in a standby configuration for this purpose.
- **Quality Assurance and Evaluation (QSE) :-** This subsystem consists of two units that are arranged in standby mode, wherein the plywood's quality and grade are checked and evaluated. This part of the system contains the methodical evaluation of the plywood products according to specified requirements and standards. By conducting a thorough examination, flaws such as empty spaces, separation, blemishes, and other irregularities are detected and categorized.
- **Finishing :-** This subsystem entails creating art, coloring, varnishing, or laminating the plywood sheets with a variety of coatings and finishes. By enhancing plywood's aesthetic appeal and functional qualities, the finishing subsystem plays a crucial role in raising the value of the finished product and enhancing its overall quality. There is only one unit in this subsystem, and it is arranged in series with other subsystems.

### 3. PETRI NETS MODELING

The Petri nets has developed as an effective graphical modeling tool that encompasses both the static and dynamic behavior of systems. Petri nets are essentially directed graphs that are bipartite and have powerful mathematical representations that allow allocation, timing, and concurrency. They are made up of places, transitions, arcs, and tokens, which are denoted by circles, rectangular bars, arrows, and large circular dots that are centered in the places, respectively.

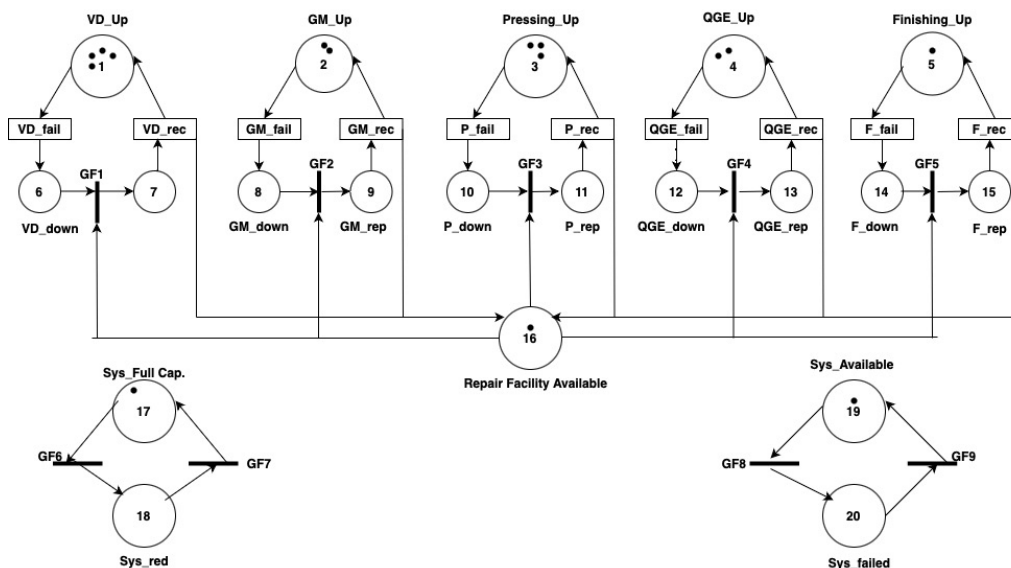


Figure 2: Petri Nets Model of Plywood Manufacturing System

Figure 2 displays the Petri net structure model of the plywood manufacturing system. The black dots, referred to as Tokens in this model, serve to represent the status and availability of the system/sub-systems and maintenance facility, correspondingly.

**Places:** In the Petri net are denoted by the circle at  $P=(P_1 to P_{12})$ ; these are the only tokens that are

provided based on the subsystem’s unit number.

**Timed Transition:** This model represents VD-fail VD-rec, GM-fail GM-rec, P-fail P-rec, QGE-fail QGE-rec, and F-fail F-rec. The timed transitions are connected to the variable failure  $\Omega_i$ , which follows an exponential distribution.

**Immediate Transition:** Immediate transitions possess a personal guard function. In this model represents GF1, GF2, GF3, GF4, GF5, GF6, GF7, GF8 and GF9. The immediate transitions are connected to the variable repair  $\Psi_i$ , which follows an exponential distribution.

The following are special notations used in the plywood Petri net model:

- **Sys\_Full Cap.** indicate that the plywood system is in a fully functional state.
- **Sys\_red** indicate the lower capacity state of the system.
- **Sys\_Available** represent the system available for working.
- **Sys\_failed** indicate that the plywood system is in a failed state.

#### 4. PERFORMANCE OPTIMIZATION

Appropriate performance optimization is essential in the plywood industry since it directly affects manufacturing expenses and the utilization of resources. An efficiently optimized system facilitates improved coordination across different phases of production, ranging from the acquisition of raw materials to the distribution of the final product. This leads to enhanced overall efficiency and strength in the market.

This section evaluates the system’s availability and provides information on the system’s behavior or performance. To assess the system’s availability or performance, a mathematical model of the system has been developed using stochastic Petri nets simulation. This modeling was conducted using the licensed GRIF2023.7 the program’s software. Failure and repair rate parameters of each subsystem affect the availability of the system.

The parameters were obtained from the maintenance history sheet, and their range is defined by  $\Omega_1 \in (0.0035, 0.0173)$   $\Psi_1 \in (0.51, 1.1)$ ,  $\Omega_2 \in (0.005, 0.014)$   $\Psi_2 \in (0.7, 2)$ ,  $\Omega_3 \in (0.0028, 0.0058)$   $\Psi_3 \in (0.21, 0.82)$ ,  $\Omega_4 \in (0.0059, 0.0128)$   $\Psi_4 \in (0.57, 1.19)$ ,  $\Omega_5 \in (0.007, 0.01)$   $\Psi_5 \in (0.61, 1.2)$ . The influence of the failure rate ( $\Omega_i$ ) and repair rate ( $\Psi_i$ ) of each subsystem on the availability of the system is illustrated in Table 1 to Table 5.

The effect of the system performance due to variation in failure and repair rate parameters of the veneer drying subsystem is shown in Table 1.

**Table 1:** The Effects of Veneer Drying Subsystem Repair and Failure Rates on System Availability

Failure rate $\Omega_1$	Repair rate $\Psi_1$				Constant Parameters
	0.51	0.71	0.91	1.1	
0.0035	0.9615	0.9635	0.9646	0.9652	$\Omega_2 = 0.005, \Omega_3 = 0.0028$
0.0081	0.9604	0.9623	0.9636	0.9644	$\Omega_4 = 0.0059, \Omega_5 = 0.0071$
0.0127	0.9514	0.9559	0.9585	0.9600	$\Psi_2 = 0.7, \Psi_3 = 0.21$
0.0173	0.9366	0.9457	0.9507	0.9539	$\Psi_4 = 0.57, \Psi_5 = 0.61$

Upon analyzing the data in Table 1, it is evident that when the failure rate of the veneer drying subsystem increases from 0.0035 to 0.0173, the availability of the plywood system reduces from 0.9615 to 0.9366, in a comparable way the repair rate of this subsystem increases from 0.51 to 1.1, and the availability of the system increases from 0.9615 to 0.9652. The system’s availability is reduced by 2.49% due to fluctuations in the failure rate of the veneer drying subsystem. Conversely, the system’s availability is increased by 0.37% due to variations in the repair rate of this subsystem. The graphical representation of the effect of variations in the parameter of the

veneer drying subsystem on system availability is depicted in Figure 3.

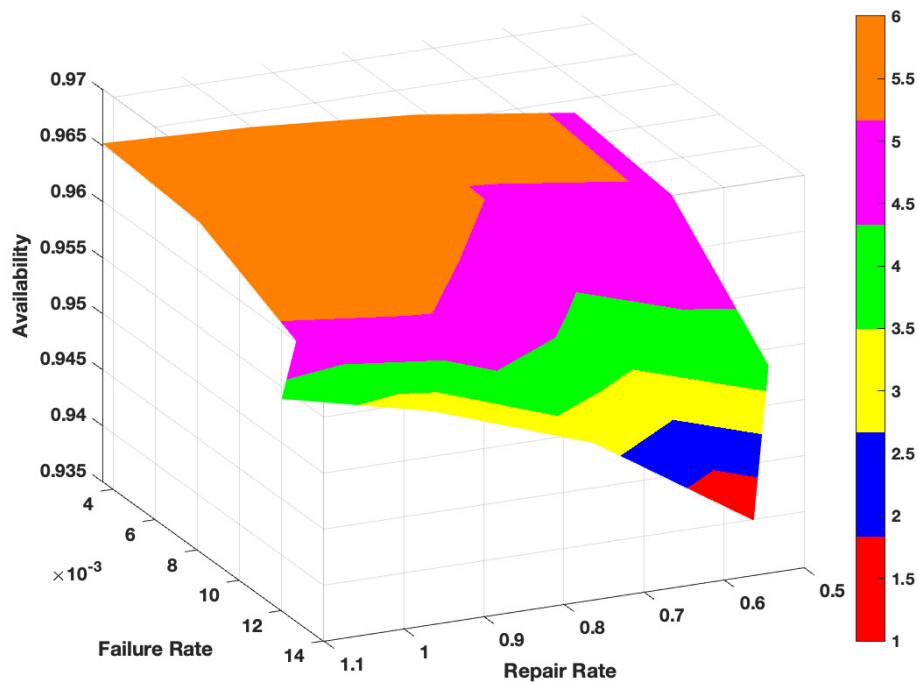


Figure 3: Effect of Veneer Drying Variation in FRR on Plywood System Performance

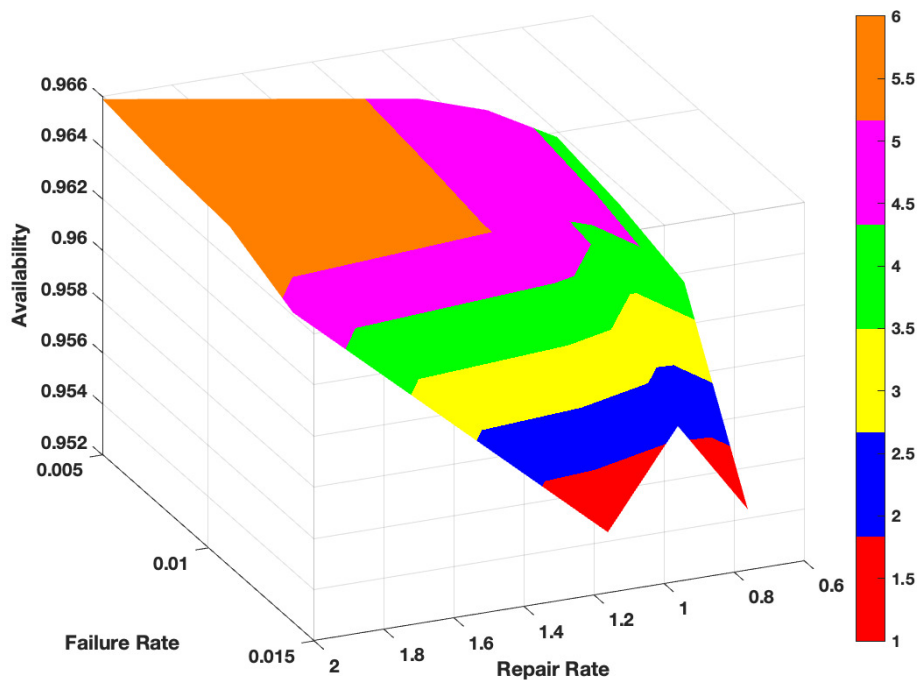
The impact of changes in the failure and repair rate parameter of the glue mixing subsystem on the system availability is demonstrated in Table 2.

Table 2: The Effects of Glue Mixing Subsystem Repair and Failure Rates on System Availability

Failure rate $\Omega_2$	Repair rate $\Psi_2$				Constant Parameters
	0.7	0.9	1.1	2	
0.005	0.9615	0.9630	0.9639	0.9659	$\Omega_1 = 0.0035, \Omega_3 = 0.0028$
0.008	0.9610	0.9626	0.9635	0.9655	$\Omega_4 = 0.0059, \Omega_5 = 0.0071$
0.011	0.9602	0.9619	0.9630	0.9653	$\Psi_1 = 0.51, \Psi_3 = 0.21$
0.0173	0.9535	0.9572	0.9595	0.9641	$\Psi_4 = 0.57, \Psi_5 = 0.61$

Upon observation, it has been determined that when the failure rate of this subsystem increases from 0.005 to 0.0173, the system’s availability reduces from 0.9615 to 0.9535. This corresponds to a decrease in system availability of 0.8% as a result of the fluctuation in the failure rate of this subsystem. The system’s availability is enhanced from 0.9615 to 0.9659 as a result of the repair rate of this subsystem increasing from 0.7 to 2. This corresponds to a 0.44% gain in availability due to the higher repair rate of this subsystem. Figure 4 displays a graphical depiction of the relationship between the system’s availability and changes in the parameter of the glue mixing subsystem.

The influence of variations in the Pressing subsystem on the failure and repair rate parameters has been demonstrated in Table 3, illustrating its effect on the system availability.



**Figure 4:** Effect of Glue Mixing Variation in FRR on Plywood System Performance

**Table 3:** The Effects of Pressing Subsystem Repair and Failure Rates on System Availability

Failure rate $\Omega_3$	Repair rate $\Psi_3$				Constant Parameters
	0.21	0.42	0.61	0.82	
0.0028	0.9615	0.9696	0.9722	0.9736	$\Omega_1 = 0.0035, \Omega_2 = 0.005$
0.0038	0.9587	0.9682	0.9712	0.9729	$\Omega_4 = 0.0059, \Omega_5 = 0.0071$
0.0048	0.9587	0.9682	0.9712	0.9729	$\Psi_1 = 0.51, \Psi_2 = 0.7$
0.0058	0.9570	0.9674	0.9706	0.9724	$\Psi_4 = 0.57, \Psi_5 = 0.61$

The analysis of Table 3 reveals that when the repair rate of this subsystem is increased from 0.21 to 0.82, the system’s availability increases from 0.9615 to 0.9736. Similarly, when the failure rate of this subsystem increases from 0.0028 to 0.0058, the availability of this subsystem decreases from 0.9615 to 0.9570.

The pressing subsystem’s improved repair rate resulted in a 1.21% gain in system availability, while this subsystem’s increased failure rate caused a 0.45% drop in system availability. Figure 5 provides a graphical depiction of how variations in the pressing subsystem parameter impact system availability.

Table 4 demonstrates that the system’s availability is influenced by the fluctuation in the failure and repair rate parameter of the Quality Assurance and Evaluation subsystem.

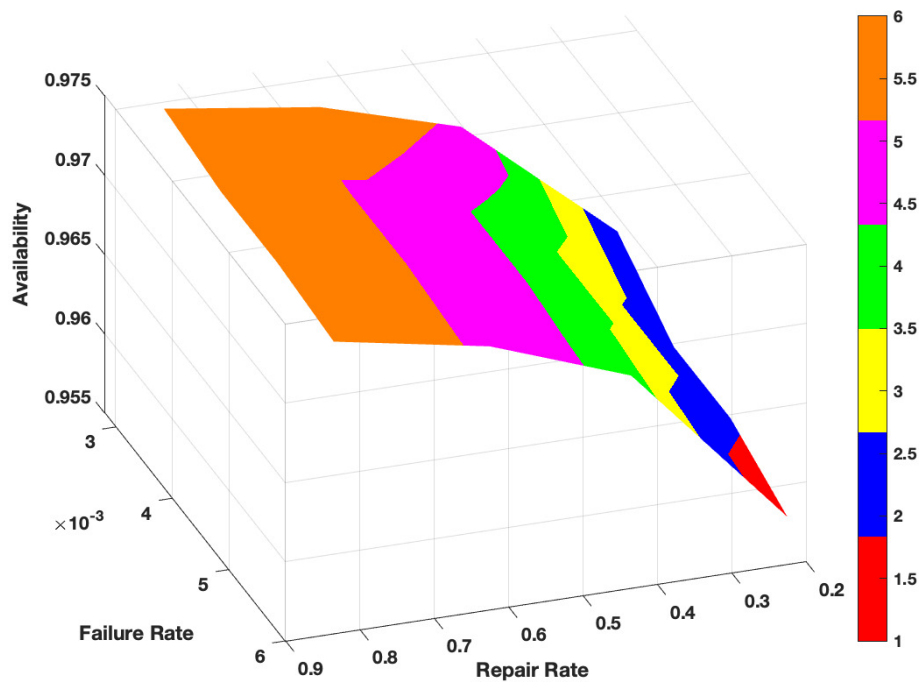


Figure 5: Effect of Pressing Variation in FRR on Plywood System Performance

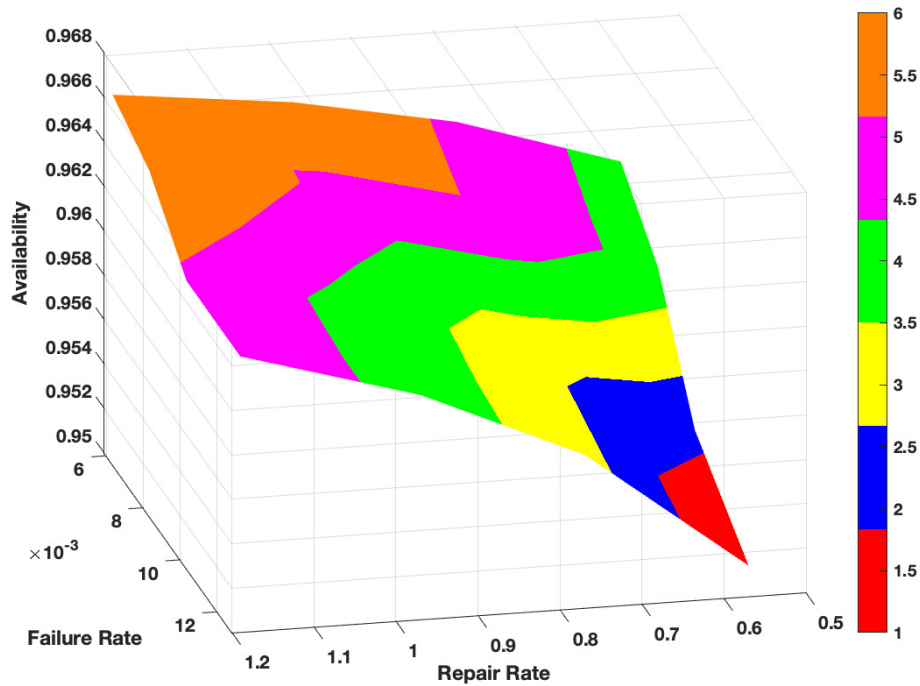
Table 4: The Effects of Quality Assurance and Evaluation Subsystem Repair and Failure Rates on System Availability

Failure rate $\Omega_4$	Repair rate $\Psi_4$				Constant Parameters
	0.57	0.77	0.97	1.19	
0.0059	0.9615	0.9638	0.9652	0.9661	$\Omega_1 = 0.0035, \Omega_2 = 0.005$
0.0079	0.9591	0.9620	0.9638	0.9650	$\Omega_3 = 0.0028, \Omega_5 = 0.0071$
0.0099	0.9541	0.9582	0.9607	0.9624	$\Psi_1 = 0.51, \Psi_2 = 0.7$
0.0128	0.9514	0.9569	0.9601	0.9624	$\Psi_3 = 0.21, \Psi_5 = 0.61$

Upon observation, it has been determined that when the failure rate of this subsystem grows from 0.0059 to 0.0128, the system availability reduces from 0.9615 to 0.9514, representing a 1.01% decrease. Similarly, when the repair rate of the subsystem grows from 0.57 to 1.19, the system availability improves from 0.9615 to 0.9661, representing a 0.46% increase. The graphical representation of this observation is depicted in Figure 6.

The finishing subsystem is a crucial component of the plywood system. The probability of failure for this subsystem is low due to its singular unit configuration. Table 5 displays the impact of system availability as a result of changes in subsystem E's failure and repair rate parameters.

Analysis is based on the finding that when the Finishing subsystem failure rate rises from 0.0071 to 0.01 the system's availability falls from 0.9944 to 0.9970; likewise, when this subsystem's repair rate rises from 0.61 to 1.2 the system's availability rises from 0.9944 to 0.9977. The availability of the system is lowered by 0.74% due to the variability in the failure rate of subsystem D, while the availability is increased by 0.27% due to the variability in its repair rate. Figure 7 depicts the graphical representation of the fluctuations in the failure and repair rate of this subsystem, and how these fluctuations impact the availability of the system.



**Figure 6:** Effect of Quality Assurance and Evaluation Variation in FRR on Plywood System Performance

**Table 5:** The Effects of Finishing Subsystem Repair and Failure Rates on System Availability

Failure rate $\Omega_5$	Repair rate $\Psi_5$				Constant Parameters
	0.61	0.73	0.91	1.2	
0.0071	0.9944	0.9953	0.9962	0.9971	$\Omega_1 = 0.0035, \Omega_2 = 0.005$
0.0083	0.9935	0.9946	0.9956	0.9967	$\Omega_3 = 0.0028, \Omega_4 = 0.0059$
0.0091	0.9870	0.9891	0.9913	0.9934	$\Psi_1 = 0.51, \Psi_2 = 0.7$
0.01	0.9970	0.9891	0.9913	0.9934	$\Psi_3 = 0.21, \Psi_4 = 0.57$

In the Petri nets simulations approach, the availability of the system is also influenced by the presence of a repairman or repair facility. System availability fluctuates with the presence of repair personnel, making it a critical factor. Based on this, we recommend that industries or engineers determine the required number of repair personnel. Table 6 illustrates the correlation or impact of system availability on the presence of a repair technician.

**Table 6:** The Effects of Overall Performance of the System due to Repair Man

Repairman	1	2	3	4	5
Availability	0.9615	0.9631	0.9631	0.9631	0.9631

By examining the correlation between the availability of the system and the number of repairmen, it is observed that the initial availability of the system is 0.9615 when there is only one repairman. As the number of repairmen increases, the availability of the system also increases. However, after a certain threshold, the availability remains constant. The plywood manufacturing system's availability is now steady at 0.9631 after two repairmen. The impact of repairmen on the

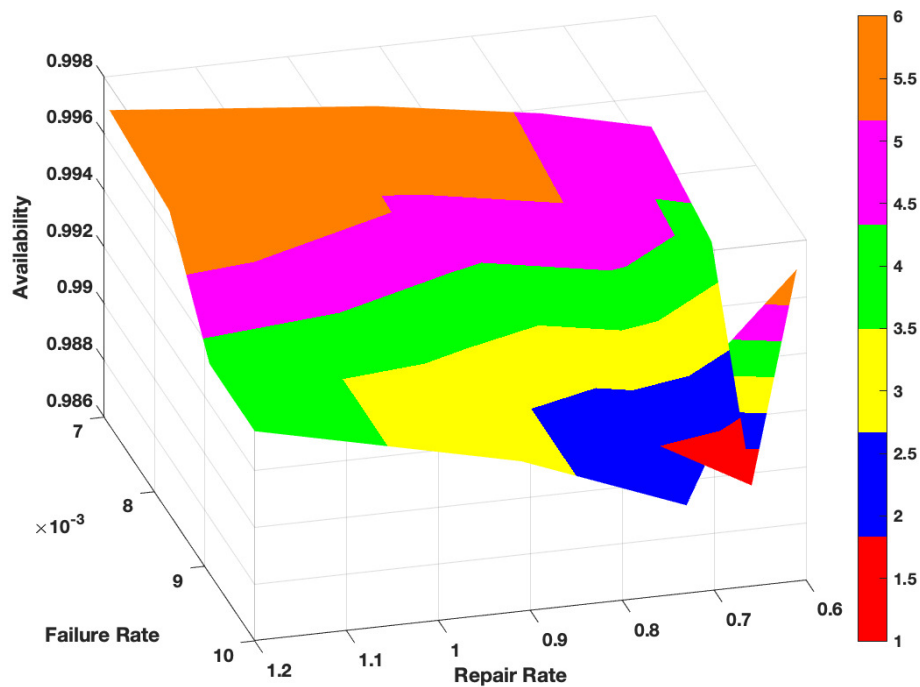


Figure 7: Effect of Finishing Variation in FRR on Plywood System Performance

system's availability is visually depicted in Figure 8.

Based on their observation, it is advised to the engineer that two repairmen are sufficient for repairing the plywood manufacturing system.

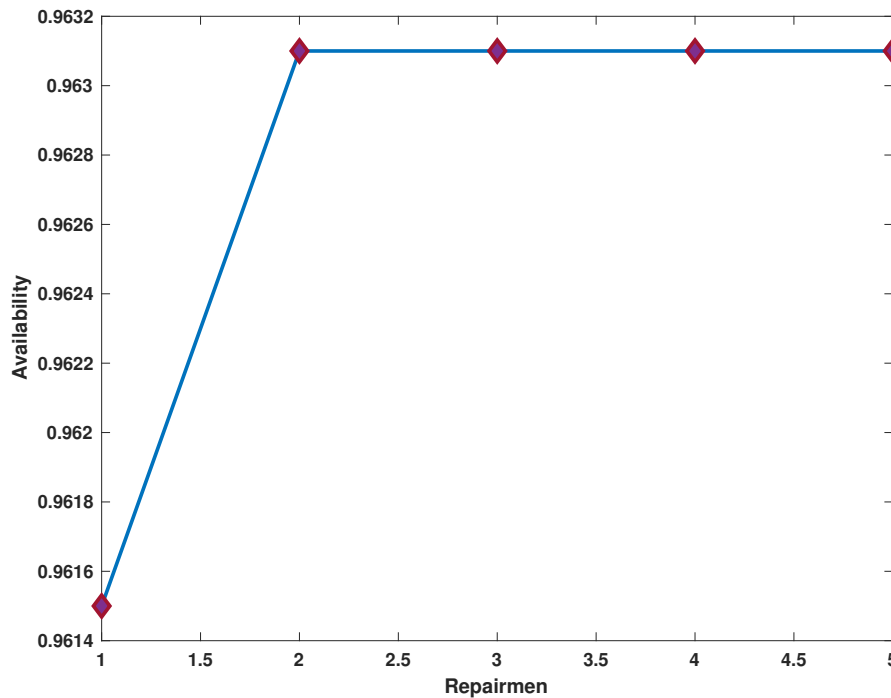
## 5. CONCLUSION

The current case study's findings indicate that the plywood manufacturing process is the most important component of the wood industry system and needs the highest maintenance level. Petri nets simulation modeling is utilized for this purpose in order to examine the system's performance behavior.

Performance matrices 1 through 4 present the findings about the impact of varying failure and repair periods on system availability throughout a range of system operating capacities. Based on the observation, it is determined that the veneer drying subsystem is the most crucial component of this system. This is because any fluctuations in the failure rate of this subsystem result in a reduction of the system's availability by 2.49%. This subsystem must give priority to maintenance with a greater level of importance. The order of maintenance priority of all the subsystems in this plywood system is listed in the following order:

1. Veneer Drying (VD) Subsystem.
2. Quality Assurance and Evaluation (QSE) Subsystem.
3. Glue Mixing (GM) Subsystem.
4. Finishing Subsystem.
5. Pressing Subsystem.





**Figure 8:** *Improving Availability with Additional Repair Capacity*

Repairmen impact the system's availability. An analysis reveals that the initial availability of the system grows as the number of repairmen increases. When there are two repairmen, the availability of the system is measured at 0.9631. If there is a rise in the number of repairs beyond two, the system availability remains at 0.9631. The premise of this study is that it is advisable for the engineer to employ a team of only two repairmen to address any issues with the plywood manufacturing system.

## REFERENCES

- [1] Tan, Jonathan S., and Mark A. Kramer. "A general framework for preventive maintenance optimization in chemical process operations." *Computers & Chemical Engineering* 21.12 (1997): 1451-1469. [https://doi.org/10.1016/S0098-1354\(97\)88493-1](https://doi.org/10.1016/S0098-1354(97)88493-1)
- [2] Adamyan, Angela, and David He. "Analysis of sequential failures for assessment of reliability and safety of manufacturing systems." *Reliability Engineering & System Safety* 76.3 (2002): 227-236. [https://doi.org/10.1016/S0951-8320\(02\)00013-3](https://doi.org/10.1016/S0951-8320(02)00013-3)
- [3] Regattieri, Alberto, and Giacomo Bellomi. "Innovative lay-up system in the plywood manufacturing process." *European Journal of Wood and Wood Products* 67.1 (2009): 55-62. DOI 10.1007/s00107-008-0282-0.
- [4] Bansal, Shikha, and Sohan Tyagi. "Reliability analysis of screw manufacturing plant using orthogonal matrix method." *Pertanika Journal of Science & Technology* 26.4 (2018): 1789-1800.
- [5] Kumar, Amit, Vinod Kumar, and Vikas Modgil. "Performance optimisation for ethanol manufacturing system of distillery plant using particle swarm optimisation algorithm." *International Journal of Intelligent Enterprise* 5.4 (2018): 345-364. <https://doi.org/10.1504/IJIE.2018.095723>
- [6] Kumar, Narendra, P. C. Tewari, and Anish Sachdeva. "Performance modelling and availability analysis of a milk pasteurising system using Petri nets formal-

- ism." *International Journal of Simulation and Process Modelling* 15.5 (2020): 401-409. <https://doi.org/10.1504/IJSPM.2020.110915>
- [7] Kumar, Narendra, P. C. Tewari, and Anish Sachdeva. "Petri Nets modeling and analysis of the Veneer Layup system of plywood manufacturing plant." *International Journal for Engineering Modelling* 33.1-2 Regular Issue (2020): 95-107. <https://doi.org/10.31534/engmod.2020.1-2.ri.07v>
- [8] Malik, Subhash, and P. C. Tewari. "Optimization of coal handling system performability for a thermal power plant using PSO algorithm." *Grey Systems: Theory and Application* 10.3 (2020): 359-376. <https://doi.org/10.1108/GS-01-2020-0002>
- [9] Tyagi, Sohan Lal, et al. "Mathematical Modeling and Availability Analysis of Leaf Spring Manufacturing Plant." *Pertanika Journal of Tropical Agricultural Science* 29.2 (2021). <https://doi.org/10.47836/pjst.29.2.18>
- [10] Bansal, Shikha, Sohan Lal Tyagi, and Vipin Kumar Verma. "Performance Modeling and Availability Analysis of Screw Manufacturing Plant." *Materials Today: Proceedings* 57 (2022): 1985-1988. <https://doi.org/10.1016/j.matpr.2021.10.170>
- [11] Kalaivani, M., and R. Kannan. "Estimation of reliability characteristics for linear consecutive k-out-of-n: F systems based on exponentiated Weibull distribution." *Reliability: Theory & Applications* 17.3 (69) (2022): 59-71. <https://doi.org/10.24412/1932-2321-2022-369-52-58>.
- [12] P. Chaudhary and S. Bansal, "Assessment of the Reliability Performance of Hydro-Electric Power Station," 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). IEEE, 2023, pp. 148-153, doi: 10.1109/ICACITE57410.2023.10183056.
- [13] Er.Sudhir Kumar, Dr. P.C. Tewari PERFORMABILITY OPTIMISATION OF MULTISTATE COAL HANDLING SYSTEM OF A THERMAL POWER PLANT HAVING SUBSYSTEMS DEPENDENCIES USING PSO AND COMPARATIVE STUDY BY PETRI NETS. *Reliability: Theory & Applications*. 2023, March 1(72): 250-263. <https://doi.org/10.24412/1932-2321-2023-172-250-263>.
- [14] S. Bansal and U. Godara, "Prediction of Reliability Factor For Multi-State Computer System With Neural Network Approach," 2023 IEEE 2nd International Conference on Industrial Electronics: Developments & Applications (ICIDeA). IEEE, 2023, pp. 132-135, doi: 10.1109/ICIDeA59866.2023.10295179.
- [15] S. L. Tyagi and S. Bansal, "Optimization Model for Wastewater Treatment Process," 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). IEEE, 2023, pp. 165-168, doi: 10.1109/ICACITE57410.2023.10182482.
- [16] Puran Rathi, Anuradha, S.C. Malik RELIABILITY MODELLING OF A PARALLEL-COLD STANDBY SYSTEM WITH REPAIR PRIORITY. *Reliability: Theory & Applications*. 2023, December 4(76): 760-770, DOI: <https://doi.org/10.24412/1932-2321-2023-476-760-770>.
- [17] U. Godara and S. Bansal, "Performance of Reliability Factors in Steam Turbine Generator Power Plant Using Boolean Function Technique and Neural Network Approach," 2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS). IEEE, 2023, pp. 507-512, doi: 10.1109/ICAECIS58353.2023.10170451.
- [18] Urvashi, Shikha Bansal PREDICTION OF RELIABILITY CHARACTERISTICS OF THRESHER PLANT BASIS ON GENERAL AND COPULA DISTRIBUTION. *Reliability: Theory & Applications*. 2023, December 4(76): 701-715, DOI: <https://doi.org/10.24412/1932-2321-2023-476-701-715>.

# PREDICTIVE MAINTENANCE SCHEME FOR PHASED MISSION SYSTEMS

<sup>1</sup>PREETI WANTI SRIVASTAVA AND <sup>2</sup>SATYA RANI



Department of Operational Research University of Delhi, Delhi-7, India

<sup>1</sup>preetisrivastava.saxena@gmail.com

## Abstract

*In both industrial and military fields, many systems are phase mission systems (PMSs) which execute mission composed of different phases in sequence. The structure, failure behaviour, and working condition of such a system may change from phase to phase. Maintenance actions comprising corrective and preventive maintenance schemes studied in the literature are aimed at retaining the maintained system in a proper condition and improving its availability and extending its life. The present paper deals with finding optimal periodic inspection time using multi-objective criteria comprising objectives of minimizing expected maintenance cost incurred due to predictive, breakdown and periodic maintenance of a PMS, and maximizing its expected residual lifetime. The predictive maintenance is condition-based preventive maintenance that anticipates system failures in order to plan timely interventions on the system and hence improve its performance. The dependency is modelled using Gumbel-Haugaaard copula.*

*An aircraft flight PMS comprising Taxiing phase, Take-Off phase, Cruising phase and Landing phase has been used to illustrate the method developed.*

**Keywords:** phased mission system, reliability, Gumbel-Haugaaard copula, predictive maintenance, periodic maintenance, mean residual life, cost optimization, cumulative exposure model

## 1. INTRODUCTION

The reliability of a supply chain depends on the reliability of all the equipment involved including transportation vehicles, sophisticated machines and computer-based information systems in network of suppliers, manufacturers and distributors whose sole aim is to provide goods and services in a timely manner. The reliability of such equipment in turn depends on their design, maintenance and subsequent repairs. Reliability engineering is therefore part and parcel of operations management.

In real life, systems such as coal transportation systems [1][2], aircrafts [3], avionic parts of airborne weapon systems [4], machining line [5], and nuclear plants are required to execute missions sequentially. Such systems called phased mission systems (PMSs) are subject to multiple, consecutive, non-overlapping operation phases. Failures of these systems during the mission may cause great economic losses to enterprises, serious security threats to personnel, or extensive damage to the environment. Some maintenance activities need to be undertaken during the mission break to reduce the probability of system failure of a PMS in the succeeding mission.

Unlike a non-repairable PMS in a repairable PMS, the state of the system depends not only on failure characteristics of its components but also on maintenance conducted during the mission.

Further, the system reliability depends on its age and the maintenance policy applied. It usually decreases as components deteriorate. Performing proper maintenance actions is necessary to keep the reliability of a system at a desired level. Maintenance is classified into two main categories: corrective maintenance (CM) and preventive maintenance (PM). Corrective maintenance

is generally performed after the system breakdown. Preventive maintenance corresponds to the scheduled actions which are performed while the system is still operational. It aims at keeping the system in available state by improving the condition of its components. Usually, preventive maintenance is more advantageous as it may prevent catastrophic losses due to unpredicted failures [6][7][8][9][10][11][12][13]. The PM actions are usually performed at predetermined points in time to keep the reliability of the system at a desired level.

Predictive maintenance (PdM) also known as condition-based maintenance is meant to minimize unscheduled equipment failures, lost production, and maintenance costs. It involves the use of information such as maintenance logs and sensor data to predict maintenance needs in advance. PdM plays a very important role in the airline industry by helping in reducing delays and costs, while improving and maintaining aircraft operational reliability.

The aim of this paper is to determine optimal periodic inspection time using multi-objective criteria of minimizing the expected maintenance cost due to predictive, breakdown and periodic maintenances, of the PMS, and maximizing its mean residual lifetime. The decision variable is the length of the periodic interval,  $T$ , subject to the constraints that the reliability of each phase does not exceed the pre-specified values.

The paper is organized as follows: Section 2 is a brief literature review. The model of Predictive maintenance cost is explained in Section 3. In Section 4 the phased mission system is explained. Traditional maintenance models involving periodic and breakdown maintenances, and integrated models involving predictive maintenance besides periodic and breakdown maintenances are discussed in Section 5. The concept of Remaining Useful Life (RUL) is highlighted in Section 6, and multi-objective optimization problem is formulated in Section 7. In Section 8, the proposed method is explained using an aircraft flight PMS. The concluding remarks have been made in the last section.

## 2. LITERATURE REVIEW

The maintenance models used in the literature predict problems that can help timely replacement or repair of an equipment before it fails for a single system. The researchers have used knowledge about degradation state of the equipment for prediction purpose [14] out-of-control condition using statistical process control [15][16][17][18][19][20] and on-line sensors [21][22] for prediction purpose for a single system. Maintenance at system-level of a PMS without considering predictive maintenance has been studied by [23]. The present paper deals with maintenance of a PMS taking into account predictive, periodic and breakdown maintenances along with its mean residual lifetime. It is assumed that the components are dependent within a phase, and all the phases involved are dependent. The dependency is modelled using Gumbel-Hougaard copula.

## 3. PREDICTIVE MAINTENANCE MODEL

Define  $f_{PMS}(t)$  as the density function that specifies the probability of failure of a PMS at time  $t$  and  $g(s|t)$  as the conditional density function that specifies the probability that the signal of a potential failure is received at time  $s$  given that the actual failure would have occurred at time  $t$ . The conditional density,  $g(s|t)$ , defines the capability (i.e., accuracy and precision) of the prediction system.

The choice of the distribution form for the prediction signal, conditional on the equipment failure, is based on the concept of "P-F curves" for prediction systems [24] as well as diagnosis of the sensor equipment by the concerned technician(s).

Thus

$$g(s|t) = \begin{cases} k(1-\beta)s^{k-1}t^{-k} & 0 \leq s \leq t \\ \beta & s > t \end{cases}$$

$$G(s|t) = \begin{cases} (1-\beta)\left(\frac{s}{t}\right)^k & 0 \leq s \leq t \\ 1 & s > t \end{cases},$$

where  $s$  is the time of the signal,  $t$  is the time of failure if no replacement is made,  $k$  is the prediction precision,  $(1 - \beta)$  is the prediction accuracy and  $k \geq 1$ ,  $0 \leq \beta \leq 1$ , are respectively, the conditional density and distribution function used for the purpose. This form of the conditional distribution function characterizes the features of typical signal and failure times seen in industry [22]).

In this paper, the objective is to minimize the expected maintenance cost of a PMS per unit time. The maintenance costs include costs of periodic and predictive replacements and that of failures. It is assumed that the PMS will go for maintenance after completing the mission and restored to “as good as new” condition, therefore, using renewal reward process the expected maintenance cost per period is:

$$\frac{E[\text{Predictive Maintenance cost} + \text{Breakdown cost} + \text{Periodic maintenance cost}]}{E[\text{Time until maintenance}]}$$

(See for reference [25]).

#### 4. PHASED MISSION SYSTEM (PMS)

A phased mission system (PMS) is defined as a system comprising multiple, consecutive, and non-overlapping phases. During each phase, a PMS needs to complete a specific task without failure. In these phases, the system may be subject to different working conditions and environmental stresses, as well as different performance requirements. For example, in a twin-engine airplane with two phases, namely, taxiing phase and take-off phase, one engine is required in the former phase, and both the engines are necessary in the latter phase. In contrast to the other phases of the flight profile the engines are more prone to failure during the take-off period due to enormous pressure they undergo during this period [26][27]. So in different phases, the system configurations and the components, failure rates and even failure criteria could be vastly different.

Let  $T_{mn}$  denote lifetime of component  $m$  of phase  $n$  with reliability  $\bar{H}_{mn}(t)$ . Let  $\bar{F}_{m1}(t)$ ,  $\bar{F}_{m2}(t)$ , ... and  $\bar{F}_{mn}(t)$  be the reliability of phase 1, phase 2, ... and phase  $n$ , respectively. Then, reliability of PMS is:

$$\bar{F}_{PMS}(t) = \begin{cases} \bar{F}_{m1}(t), & 0 \leq t \leq \tau_1 \\ \bar{F}_{m2}(t), & \tau_1 \leq t \leq \tau_2 \\ \vdots \\ \bar{F}_{mn}(t), & \tau_{n-1} \leq t \leq \tau_n, \end{cases} \quad (1)$$

where  $(\tau_{n-1}, \tau_n)$  represents time-duration of functioning of phase  $n$  of the phased mission system  $n = 1, 2, 3, 4, \dots, n$ ,  $\tau_0 = 0$ .

Since considering phase  $n$  has  $m$  dependent components and reliability of phase  $n$  denoted by  $\bar{F}_{mn}(t)$  so dependency is modelled using Gumbel-Haugaaard copula [28] gives,

$$\bar{F}_{mn}(t) = C(\bar{H}_{1n}(t), \bar{H}_{2n}(t), \dots, \bar{H}_{mn}(t)). \quad (2)$$

And, reliability of PMS is:

$$\bar{F}_{PMS}(t) = C(\bar{F}_{m1}(\tau_1), \bar{F}_{m2}(\tau_2), \bar{F}_{m3}(\tau_3), \dots, \bar{F}_{mn}(\tau_n)). \quad (3)$$

The cumulative exposure model [29] is used in equation (2), to obtain the reliability of phase  $n$  at  $\tau_n$ . We obtain,

$$\bar{F}_{mn}(\tau_n) = C(\bar{H}_{1n}(\tau_n - \tau_{n-1} + l_{1n}), \bar{H}_{2n}(\tau_n - \tau_{n-1} + l_{2n}), \dots, \bar{H}_{mn}(\tau_n - \tau_{n-1} + l_{mn})) \quad (4)$$

$l_{mn}$ , where  $m$  denotes the components and  $n$  denotes the phase of the system,  $m = 1, \dots, m$ , &  $n = 1, \dots, n$ , is determined in such a way that [30])

$\bar{H}_{mn}(l_{mn}) = \bar{H}_{mn-1}(\tau_{n-1} - \tau_{n-2} + l_{mn-1})$ , and  $l_{1n-1} = 0$ ,  
 where  $C(\bar{H}_{1n}(t), \bar{H}_{2n}(t), \bar{H}_{3n}(t), \dots, \bar{H}_{mn}(t))$  is the  $m$ -dimensional Gumbel-Hougaard.  
 Thus,

$$C(\bar{H}_{1n}(t), \bar{H}_{2n}(t), \bar{H}_{3n}(t), \dots, \bar{H}_{mn}(t)) = \exp \left[ - \left( (-\log(\bar{H}_{1n}(t)))^\theta + (-\log(\bar{H}_{2n}(t)))^\theta + (-\log(\bar{H}_{3n}(t)))^\theta + \dots + (-\log(\bar{H}_{mn}(t)))^\theta \right)^{1/\theta} \right].$$

## 5. MAINTENANCE MODEL

The present section focuses on the traditional periodic maintenance model (TM) and integrated model (IM).

### 5.1. Traditional Model

In TM no predictive maintenance is used, periodic maintenance is conducted if there has been no failure prior to time  $T$ , and breakdown maintenance is conducted if the equipment fails prior to time  $T$ .

For the TM, the decision variable is the periodic interval  $T$  and the objective function value is as follows:

$$C_{TM}(T) = \frac{E[C_{BP}]}{E[C_{T1}]}, \tag{5}$$

where

$$E[C_{BP}] = M_b \left[ \int_0^T f_{PMS}(t) dt \right] + M_p \left[ \int_T^\infty f_{PMS}(t) dt \right],$$

is sum of expectation of breakdown maintenance costs and periodic maintenance cost, and

$$E[C_{T1}] = \left[ \int_0^T t f_{PMS}(t) dt \right] + T \left[ \int_T^\infty f_{PMS}(t) dt \right],$$

is mean time between failure (replacement).

### 5.2. Integrated Model (IM)

The second model utilizes both predictive and periodic maintenance and is referred to as the Integrated Model. For IM, the decision variable is the periodic interval,  $T$ , and the objective function is:

$$C_{IM}(T) = \frac{E[C_{PdBP}]}{E[C_{T2}]}, \tag{6}$$

where,

$$E[C_{PdBP}] = M_{pd} \left[ (1 - \beta) \int_0^T f_{PMS}(t) dt + \int_T^\infty G(T | t) f_{PMS}(t) dt \right] + M_b \left[ \beta \int_0^T f_{PMS}(t) dt \right] + M_p \left[ \int_T^\infty [1 - G(T | t)] f_{PMS}(t) dt \right],$$

is sum of expectation of predictive maintenance cost, breakdown maintenance costs and periodic maintenance cost,  $M_{pd}$  is predictive maintenance cost,  $M_b$  is breakdown maintenance and  $M_p$  is periodic maintenance, and

$$E[C_{T2}] = \left[ \int_0^T \int_0^t sg(s|t) f_{PMS}(t) ds dt + \int_T^\infty \int_0^T sg(s|t) f_{PMS}(t) ds dt \right] + \left[ \beta \int_0^T t f_{PMS}(t) dt \right] + \left[ T \int_T^\infty [1 - G(T|t)] f_{PMS}(t) dt \right],$$

is sum of expected time between replacement with signal and without signal.

### 6. REMAINING USEFUL LIFE (RUL)

RUL is the residual life time of a system used to perform its functional capabilities before failure. It is a key metric and critical for predicting the failure of a machine in the production line, and is used by engineers to decide whether to do maintenance or delay it due to production requirements [31].

Let  $T_{PMS}$  be the time to failure of the phased mission system, and suppose the phased mission system has survived until time  $t$ . Then the “conditional” random variable

$$X_{PMS} = T_{PMS} - t(T_{PMS} > t),$$

i.e., the remaining time to failure, is called “RUL” of the phased mission system.

The conditional reliability function

$$\bar{F}_{PMS}(t) = P_{PMS}(X_{PMS} > x) = P(T_{PMS} - t > x | T_{PMS} - t), x \geq 0,$$

incorporates all the information relevant for prediction and future planning. The mean residual life (MRL) used as a point estimate of RUL or a prediction interval for RUL is defined as:

$$\mu_{PMS}(t) = E_{PMS}[X_{PMS}] = E[[T_{PMS} - t | T_{PMS} > t]].$$

Then,  $\mu_{PMS}(0) = \mu_{PMS} = E[T]$  and

$$\mu_{PMS}(t) = \int_0^\infty \bar{F}_{PMS}(x) dx = \frac{\int_t^\infty \bar{F}_{PMS}(x) dx}{\bar{F}_{PMS}(t)}. \tag{7}$$

### 7. OPTIMIZATION PROBLEMS

Amongst various approaches used to solve a multi-objective optimization problem, one of the commonly used approach is to combine the objectives involved into one single composite objective so that the traditional mathematical programming method can be used for the propose.

In this paper, the weighted sum multi-objective optimization problem is used to minimize the expected maintenance cost per unit time and maximize mean residual lifetime function for the PMS subject to the constraints that the reliability of the each phase does not exceed the pre-specified values,  $R_i, i = 1, 2, \dots, n$ .

Let  $T1$  be the periodic inspection time for the traditional model and  $T2$  be that for the Integrated Model.

The optimization problem is formulated as:

### 7.1. Optimizing $C_{TM}$

$$\text{Min } Z_1 = w_1 C_{TM}(T_1) + w_2(-\mu_{PMS}(T_1))$$

subject to,  $T_1 \geq \tau_n$ ,

$$1 \geq \bar{F}_{mi}(T_1) \geq R_i, i = 1, 2, \dots, n C_{TM} \geq 0.$$

### 7.2. Optimizing $C_{IM}$

$$\text{Min } Z_2 = w_1 C_{IM}(T_2) + w_2(-\mu_{PMS}(T_2))$$

subject to,  $T_2 \geq \tau_n$ ,

$$1 \geq \bar{F}_{mi}(T_2) \geq R_i, i = 1, 2, \dots, n C_{IM} \geq 0.$$

*Mathematica 11.0* has been used for solve the optimization problem.

## 8. NUMERICAL ILLUSTRATIONS

In this section aircraft flight PMS used for the illustrative purpose, Figure 1(a)-(d), shows reliability block diagrams for the four-phase aircraft flight comprises. The first phase is taxiing in which the navigation system, one out of the four engines and all three landing gears are needed, the second phase is take-off where in all four engines, the navigation system and all three landing gears are needed, the third phase is cruising in which the navigation system and three of the four engines are required. Finally, the fourth phase is landing comprising the navigation system, two of the four engines and all three landing gears.

### 8.1. Reliability of Aircraft Flight PMS system

Let  $T_1$  denote lifetimes of navigation with reliability  $\bar{H}_{1n}(t)$ .  $T_2, T_3, T_4$  and  $T_5$  denote lifetimes of the four engines  $E_1, E_2, E_3$  and  $E_4$  with reliabilities  $\bar{H}_{2n}(t), \bar{H}_{3n}(t), \bar{H}_{4n}(t)$  and  $\bar{H}_{5n}(t)$ , respectively, and  $T_6, T_7$  and  $T_8$  denote lifetimes of landing gear 1 ( $G_1$ ), landing gear 2 ( $G_2$ ) and landing gear 3 ( $G_3$ ) with reliabilities  $\bar{H}_{6n}(t), \bar{H}_{7n}(t)$  and  $\bar{H}_{8n}(t)$ , respectively. Let  $\bar{F}_{p1}(t), \bar{F}_{p2}(t), \bar{F}_{p3}(t)$ , and  $\bar{F}_{p4}(t)$  be the reliability of subsystems in phase 1, phase 2, phase 3 and phase 4, respectively. Then, reliability of 4-PMS is:

$$\bar{F}_{PMS}(t) = \begin{cases} \bar{F}_{p1}(t), & 0 \leq t \leq \tau_1 \\ \bar{F}_{p2}(t), & \tau_1 \leq t \leq \tau_2 \\ \bar{F}_{p3}(t), & \tau_2 \leq t \leq \tau_3 \\ \bar{F}_{p4}(t), & \tau_3 \leq t \leq \tau_4. \end{cases} \quad (8)$$

#### PHASE-1(Taxiing Phase)

Let  $H_{11}(t)$  be life distribution of navigation,  $H_{21}(t), H_{31}(t), H_{41}(t)$  &  $H_{51}(t)$  be life distribution of components ' $E'_1, E'_2, E'_3$  & ' $E'_4$ ', respectively further let  $H_{61}(t), H_{71}(t)$  &  $H_{81}(t)$  be life distribution of components ' $G'_1, G'_2$  & ' $G'_3$ ', respectively.

Reliability of navigation,

$$\bar{F}_{11}(t) = p [T_1 > t].$$

Reliability of engines,

$$\begin{aligned} \bar{F}_{21}(t) = & p [T_2 > t, T_3 \leq t, T_4 \leq t, T_5 \leq t] + p [T_2 \leq t, T_3 > t, T_4 \leq t, T_5 \leq t] + p [T_2 \leq t, T_3 \leq t, T_4 > t, T_5 \leq t] + \\ & p [T_2 \leq t, T_3 \leq t, T_4 \leq t, T_5 > t] + p [T_2 > t, T_3 > t, T_4 \leq t, T_5 \leq t] + p [T_2 > t, T_3 \leq t, T_4 > t, T_5 \leq t] + \\ & p [T_2 > t, T_3 \leq t, T_4 \leq t, T_5 > t] + p [T_2 \leq t, T_3 > t, T_4 > t, T_5 \leq t] + p [T_2 \leq t, T_3 > t, T_4 \leq t, T_5 > t] + \\ & p [T_2 \leq t, T_3 \leq t, T_4 > t, T_5 > t] + p [T_2 > t, T_3 > t, T_4 > t, T_5 \leq t] + p [T_2 > t, T_3 > t, T_4 \leq t, T_5 > t] + \\ & p [T_2 > t, T_3 \leq t, T_4 > t, T_5 > t] + p [T_2 \leq t, T_3 > t, T_4 > t, T_5 > t] + p [T_2 > t, T_3 > t, T_4 > t, T_5 > t]. \end{aligned}$$



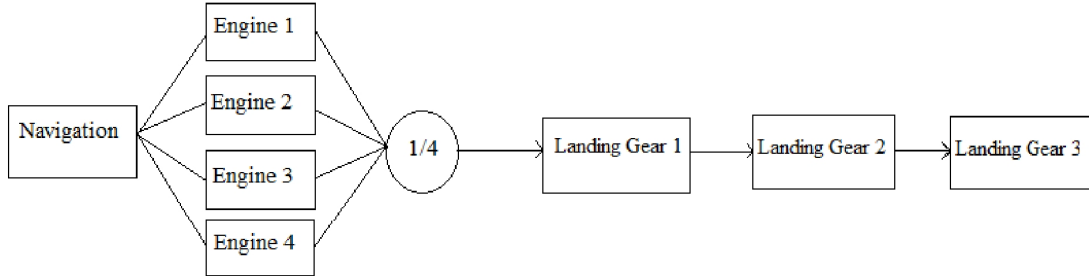


Figure 1(a): Taxiing Phase

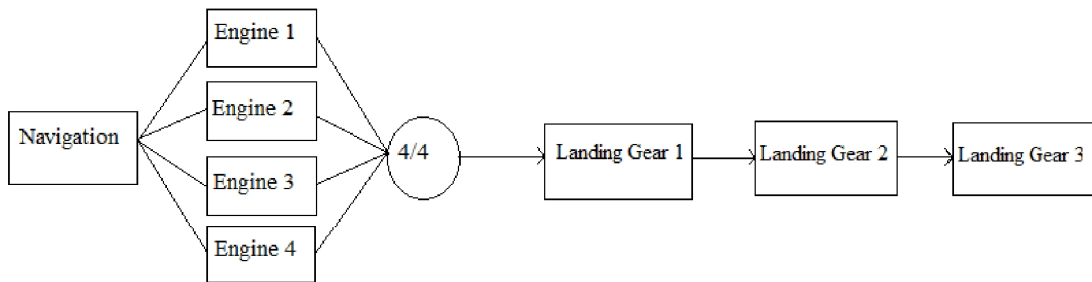


Figure 1(b): Take-Off Phase

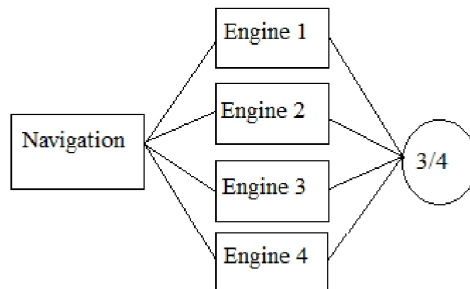


Figure 1(c): Cruising Phase

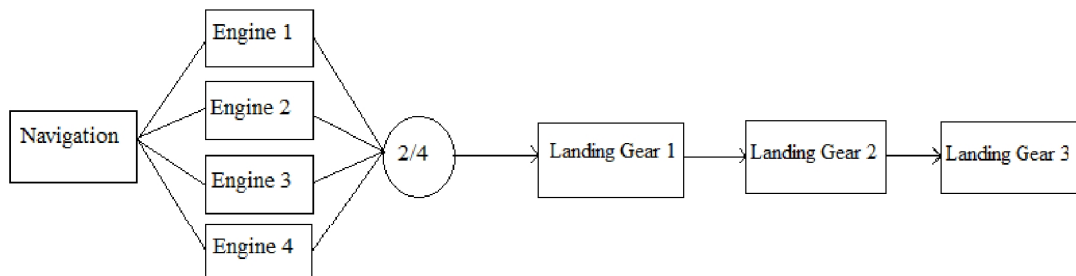


Figure 1(d): Landing Phase

Figure 1: 1(a)-(d) Reliability Block Diagrams for the four- phase aircraft flight[32]

Reliability of landing gear,

$$\bar{F}_{31}(t) = p [T_6 > t, T_7 > t, T_8 > t].$$

Thus, Reliability of phase-1,

$$\bar{F}_{p1}(t) = \bar{F}_{11}(t) \cdot \bar{F}_{21}(t) \cdot \bar{F}_{31}(t). \quad (9)$$

**PHASE-2 (Take-Off Phase)**

Let  $H_{12}(t)$  be life distribution of navigation,  $H_{22}(t), H_{32}(t), H_{42}(t)$  &  $H_{52}(t)$  be life distribution of components ' $E'_1, E'_2, E'_3$  & ' $E'_4$ ', respectively further let  $H_{62}(t), H_{72}(t)$  &  $H_{82}(t)$  be life distribution of components ' $G'_1, G'_2$  & ' $G'_3$ ', respectively.

Reliability of navigation,

$$\bar{F}_{12}(t) = p [T_1 > t].$$

Reliability of engines,

$$\bar{F}_{22}(t) = p [T_2 > t, T_3 > t, T_4 > t, T_5 > t].$$

Reliability of landing gear,

$$\bar{F}_{32}(t) = p [T_6 > t, T_7 > t, T_8 > t].$$

Thus, Reliability of phase-2,

$$\bar{F}_{p2}(t) = \bar{F}_{12}(t) \cdot \bar{F}_{22}(t) \cdot \bar{F}_{32}(t). \quad (10)$$

**PHASE-3 (Cruising Phase)**

Let  $H_{13}(t)$  be life distribution of navigation,  $H_{23}(t), H_{33}(t), H_{43}(t)$  &  $H_{53}(t)$  be life distribution of components ' $E'_1, E'_2, E'_3$  & ' $E'_4$ ', respectively further let  $H_{63}(t), H_{73}(t)$  &  $H_{83}(t)$  be life distribution of components ' $G'_1, G'_2$  & ' $G'_3$ ', respectively.

Reliability of navigation,

$$\bar{F}_{13}(t) = p [T_1 > t].$$

Reliability of engines,

$$\bar{F}_{23}(t) = p [T_2 > t, T_3 > t, T_4 > t, T_5 \leq t] + p [T_2 > t, T_3 > t, T_4 \leq t, T_5 > t] + p [T_2 > t, T_3 \leq t, T_4 > t, T_5 > t] + p [T_2 \leq t, T_3 > t, T_4 > t, T_5 > t] + p [T_2 > t, T_3 > t, T_4 > t, T_5 > t].$$

Thus, Reliability of phase-3,

$$\bar{F}_{p3}(t) = \bar{F}_{13}(t) \cdot \bar{F}_{23}(t). \quad (11)$$

**PHASE-4 (Landing Phase)**

Let  $H_{14}(t)$  be life distribution of navigation,  $H_{24}(t), H_{34}(t), H_{44}(t)$  &  $H_{54}(t)$  be life distribution of components ' $E'_1, E'_2, E'_3$  & ' $E'_4$ ', respectively further let  $H_{64}(t), H_{74}(t)$  &  $H_{84}(t)$  be life distribution of components ' $G'_1, G'_2$  & ' $G'_3$ ', respectively.

Reliability of navigation,

$$\bar{F}_{14}(t) = p [T_1 > t].$$

Reliability of engines,

$$\bar{F}_{24}(t) = p [T_2 > t, T_3 > t, T_4 \leq t, T_5 \leq t] + p [T_2 > t, T_3 \leq t, T_4 > t, T_5 \leq t] + p [T_2 > t, T_3 \leq t, T_4 \leq t, T_5 > t] + p [T_2 \leq t, T_3 > t, T_4 > t, T_5 \leq t] + p [T_2 \leq t, T_3 > t, T_4 \leq t, T_5 > t] + p [T_2 \leq t, T_3 \leq t, T_4 > t, T_5 > t] + p [T_2 > t, T_3 > t, T_4 > t, T_5 \leq t] + p [T_2 > t, T_3 > t, T_4 \leq t, T_5 > t] + p [T_2 > t, T_3 \leq t, T_4 > t, T_5 > t] + p [T_2 \leq t, T_3 > t, T_4 > t, T_5 > t] + p [T_2 > t, T_3 > t, T_4 > t, T_5 > t].$$

Reliability of landing gear,

$$\bar{F}_{34}(t) = p [T_6 > t, T_7 > t, T_8 > t].$$

Thus, Reliability of phase-2,

$$\bar{F}_{p4}(t) = \bar{F}_{14}(t) \cdot \bar{F}_{24}(t) \cdot \bar{F}_{34}(t). \tag{12}$$

Reliability of Aircraft flight PMS system using equation (3), we have

$$\bar{F}(t) = C(\bar{F}_{p1}(t), \bar{F}_{p2}(t), \bar{F}_{p3}(t), \bar{F}_{p4}(t)) \tag{13}$$

equations(9), (10), (11) and (12) give reliability of the four phases in PMS.

After using  $p[AB] + p[AB^c] = p[A]$  and Gumbel-Hougaard copula equation (2) in above equations we get,

$$\bar{F}_{11}(t) = \bar{H}_{11}(t),$$

$$\begin{aligned} \bar{F}_{21}(t) &= C(\bar{H}_{21}(t), 1, 1, 1) + C(1, \bar{H}_{31}(t), 1, 1) + C(1, 1, \bar{H}_{41}(t), 1) + C(1, 1, 1, \bar{H}_{51}(t)) \\ &- C(\bar{H}_{21}(t), \bar{H}_{31}(t), 1, 1) - C(\bar{H}_{21}(t), 1, \bar{H}_{41}(t), 1) - C(\bar{H}_{21}(t), 1, 1, \bar{H}_{51}(t)) - C(1, 1, \bar{H}_{41}(t), \bar{H}_{51}(t)) \\ &- C(1, \bar{H}_{31}(t), 1, \bar{H}_{51}(t)) - C(1, \bar{H}_{31}(t), \bar{H}_{41}(t), 1) + C(\bar{H}_{21}(t), \bar{H}_{31}(t), \bar{H}_{41}(t), 1) \\ &+ C(\bar{H}_{21}(t), \bar{H}_{31}(t), 1, \bar{H}_{51}(t)) + C(\bar{H}_{21}(t), 1, \bar{H}_{41}(t), \bar{H}_{51}(t)) + C(1, \bar{H}_{31}(t), \bar{H}_{41}(t), \bar{H}_{51}(t)) \\ &- C(\bar{H}_{21}(t), \bar{H}_{31}(t), \bar{H}_{41}(t), \bar{H}_{51}(t)), \end{aligned}$$

$$\bar{F}_{31}(t) = C(\bar{H}_{61}(t), \bar{H}_{71}(t), \bar{H}_{81}(t)),$$

$$\bar{F}_{12}(t) = \bar{H}_{12}(t),$$

$$\bar{F}_{22}(t) = C(\bar{H}_{22}(t), \bar{H}_{32}(t), \bar{H}_{42}(t), \bar{H}_{52}(t)),$$

$$\bar{F}_{32}(t) = C(\bar{H}_{62}(t), \bar{H}_{72}(t), \bar{H}_{82}(t)),$$

$$\bar{F}_{13}(t) = \bar{H}_{13}(t),$$

$$\begin{aligned} \bar{F}_{23}(t) &= C(\bar{H}_{23}(t), \bar{H}_{33}(t), \bar{H}_{43}(t), 1) + C(\bar{H}_{23}(t), \bar{H}_{33}(t), 1, \bar{H}_{53}(t)) \\ &+ C(\bar{H}_{23}(t), 1, \bar{H}_{43}(t), \bar{H}_{53}(t)) + C(1, \bar{H}_{33}(t), \bar{H}_{43}(t), \bar{H}_{53}(t)) - 3C(\bar{H}_{23}(t), \bar{H}_{33}(t), \bar{H}_{43}(t), \bar{H}_{53}(t)), \end{aligned}$$

$$\bar{F}_{14}(t) = \bar{H}_{14}(t),$$

$$\begin{aligned} \bar{F}_{24}(t) &= C(\bar{H}_{24}(t), \bar{H}_{34}(t), 1, 1) + C(\bar{H}_{24}(t), 1, \bar{H}_{44}(t), 1) + C(\bar{H}_{24}(t), 1, 1, \bar{H}_{54}(t)) \\ &+ C(1, 1, \bar{H}_{44}(t), \bar{H}_{54}(t)) + C(1, \bar{H}_{34}(t), 1, \bar{H}_{54}(t)) + C(1, \bar{H}_{34}(t), \bar{H}_{44}(t), 1) \\ &- 2C(\bar{H}_{24}(t), \bar{H}_{34}(t), \bar{H}_{44}(t), 1) - 2C(\bar{H}_{24}(t), \bar{H}_{34}(t), 1, \bar{H}_{54}(t)) - 2C(\bar{H}_{24}(t), 1, \bar{H}_{44}(t), \bar{H}_{54}(t)) \\ &- 2C(1, \bar{H}_{34}(t), \bar{H}_{44}(t), \bar{H}_{54}(t)) + 3C(\bar{H}_{24}(t), \bar{H}_{34}(t), \bar{H}_{44}(t), \bar{H}_{54}(t)), \end{aligned}$$

$$\bar{F}_{34}(t) = p [T_6 > t, T_7 > t, T_8 > t] = C(\bar{H}_{64}(t), \bar{H}_{74}(t), \bar{H}_{84}(t)).$$

The cumulative exposure model is used in above equations, to obtain the reliability of subsystems in phase 1, phase 2, phase 3 and phase 4 at  $\tau_1, \tau_2, \tau_3$  and  $\tau_4$ , respectively. Thus,

$$\bar{F}_{11}(\tau_1) = \bar{H}_{11}(\tau_1),$$

$$\begin{aligned} \bar{F}_{21}(\tau_1) = & C(\bar{H}_{21}(\tau_1), 1, 1, 1) + C(1, \bar{H}_{31}(\tau_1), 1, 1) + C(1, 1, \bar{H}_{41}(\tau_1), 1) + C(1, 1, 1, \bar{H}_{51}(\tau_1)) \\ & - C(\bar{H}_{21}(\tau_1), \bar{H}_{31}(\tau_1), 1, 1) - C(\bar{H}_{21}(\tau_1), 1, \bar{H}_{41}(\tau_1), 1) - C(\bar{H}_{21}(\tau_1), 1, 1, \bar{H}_{51}(\tau_1)) \\ & - C(1, 1, \bar{H}_{41}(\tau_1), \bar{H}_{51}(\tau_1)) - C(1, \bar{H}_{31}(\tau_1), 1, \bar{H}_{51}(\tau_1)) - C(1, \bar{H}_{31}(\tau_1), \bar{H}_{41}(\tau_1), 1) \\ & + C(\bar{H}_{21}(\tau_1), \bar{H}_{31}(\tau_1), \bar{H}_{41}(\tau_1), 1) + C(\bar{H}_{21}(\tau_1), \bar{H}_{31}(\tau_1), 1, \bar{H}_{51}(\tau_1)) + C(\bar{H}_{21}(\tau_1), 1, \bar{H}_{41}(\tau_1), \bar{H}_{51}(\tau_1)) \\ & + C(1, \bar{H}_{31}(\tau_1), \bar{H}_{41}(\tau_1), \bar{H}_{51}(\tau_1)) - C(\bar{H}_{21}(\tau_1), \bar{H}_{31}(\tau_1), \bar{H}_{41}(\tau_1), \bar{H}_{51}(\tau_1)), \end{aligned}$$

$$\bar{F}_{31}(\tau_1) = C(\bar{H}_{61}(\tau_1), \bar{H}_{71}(\tau_1), \bar{H}_{81}(\tau_1)),$$

$$\bar{F}_{12}(\tau_2) = \bar{H}_{12}(\tau_2 - \tau_1 + l_{12}),$$

$$\bar{F}_{22}(\tau_2) = C(\bar{H}_{22}(\tau_2 - \tau_1 + l_{22}), \bar{H}_{32}(\tau_2 - \tau_1 + l_{32}), \bar{H}_{42}(\tau_2 - \tau_1 + l_{42}), \bar{H}_{52}(\tau_2 - \tau_1 + l_{52})),$$

$$\bar{F}_{32}(\tau_2) = C(\bar{H}_{62}(\tau_2 - \tau_1 + l_{62}), \bar{H}_{72}(\tau_2 - \tau_1 + l_{72}), \bar{H}_{82}(\tau_2 - \tau_1 + l_{82})),$$

$$\bar{F}_{13}(\tau_3) = \bar{H}_{13}(\tau_3 - \tau_2 + l_{13}),$$

$$\begin{aligned} \bar{F}_{23}(\tau_3) = & C(\bar{H}_{23}(\tau_3 - \tau_2 + l_{23}), \bar{H}_{33}(\tau_3 - \tau_2 + l_{33}), \bar{H}_{43}(\tau_3 - \tau_2 + l_{43}), 1) \\ & + C(\bar{H}_{23}(\tau_3 - \tau_2 + l_{23}), \bar{H}_{33}(\tau_3 - \tau_2 + l_{33}), 1, \bar{H}_{53}(\tau_3 - \tau_2 + l_{53})) \\ & + C(\bar{H}_{23}(\tau_3 - \tau_2 + l_{23}), 1, \bar{H}_{43}(\tau_3 - \tau_2 + l_{43}), \bar{H}_{53}(\tau_3 - \tau_2 + l_{53})) \\ & + C(1, \bar{H}_{33}(\tau_3 - \tau_2 + l_{33}), \bar{H}_{43}(\tau_3 - \tau_2 + l_{43}), \bar{H}_{53}(\tau_3 - \tau_2 + l_{53})) \\ & - 3C(\bar{H}_{23}(\tau_3 - \tau_2 + l_{23}), \bar{H}_{33}(\tau_3 - \tau_2 + l_{33}), \bar{H}_{43}(\tau_3 - \tau_2 + l_{43}), \bar{H}_{53}(\tau_3 - \tau_2 + l_{53})), \end{aligned}$$

$$\bar{F}_{14}(\tau_4) = \bar{H}_{14}(\tau_4 - \tau_3 + l_{14}),$$

$$\begin{aligned} \bar{F}_{24}(\tau_4) = & C(\bar{H}_{24}(\tau_4 - \tau_3 + l_{24}), \bar{H}_{34}(\tau_4 - \tau_3 + l_{34}), 1, 1) + C(\bar{H}_{24}(\tau_4 - \tau_3 + l_{24}), 1, \bar{H}_{44}(\tau_4 - \tau_3 + l_{44}), 1) \\ & + C(\bar{H}_{24}(\tau_4 - \tau_3 + l_{24}), 1, 1, \bar{H}_{54}(\tau_4 - \tau_3 + l_{54})) + C(1, 1, \bar{H}_{44}(\tau_4 - \tau_3 + l_{44}), \bar{H}_{54}(\tau_4 - \tau_3 + l_{54})) \\ & + C(1, \bar{H}_{34}(t), 1, \bar{H}_{54}(t)) + C(1, \bar{H}_{34}(t), \bar{H}_{44}(t), 1) - 2C(\bar{H}_{24}(t), \bar{H}_{34}(t), \bar{H}_{44}(t), 1) \\ & - 2C(\bar{H}_{24}(t), \bar{H}_{34}(t), 1, \bar{H}_{54}(t)) - 2C(\bar{H}_{24}(t), 1, \bar{H}_{44}(t), \bar{H}_{54}(t)) - 2C(1, \bar{H}_{34}(t), \bar{H}_{44}(t), \bar{H}_{54}(t)) \\ & + 3C(\bar{H}_{24}(t), \bar{H}_{34}(t), \bar{H}_{44}(t), \bar{H}_{54}(t)) \end{aligned}$$

$$\bar{F}_{34}(\tau_4) = p[T_6 > t, T_7 > t, T_8 > t] = C(\bar{H}_{64}(t), \bar{H}_{74}(t), \bar{H}_{84}(t)).$$

It is assumed that a component's life distribution in a phase is Weibull with reliability function:

$$\bar{H}_{mn}(t) = \exp[-(t/\alpha_{mn})^\gamma], t > 0; \alpha_{mn} > 0; \gamma > 0; n = 1, 2, 3, 4, m = 1, 2, 3, 4, 5, 6, 7, 8.$$

To illustrate the above model, assume that each of the phase- Taxiing and Take-Off has duration of 15 minutes, cruising phase has duration of 130 minutes and landing phase has duration of 20 minutes. Components of the aircraft follow weibull distribution with  $\gamma = 1.8$  with  $\alpha_{mn} = 1000$  hours for navigation system,  $\alpha_{mn} = 950$  hours for engines and  $\alpha_{mn} = 925$  hours for the landing gear. The value of  $M_p=10000$ ,  $M_{pd}=M_p$ ,  $M_b = 5.500 * M_p$ ,  $\beta = 0.260$ ,  $k = 2.00$  [22].

Tables 1.1- 1.4 are obtained using these data for both the optimization problems formulated in Section 7, with  $R_i = 0.995, i = 1, 2, \dots, 4$

**Table 1.1:** Values of Multi-objective functions and T (in minutes) for  $\theta = 1.0$  with different weights

w1	w2	$C_{TM} (T1)$	RUL1	$C_{IM} (T2)$	RUL2	T1	T2
1/2	1/2	22309.9	123841	7091.89	123841	4033.45	4033.45
1/3	2/3	22309.9	123841	7729.81	124301	4033.45	3894.98
1/4	3/4	22309.9	123841	8005.13	124416	4033.45	3839.89
2/3	1/3	22309.9	123841	7091.89	123841	4033.45	4033.45
3/4	1/4	22309.9	123841	7091.89	123841	4033.45	4033.45

**Table 1.2:** Values of Multi-objective functions and T (in minutes) for  $\theta = 1.182$  with different weights

w1	w2	$C_{TM} (T1)$	RUL1	$C_{IM} (T2)$	RUL2	T1	T2
1/2	1/2	12647.6	121420	4352.71	123822	4832.62	3917
1/3	2/3	15018.8	123150	5135.21	124046	4085.97	3817.32
1/4	3/4	16155.5	123620	5250.38	124094	3977.96	3781.49
2/3	1/3	12082.4	120810	4360.13	123125	4426.73	4091.04
3/4	1/4	12082.4	120810	3998.01	122226	4426.73	4242.69

**Table 1.3:** Values of Multi-objective functions and T (in minutes) for  $\theta = 2.182$  with different weights

w1	w2	$C_{TM} (T1)$	RUL1	$C_{IM} (T2)$	RUL2	T1	T2
1/2	1/2	4207.26	121771	2847.66	123458	4207.26	3818.37
1/3	2/3	11723.3	122956	3755.75	123547	3986.95	3755.75
1/4	3/4	12465.1	123263	3012.93	123566	3899.44	3733.48
2/3	1/3	8159.79	118946	2652.79	123170	4546.32	3929
3/4	1/4	6953.79	115971	2499.36	122789	4818.2	4025.64

**Table 1.4:** Values of Multi-objective functions and T (in minutes) for  $\theta = 3.182$  with different weights

w1	w2	$C_{TM} (T1)$	RUL1	$C_{IM} (T2)$	RUL2	T1	T2
1/2	1/2	9877.32	121727	2649.01	123346	3814.62	4192.79
1/3	2/3	11438.3	122870	2769.67	123436	3976.54	3751.94
1/4	3/4	12150.9	123165	2814.46	123165	3890.79	3729.71
2/3	1/3	7992.92	118990	2453.59	123057	4526.04	3925.37
3/4	1/4	6821.02	116098	2299.76	122675	4793.23	4022.06

Table 1.1- Table 1.4 gives optimal cost and optimal residual useful life for the two models using different weight combinations. It is observed that the integrated model in almost all the cases yields lower cost and higher RUL with smaller periodic inspection time. Table 1.1 shows that for IM the minimum cost is obtained when  $w1 = 1/4$  and  $w2 = 3/4$  and the optimal periodic inspection time is  $T2 = 3839.89$  implying that four-phase aircraft flight needs to be send for maintenance after every 21 cycles. Similar interpretation holds for data depicted in Table- 1.2 to Table 1.4

## 9. CONCLUSION

In this paper predictive maintenance framework is proposed for a phased mission system. The multi-objective problem is used when weighted sum of expected maintenance cost and mean residual life function of the PMS is minimized subject to the constraints that the reliability of each phase doesn't exceed the pre-specified values. The decision variable is the length of the periodic interval. The optimal solution obtained using IM model is compared with traditional model (TM). For illustrative purpose aircraft flight PMS composed of four phases, namely; taxiing, take-off, cruising, and landing is used with dependency between components of each phase modelled using Gumbel-Haugaar d copula. The cumulative exposure model is used to determine the reliability of the PMS. It is found that the integrated model yields lower cost and higher RUL with smaller periodic inspection time. Thus, the use of predictive tools with

periodic maintenance reduces overall equipment maintenance costs with higher mean residual life.

#### ACKNOWLEDGEMENT

This research work is financially supported by University of Delhi, Delhi-7, INDIA.

#### DECLARATION OF CONFLICT INTEREST

The authors have declared that no conflict of interests exist.

#### REFERENCES

- [1] Liu, Y., Huang, H. Z. (2010). Optimal selective maintenance strategy for multi-state systems under imperfect maintenance. *IEEE Transactions on Reliability*, 59(2), 356-367.
- [2] Pandey, M., Zuo, M. J., and Moghaddass, R. (2013). Selective maintenance modeling for a multistate system with multistate components under imperfect maintenance. *Iie Transactions*, 45(11), 1221-1234.
- [3] Pandey, M., Zuo, M. J., Moghaddass, R., and Tiwari, M. K. (2013). Selective maintenance for binary systems under imperfect repair. *Reliability Engineering & System Safety*, 113, 42-51.
- [4] Guo, C., Wang, W., Guo, B., and Si, X. (2013). A maintenance optimization model for mission-oriented systems based on Wiener degradation. *Reliability Engineering & System Safety*, 111, 183-194.
- [5] Zhu, H., Liu, F., Shao, X., Liu, Q., and Deng, Y. (2011). A costbased selective maintenance decisionmaking method for machining line. *Quality and Reliability Engineering International*, 27(2), 191-201, 2011.
- [6] Wang, H. (2002) A survey of maintenance policies of deteriorating systems. *European journal of operational research*, 139(3), 469-489.
- [7] Dekker, R., Wildeman, R. E., and Van der Duyn Schouten, F. A. (1997). A review of multi-component maintenance models with economic dependence. *Mathematical methods of operations research*, 45(3), 411-435.
- [8] Tsai, Y. T., Wang, K. S., and Tsai, L. C. (2004). A study of availability-centered preventive maintenance for multi-component systems. *Reliability Engineering & System Safety*, 84(3), 261-270.
- [9] Mobley, R. K. (2002). *An introduction to predictive maintenance*. Elsevier.
- [10] Nakagawa, T. (1986). Periodic and sequential preventive maintenance policies. *Journal of Applied probability*, 23(2), 536-542.
- [11] Dhillon, B. S. (2006). *Maintainability, maintenance, and reliability for engineers*. CRC press.
- [12] Barone, G., and Frangopol, D. M. (2014). Reliability, risk and lifetime distributions as performance indicators for life-cycle maintenance of deteriorating structures. *Reliability Engineering & System Safety*, 123, 21-37.
- [13] Nouralfath, M., Châtelet, E., and Nahas, N. (2012). Joint redundancy and imperfect preventive maintenance optimization for series-parallel multi-state degraded systems. *Reliability Engineering & System Safety*, 103, 51-60.
- [14] Tapiero, C. S. (1986). Continuous quality production and machine maintenance. *Naval Research Logistics Quarterly*, 33(3), 489-499.
- [15] Pate-Cornell, M. E., Lee, H. L., and Tagaras, G. (1987). Warnings of malfunction: the decision to inspect and maintain production processes on schedule or on demand. *Management Science*, 33(10), 1277-1290.
- [16] TAGARAS, G. E. (1988). An Integrated Cost Model for the Joint Optimization of Process Control and Maintenance, *Journal of the Operational Research Society*, 39, 8, 757-767.
- [17] Rahim, M. A. (1994). Joint determination of production quantity, inspection schedule, and control chart design. *IIE transactions*, 26(6), 2-11.
- [18] Ben-Daya, M., and Makhdom, M. (1998). Integrated production and quality model under various preventive maintenance policies. *Journal of the operational Research Society*, 49(8), 840-853.

- [19] Ben-Daya, M. (1999). Integrated production maintenance and quality model for imperfect processes. *IIE transactions*, 31(6), 491-501.
- [20] Cassady, C. R., Bowden, R. O., Liew, L., and Pohl, E. A. (2000). Combining preventive maintenance and statistical process control: a preliminary investigation. *IIE transactions*, 32, 471-478.
- [21] Gong, L., Jwo, W., and Tang, K. (1997). Using on-line sensors in statistical process control. *Management Science*, 43(7), 1017-1028.
- [22] McKone, K. E., & Weiss, E. N. (2002). Guide lines for implementing predictive maintenance. *Production and Operations Management*, 11(2), 109-124.
- [23] Qi, X., Guo, B., and Liu, Y. (2012). Simulation based reliability analysis of phased-mission system considering maintenance in system-level. In *2012 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering* (pp. 1435-1438). IEEE.
- [24] Moubrey, J. (2001). *Reliability-centered maintenance*. Industrial Press Inc.
- [25] Nakagawa, T. (2005). *Maintenance theory of reliability*. Springer Science & Business Media.
- [26] Somani, AK., Ritcey, JA., and SHL, AU. (1992). Computationally-efficient phased-mission reliability analysis for systems with variable configurations *IEEE Transactions on Reliability* 1992; 41: 504-511.
- [27] Xing, L. (2007). Reliability Evaluation of Phased-Mission Systems with Imperfect Fault Coverage and Common-Cause Failures, *IEEE Trans. Reliability* 2007; 56: 58-68.
- [28] Nelsen, R. B. (2006). *An Introduction to Copulas*, 2nd edition. Springer Science + Business Media, Inc., New York.
- [29] Nelson W. (1990). *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*, John Wiley & Sons, New York, 1990.
- [30] Srivastava, P. W., and Rani, S. (2022). COPULA-BASED APPROACH TO RELIABILITY ANALYSIS OF PHASED-MISSION SYSTEMS. *International Journal of Reliability, Risk and Safety: Theory and Application*.
- [31] Banjevic D. (2009). Remaining useful life in theory and practice. *Metrika*. Mar; 69:337-49.
- [32] [https://help.reliasoft.com/reference/system\\_analysis/sa/reliability\\_phase\\_diagrams\\_r\\_pds.html](https://help.reliasoft.com/reference/system_analysis/sa/reliability_phase_diagrams_r_pds.html)

# PROFIT ANALYSIS OF REPAIRABLE JUICE PLANT

Rahul<sup>1</sup>, Mohit Yadav<sup>2\*</sup> and Hemant Kumar<sup>3</sup>

<sup>1,2</sup>Department of Mathematics, University Institute of Sciences  
Chandigarh University, Mohali, Chandigarh

<sup>3</sup>SOET, Raffles University, Neemrana, Rajasthan

\*Corresponding Author

[22msm40223@cuchd.in](mailto:22msm40223@cuchd.in), [mohit.e15793@cumail.in](mailto:mohit.e15793@cumail.in),  
[hemantkumar@rafflesuniversity.edu.in](mailto:hemantkumar@rafflesuniversity.edu.in)

## Abstract

*Juice is a non-fermented beverage that is obtained by squeezing fruits to increase immunity. 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 a juice plant are discussed. The juice plant has 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 reliability measures.*

**Keywords:** Reliability measures, juice plant, evaporation and pasteurization.

## I. Introduction

Manufacturers must constantly innovate their products in order to keep up with the rising demand for their products, which is made feasible by optimizing their manufacturing processes. The MTSE, availability and profitability of a juice factory with priority in repair are discussed in this study by utilizing the regenerating point graphical technique under specific circumstances.

Barlow *et al.* [2] investigated the reliability theory with redundancy and system availability while taking into account the significance of individual system components. The reliability study of a single unit system with non-repairable spare units and its optimization applications was covered by Nakagawa and Osaki [12]. Balagurusamy [1] described the terms related to the system's meantime, failure, repair, redundancy, maintainability, availability, etc. Tuteja and Malik [16] examined the dependability of two distinct single-unit models with three operating modes and various repair procedures applied to the repairman. Malik [11] examined a single-unit system with a server under inspection. Pawar *et al.* [13] threw light on an operating system under different climates having repair at varying levels of damages subject to inspection.

Gupta [4] talked about employing a base state to analyze a single-unit system. The reliability analysis of a one-unit system with finite vacations was examined by Liu and Liu [10]. The dependability metrics of a repairable stochastic model on the production of printed circuit boards were given by Kumar and Batra [9]. Chaudhary *et al.* [3] studied the valuable parameters for the nature of the distillery system having three distinct units and a single server facility using the



regenerative point graphical technique. Kumar *et al.* [8] threw light on the preventive maintenance of a sustainable one unit system under degradation facilities. Sharma and Goel [15] described the nature of whole-grain flour mills having two units using base state and regenerative point techniques. Kumar and Saini [5] described the fault detection concept in stochastic computing device under repair and replacement by an expert repairman. A redundant system with a first come, first served repair policy was examined by Kumar *et al.* [7] under different weather. Sengar and Mangey [14] analyzed the reliability measures of a complex manufacturing system with an inspection facility. Kumar *et al.* [6] analyzed the reliability and performance of two unit system under inspection facility.

## II. System Assumptions

To describe the juice plant, there are following assumptions

- The juice plant consists of three distinct units  $A$ ,  $B$  and  $C$ .
- 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.
- 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.

## III. System Notations

To explain the juice plant, there are following notations

$i \xrightarrow{Sr} j$	$r^{\text{th}}$ directed simple path from state ' $i$ ' to state ' $j$ ' where ' $r$ ' takes the positive integral values for different directions from state ' $i$ ' to state ' $j$ '.
$\xi \xrightarrow{sf} i$	A directed simple failure free path from state $\xi$ to state ' $i$ '.
$m - \text{cycle}$	A circuit (may be formed through regenerative or non regenerative / failed state) whose terminals are at the regenerative state ' $m$ '.
$\overline{m - \text{cycle}}$	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.
$\overline{U_{k,k}}$	The <u>probability</u> factor of state ' $k$ ' reachable from the terminal state ' $k$ ' of $k$ cycle.
$\mu_i$	Mean sojourn time spent in the state ' $i$ ' before visiting any other states.
$\mu'_i$	Total unconditional time spent before transiting to any other regenerative state while the system entered regenerative state ' $i$ ' at $t=0$ .
$\eta_i$	Expected waiting time spent while doing a job given that the system entered to the regenerative state ' $i$ ' at $t=0$ .
$A/\overline{A}/a$	First unit is in the operative state/reduced state/failed state.
$B/\overline{B}/b$	Second unit is in the operative state/reduced state/failed state.
$C/\overline{C}/c$	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.
$\lambda_4, \lambda_5$	Fixed complete failure rate of the unit A/B respectively.

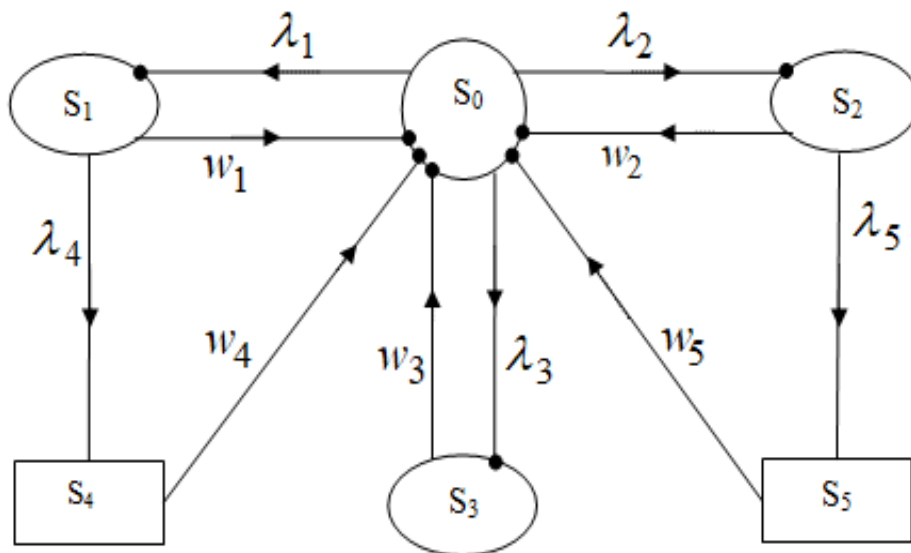
- $w_1, w_2, w_3$  Fixed repair rate of the unit A/B/C after partial failure respectively.  
 $w_4, w_5$  Fixed repair rate of unit A/B after the complete failure respectively.  
 $\circ$   $\ominus$   $\square$  Upstate/ reduced state/ failed state.

#### IV. Circuits Descriptions

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)	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)



**Figure 1** State Transition Diagram

where,  $S_0 = ABC$ ,  $S_1 = \bar{A}BC$ ,  $S_2 = A\bar{B}C$   
 $S_3 = ABC\bar{C}$ ,  $S_4 = aBC$ ,  $S_5 = AbC$

### V. Transition Probabilities

There are following transition probabilities

$$\begin{aligned}
 p_{0,1} &= \lambda_1 / (\lambda_1 + \lambda_2 + \lambda_3), p_{0,2} = \lambda_2 / (\lambda_1 + \lambda_2 + \lambda_3), p_{0,3} = \lambda_3 / (\lambda_1 + \lambda_2 + \lambda_3) \\
 p_{1,0} &= w_1 / (w_1 + \lambda_4), p_{1,4} = \lambda_4 / (w_1 + \lambda_4), p_{2,0} = w_2 / (w_2 + \lambda_5) \\
 p_{2,5} &= \lambda_5 / (w_2 + \lambda_5), p_{3,0} = p_{4,0} = p_{5,0} = 1
 \end{aligned} \tag{1}$$

It has been conclusively established that

$$\begin{aligned}
 p_{01} + p_{03} &= 1, p_{10} + p_{12} + p_{14} = 1, p_{21} + p_{27} = 1, p_{31} + p_{38} = 1 \\
 p_{41} + p_{45} &= 1, p_{56} = p_{76} = p_{86} = 1, p_{31} + p_{31.8(65)^n} = 1 \\
 p_{10} + p_{12} + p_{11.4} + p_{11.4(56)^n} &= 1, p_{21} + p_{21.7(65)^n} = 1
 \end{aligned} \tag{2}$$

### VI. Mean Sojourn Time

Time taken by a system in a particular state becomes,  $\mu_i = \sum_j m_{i,j} = \int_0^{\infty} P(T > t) dt$ .

$$\begin{aligned}
 \mu_0 &= 1 / (\lambda_1 + \lambda_2 + \lambda_3) \\
 \mu_1 &= 1 / (w_1 + \lambda_4), \mu_2 = 1 / (w_2 + \lambda_5) \\
 \mu_3(t) &= 1 / (w_3), \mu_4 = 1 / (w_4), \mu_5 = 1 / (w_5)
 \end{aligned} \tag{3}$$

### 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

$$\begin{aligned}
 U_{0,0} &= (0,1,0) + (0,2,0) + (0,3,0) = 1, U_{0,1} = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3}, \\
 U_{0,2} &= \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_3}, U_{0,3} = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} \\
 U_{0,4} &= \frac{\lambda_1 \lambda_4}{(\lambda_1 + \lambda_2 + \lambda_3)(w_1 + \lambda_4)}, U_{0,5} = \frac{\lambda_2 \lambda_5}{(\lambda_1 + \lambda_2 + \lambda_3)(w_2 + \lambda_5)}
 \end{aligned}$$

#### 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  $\xi=0$ ) then MTSF becomes

$$T_0 = \left[ \sum_{i=0}^3 Sr \left\{ \frac{\left\{ pr(0 \xrightarrow{Sr(sff)} \rightarrow i) \right\} \cdot \mu_i}{\prod_{k_1 \neq 0} \left\{ 1 - V_{k_1 k_1} \right\}} \right\} \right] \div \left[ 1 - \sum Sr \left\{ \frac{\left\{ pr(0 \xrightarrow{Sr(sff)} \rightarrow 0) \right\}}{\prod_{k_2 \neq 0} \left\{ 1 - V_{k_2 k_2} \right\}} \right\} \right]$$

$$T_0 = \frac{\left[ \begin{array}{l} (w_1 + \lambda_4)(w_2 + \lambda_5)(w_3 + \lambda_3) \\ + w_3[\lambda_1(w_2 + \lambda_5) + \lambda_2(w_1 + \lambda_4)] \end{array} \right]}{\left[ \begin{array}{l} w_3[(\lambda_1 + \lambda_2 + \lambda_3)(w_1 + \lambda_4)(w_2 + \lambda_5) \\ - \lambda_1 w_1(w_2 + \lambda_5) - \lambda_2 w_2(w_1 + \lambda_4)] \end{array} \right]} \quad (4)$$

## II. Availability of the system

The system is available for use at regenerative states  $j=0, 1, 2, 3$  with  $\xi=0$  then the availability of system is defined as

$$A_0 = \left[ \begin{array}{l} \sum_{j=0}^3 Sr \left\{ \frac{\{pr(0 \xrightarrow{Sr} j)\} \cdot f_j \cdot \mu_j}{\prod_{k_1 \neq 0} \left\{ 1 - V_{\frac{1}{k_1 k_1}} \right\}} \right\} \end{array} \right] \div \left[ \begin{array}{l} \sum_{i=0}^5 Sr \left\{ \frac{\{pr(0 \xrightarrow{Sr} i)\} \cdot \mu'_i}{\prod_{k_2 \neq 0} \left\{ 1 - V_{\frac{1}{k_2 k_2}} \right\}} \right\} \end{array} \right]$$

$$A_0 = \frac{\left[ \begin{array}{l} w_4 w_5 [(w_1 + \lambda_4)(w_2 + \lambda_5)(w_3 + \lambda_3) \\ + w_3 \{\lambda_1(w_2 + \lambda_5) + \lambda_2(w_1 + \lambda_4)\}] \end{array} \right]}{\left[ \begin{array}{l} w_4 w_5 (w_1 + \lambda_4)(w_2 + \lambda_5)(w_3 + \lambda_3) \\ + \lambda_1 w_3 w_5 (w_2 + \lambda_5)(w_4 + \lambda_4) \\ + \lambda_2 w_3 w_4 (w_1 + \lambda_4)(w_5 + \lambda_5) \end{array} \right]} \quad (5)$$

## III. Busy Period of the Technician

The Technician is busy due to repair of the failed unit at regenerative states  $j=1, 2, 3, 4, 5$  with  $\xi = 0$  then the fraction of time for which the server remains busy is defined as

$$B_0 = \left[ \begin{array}{l} \sum_{j=1}^5 Sr \left\{ \frac{\{pr(0 \xrightarrow{Sr} j)\} \cdot \eta_j}{\prod_{k_1 \neq 0} \left\{ 1 - V_{\frac{1}{k_1 k_1}} \right\}} \right\} \end{array} \right] \div \left[ \begin{array}{l} \sum_{i=0}^5 Sr \left\{ \frac{\{pr(0 \xrightarrow{Sr} i)\} \cdot \mu'_i}{\prod_{k_2 \neq 0} \left\{ 1 - V_{\frac{1}{k_2 k_2}} \right\}} \right\} \end{array} \right]$$

$$B_0 = \frac{\left[ \begin{array}{l} w_4 w_5 \lambda_3 (w_1 + \lambda_4)(w_2 + \lambda_5) \\ + \lambda_1 w_3 w_5 (w_2 + \lambda_5)(w_4 + \lambda_4) \\ + \lambda_2 w_3 w_4 (w_1 + \lambda_4)(w_5 + \lambda_5) \end{array} \right]}{\left[ \begin{array}{l} w_4 w_5 (w_1 + \lambda_4)(w_2 + \lambda_5)(w_3 + \lambda_3) \\ + \lambda_1 w_3 w_5 (w_2 + \lambda_5)(w_4 + \lambda_4) \\ + \lambda_2 w_3 w_4 (w_1 + \lambda_4)(w_5 + \lambda_5) \end{array} \right]} \quad (6)$$

#### IV. Estimated number of visits made by the Technician

The technician visits at regenerative states  $j = 1, 2, 3$  with  $\xi=0$  then the number of visits by the repairman is defined as

$$V_0 = \left[ \sum_{j=1}^3 Sr \left\{ \frac{\{pr(0 \xrightarrow{Sr} j)\}}{\prod_{k_1 \neq 0} \left\{ 1 - V_{\frac{k_1}{k_1}} \right\}} \right\} \right] \div \left[ \sum_{i=0}^5 Sr \left\{ \frac{\{pr(0 \xrightarrow{Sr} i)\} \cdot \mu_i'}{\prod_{k_2 \neq 0} \left\{ 1 - V_{\frac{k_2}{k_2}} \right\}} \right\} \right]$$

$$V_0 = \frac{\left[ \begin{array}{l} w_4 w_5 \lambda_3 (w_1 + \lambda_4)(w_2 + \lambda_5) \\ + \lambda_1 w_3 w_4 w_5 (w_2 + \lambda_5) \\ + \lambda_2 w_3 w_4 w_5 (w_1 + \lambda_4) \end{array} \right]}{\left[ \begin{array}{l} w_4 w_5 (w_1 + \lambda_4)(w_2 + \lambda_5)(w_3 + \lambda_3) \\ + \lambda_1 w_3 w_5 (w_2 + \lambda_5)(w_4 + \lambda_4) \\ + \lambda_2 w_3 w_4 (w_1 + \lambda_4)(w_5 + \lambda_5) \end{array} \right]} \quad (7)$$

#### V. Profit Analysis

The profit function may be used to do a profit analysis of the system and it is given by

$$P = E_0 A_0 - E_1 B_0 - E_2 V_0 \quad (8)$$

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

Tables 2, 3 and 4 described the nature of mean time to system failure, availability and profit values

**Table 2:** MTSF vs. Repair Rate ( $w_2$ )

$w_2$ ↓	$\lambda_1=0.3, \lambda_2=0.4$ $\lambda_3=0.25, \lambda_4=0.35$ $\lambda_5=0.5, w_1=0.4$ $w_3=0.5, w_4=0.5$ $w_5=0.6$	$\lambda_1=0.4$	$\lambda_2=0.5$	$\lambda_3=0.3$
0.4	3.628692	3.333333	3.090278	3.037974
0.45	3.675035	3.368794	3.132184	3.062553
0.5	3.720609	3.403509	3.173516	3.086409
0.55	3.765432	3.4375	3.214286	3.109568
0.6	3.809524	3.47079	3.254505	3.197278
0.65	3.852901	3.503401	3.294183	3.153901
0.7	3.895582	3.535354	3.333333	3.175126
0.75	3.937583	3.566667	3.371965	3.195751
0.8	3.97892	3.59736	3.410088	3.2158
0.85	4.019608	3.627451	3.447712	3.235294

**Table 3:** Availability vs. Repair Rate ( $w_2$ )

$w_2$ ↓ ▼	$\lambda_1=0.3, \lambda_2=0.4$ $\lambda_3=0.25, \lambda_4=0.35$ $\lambda_5=0.5, w_1=0.4$ $w_3=0.5, w_4=0.5$ $w_5=0.6$	$\lambda_1=0.4$	$\lambda_2=0.5$	$\lambda_3=0.3$
0.4	0.623324	0.604782	0.542904	0.58624
0.45	0.628307	0.609813	0.547959	0.591319
0.5	0.633159	0.614717	0.552904	0.596275
0.55	0.637887	0.619499	0.557741	0.601111
0.6	0.642494	0.624164	0.562476	0.605834
0.65	0.646985	0.628716	0.56711	0.610447
0.7	0.651365	0.633159	0.571646	0.614953
0.75	0.655637	0.637497	0.576089	0.619357
0.8	0.659806	0.641734	0.58044	0.623662
0.85	0.663876	0.645873	0.584703	0.62787

**Table 4:** Profit vs. Repair Rate ( $w_2$ )

$w_2$ ↓ ▼	$\lambda_1=0.3, \lambda_2=0.4$ $\lambda_3=0.25, \lambda_4=0.35$ $\lambda_5=0.5, w_1=0.4$ $w_3=0.5, w_4=0.5$ $w_5=0.6$	$\lambda_1=0.4$	$\lambda_2=0.5$	$\lambda_3=0.3$
0.4	2438.338	2386.076	2019.52	2333.814
0.45	2467.262	2415.876	2049.467	2364.49
0.5	2495.431	2444.927	2078.759	2394.423
0.55	2522.874	2473.257	2107.417	2423.64
0.6	2549.618	2500.892	2135.461	2452.166
0.65	2575.691	2527.857	2162.912	2480.023
0.7	2601.117	2554.178	2189.787	2507.239
0.75	2625.919	2579.875	2216.104	2533.831
0.8	2650.121	2604.972	2241.881	2559.823
0.85	2673.744	2629.49	2267.135	2584.836

of the juice plant having an increasing trend corresponding to repair rate ( $w_2$ ). In these tables, the values of parameters  $\lambda_1=0.3, \lambda_2=0.4, \lambda_3=0.25, \lambda_4=0.35, \lambda_5=0.5, w_1=0.4, w_3=0.5, w_4=0.5, w_5=0.6$  respectively taking as constant for the simplicity. When  $\lambda_1=0.3$  changing into  $\lambda_1=0.4$ ;  $\lambda_2=0.4$  changing into  $\lambda_2=0.5$  and  $\lambda_3=0.25$  changing into  $\lambda_3=0.3$  then MTSF, availability and profit values have decreasing trends.

### IX. 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, system's 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 analyzed 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.

### References

- [1] Balagurusamy, E. (1984). *Reliability Engineering*. Tata McGraw-Hill Education.
- [2] Barlow, R. E., Proschan, F. and Hunter, L. C. (1965). *Mathematical Theory of Reliability* John Wiley and Sons Inc. *New York*, 4, 927-929.
- [3] Chaudhary, N., Goel, P. and Kumar, S. (2013). Developing the reliability model for availability and behaviour analysis of a distillery using Regenerative Point Graphical Technique. *International Journal of Informative and Futuristic Research*, 1(4), 26-40.
- [4] Gupta, V. K. (2011). Analysis of a single unit system using a base state. *Aryabhatta Journal of Mathematics & Informatics*, 3(1), 59-66.
- [5] Kumar, A. and Saini, M. (2018). Stochastic modeling and cost-benefit analysis of computing device with fault detection subject to expert repair facility. *International Journal of Information Technology*, 10, 391-401.
- [6] 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.
- [7] Kumar, A., Pawar, D. and Malik, S. C. (2020). Reliability analysis of a redundant System with 'FCFS' repair policy subject to weather conditions. *International Journal of Advanced Science and Technology*, 29(3), 7568-7578.
- [8] Kumar, J., Kadyan, M. S., Malik, S. C. and Jindal, C. (2014). Reliability measures of a single-unit system under preventive maintenance and degradation with arbitrary distributions of random variables. *Journal of Reliability and Statistical Studies*, 77-88.
- [9] Kumar, R. and Batra, S. (2012). Economic and reliability analysis of a stochastic model on printed circuit boards manufacturing system considering two types of repair facilities. *International Journal of Electrical Electronics and Telecommunication Engineering*, 43(10), pp. 432-435.
- [10] Liu, R. and Liu, Z. (2011). Reliability analysis of a one-unit system with finite vacations. In *MSIE*, 248-252.
- [11] Malik, S. C., Chand, P. and Singh, J. (2008). Stochastic analysis of an operating system with two types of inspection subject to degradation. *Journal of Applied Probability and Statistics*, 3(2), 227-241.
- [12] Nakagawa, T. and Osaki, S. (1976). Reliability analysis of a one unit system with unrepairable spare units and its optimization applications. *Journal of the Operational Research society*, 27(1), 101-110.
- [13] Pawar, D., Malik, S. C. and Bahl, S. (2010). Steady state analysis of an operating system with repair at different levels of damages subject to inspection and weather conditions. *International Journal of Agriculture and Statistical Sciences*, 6(1), 225-234.
- [14] Sengar S. and Mangey R. (2022). Reliability and performance analysis of a complex manufacturing system with inspection facility using copula methodology. *Reliability Theory & Applications*, 17(71): 494-508.
- [15] Sharma S. and Goel, P. (2015). Behavioral Analysis of Whole Grain Flour Mill Using RPGT. *International Journal of Engineering Technology, Management and Applied Sciences*, 3, 194-201.
- [16] Tuteja, R. K. and Malik, S. C. (1992). Reliability and profit analysis of two single unit models with three modes and different repair policies of repairmen who appear and disappear randomly. *Microelectronics Reliability*, 32(3), 351-356.

# A LITERATURE REVIEW ON DEVELOPMENT OF QUEUEING NETWORKS

V. Narmadha<sup>1</sup>, P. Rajendran<sup>2,\*</sup>



<sup>1,\*</sup> Department of Mathematics, School of Advanced Sciences, Vellore Institute of Technology,  
Vellore, Tamil Nadu 632014, India.

<sup>1</sup>narmadha.v@vit.ac.in, <sup>2,\*</sup>prajendran@vit.ac.in

## Abstract

*This study conducts a quantitative research survey on the development of queueing networks over years. Development is a process of gradual change that takes place over many years, during which a theory slowly progress and attain a good state. Queueing theory has been through many developments which made its existence inevitable in every field. Queueing networks can be considered as a collection of nodes, where each node stands for a service facility. It has been proved to be a powerful and versatile tool for modelling facilities in manufacturing units and telecommunication networks. This paper presents the development in Queueing networks and its types over years. This paper's main objective is to give all the analysts and researchers the knowledge about the evolution that happened in Queueing networks over years.*

**Keywords:** Quazi-reversible Queueing networks (QRQN), Stationary distribution, Automated Manufacturing Systems(AMS), Recurrent neural networks(RNN).

## 1. Introduction

The goal of queueing theory is to create efficient, cost-effective systems that can serve customers promptly and effectively. Agner Krarup Erlang, a Danish mathematician, statistician and engineer conducted an analysis of the Copenhagen telephone exchange in the early 1900s, which is where queueing theory first emerged. His work paved the way for the development of telephone network assessment and the Erlang concept of effective networks. The notion of queues is used to locate and eliminate bottlenecks in a process. Owing to the fact that queueing models only need a little amount of data and are easy to implement, it is a very effective and useful technique. They can be used to instantly examine and compare different service delivery solutions because of their simplicity and speed. Queueing models can be effective in obtaining insights on the degree of specialisation or flexibility for the utilization of resources in an organisation, which goes beyond the most fundamental task of predicting how much resource is required to accomplish a specific service level. Because of this, there are several prospects for its implementation in several industries.



## 2. Queueing networks (QN) and its types

In simple words QN are nothing but the jobs moving between interconnected queues in a continuous flow. In some cases, it may be simpler to describe a complex service environment as a queueing network in order to more accurately represent how the service is actually delivered. Telecommunication networks, machine shop problems and computer system are some of the instances of QN.

### 2.1. Types of QN

Open queueing networks	Closed queueing networks	Mixed queueing networks
<ul style="list-style-type: none"><li>•The network must be open for each job class if it has multiple job classes.</li></ul>	<ul style="list-style-type: none"><li>•The network must be closed for each job class if it has several job classes.</li></ul>	<ul style="list-style-type: none"><li>•The network contains a variety of job classes, some of which are open while the others are closed.</li></ul>

Based on the capacity of the Queue's component there are two types of QN

<b>Blocking</b> <ul style="list-style-type: none"><li>•This happens in a network, when there are one or perhaps more queues with finite capacities.</li></ul>
<b>No blocking</b> <ul style="list-style-type: none"><li>•This happens in a network, when there are many queues of infinite capacity.</li></ul>

Some of the blocking models in QN are listed below

- **Rejection blocking**  
The blocked jobs will be forced to leave the system and it is only applicable for open networks.
- **Transfer blocking**  
The blocked job will wait at source line  $K_i$  until the job is accepted at destined line  $K_j$ .
- **Repetitive service blocking**  
The blocked job will again dwell in  $K_i$  for another service and the process will be repeated until the job can move out of  $K_i$ .
- **Blocking before service**  
At  $K_i$ , the service starts for the job only when the destined line  $K_j$  is free and is ready to accept the jobs from  $K_i$ .

## 3. Literature Review

The queueing theory has undergone numerous developments, which draws scholars to use it in the best way possible. The Erlang formula, which became a cornerstone of contemporary

telecommunication network studies, was developed as a result of Erlang's 1920 publication of telephone waiting times [1], which examined the use of local, exchange, and trunk telephone lines in a close knit community to interpret the empirical needs of an efficient network. One model combines and extends a number of distinct results from open, closed, and mixed network of queues with various kinds of clients [2]. To determine the generic model's equilibrium state probability, this study integrates past results from networks of queues with different customer classes across a range of service areas and a large spectrum of estimated time distributions. QN with individuals from various groups [3] are viewed as a generalisation of mechanisms that enable customers in a particular queue to be heterogeneous. The robustness of QN where customers may be of various types is the subject of this study. QN [4] examines how networks behave when they are in an equilibrium, and in some instances it is demonstrated that the status of one queue does not rely on the condition of the overall network. Since the processes in this study are irreversible, the restrictions placed on the potential customer paths by earlier writers are further loosened. The link between regional balance and product form in QN [5] seems to provide insight into why some domains yield product form answers to problems for queues and networks using nonexponential service domains compared to others. A queueing regimen satisfies station balance if the pace at which customers receive service at every position in the queue is proportionate to the possibility that a customer also could appear at that position. The paper's conclusions extrapolate past research on local balance to any stochastic, differentiable service distribution endeavours that would result in regional balance and product form characterization. In steady state, the state's distribution has the product form, and interconnections of markov chains and QRQN [6] demonstrate that a network created by joining queues, each of which is QR when taken separately, is also QR. An aspect of convergence of the source performance function component in a closed Jackson network(JN) is enabled by the sample's convergence property in closed Jackson queuing networks [7] research. This finding offers some fresh perspectives on QN theory that might not be found in the well-known product form solution. A category of QN with rejection inhibition has a product form steady state flow pattern, and the overall population is insensitive, according to Exact Solutions for Open, Closed, and Mixed QN with Rejection Blocking [8]. The outcomes are startlingly comparable to those for conventional (non-blocking) networks. The stability of open QN [9] only requires the additional assumption that service time ranges have finite first moments in order to prove stability for the open network. It is permitted for the inter - arrival time distribution to have an infinite first instant. The results are expanded to include multi-server nodes, non-Markovian routing, and Markov modulated arrivals. Recent developments in QN: a survey with applications to AMS [10] highlighted the use of QN models for the performance assessment of AMSs, separately addressing the problems of computing larger parts of performance metrics, blocking events, and analysing an open network of AMS model's multiqueues.

Using the Right-continuous Markov processes(MP) with values theory which provides a single method for finding both optimal and suboptimal feedback control laws in some QN [11]. This method can be used with QN made up of machines and buffers. The results of single-server QN under optimal control [12] provide effective ways to compute the indices. The greatest remaining index approach is presented in its general form in this publication. We can now locate every index for our universal single server QN model. Unique features of the optimum static routing solution in open BCMP QN [13] determine the relay nodes of the underlying optimal policy and demonstrate that they may not be strange, but the overall determination of the usage of each repair facility is unique. We also take into account a policy that is individually optimal and routes jobs. If each job is aware of the average time delay for each path, it can feel as though its own anticipated response time is decreased The iterative process for a class of Batch-movement QN, which is a natural generalisation of the mean-value analysis of JN, was illustrated in [14]. The recurrence relations used in this approach can be easily extended to the generic group of product-

form batch-movement QN/petrinets with equilibrium probability distribution. Using the Z-transform, computing the normalisation constants in QN is made easier [15].

The computation required by the proposed strategy in [16] is relatively simpler than that in Gordon's paper in the scenarios of networks with numerous repair queues or with a single server queue and equal traffic levels. Numerous forms of monotone routing strategies for limited capacity QN have successfully used perfect simulation of index-based routing QN. This research could be expanded to batch arrival or batch services as well as monotone network events more broadly. Through the use of current approximate inference methods from graphical models, probabilistic inference in QN [17] provided a new family of tools for studying queueing models. Networks and queueing systems, models, and applications in [18] uses traditional Markovian systems in queueing systems that combines individual service, an exponential service times and a Poisson arrival procedure. Given that academics are interested in using queueing systems and QN to modelling human performance, it provides an architecture known as the Queueing Network-Model Human Processor. In order to represent an industrial system, a four-input, three-stage QN technique was used [19]. This approach computed the best route that results in the shortest reaction times for the delivery of products to the end destination along the three phases of the network. Modeling a supply chain using a QN, a supply chain is shown as a two-input, three-stage QN [20]. The goal of this study is to determine the minimal response time required to deliver products to their destination along the network's three stages. The total number of products that can be distributed with this quickest response period makes up the QN's maximum capacity.

QN and graphical models are combined in the innovative perspective of reasoning and acquiring knowledge in networks of queues [21], which enables the use of Markov chain Monte Carlo. We use actual data from a standard web application to show how successful the sample is. In order to maximise the throughput of single server, generic QN, a multiobjective technique was devised, called throughput maximisation of QN with concurrent reduction of service rates and buffers [22]. It should be investigated further to see whether more optimum conditions in finite QN can be found using this methodology. Stability in constrained network architectures with queueing lags, queue-storage, blocking back, and control [23] has presented numerous techniques for spatial queuing appropriately without using dynamic assignment, hence the strategy is alternative to the methodology used by Bliemer et al (2012). QN with a single shared server: light and high traffic [24] provide a significant two-fold contribution. First, we examine the system under consideration's precise heavy-traffic asymptotics. Second, based on an approximation between the light-traffic and high-traffic limitations, we construct a closedform approximation for the average lag for random loads. The analysis presented in this work can be expanded in a number of ways, for as by considering various server configuring policies or service standards. However, these results are not explored because of compactness. An approximation technique for the assessment of a finite open QN with Transfer blocking and feedback was described in a restricted open QN application to healthcare systems [25]. An unbounded topology network with a focus on a single server finite capacity model based fertility clinic healthcare system is discussed which uses an expansion approach to determine each node's performance measures and throughput. Deadlock in open restricted QNs has been studied in [26]. It has been demonstrated that analysing the corresponding state digraph of a QN is sufficient to identify stalemate. Three deadlocking QN Markov models have been created. The open two-node, multi-server restricted queueing network requires the development of a Markov model with paths between nodes and feedback loops. Modeling urban taxi services with e-hailings: a QN strategy [27] places an emphasis on the macro-interactions between the urban roadway and taxi systems, but it leaves out the intention of changing speed of the individuals and how they react to the taxi prices. Future research will generalise the suggested QN to take into account the complete dynamics of the taxi market and individual behaviour, giving us keen insight into system control. In a single-class open

QN with Markovian routing, infinite waiting space, and the first-come, first-served, A Robust Queueing Network Analyzer Based on Indices of Dispersion(RQNA-IDC) [28] offers practical algorithms to approximate the performance metrics of the steady state. Future research should focus on a number of excellent directions, such as (i) approximations for flows that use multi-dimensional robust models than one dimensional robust models and (ii) expanding RQNA-IDC to additional open QN models, such as models with more than one servers and other service domains.

It has been demonstrated that there is a direct correlation between the architecture of the QN fluid estimation and the typical activation functions and layers of an recurrent neural networks(RNN) in [29]. As far as we are aware, this is the first method that formally unites the vividness of quantitative performance models with the learning potential of machine learning, favourably contributing to the discussion of whether 'AI will be at the centre of performance engineering'. Using stock critical intensities in a QN, handling shared mobility systems [30] investigates a closed JN taking into account nodes for stops and paths. Focusing on the Mean Value Analysis(MVA) approach, a genetic algorithm was created to solve the issue, and an approximation method was offered to determine the crucial parts from the answer. The model can be expanded to take into account static or dynamic equilibrium techniques as part of additional studies. Hospitals can utilise workflow forecasting to manage healthcare systems in practise, as demonstrated by Simulation and betterment of Patients' Workload in cardiac clinics during COVID-19 pandemic using Timed Colored petri nets [31]. This method would be helpful in these trying times because nosocomial transmission puts the health of the personnel and other patients at danger. A new approach to reducing flight delay rates in airports was presented in Reduction of Delay Rate in Open QN[32] in conjunction with deterministic timed petri nets (DTPN) and open QN. The Federal Aviation Administration's performance is evaluated using the flight delay information gathered by the Operations Network (ON). A numerical example is used to demonstrate the significant reduction in the delay rate. A fresh approach to the problem of multistage semi-open queueing networks(SOQN) i.e., A innovative and all-encompassing method for estimating the work departure process parameters from a SOQN is shown with an application in shuttle-based compact storage systems [33], which adds to the body of knowledge in this area. An accurate assessment of the work departure process from the SOQN is very difficult to perform when the work inter-arrival and service times exhibit broad distributions. As a result, it suggest a practical two-moment approximation method in this study.

#### 4. Developmental analyses

It all began with an infinite series to determine whether a call has to be shut out or allowed. Then a number of queues were taken into consideration which paved way for the rise of QN. When same kind of queues are taken into account for research, over time researchers started combining different types of queues which resulted in open, closed and mixed networks with distinct clients. It further paved way for letting customers within a particular queue to be heterogeneous. The next noteworthy development in QN is the product form networks where the state probabilities are given by the products of functions of number of jobs in the queues. This gave rise to Jackson networks which showed that any arbitrary open QN with  $k$  servers that follows an exponentially distributed service time has a product form solution. Following this BCMP network was described by Baskett, Chandy, Muntz and Palacios which is an extension of Jackson networks. Based on the capacity of queue's components it is further classified into Blocking and no blocking QN. Supply chain has been combined with QN to efficiently manage organisations. QN's development has paved a way for its application in various fields which includes healthcare sector, transport system and banking sector etc.

## 5. Applications

QN's development has paved a way for its application in various fields which includes healthcare sector, transport system and banking sector etc.

- QN is applied in a variety of different domains, including computer science, civil engineering, and operations research.
- It is also utilised in computer science to optimise communication network performance and to model the behaviour of computer systems.
- QN is used in civil engineering to optimise traffic flow and model the behaviour of traffic networks.
- On the other hand, QN is applied in operations research to enhance the effectiveness of business processes and optimise resource allocation.

**Declaration of conflicting interest:** The authors declare that there is no conflict of interest.

## References

- [1] Matematisk Tidsskrift, B. (1920). *Telephone waiting times*, 31(1):25.
- [2] Baskett, F., Chandy, K.M., Muntz, R., and Palacios, M.G. (1975). Open, Closed, and Mixed Networks of Queues with Different Classes of Customers, *Journal of the Association for Computing Machinery*, 22(2):248-260.
- [3] Kelly, F.P. (1975). Networks of Queues with Customers of Different Types, *Journal of Applied Probability*, 12(3):542-554.
- [4] Kelly, F.P. (1976). Networks of queues, *Adv. Appl. Prob.*, (8):416-432.
- [5] Mani Chandy, K., John H. Howard Jr., And Donald F. Towsley. (1977). Product Form and Local Balance in Queueing Networks, *Journal of the Association for Computing Machinery*, 24(2):250-263.
- [6] Jean Walrand and Pravin Varaiya. (1980). Interconnections Of Markov Chains And Quazi-Reversible Queueing Networks, *Stochastic Processes and their Applications*, 10:209-219.
- [7] Xi-Ren Cao. (1989). The Convergence Property Of Sample In Closed Jackson Queueing Networks, *Stochastic Processes and their Applications*, 33:105-122.
- [8] Akyildiz, I.F., and Von Brand, H. (1989). Exact Solutions For Open, Closed And Mixed Queueing Networks With Rejection Blocking , *Theoretical Computer Science*, 64:203-219.
- [9] Karl Sigman. (1990). The Stability Of Open Queueing Networks, *Stochastic Processes and their Applications*, 35:11-25.
- [10] Ram, R. and Viswanadham, N. (1990). Recent Advances In Queueing Networks: A Survey With Applications To Automated Manufacturing.
- [11] Yawn, Y. and Frangos, C. (1992). Optimal Control Of Some Queueing Networks, *Mathl. Comput. Modelling*, 16(5):3-12.
- [12] Svend-Holger Friis, Ulrich Rieder And Jorgen Weishaupt. (1993). Optimal Control of Single-Server Queueing Networks, *ZOR - Methods and Models of Operations Research*, (37):187-205.
- [13] Kameda, H. and Zhang, Y. (1995). Uniqueness of the Solution for Optimal Static Routing in Open BCMP Queueing Networks, *Mathl. Comput. Model*, 22(12): 119-130.
- [14] Coyle, A.J., Henderson, W., Pearce, C.E.M., and Taylor, P.G. (1995). Mean-Value Analysis for a Class of Petri Nets and Batch-Movement Queueing Networks with Product-Form Equilibrium Distributions, *Mathl. Comput. Modelling*, 22(10):27-34.
- [15] Yang, H. and Gong, W.B. (1998). Calculating the Normalization Constants in Queueing

Networks , *Appl. Math. Lett.*, 11(6):87-91.

[16] Jean-Marc Vincent and Jerome Vienne. (2005). Perfect simulation of index based routing queueing networks.

[17] Charles Sutton and Michael I. Jordan. (2005). Probabilistic Inference in Queueing Networks.

[18] Filipowicz, B. and Kwecien, J. (2008). Queueing systems and networks. Models and applications, *Bulletin Of The Polish Academy Of Sciences*, 56(4).

[19] Vidhyacharan Bhaskar and Patrick Lallement. (2009). A four-input three-stage queueing network approach to model an industrial system, *Applied Mathematical Modelling* (33):3465–3487.

[20] Vidhyacharan Bhaskar and Patrick Lallement. (2010). Modeling a supply chain using a network of queues, *Applied Mathematical Modelling*, (34):2074–2088.

[21] Charles Sutton and Michael I. Jordan. (2010). Inference and Learning in Networks of Queues, *Artificial Intelligence and Statistics (AISTATS)*.

[22] Cruz, F.R.B., Kendall, G., While, L., Duarte, A.R. and Brito, N.L.C. (2012). Throughput Maximization of Queueing Networks with Simultaneous Minimization of Service Rates and Buffers, *Mathematical Problems in Engineering*.

[23] Mike Smitha, Wei Huangb and Francesco Vitib. (2013). Equilibrium in capacitated network models with queueing delays, queue-storage, blocking back and control , *Procedia - Social and Behavioral Sciences*, 80:860 – 879.

[24] Boon, M.A.A., Van Der Mei, R.D. and Winands, E.M.M. (2014). Queueing networks with a single shared server: Light and heavy traffic.

[25] Sreekala, M.S. and Manoharan, M. (2016). An Application of Restricted Open Queueing Networks to Healthcare System, *International Journal of Latest Trends in Engineering and Technology (IJLTET)*, 7(1).

[26] Geraint I. Palmer, Paul R. Harper and Vincent A. Knight. (2018). Modelling deadlock in open restricted queueing networks, *European Journal of Operational Research*, 266:609–621.

[27] Wenbo Zhanga, C., Harsha Honnappab and Satish, V. (2019). Modeling Urban Taxi Services with E-Hailings: A Queueing Network Approach, *Transportation Research Procedia*,(38):751–771.

[28] Ward Whitt and Wei You. (2020). A Robust Queueing Network Analyzer Based on Indices of Dispersion, (6).

[29] Giulio Garbi, Emilio Inserto and Micro Tribastone. (2020). Learning Queueing Networks By Recurrent Neural Networks.

[30] Behzad Maleki Vishkaeia, Iraj Mahdavia, Nezam Mahdavi-Amirib and Esmail Khorramc. (2020). Balancing public bicycle sharing system using inventory critical levels in queueing network, *Computers & Industrial Engineering*.

[31] Masoomah Zeinalnezhad, Abdoulmohammad Gholamzadeh Chofreh, Feybi Ariani Goni, Jaromir Klemes and Emelia Sari. (2020). Simulation and Improvement of Patients Workflow in Heart Clinics during COVID-19 Pandemic Using Timed Coloured Petri Nets.

[32] Banu Priya, K. and Rajendran, P. (2019). Reduction of Delay Rate in Open Queueing Network, *International Journal Of Scientific & Technology Research*, 8(12).

[33] Govind Lal Kumawat and Debjit Roy. (2021). A new solution approach for multi- stage semi-open queueing networks: An application in shuttle-based compact storage systems, *Computers and Operations Research*.

# MODELING AND ANALYSIS OF SINE POWER RAYLEIGH DISTRIBUTION : PROPERTIES AND APPLICATIONS

AADIL AHMAD MIR

•  
Department of Statistics, University of Kashmir, Srinagar, India  
stataadil29@gmail.com

S.P.AHMAD\*

•  
Department of Statistics, University of Kashmir, Srinagar, India  
sprvz@yahoo.com

## Abstract

*In this manuscript, a new probability model named as Sine Power Rayleigh distribution (SPRD) is proposed using a Sine-G function as generator. Various statistical properties of this new distribution were investigated, including the survival function, hazard function, reverse hazard rate, cumulative hazard function, mills ratio, quantile function, moments, moment generating function, conditional moments, entropy, and order statistics. The parameters of the proposed distribution were estimated using the method of maximum likelihood estimation. To assess the model's versatility and applicability, we conduct analyses on two real life data sets. The outcomes affirm the superior performance of the newly proposed model SPRD as compared to existing models.*

**Keywords:** Sine G family, Rayleigh distribution, Sine Rayleigh distribution, Reliability Analysis, Entropy, Order Statistics, Maximum Likelihood Estimation.

## 1. INTRODUCTION

The concept of probability distribution has shown to be quite helpful in managing both small and large data sets. Probability distribution models are essential and widely utilised in many domains, including as physics, medicine, business management, engineering, and food. The field of probability distributions has advanced steadily due to the wide range of domains in which they are applied. Over the past few decades, researchers have used a variety of ways to introduce numerous novel probability distributions. New distributions are needed to address the problem more precisely and effectively, even though there are numerous existing ways for handling real-world data. From an applied and practical perspective, the new family of distributions modifies some of the current distributions to make them more flexible, which serves key purposes in the generalisation of distributions. There are several ways to create new models, including exponentiation, compounding, and changing and adding constants to well-known distributions.

The Rayleigh distribution (RD), named after Lord Rayleigh [15] is prominent lifetime probability model concerned with describing skewed data. The probability density function (PDF) associated with random variable  $x > 0$  having RD with scale parameter  $\theta$  is given by

$$f(x;\theta) = \frac{x}{\theta^2} \exp\left(-\frac{x^2}{2\theta^2}\right); \quad x > 0, \quad \theta > 0$$

and the corresponding cumulative distribution function (CDF) is given as

$$F(x; \theta) = 1 - \exp\left(-\frac{x^2}{2\theta^2}\right); \quad x > 0, \quad \theta > 0$$

In the statistical literature, numerous extensions of Rayleigh distribution (RD) have been proposed. Surless and Padgett[17] introduced the two parameter Burr type X distribution and named it as exponentiated Rayleigh distribution (ERD) or generalized Rayleigh distribution. Kundu and Raqab [11] studied and estimated the parameters of the generalized Rayleigh distribution using different estimation techniques. Ahmed et al. [2] used the square error loss function and Al-Bayyati's loss function to perform a Bayesian analysis of RD. Ajami and Jhansi [3] discussed the parameter estimation of weighted Rayleigh distribution. Ahmad et al. [1] proposed the Weibull-Rayleigh distribution and studied its characterization and parameter estimation using the transformed transformer technique. Bhat and Ahmad [6] proposed a new extension of exponentiated Rayleigh distribution and studied its various properties and demonstrated its applicability by considering different datasets. Bhat and Ahmad [5] studied mathematical properties of mixture of Gamma and Rayleigh distributions. Kilai et al. [8] proposed a new versatile modification of the Rayleigh distribution for modeling COVID-19 mortality rates. Various researchers have introduced generalised distributions and their applications, see Mahmood et al. [12] , Muse et al. [13] and Ahmed et al. [15]. Bhat et al. [7] proposed a new extension of odd lindley power rayleigh distribution, studied its properties and evaluated parameter estimation techniques using both classical and Bayesian methods. Bhat and Ahmad [4] recently introduced a new generalization of the Rayleigh distribution using power transformation technique with PDF and CDF respectively given by

$$g(x; \beta, \theta) = \frac{\beta}{\theta^2} x^{2\beta-1} \exp\left(-\frac{x^{2\beta}}{2\theta^2}\right); \quad x > 0, \quad \beta, \theta > 0 \quad (1)$$

and the corresponding cumulative distribution function (CDF) is given as

$$G(x; \beta, \theta) = 1 - \exp\left(-\frac{x^{2\beta}}{2\theta^2}\right); \quad x > 0, \quad \beta, \theta > 0 \quad (2)$$

In the present manuscript, we proposed a new extension of Power Rayleigh distribution (PRD) using the Sine G family of generated distributions. The proposed distribution is named as Sine Power Rayleigh distribution (SPRD). It is more flexible and exhibits more complex shapes of density and hazard rate functions. Also, the proposed model outclass some well established models in terms of two real life data sets. The rest of the article is unfolded as : In section 2, the Ratio Transformation (RT) method is discussed. In Section 3, the PDF and CDF of the proposed model i.e., SPRD are defined. Section 4 deals with the reliability measures of the SPRD. The expansion of PDF and CDF is discussed in Section 5. Some of important statistical properties are explored in Section 6. The parameter estimation is discussed in Section 7. The simulation study and applicability of the model is debated in section 8 and 9 respectively. Finally, some conclusion are provided in Section 10.

## 2. SINE G FAMILY OF DISTRIBUTIONS

The CDF and PDF of the Sine G family of distributions proposed by [10] are defined by the following equations respectively:

$$F(x; \zeta) = \sin\left[\frac{\pi}{2}G(x; \zeta)\right]; \quad x \in \mathbb{R} \quad (3)$$

$$f(x; \zeta) = \frac{\pi}{2}g(x; \zeta) \cos\left[\frac{\pi}{2}G(x; \zeta)\right]; \quad x \in \mathbb{R} \quad (4)$$



Where  $G(x; \zeta)$  and  $g(x; \zeta)$  in equation (3) and (4) are the CDF and PDF of the base line distribution with parameter vector  $\zeta$ , respectively.

### 3. SINE POWER RAYLEIGH DISTRIBUTION (SPRD)

The PDF of the newly proposed probability distribution Sine Power Rayleigh Distribution (SPRD) is obtained as

$$f(x; \beta, \theta) = \frac{\pi}{2} \frac{\beta}{\theta^2} x^{2\beta-1} e^{-\frac{x^{2\beta}}{2\theta^2}} \cos \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right]; \quad x \in \mathbb{R}^+, \quad \beta, \theta > 0 \quad (5)$$

The CDF of the newly proposed probability distribution Sine Power Rayleigh Distribution (SPRD) is obtained as

$$F(x; \beta, \theta) = \sin \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right]; \quad x \in \mathbb{R}^+, \quad \beta, \theta > 0 \quad (6)$$

The plots of density function of SPRD for different parameter combinations are presented in Figure 1 . It is clear from the density function plots that the proposed distribution is unimodal, decreasing, symmetric and positively skewed.

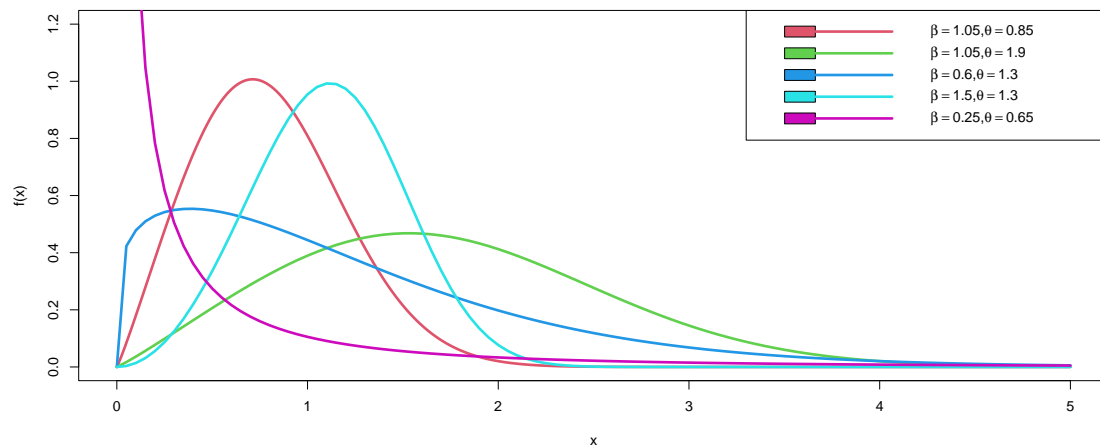


Figure 1: Density plots of SPRD for different combinations of  $\beta$  and  $\theta$ .

### 4. RELIABILITY ANALYSIS OF THE SINE POWER RAYLEIGH DISTRIBUTION (SPRD)

This section focuses on obtaining the reliability (survival function), hazard rate (failure rate), reverse hazard function, cumulative hazard function and mills ratio expressions respectively for SPRD.

#### 4.1. Survival function

The survival function or reliability function is defined as the probability that a system will survive beyond a specified time and is obtained for the SPRD as

$$R(x; \beta, \theta) = 1 - F(x; \beta, \theta) = 1 - \sin \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right] \quad (7)$$

### 4.2. Hazard Rate

The Hazard rate evaluates a lifetime component’s likelihood of failure or expiration based on the completed portion of its life, and consequently, it finds diverse applications in the analysis of lifetime distributions. Using equation (5) and (7), the expression for the hazard rate of SPRD is obtained as

$$h(x; \beta, \theta) = \frac{f(x; \beta, \theta)}{R(x; \beta, \theta)} = \frac{\frac{\pi}{2} \frac{\beta}{\theta^2} x^{2\beta-1} e^{-\frac{x^{2\beta}}{2\theta^2}} \cos \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right]}{1 - \sin \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right]} \tag{8}$$

Figure 2 depicts graphs of the hazard rate of the SPRD for different parameter values. Figure 2 suggests that the proposed distribution is quite flexible in nature and can exhibit variety of shapes such as constant, decreasing, increasing and j-shaped shaped over the parameter space.

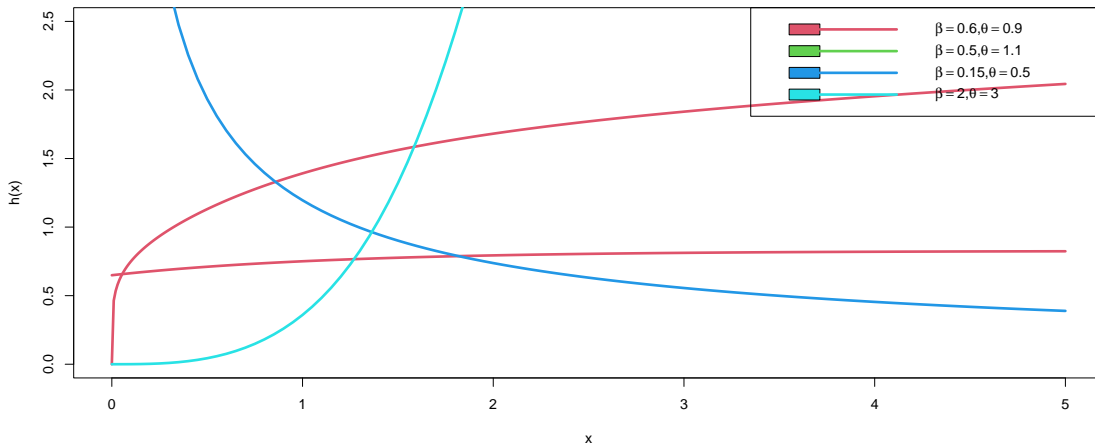


Figure 2: Hazard rate plots of SPRD for different combinations of  $\beta$  and  $\theta$ .

### 4.3. Reverse Hazard function

The concept of reversed hazard rate of a random life is defined as the ratio between the life probability density to its distribution function . It is expressed as

$$h_r(x; \beta, \theta) = \frac{f(x; \beta, \theta)}{F(x; \beta, \theta)} = \frac{\frac{\pi}{2} \frac{\beta}{\theta^2} x^{2\beta-1} e^{-\frac{x^{2\beta}}{2\theta^2}} \cos \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right]}{\sin \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right]}$$

### 4.4. Cumulative Hazard function

The cumulative hazard function can be thought of as providing the total accumulated risk of experiencing the event of interest that has been gained by progressing to time t. The cumulative hazard function for the SPRD is defined as

$$\Lambda_{SPRD}(x; \beta, \theta) = -\log R(x; \beta, \theta) = -\log \left\{ 1 - \sin \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right] \right\}$$

#### 4.5. Mills Ratio

The mills ratio for the SPRD is defined as

$$M.R = \frac{F(x; \beta, \theta)}{R(x; \beta, \theta)} = \frac{\sin \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right]}{1 - \sin \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right]} \quad (9)$$

#### 4.6. Quantile function

The quantile function for the SPRD is given by

$$x = \left[ -2\theta^2 \log \left( 1 - \frac{2}{\pi} \sin^{-1} u \right) \right]^{\frac{1}{2\beta}} \quad (10)$$

The first quartile ( $Q_1$ ), median ( $Q_2$ ), and third quartile ( $Q_3$ ) can be derived by setting  $u = \frac{1}{4}, \frac{1}{2},$  and  $\frac{3}{4}$  in equation (10) respectively.

### 5. EXPANSION OF PDF AND CDF

Various statistical properties can be easily deduced by using mixture representation of PDF and CDF of the proposed model.

expansion of  $\cos \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right]$  can be expressed as

$$\cos \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right] = \sum_{l=0}^{\infty} \frac{(-1)^l}{2l!} \frac{\pi^{2l}}{2^{2l}} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right)^{2l}$$

Also  $\left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right)^{2l}$  can be expressed as

$$\left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right)^{2l} = \sum_{m=0}^{\infty} (-1)^m \binom{2l}{m} e^{-\frac{mx^{2\beta}}{2\theta^2}}$$

expansion of  $\sin \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right]$  can be expressed as

$$\sin \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right] = \sum_{p=0}^{\infty} \frac{(-1)^p}{(2p+1)!} \frac{\pi^{2p+1}}{2^{2p+1}} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right)^{2p+1}$$

Also  $\left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right)^{2p+1}$  can be expressed as

$$\left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right)^{2p+1} = \sum_{q=0}^{\infty} (-1)^q \binom{2p+1}{q} e^{-\frac{qx^{2\beta}}{2\theta^2}}$$

Thus, the PDF and CDF of the proposed model can be written in the mixture representation respectively as

$$f(x; \beta, \theta) = \frac{\beta}{\theta^2} x^{2\beta-1} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} e^{-\frac{(m+1)x^{2\beta}}{2\theta^2}} \quad (11)$$

$$F(x; \beta, \theta) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(-1)^{p+q}}{(2p+1)!} \binom{2p+1}{q} \frac{\pi^{2p+1}}{2^{2p+1}} e^{-\frac{qx^{2\beta}}{2\theta^2}} \quad (12)$$

## 6. STATISTICAL PROPERTIES OF SPRD

Some of the mathematical properties such as the  $r^{th}$  moment, moment generating function, conditional moments and associated measures, the entropy and order statistics are derived.

### 6.1. Moments

The  $r^{th}$  moment of the SPRD can be evaluated directly by extending the PDF given in equation (11)

$$E(X^r) = \int_0^{\infty} x^r f(x; \beta, \theta) dx \quad , r = 1, 2, ..$$

where  $f(x)$  is the PDF of the SPRD given in equation (11), thus

$$E(X^r) = \frac{\beta}{\theta^2} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \int_0^{\infty} x^{r+2\beta-1} e^{-\frac{(m+1)x^{2\beta}}{2\theta^2}} \quad (13)$$

Using integration via substitution method in equation (13), we perform the following operations.

$$\text{let } \frac{(m+1)x^{2\beta}}{2\theta^2} = z \implies x = \left(\frac{2\theta^2 z}{m+1}\right)^{\frac{1}{2\beta}}, \text{ such that } dx = \frac{1}{2\beta} \left(\frac{2\theta^2}{m+1}\right)^{\frac{1}{2\beta}} (z)^{\frac{1}{2\beta}-1}$$

Thus, simplifying equation (13) yields

$$E(X^r) = (2\theta^2)^{\frac{r}{2\beta}} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{r}{2\beta}+1} \Gamma\left(\frac{r}{2\beta} + 1\right) \quad (14)$$

where,

$$\Gamma\left(\frac{r}{2\beta} + 1\right) = \int_0^{\infty} z^{\left(\frac{r}{2\beta}+1\right)-1} e^{-z} dz$$

setting  $r = 1$  in equation (14) the mean of the model is computed as

$$E(X) = (2\theta^2)^{\frac{1}{2\beta}} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{1}{2\beta}+1} \Gamma\left(\frac{1}{2\beta} + 1\right) \quad (15)$$

Similarly for  $r = 2, 3$  and  $4$  in equation (14), the second, third and fourth moment about origin are respectively calculated as

$$E(X^2) = (2\theta^2)^{\frac{2}{\beta}} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{2}{\beta}+1} \Gamma\left(\frac{2}{\beta} + 1\right) \quad (16)$$

$$E(X^3) = (2\theta^2)^{\frac{3}{2\beta}} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{3}{2\beta}+1} \Gamma\left(\frac{3}{2\beta} + 1\right) \quad (17)$$

$$E(X^4) = (2\theta^2)^{\frac{2}{\beta}} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{2}{\beta}+1} \Gamma\left(\frac{2}{\beta} + 1\right) \quad (18)$$

### 6.2. Moment Generating function of SPRD

we can calculate moment generating function based on the  $r^{th}$  moment of SPRD as given by

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} E(X^r) \tag{19}$$

$$M_X(t) = (2\theta^2)^{\frac{r}{2\beta}} \sum_{r=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{t^r}{r!} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{r}{2\beta}+1} \Gamma\left(\frac{r}{2\beta} + 1\right) \tag{20}$$

### 6.3. Conditional moments and associated measures

In this section, the expression for conditional moments is acquired. But first we will introduce an important lemma which will be applied in the next section.

**Lemma 1.** Let us suppose a random variable X follows SPRD  $(\beta, \theta)$  with PDF given in equation (11) and let  $\varphi_r(z) = \int_0^z x^r f(x; \beta, \theta) dx$  denotes the  $r^{th}$  incomplete moment, then we have

$$\varphi_r(z) = (2\theta^2)^{\frac{r}{2\beta}} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{r}{2\beta}+1} \gamma\left(\left(\frac{r}{2\beta} + 1\right), \frac{(m+1)z^{2\beta}}{2\theta^2}\right) \tag{21}$$

where  $\gamma(a, b) = \int_0^b z^{a-1} e^{-z} dz$  denotes the lower incomplete gamma function.

**Proof:** Using the PDF of SPRD given in equation (11), we have

$$\varphi_r(z) = \int_0^z x^r f(x; \beta, \theta) dx = \frac{\beta}{\theta^2} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \int_0^z x^{r+2\beta-1} e^{-\frac{(m+1)x^{2\beta}}{2\theta^2}} \tag{22}$$

On Simplification, we obtain

$$\varphi_r(z) = (2\theta^2)^{\frac{r}{2\beta}} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{r}{2\beta}+1} \gamma\left(\left(\frac{r}{2\beta} + 1\right), \frac{(m+1)z^{2\beta}}{2\theta^2}\right) \tag{23}$$

Setting  $r=1$  in equation (23) will yield first incomplete moment as given by

$$\varphi_1(z) = (2\theta^2)^{\frac{1}{2\beta}} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{1}{2\beta}+1} \gamma\left(\left(\frac{1}{2\beta} + 1\right), \frac{(m+1)z^{2\beta}}{2\theta^2}\right) \tag{24}$$

#### 6.3.1 Lorenz and Bonferroni inequality Curves

The Lorenz and Bonferroni inequality curves are an important application of the first incomplete moment. For a given probability distribution, they are defined by

$$L_p = \frac{1}{E(X)} \int_0^t x f(x; \beta, \theta) dx = \frac{\varphi_1(t)}{E(X)}$$

$$L_p = \frac{\sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{1}{2\beta}+1} \gamma\left(\left(\frac{1}{2\beta} + 1\right), \frac{(m+1)t^{2\beta}}{2\theta^2}\right)}{\sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{1}{2\beta}+1} \Gamma\left(\frac{1}{2\beta} + 1\right)}$$

Similarly,

$$B_p = \frac{1}{pE(X)} \int_0^t x f(x; \beta, \theta) dx = \frac{\varphi_1(t)}{pE(X)}$$

$$B_p = \frac{\sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{1}{2\beta}+1} \gamma\left(\left(\frac{1}{2\beta} + 1\right), \frac{(m+1)t^{2\beta}}{2\theta^2}\right)}{p \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{1}{2\beta}+1} \Gamma\left(\frac{1}{2\beta} + 1\right)}$$

### 6.3.2 $r^{th}$ Conditional Moment and $r^{th}$ Reversed Conditional Moment of SPRD

The  $r^{th}$  conditional moment of the SPRD is calculated by

$$E[X^r | x > t] = \frac{1}{R(t)} \int_t^{\infty} x^r f(x; \beta, \theta) dx = \frac{1}{R(t)} [E(X^r) - \varphi_r(t)]$$

where  $R(t)$  is the reliability of SPRD at time  $t$ .

Inserting the value of equation (7), (14) and (23), we obtain

$$E[X^r | x > t] = \frac{(2\theta^2)^{\frac{r}{2\beta}} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{r}{2\beta}+1} \left[ \Gamma\left(\frac{r}{2\beta} + 1\right) - \gamma\left(\left(\frac{r}{2\beta} + 1\right), \frac{(m+1)t^{2\beta}}{2\theta^2}\right) \right]}{1 - \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(-1)^{p+q}}{(2p+1)!} \binom{2p+1}{q} \frac{\pi^{2p+1}}{2^{2p+1}} e^{-\frac{qt^{2\beta}}{2\theta^2}}}$$

Similarly, the  $r^{th}$  reversed conditional moment of the SPRD is defined by

$$E[X^r | x \leq t] = \frac{1}{F(t)} \int_0^t x^r f(x; \beta, \theta) dx = \frac{\varphi_r(t)}{F(t)}$$

$$E[X^r | x \leq t] = \frac{(2\theta^2)^{\frac{r}{2\beta}} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{r}{2\beta}+1} \gamma\left(\left(\frac{r}{2\beta} + 1\right), \frac{(m+1)t^{2\beta}}{2\theta^2}\right)}{\sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(-1)^{p+q}}{(2p+1)!} \binom{2p+1}{q} \frac{\pi^{2p+1}}{2^{2p+1}} e^{-\frac{qt^{2\beta}}{2\theta^2}}}$$

### 6.3.3 Mean Residual Life (MRL) and Mean Waiting Time (MWT)

The MRL is defined as

$$\mu(t) = \frac{1}{R(t)} \left[ E(t) - \int_0^t x f(x; \beta, \theta) dx \right] - t = \frac{1}{R(t)} [E(t) - \varphi_1(t)] - t$$

After inserting the value of equation (7), (15) and (24), we obtain the required expression for mean residual life as

$$\mu(t) = \frac{(2\theta^2)^{\frac{1}{2\beta}} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{1}{2\beta}+1} \left[ \Gamma\left(\frac{1}{2\beta} + 1\right) - \gamma\left(\left(\frac{1}{2\beta} + 1\right), \frac{(m+1)t^{2\beta}}{2\theta^2}\right) \right]}{1 - \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(-1)^{p+q}}{(2p+1)!} \binom{2p+1}{q} \frac{\pi^{2p+1}}{2^{2p+1}} e^{-\frac{qt^{2\beta}}{2\theta^2}}} - t$$

The MWT is defined as

$$\bar{\mu}(t) = t - \frac{1}{F(t)} \int_0^t x f(x; \beta, \theta) dx = t - \frac{\phi_1(t)}{F(t)}$$

$$\bar{\mu}(t) = t - \frac{(2\theta^2)^{\frac{1}{2\beta}} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}} \left(\frac{1}{m+1}\right)^{\frac{1}{2\beta}+1} \gamma\left(\left(\frac{1}{2\beta} + 1\right), \frac{(m+1)t^{2\beta}}{2\theta^2}\right)}{\sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(-1)^{p+q}}{(2p+1)!} \binom{2p+1}{q} \frac{\pi^{2p+1}}{2^{2p+1}} e^{-\frac{qt^{2\beta}}{2\theta^2}}}$$

### 6.4. Renyi entropy

The entropy of a random variable is defined as the average amount of information lost during a random experiment. The Renyi entropy, which Alfred Renyi introduced [16] and generalises Shannon’s measure of information, is defined as

$$R_\eta = \frac{1}{1-\eta} \log \int_{-\infty}^{\infty} f^\eta(x; \beta, \theta) dx, \quad \eta > 0, \quad \eta \neq 1$$

Using the PDF given in equation (11), we have

$$R_\eta = \frac{1}{1-\eta} \log \left(\frac{\beta}{\theta}\right)^\eta \left(\sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}}\right)^\eta \int_0^{\infty} x^{\eta(2\beta-1)} e^{-\frac{\eta(m+1)x^{2\beta}}{2\theta^2}}$$

$$R_\eta = \frac{1}{1-\eta} \log \left(\frac{\beta}{\theta}\right)^\eta \frac{1}{2\beta} \left(\sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m}}{2l!} \binom{2l}{m} \frac{\pi^{2l+1}}{2^{2l+1}}\right)^\eta \left(\frac{2\theta^2}{\eta(m+1)}\right)^{\frac{\eta(2\beta-1)+1}{2\beta}} \Gamma\left(\frac{\eta(2\beta-1)+1}{2\beta}\right)$$

### 6.5. Order Statistics of SPRD

The order statistics connected to the SPRD is devoted in this section. Let  $x_{(r;n)}$  be the  $r^{th}$  order statistics with the random sample  $x_{(1)}, x_{(2)}, x_{(3)}, \dots, x_{(n)}$  derived from the SPRD having the PDF  $f(X; \beta, \theta)$  and CDF  $F(X; \beta, \theta)$ . Therefore, the PDF and CDF of  $x_{(r;n)}$  say  $f_{(r;n)}(x)$  and  $F_{(r;n)}(x)$  are respectively defined as

$$f_{(r;n)}(x) = \frac{1}{B(n, n-r+1)} [F(x; \beta, \theta)]^{r-1} [1 - F(x; \beta, \theta)]^{n-r} f(x; \beta, \theta) \tag{25}$$

$$F_{(r;n)}(x) = \sum_{j=r}^n \binom{n}{j} [F(x; \beta, \theta)]^j [1 - F(x; \beta, \theta)]^{n-j} \tag{26}$$

Using equation (5) and equation (6) in equation (25) and equation (26), the PDF and CDF of  $r^{th}$  ordered statistics for the SPRD are derived and are expressed as

$$f_{(r;n)}(x) = \frac{\frac{\pi}{2} \frac{\beta}{\theta^2} x^{2\beta-1} e^{-\frac{x^{2\beta}}{2\theta^2}} \cos\left[\frac{\pi}{2} \left(1 - e^{-\frac{x^{2\beta}}{2\theta^2}}\right)\right]}{B(n, n-r+1)} \left\{ \sin\left[\frac{\pi}{2} \left(1 - e^{-\frac{x^{2\beta}}{2\theta^2}}\right)\right] \right\}^{r-1} \left\{ 1 - \sin\left[\frac{\pi}{2} \left(1 - e^{-\frac{x^{2\beta}}{2\theta^2}}\right)\right] \right\}^{n-r}$$

$$F_{(r,n)}(x) = \sum_{j=r}^n \binom{n}{j} \left\{ \sin \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right] \right\}^j \left\{ 1 - \sin \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x^{2\beta}}{2\theta^2}} \right) \right] \right\}^{n-j}$$

where  $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$  is the beta function.

## 7. ESTIMATION OF PARAMETERS

The goal of this study is to estimate the unknown parameters  $\beta$  and  $\theta$  of the SPRD using Maximum Likelihood Estimation (MLE). we assume that  $x_1, x_2, \dots, x_n$  be a random sample of  $n$  observations drawn from the SPRD  $(\beta, \theta)$  with unknown parametric vector  $\Theta = (\beta, \theta)^T$ .

### 7.1. Maximum Likelihood Estimation (MLE)

Here, Maximum Likelihood Estimation (MLE) approach is used to obtain the estimators of the unknown parameters of SPRD  $(\beta, \theta)$ . The likelihood function is given by

$$L(\Theta) = \left[ \frac{\pi\beta}{2\theta^2} \right]^n e^{-\sum_{i=1}^n \frac{x_i^{2\beta}}{2\theta^2}} \prod_{k=1}^n x_k^{2\beta-1} \cos \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x_k^{2\beta}}{2\theta^2}} \right) \right]$$

For the parametric vector  $(\Theta) = (\beta, \theta)^T$ , the logarithm likelihood function is expressed as

$$\begin{aligned} \ell = n \log \left( \frac{\pi}{2} \right) + n \log(\beta) - 2n \log(\theta) - \frac{1}{2\theta^2} \sum_{k=1}^n x_k^{2\beta} + (2\beta - 1) \sum_{k=1}^n \log x_k \\ + \sum_{k=1}^n \log \cos \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x_k^{2\beta}}{2\theta^2}} \right) \right] \end{aligned} \quad (27)$$

The elements of the score vector  $U(\Theta) = (U_\beta, U_\theta)$  are obtained by partially differentiating Equation (27) with respect to the model parameters and are given by

$$\frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} + 2 \sum_{k=1}^n \ln(x_k) - \frac{1}{2\theta^2} \sum_{k=1}^n x_k^{2\beta} \ln(x_k) - \frac{\pi}{4\theta^2} \sum_{k=1}^n \tan \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x_k^{2\beta}}{2\theta^2}} \right) \right] e^{-\frac{x_k^{2\beta}}{2\theta^2}} x_k^{2\beta} \ln(x_k)$$

$$\frac{\partial \ell}{\partial \theta} = \frac{-2n}{\theta} + \frac{1}{\theta^3} \sum_{k=1}^n x_k^{2\beta} + \frac{\pi}{2\theta^3} \sum_{k=1}^n \tan \left[ \frac{\pi}{2} \left( 1 - e^{-\frac{x_k^{2\beta}}{2\theta^2}} \right) \right] e^{-\frac{x_k^{2\beta}}{2\theta^2}} x_k^{2\beta}$$

The likelihood estimates of the model parameters can be obtained by setting the score vector  $U(\Theta) = 0$ . Since, the above equations are non-linear and hence the model parameters are estimated using Newton-Raphson algorithm.

## 8. SIMULATION ILLUSTRATION

In this section, we carry out simulation study using R software to examine the behaviour of MLE's for various sample sizes. We generate the random samples of size 25, 75, 150, 300 and 500 from



SPRD and repeat the process for 1000 times in R software. Various combinations of parameters are chosen as (1.5, 1.35) and (0.5, 2.2) with relation to the standard order  $(\beta, \theta)$ . The average MLE values, bias, and related empirical mean squared errors (MSEs) were determined for each scenario. Tables 1 exhibits the ML estimates, bias and MSE. We observe from table 1 that the agreement between theory and practice improves as the sample size  $n$  increases. MSE and bias of the estimators suggest that the estimators are consistent and the maximum likelihood estimator of the parameters perform quite well and the results are precise and accurate. The MSE decreases with increasing sample size under all conditions.

**Table 1:** MLE, Bias and MSE for the parameters  $\beta$  and  $\theta$

sample size n	Parameters		MLE		Bias		MSE	
	$\beta$	$\theta$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\beta}$	$\hat{\theta}$
25	1.5	1.35	1.58963	1.38116	0.21193	0.15622	0.07685	0.04192
75			1.52863	1.36292	0.11586	0.08563	0.02170	0.01211
150			1.51474	1.35744	0.07911	0.05752	0.00999	0.00543
300			1.50528	1.35236	0.05462	0.03945	0.00459	0.00248
500			1.50487	1.35130	0.04267	0.03108	0.00278	0.00153
25	0.5	2.2	0.53233	2.39244	0.07177	0.40725	0.00960	0.36239
75			0.50767	2.24504	0.03767	0.20412	0.00222	0.06987
150			0.50439	2.22579	0.02799	0.14628	0.00126	0.03659
300			0.50299	2.21458	0.01852	0.10195	0.00054	0.01658
500			0.50085	2.20587	0.01432	0.07776	0.00034	0.00967

## 9. APPLICATION

This section is devoted to illustrate the flexibility, adaptability, and suitability of the SPRD, by means of two real data sets . We compare the proposed distribution with the following models :

- Power Rayleigh distribution (PRD) With PDF given as

$$f(x; \beta, \theta) = \frac{\beta}{\theta^2} x^{2\beta-1} \exp\left(-\frac{x^{2\beta}}{2\theta^2}\right); \quad \beta, \theta > 0$$

- Weighted Rayleigh Distribution (WRD) with PDF given as

$$f(x; \beta, \theta) = \frac{x^{\beta+1} \exp\left(-\frac{x^2}{2\theta^2}\right)}{\theta^{\beta+2} 2^{\frac{\beta}{2}} \Gamma\left(\frac{\beta}{2} + 1\right)}; \quad \beta, \theta > 0$$

- Rayleigh distribution (RD) with PDF given as

$$f(x; \theta) = \frac{x}{\theta^2} \exp\left(-\frac{x^2}{2\theta^2}\right); \quad \theta > 0$$

Here, several goodness of fit criterion such as -2ll, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Akaike Information Criterion Corrected (AICC) , Kolmogorov-Smirnov (KS) and P value statistics are used. The statistic with the lowest value of -2ll, AIC, BIC, AICC, K-S and maximum value of P value is considered the best fit.

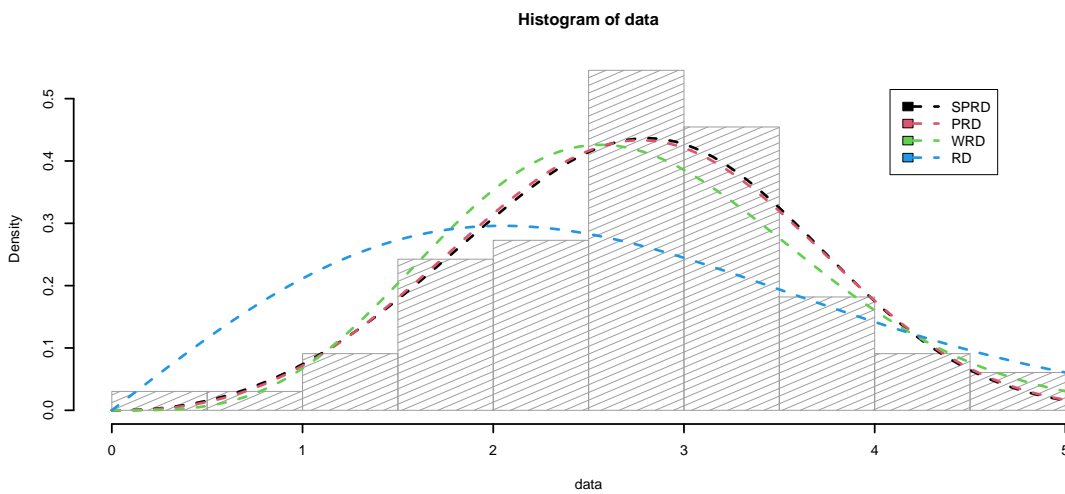
### 9.1. Data Set 1

Data set 1: The first data is on the breaking stress of carbon fibres of 50 mm length (GPa). The data has been previously used by [4] and [14]. The data is as follows:

0.39, 0.85, 1.08, 1.25, 1.47, 1.57, 1.61, 1.61, 1.69, 1.80, 1.84, 1.87, 1.89, 2.03, 2.03, 2.05, 2.12, 2.35, 2.41, 2.43, 2.48, 2.50, 2.53, 2.55, 2.55, 2.56, 2.59, 2.67, 2.73, 2.74, 2.79, 2.81, 2.82, 2.85, 2.87, 2.88, 2.93, 2.95, 2.96, 2.97, 3.09, 3.11, 3.11, 3.15, 3.15, 3.19, 3.22, 3.22, 3.27, 3.28, 3.31, 3.31, 3.33, 3.39, 3.39, 3.56, 3.60, 3.65, 3.68, 3.70, 3.75, 4.20, 4.38, 4.42, 4.70, 4.90

**Table 2:** Estimates (standard errors),  $-2ll$ , AIC, BIC, AICC, K-S statistic and P-value for Data-set 1.

Model	$\hat{\beta}$	$\hat{\theta}$	$-2ll$	AIC	BIC	AICC	K-S	P-value
SPRD	1.6366 (0.1595)	5.8515 (1.2057)	171.6825	175.6825	180.0618	175.8730	0.0791	0.8029
PRD	1.7205 (0.1654)	4.8502 (1.0369)	172.1352	176.1352	180.5145	176.3256	0.0823	0.7625
WRD	2.5727 (0.7452)	1.3551 (0.1234)	175.7107	179.7107	184.0900	179.9012	0.1104	0.3963
RD		2.0491 (0.1261)	196.4168	198.4168	200.6065	198.4793	0.2265	0.0022



**Figure 3:** Fitted density plots for dataset 1

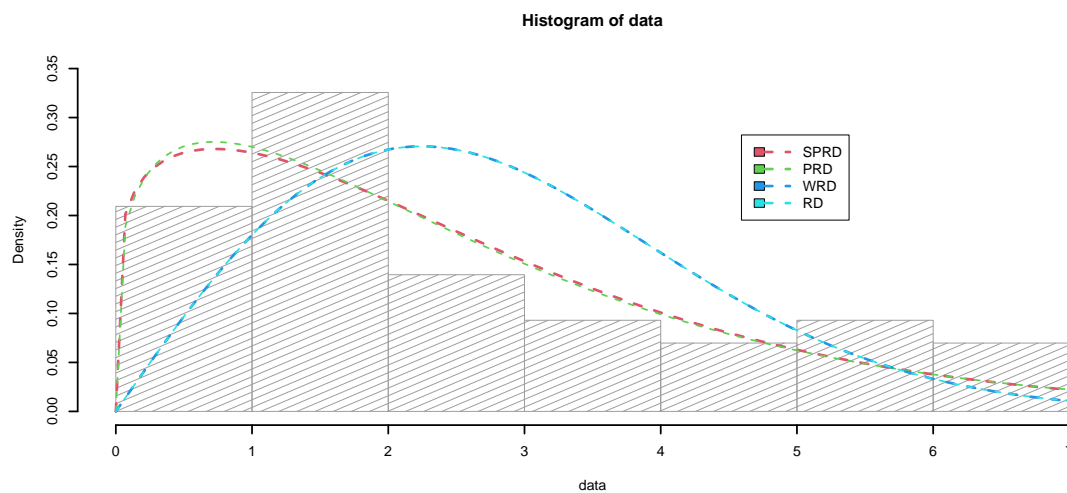
### 9.2. Data set 2

Data set 2: Consider the following data set in Johnson and Kotz [9] and represent the survival times (in years) after diagnosis of 43 patients with a certain kind of leukemia.

0.019, 0.129, 0.159, 0.203, 0.485, 0.636, 0.748, 0.781, 0.869, 1.175, 1.206, 1.219, 1.219, 1.282, 1.356, 1.362, 1.458, 1.564, 1.586, 1.592, 1.781, 1.923, 1.959, 2.134, 2.413, 2.466, 2.548, 2.652, 2.951, 3.038, 3.6, 3.655, 3.745, 4.203, 4.690, 4.888, 5.143, 5.167, 5.603, 5.633, 6.192, 6.655, 6.874

**Table 3:** Estimates (standard errors),  $-2ll$ , AIC, BIC, AICC, K-S statistic and P-value for Data-set 2.

Model	$\hat{\beta}$	$\hat{\theta}$	$-2ll$	AIC	BIC	AICC	K-S	P-value
SPRD	0.5887 (0.0736)	1.6864 (0.2041)	162.9906	166.9906	170.5130	167.2906	0.0869	0.901
PRD	0.6198 (0.0766)	1.3094 (0.1647)	163.2203	167.2203	170.7427	167.5203	0.0903	0.8744
WRD	0.0010 (0.3799)	2.2409 (0.2728)	181.9592	185.9592	189.4816	186.2592	0.2423	0.0128
RD		2.2415 (0.1709)	181.9277	183.9277	185.6889	184.0252	0.2421	0.0128



**Figure 4:** Fitted density plots for dataset 2

The results obtained in Table 2 and Table 3 reveal that SPRD has the least value of all the comparison criterions, hence SPRD can be considered a strong competitor to other distributions compared here for fitting data. The plots of the fitted models are displayed in figure 3 and 4. Also, from these plots , it is evident that SPRD provides a close fit to the two data sets.

## 10. CONCLUSION

In this paper, a new life time distribution namely Sine Power Rayleigh distribution (SPRD) is proposed and studied. The SPRD model is an expansion that incorporates the Sine-G family of distributions introduced by [10] resulting in a novel trigonometric distribution. The new distribution is more flexible and its hazard rate function exhibits complex shapes. The study derives various properties of the proposed distribution, including the survival function, hazard rate function, reverse hazard function, cumulative hazard function, moments, moment generating function, quantile function, Lorenz and Bonferroni inequality curves, Renyi entropy and order statistics. The parameters of the proposed distribution are estimated using the maximum likelihood method and a simulation study is conducted to assess the performance of the maximum likelihood estimators (MLEs) for these parameters. Furthermore, the effectiveness of the proposed

distribution is evaluated by applying it to two distinct real life datasets and comparing it with well known standard distributions such as the Rayleigh distribution, Power Rayleigh distribution and Weighted Rayleigh distribution. The results demonstrate that the Sine Power Rayleigh distribution (SPRD) surpasses its competitors in terms of fitting the two datasets.

## REFERENCES

- [1] A. Ahmad, S. P. Ahmad, and A. Ahmed. Characterization and estimation of weibull-rayleigh distribution with applications to life time data. *Appl. Math. Inf. Sci. Lett*, 5:71–79, 2017.
- [2] A. Ahmed, S. P. Ahmad, and J. Reshi. Bayesian analysis of rayleigh distribution. *International Journal of Scientific and Research Publications*, 3(10):1–9, 2013.
- [3] M. Ajami and S. Jahanshahi. Parameter estimation in weighted rayleigh distribution. *Journal of Modern Applied Statistical Methods*, 16(2):14, 2017.
- [4] A. A. Bhat and S. P. Ahmad. A new generalization of rayleigh distribution: Properties and applications. *Pakistan journal of statistics*, 36(3), 2020.
- [5] A. A. Bhat and S. P. Ahmad. Mixture of gamma and rayleigh distributions: Mathematical properties and applications. *Journal of Applied Probability*, 16(2):81–97, 2021.
- [6] A. A. Bhat and S. P. Ahmad. An extension of exponentiated rayleigh distribution: Properties and applications. *Thailand Statistician*, 21(1):209–227, 2023.
- [7] A. A. Bhat, S. P. Ahmad, E. M. Almetwally, N. Yehia, N. Alsadat, and A. H. Tolba. The odd lindley power rayleigh distribution: properties, classical and bayesian estimation with applications. *Scientific African*, 20:e01736, 2023.
- [8] M. Kilai, G. A. Waititu, W. A. Kibira, M. Abd El-Raouf, and T. A. Abushal. A new versatile modification of the rayleigh distribution for modeling covid-19 mortality rates. *Results in Physics*, 35:105260, 2022.
- [9] S. Kotz, N. Balakrishnan, C. B. Read, and B. Vidakovic. *Encyclopedia of Statistical Sciences, Volume 1*. John Wiley & Sons, 2005.
- [10] D. Kumar, U. Singh, and S. K. Singh. A new distribution using sine function-its application to bladder cancer patients data. *Journal of Statistics Applications & Probability*, 4(3):417, 2015.
- [11] D. Kundu and M. Z. Raqab. Generalized rayleigh distribution: different methods of estimations. *Computational statistics & data analysis*, 49(1):187–200, 2005.
- [12] Z. Mahmood, T. M Jawa, N. Sayed-Ahmed, E. Khalil, A. H. Muse, A. H. Tolba, et al. An extended cosine generalized family of distributions for reliability modeling: Characteristics and applications with simulation study. *Mathematical Problems in Engineering*, 2022, 2022.
- [13] A. H. Muse, A. H. Tolba, E. Fayad, O. A. Abu Ali, M. Nagy, M. Yusuf, et al. Modelling the covid-19 mortality rate with a new versatile modification of the log-logistic distribution. *Computational Intelligence and Neuroscience*, 2021, 2021.
- [14] S. U. Rasool and S. P. Ahmad. Power length biased weighted lomax distribution. *Reliability: Theory & Applications*, 17(4 (71)):543–558, 2022.
- [15] L. Rayleigh. Xii. on the resultant of a large number of vibrations of the same pitch and of arbitrary phase. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 10(60):73–78, 1880.
- [16] 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.
- [17] J. Surles and W. Padgett. Inference for reliability and stress-strength for a scaled burr type x distribution. *Lifetime data analysis*, 7:187–200, 2001.

# NEW DISCRETE DISTRIBUTION FOR ZERO-INFLATED COUNT DATA

Peer Bilal Ahmad<sup>1</sup> Mohammad Kafeel Wani<sup>2</sup>

•

<sup>1,2</sup>Department of Mathematical Sciences, Islamic University of Science & Technology, Kashmir,  
INDIA, 192122

<sup>1,\*</sup>[bilalahmadpz@gmail.com](mailto:bilalahmadpz@gmail.com) , <sup>2</sup>[wanimk5@gmail.com](mailto:wanimk5@gmail.com)

## Abstract

*Over-dispersed models are commonly utilized when the variation is more than what the model actually predicts. Since one of the reasons for over-dispersion is the large number of zeros, we employ zero-inflated models instead of more traditional ones to handle this observed occurrence. We present a zero-inflated version of a discrete distribution that was developed in 2021 in our research. Significant statistical characteristics of the suggested model have been identified, such as moments, the over-dispersion feature, generating functions, and related measures, among others. We have carried the parametric estimation using the maximum likelihood estimate. Maximum likelihood estimates are checked for usefulness in a simulation exercise. We evaluated the applicability of our developed model using three real-world data sets,*

**Keywords:** Over-dispersion, Zero-inflation, Discrete distribution, Simulation, Goodness-of-fit, Testing of hypothesis.

## I. Introduction

To perform statistical analysis, statisticians use one of several methods, and these methods are the building blocks of statistical models. Mathematical representations of observable data are provided by statistical models. We choose statistical modeling of data for the purpose of understanding a wide range of random events across disciplines. Its applications are not limited to mathematical and statistical studies; rather, they permeate a wide variety of fields of study. Count data plays an important role in almost every scientific study, no matter how big or small. This data is used to draw inferences in relation to the population from which it is collected but, typically this data exhibits more variation than what is predicted by our hypothesized model. More precisely, this observable fact is called as over-dispersion (variance goes beyond mean). One cause of over-dispersion in count data is the presence of many more zeros than predicted by a statistical model. This phenomenon of finding excessive number of zeros is referred to as zero-inflation, and in order to model such dilemma, we use zero-inflated models rather than the more often used standard models.

Over-dispersion in count data due to zero-inflation is common, thus researchers are always developing new ideas and methods to shed light on this phenomenon. In order to deal with an excessive amount of zeros in count data, Lambert developed a new model called as zero-inflated Poisson (ZIP) regression model [7]. She used this model to investigate manufacturing flaws and

found that ZIP regression model is both simple and effective. Böhning argued that ZIP distribution is usually capable of easily dealing with the situation where there is an excessive quantity of zero counts [3]. To deal with count data containing an excessive number of zeros and ones, Melkersson & Olsson offered an improved version of the ZIP distribution, which he named as the zero-one inflated Poisson (ZOIP) distribution [9]. Yau *et al.* presented a mixed regression model of zero-inflated Negative Binomial (ZINB) to examine pancreatic disorder Length of Stay (LOS) times that account for same-day discharges [17]. Gilthroe *et al.* took into account biological count data with an excessive quantity of zeros, and he sought to address variety of factors [5]. An overview of the field of statistical modeling of over-dispersed data was provided by discussing its antecedents, motivations, pioneering contributions, major milestones, and practical uses [16]. Zhang *et al.* made an effort to investigate the characteristics and patterns of ZOIP distribution [19]. In order to evaluate the capacity to incorporate zero-inflation and over-dispersion in count data, Pittman *et al.* evaluated a number of methods, including ZIP, ZINB, and Hurdle Poisson (HUP) regression model [10]. Tüzen *et al.* analyzed the implementation of count data models using simulated data, which allowed for a wide range of outliers and zero-inflation scenarios [14]. They considered Poisson, Negative-Binomial, ZIP, ZINB, HUP and Negative-Binomial Hurdle models to check the compatibility of these models in presence of outliers and excess zeros. Bodhisuwan & Kehler proposed a new distribution called the zero-inflated Negative-Binomial-Exponential (ZINB-E) distribution [2]. To address the issue of too many zeros in count data, Rivas & Campos introduced the zero-inflated Waring (ZIW) distribution [11]. If the Waring distribution can't sufficiently characterize the behavior of the data, as is often the case when there is a large frequency of observed zeros, then the ZIW distribution is thought to be a better fit. Young, Roemmele & Shi evaluated a study that provided a snapshot of the current level of knowledge in the field of zero-inflation [18]. Ahmad & Wani introduced a compound model for handling over-dispersed count data. They used four different data sets and compared the fit with several potential models of interest. The fitting results showed the flexibility of the devised model in handling over-dispersed count data [1]. One of the recent works in zero-inflation aspect of the count data is by Wani & Ahmad [15]. They put forward the zero-inflated version of Poisson-Akash distribution. Much advancement has been made in this field of statistical modeling, yet there is still a consistent need for new models to be created. These new models are driven by the regular emergence of unexpected patterns in count data. We intend to extend this contribution by a devising a new zero-inflated model with a very clear-cut Probability function.

## 2. Zero-Inflated Discrete Distribution

A discrete distribution (DD) was proposed by Jain *et al.* in 2021 by discretization of a continuous distribution [6]. If  $Z$  follows the DD, then the probability mass function (PMF) of the Discrete Distribution is given as follows

$$P(Z = z) = \frac{\theta - 1}{\theta^{z+1}} \quad ; \theta > 1, z = 0, 1, 2, 3, \dots \quad (1)$$

The Discrete distribution (1) is itself an over-dispersed model but it also suffers at times to handle the excessive number of zeros in count data. We have thus made an effort to put forward the zero-inflated version of Discrete distribution. If  $X$  is a random variable following the Discrete distribution with parameter  $\theta > 1$  and  $\alpha$  ( $0 < \alpha < 1$ ) is the extra amount added to the proportion of point zero (zero-inflated distribution), then the probability mass function (PMF) of zero-inflated Discrete distribution (ZIDD) can be written as follows

$$P(X = x) = \begin{cases} \alpha + (1 - \alpha) \frac{\theta - 1}{\theta} & x = 0 \\ (1 - \alpha) \frac{\theta - 1}{\theta^{x+1}} & x = 1, 2, 3, \dots \end{cases} \quad (2)$$

The Cumulative distribution function (CDF) of ZIDD can be expressed as

$$F_x(X) = 1 - (1 - \alpha) \frac{1}{\theta^{x+1}} \quad (3)$$

It can be seen from the plots of PMF given in Fig. 1 that the model has mode at point zero. Furthermore, it is positively skewed and the tail shows a rapid decrease as parameters take higher values.

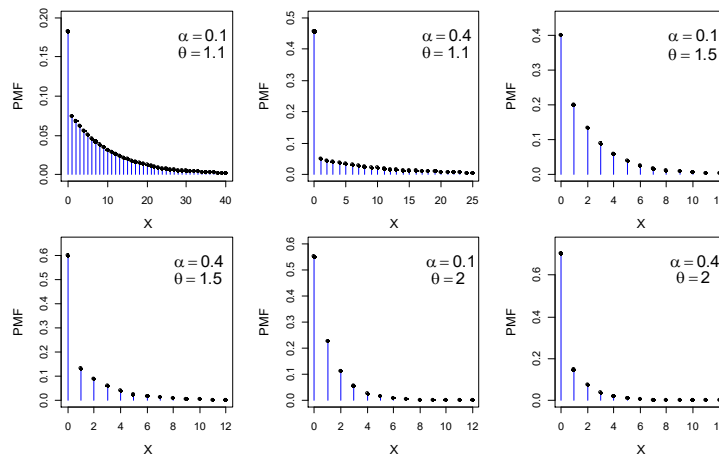


Fig. 1: Plots of Probability mass function of ZIDD for different choices of parameter values

### 3. Statistical Properties

In this section, we have derived some vital statistical characteristics of the newly developed model.

#### 3.1 Generating Functions

When dealing with discrete random variables, the Probability Generating Function (PGF) is an important tool. Its major benefit is that it makes it straightforward for us to explain the distribution of  $X+Y$  when they are independent. The PGF of ZIDD can be obtained as

$$P_x(t) = E(t^x) = \sum_{x=0}^{\infty} t^x P(X = x)$$

$$P_x(t) = t^0 \left( \alpha + (1 - \alpha) \frac{\theta - 1}{\theta} \right) + (1 - \alpha) \left( \frac{\theta - 1}{\theta} \right) \sum_{x=1}^{\infty} \left( \frac{t}{\theta} \right)^x$$

$$P_x(t) = \alpha + (1 - \alpha) \left( \frac{\theta - 1}{\theta} \right) + (1 - \alpha) \left( \frac{\theta - 1}{\theta} \right) \left( \frac{t}{\theta} \right) \left( \frac{\theta}{\theta - t} \right)$$

$$P_x(t) = \alpha + (1 - \alpha) \left( \frac{\theta - 1}{\theta - t} \right) \tag{4}$$

In equation (4), take  $t = e^t$ , that will yield the Moment Generating Function of ZIDD as follows

$$M_x(t) = \alpha + (1 - \alpha) \left( \frac{\theta - 1}{\theta - e^t} \right) \tag{5}$$

### 3.2 Moments and Related Characteristics

The  $r$ th moment about origin (Raw Moment) of ZIDD is obtained by employing its PMF (2). It follows that

$$\begin{aligned} E(X^r) &= \sum_{\forall x} x^r P(X = x) \\ E(X^r) &= 0^r [\alpha + (1 - \alpha)P_{DD}(X = 0)] + 1^r [(1 - \alpha)P_{DD}(X = 1)] + 2^r [(1 - \alpha)P_{DD}(X = 2)] + \dots \\ E(X^r) &= (1 - \alpha) \sum_{x=0}^{\infty} X^r [P_{DD}(X = x)] \\ E(X^r) &= (1 - \alpha)E_{DD}(X^r) \end{aligned}$$

With  $P_{DD}(X=x)$  and  $E_{DD}(X^r)$  representing the PMF and the  $r$ th Raw Moment of the baseline model respectively.

As a result, the mean and variance of ZIDD comes out to be as follows

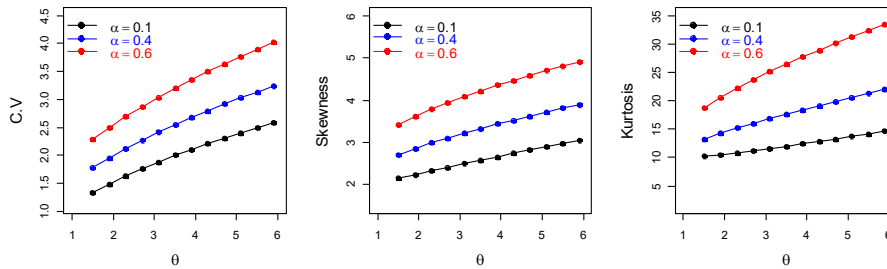
$$\text{Mean} = (1 - \alpha) \frac{1}{\theta - 1} \qquad \text{Variance} = (1 - \alpha) \frac{\alpha + \theta}{(\theta - 1)^2}$$

Some of the statistical properties of our proposed model can be expressed by means of Raw and Central moments. These properties include Index of Dispersion (IOD), Coefficient of Variation (C.V), Coefficient of Skewness and Coefficient of Kurtosis.

$$\begin{aligned} IOD &= \frac{\sigma^2}{\mu_1'} = \frac{\alpha + \theta}{\theta - 1} & C.V &= \frac{\sigma}{\mu_1'} = \left( \frac{\alpha + \theta}{1 - \alpha} \right)^{1/2} \\ Skewness &= \frac{\mu_3}{\mu_2^{3/2}} = \frac{\theta^2 + \theta + 3\alpha\theta + 2\alpha^2 - \alpha}{\sqrt{1 - \alpha}(\alpha + \theta)^{3/2}} \\ Kurtosis &= \frac{\mu_4}{\mu_2^2} = \frac{\theta^3 + 7\theta^2 + \theta + 2 + 4\alpha\theta^2 + 6\alpha^2\theta + 4\alpha\theta + \alpha^3 + 3\alpha^2 - 5\alpha}{(1 - \alpha)(\alpha + \theta)^2} \end{aligned}$$

From the plots of Coefficient of Skewness and Kurtosis given in Fig. 3, it can be noted that both of these increase monotonically for greater values of the parameter. Moreover, our proposed model possesses positive skewness and a leptokurtic shape.



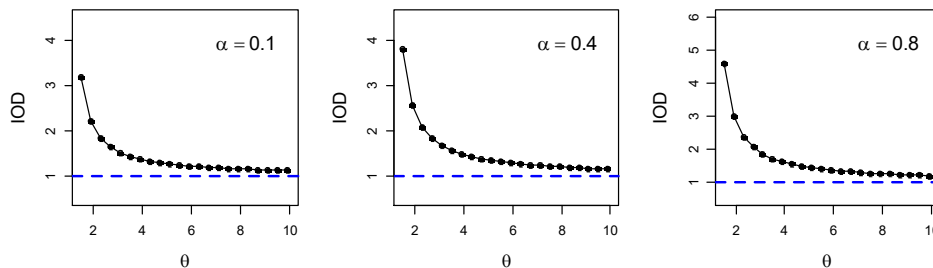


**Fig. 2:** Plots of Coefficient of Variation, Skewness and Kurtosis for some values of parameters

A significant characteristic enjoyed by our proposed model lies in the fact that it is always over-dispersed i.e., variance is always going to surpass the mean. We have

$$Varinace = (1-\alpha) \frac{\alpha + \theta}{(\theta-1)^2} = Mean + (1-\alpha) \frac{(1+\alpha)}{(\theta-1)^2}$$

The second term is obviously positive as  $\theta > 1$  and  $\alpha$  ( $0 < \alpha < 1$ ). This proves the over-dispersion property of ZIDD. The Index of Dispersion has also been plotted (see Fig. 3) for a choice of parameter values, which graphically demonstrates the over-dispersion of the model.



**Fig. 3:** Index of Dispersion plots for various values of parameters

#### 4. Parametric Inference

The foundation for estimation is in actual fact clear-cut. Knowing the parameter put forwards information concerning the entire population when sampling is done from a population that is represented by a specific distribution. So, it makes sense to carry out the estimation of parameters. The maximum likelihood estimation (MLE) is more often used for the reason that it enjoys greater efficiency and improved numerical stability. MLE is a statistical technique for estimating the parameters of a probability distribution that is assumed given some observed data.

The likelihood function of ZIDD can be defined as follows if  $x_1, x_2, \dots, x_n$  is the random sample of size  $n$  from ZIDD and  $Y$  is the number of  $x_i$ 's having value zero

$$L = \left( \alpha + (1-\alpha) \frac{\theta-1}{\theta} \right)^Y \prod_{\substack{i=1 \\ x_i \neq 0}}^n (1-\alpha) \frac{\theta-1}{\theta^{x_i+1}}$$

The log-likelihood can thus be written as

$$\log L = Y \log(\alpha + \theta - 1) - (n + Y + \sum_{\substack{i=1 \\ x_i \neq 0}}^n x_i) \log \theta + (n - Y) \log(1 - \alpha) + (n - Y) \log(\theta - 1)$$

In this study we have used *firdistrplus* in R-software to obtain the ML estimates [4].

### 5. Simulation Study

A simulation study has been undertaken in this part to evaluate the finite sample performance of the ML estimates of ZIDD. We attempt the Monte Carlo Simulation study by employing the *discrete inverse transform method*. In order to calculate Average Values (AVs), Average Biases (ABs), Mean Square Errors (MSEs), Mean Relative Estimates (MRESs), Mean Relative Errors (MRERs), and Average Dispersion Indices (AVDIs), we considered four different values for parameter and repeated the course of action N=1000 times starting from a small sample to large sample (n=25, 75, 100, 300, 600). The results are provided in Table 1. As it can be seen from Table 1, the ML estimates are asymptotically unbiased and consistent.

**Table 1:** Simulation results for maximum likelihood estimates of parameters of proposed model

	$\hat{\theta}$					$\hat{\alpha}$				
	AVs	ABs	MSEs	MRESs	MRERs	AVs	ABs	MSEs	MRESs	MRERs
n	Parameter Set 1: $\theta=1.5$ $\alpha=0.4$									
25	1.614	0.114	0.175	1.076	0.163	0.387	0.012	0.035	0.968	0.3903
75	1.508	0.008	0.011	1.005	0.055	0.401	0.001	0.010	1.004	0.2049
100	1.529	0.029	0.012	1.019	0.056	0.385	0.014	0.007	0.984	0.1713
300	1.503	0.003	0.002	1.002	0.024	0.398	0.001	0.002	0.996	0.0876
600	1.504	0.004	0.001	1.002	0.021	0.395	0.004	0.000	0.998	0.0213
n	Parameter Set 2: $\theta=1.8$ $\alpha=0.4$									
25	1.959	0.159	0.527	1.088	0.214	0.380	0.019	0.049	0.950	0.462
75	1.856	0.056	0.046	1.031	0.094	0.396	0.003	0.018	0.992	0.263
100	1.793	0.006	0.036	0.996	0.087	0.387	0.012	0.011	0.967	0.221
300	1.809	0.009	0.009	1.005	0.039	0.396	0.003	0.003	0.991	0.114
600	1.809	0.009	0.007	1.005	0.039	0.403	0.003	0.001	1.007	0.089
n	Parameter Set 3: $\theta=2.0$ $\alpha=0.2$									
25	2.238	0.238	0.574	1.110	0.235	0.172	0.027	0.034	0.862	0.819
75	2.028	0.028	0.066	1.014	0.089	0.186	0.013	0.017	0.933	0.532
100	2.022	0.022	0.064	1.011	0.090	0.205	0.005	0.015	1.027	0.528
300	2.017	0.017	0.024	1.008	0.0621	0.188	0.011	0.007	0.943	0.329
600	1.994	0.005	0.006	0.997	0.0333	0.209	0.009	0.002	1.049	0.186
n	Parameter Set 4: $\theta=2.5$ $\alpha=0.3$									
25	2.674	0.174	1.501	1.069	0.260	0.288	0.011	0.054	0.960	0.667
75	2.495	0.044	0.204	0.998	0.157	0.306	0.006	0.038	1.023	0.565
100	2.542	0.042	0.171	1.016	0.130	0.306	0.006	0.017	1.021	0.354
300	2.565	0.035	0.083	1.026	0.095	0.296	0.003	0.012	0.988	0.317
600	2.511	0.011	0.027	1.004	0.055	0.299	0.001	0.005	0.997	0.190

### 6. Data Fitting

Real life datasets from different fields have been employed to test the compatibility of our proposed model in presence of over-dispersion caused by zero-inflation. In addition to this, we compared the fitting results from our proposed model with other statistical models of competing interest. The models with which we have compared our devised models include Poisson distribution (PD), zero-inflated Poisson distribution (ZIPD), zero-inflated Negative-Binomial

distribution (ZINBD), Discrete Weibull distribution (DWD) and Discrete distribution (DD). In order to estimate the parameters of each distribution, we used maximum likelihood estimation method.

### 6.1 Data set 1

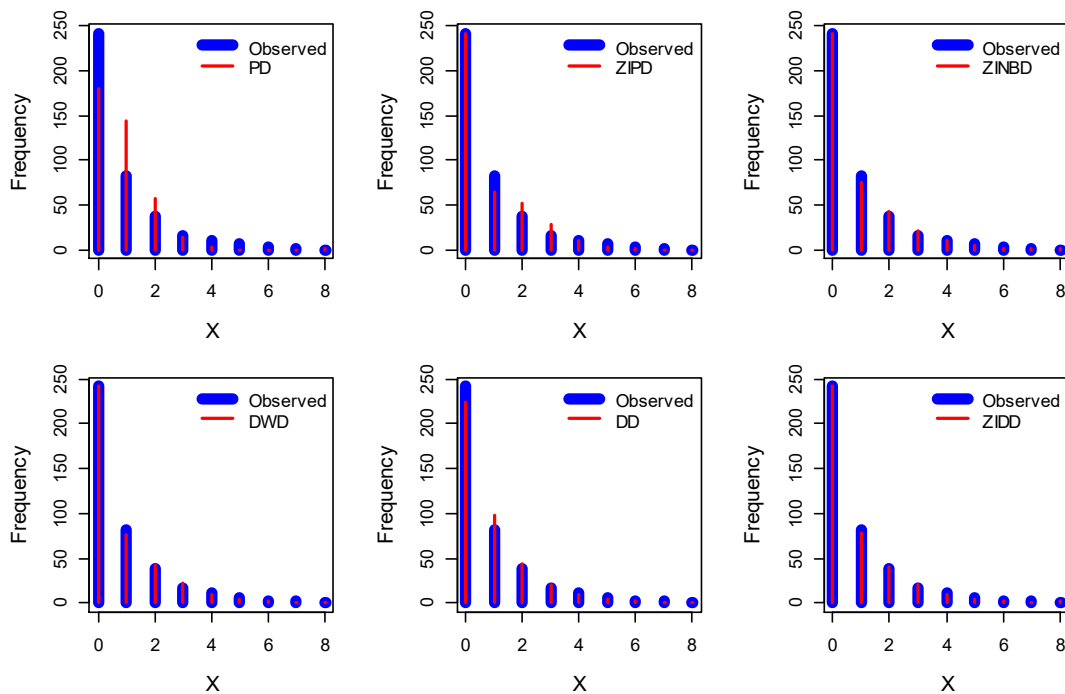
This dataset stands for the observed number of households according to total number of migrants [13]. The data is expressed in Table 2 and the performance of fitting this data is summarized in Table 3. From the fitting results, it is obvious that our model performs better than other competing models of interest. The plots for observed an expected frequencies under different models given in Fig. 4 provides a clearer view of the fitting results.

**Table 2: Data set 1**

Number of Households	0	1	2	3	4	5	6	7	8
Observed Frequency	242	82	38	17	11	7	3	2	0

**Table 3: Fitting results of Data set 1**

Model	$\chi^2$	d.f	p-value	L	AIC	BIC
PD	118.64	3	0.0001	-555.060	1112.121	1116.117
ZIPD	16.80	2	0.0002	-501.796	1007.592	1015.585
ZINBD	3.29	2	0.1930	-492.597	991.195	1003.185
DWD	1.57	4	0.8141	-492.362	988.724	996.717
DD	9.41	4	0.0516	-495.785	993.571	997.567
ZIDD	1.51	4	0.8248	-492.030	988.060	996.053



**Fig. 4: Observed and expected frequencies plots for PD, ZIPD, ZINBD, DWD, DD, and ZIDD for Data set 1**

### 6.2 Data set 2

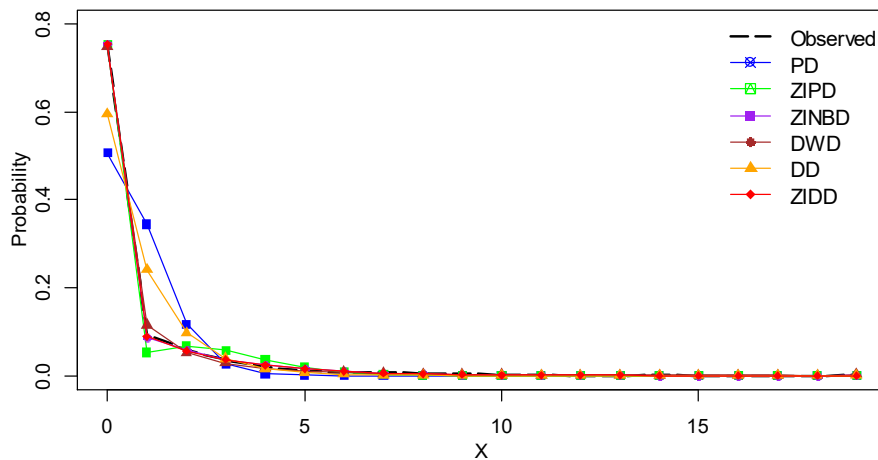
This dataset presents the number of spots in southern beetle [8]. The data is presented in Table 4 and the fitting results are given in Table 5. The performance measures indicate that our model suffers minimum loss compared to other models and the value of Chi-square is comparatively smaller. Moreover, the plots for observed and expected PMFs are given in Fig. 5.

**Table 4:** Data set 2

Number of Spots	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Observed Frequency	1169	144	92	54	29	18	10	12	6	9	3	2	0	0	1	0	0	0	0	1

**Table 5:** Fitting results of Data set 2

Model	$\chi^2$	d.f	p-value	L	AIC	BIC
PD	1361.80	4	0.0001	-2291.139	4584.277	4589.623
ZIPD	184.65	5	0.0001	-1648.989	3301.978	3312.670
ZINBD	9.94	6	0.1272	-1554.612	3115.223	3131.261
DWD	14.30	7	0.0460	-1560.532	3125.064	3135.756
DD	432.62	5	0.0001	-1757.427	3516.855	3522.201
ZIDD	7.96	7	0.3361	-1554.001	3112.002	3122.694



**Fig. 5:** Plots of observed and expected PMFs under PD, ZIPD, ZINBD, DWD, DD, and ZIDD for Data set 2

### 6.3 Data set 3

This dataset represents the number of units of Brand K of Chatfield bought by numbers of consumers over a number of weeks [12]. The data set is given in Table 6 and the Table 7 presents

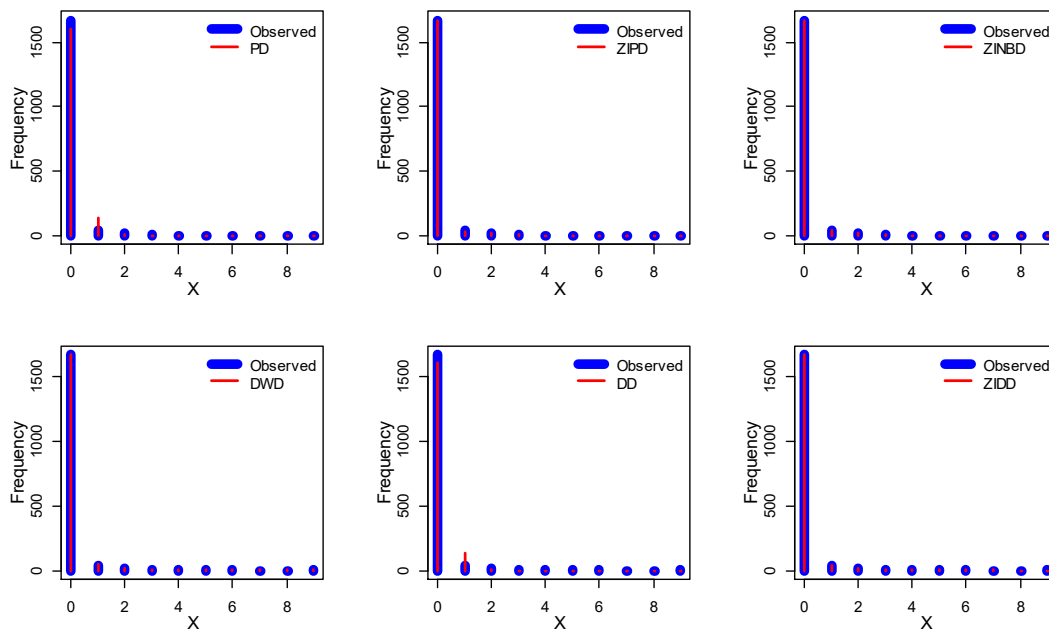
the fitting results of this data set. The results from the fitting table prove that the devised model shows better fitting compared to other competing models. In addition to this, the expected frequencies are closer to observed frequencies in case of our model (see Fig. 6).

**Table 6:** Data set 3

Brand K	0	1	2	3	4	5	6	7	8	9
Number of Consumers	1671	43	19	9	2	3	1	0	0	2

**Table 7:** Fitting results of data set 3

Model	$\chi^2$	d.f	p-value	L	AIC	BIC
PD	221.99	1	0.0001	-612.908	1227.8171	1233.2840
ZIPD	6.28	2	0.0433	-439.794	883.5894	894.5241
ZINBD	1.40	1	0.0942	-429.660	865.3218	881.7239
DWD	2.48	2	0.2893	-429.407	862.8156	873.7504
DD	118.40	1	0.0001	-537.381	1076.763	1082.230
ZIDD	0.48	2	0.7866	-429.334	862.6682	873.6030



**Fig. 6:** Observed and expected frequency plots for Data set 3 under PD, ZIPD, ZINBD, DWD, DD, and ZIDD

### 7. Testing of Hypothesis

In order to test the significance of the zero-inflation parameter of our proposed model, we take on different test statistics to test the null hypothesis given as follows

$$H_0 : \alpha = 0 \text{ vs. the alternative hypothesis } H_1 : \alpha > 0$$

#### 7.1 Likelihood Ratio test

The Likelihood ratio test (LRT) evaluates the ratio of two log-likelihood functions in order to test

the null hypothesis  $H_0$  against an alternative hypothesis  $H_1$ . In case of LRT, the test statistic is

$$LRT_{\alpha} = -2 \times [L(\hat{\theta}) - L(\hat{\alpha}, \hat{\theta})],$$

where  $L(\hat{\theta})$  and  $L(\hat{\alpha}, \hat{\theta})$  are the maximum log-likelihood under DD and ZIDD respectively. The LRT test statistic is asymptotically distributed as Chi-square with one degree of freedom. The LRT for all the three datasets is respectively given as

$$LRT_1=7.51, \quad LRT_2=406.852, \quad LRT_3=216.094.$$

## 7.2 Wald test

This test is used to determine the presence or absence of an effect. In this section, we will construct a Wald test for the effect of zero-inflation parameter in our proposed model. The test statistic under Wald test is given by

$$Wald_{\alpha} = \frac{\hat{\alpha}^2}{Var(\hat{\alpha})},$$

where  $Var(\hat{\alpha})$  represents the pertinent diagonal component of the Fisher information matrix calculated at  $\alpha = \hat{\alpha}$  and  $\theta = \hat{\theta}$ . The Wald test statistic is asymptotically distributed as Chi-square with one degree of freedom. The Wald test statistic value for all the three datasets can be correspondingly given as

$$Wald_1=9.53, \quad Wald_2=1014.12, \quad Wald_3=517.44.$$

On comparing the LRT and Wald test values from all data sets with the critical value (3.84), we reject the null hypothesis in case of all the three tests and draw the conclusion that the zero-inflation parameter in our proposed model is of significant importance.

## 8. Conclusion

In this research we have made an effort to present a new zero-inflated count data model. It was investigated how the probability mass function behaves for varied values of parameters. We discussed some important statistical properties of our proposed model. Simulation study was carried out to test the performance of maximum likelihood estimates and the results were pretty much significant. For the testing the compatibility of our proposed model, we tested the proposed distribution on real datasets using different performance metrics like Chi-square Goodness-of-fit, AIC, BIC etc. Moreover, we compared the fitting results of our devised model with other competing models. The results verified that our proposed model is adaptable and can be considered for handling over-dispersion in count data caused by zero-inflation. Finally, we carried out the Likelihood Ratio test and the Wald test on all datasets to see the significance of zero-inflation parameter and the results were significant.

## Acknowledgements

We are highly thankful to the editor-in-chief and the referees. The second author is particularly thankful to the Department of Science and Technology (Government of India) for INSPIRE fellowship (DST/INSPIRE/03/2022/002460).

## Conflict of Interest

The authors report no conflict of interest.

## References

- [1] Ahmad, P. B., & Wani, M. K. (2023). A New Compound Distribution and Its Applications in Over-dispersed Count Data. *Annals of Data Science*. <https://doi.org/10.1007/s40745-023-00478-0>
- [2] Bodhisuwan, R., & Kehler, A. (2021). The Zero-inflated Negative Binomial-Exponential Distribution and Its Application. *Lobachevskii Journal of Mathematics*, 42(2), 300–307. <https://doi.org/10.1134/s1995080221020062>
- [3] Böhning, D. (1998). Zero-Inflated Poisson Models and C.A.MAN: A Tutorial Collection of Evidence. *Biometrical Journal*, 40(7), 833–843. [https://doi.org/10.1002/\(sici\)1521-4036\(199811\)40:7<833::aid-bimj833>3.0.co;2-o](https://doi.org/10.1002/(sici)1521-4036(199811)40:7<833::aid-bimj833>3.0.co;2-o)
- [4] Delignette-Muller, M. L., & Dutang, C. (2015). *Fitdistrplus: An R Package for Fitting Distributions*. *Journal of Statistical Software*, 64(4). <https://doi.org/10.18637/jss.v064.i04>
- [5] Gilthorpe, M. S., Frydenberg, M., Cheng, Y., & Baelum, V. (2009). Modelling count data with excessive zeros: The need for class prediction in zero-inflated models and the issue of data generation in choosing between zero-inflated and generic mixture models for dental caries data. *Statistics in Medicine*, 28(28), 3539–3553. Portico. <https://doi.org/10.1002/sim.3699>
- [6] Jain, S., Siddiqui, S. A., Dwivedi, S., Siddiqui, I., & Kamal, M. (2021). A NEW DISCRETE DISTRIBUTION WITH ITS MATHEMATICAL PROPERTIES. *Int. J. Agricult. Stat. Sci. Vol*, 17(2), 693-697.
- [7] Lambert, D. (1992). Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics*, 34(1), 1. <https://doi.org/10.2307/1269547>
- [8] Lin, S. K. (1985). Characterization of lightning as a disturbance to the forest ecosystem in East Texas (Doctoral dissertation, Texas A & M University).
- [9] Melkersson, M., & Olsson, C. (1999). Is visiting the dentist a good habit?: Analyzing count data with excess zeros and excess ones. *Univeristy of Umeå*.
- [10] Pittman, B., Buta, E., Krishnan-Sarin, S., O'Malley, S. S., Liss, T., & Gueorguieva, R. (2018). Models for Analyzing Zero-Inflated and Overdispersed Count Data: An Application to Cigarette and Marijuana Use. *Nicotine & Tobacco Research*, 22(8), 1390–1398. <https://doi.org/10.1093/ntr/nty072>
- [11] Rivas, L., & Campos, F. (2021). Zero inflated Waring distribution. *Communications in Statistics – Simulation and Computation*, 1–16. <https://doi.org/10.1080/03610918.2021.1944638>
- [12] Shoukri, M. M., & Consul, P. C. (1987). Some Chance Mechanisms Generating the Generalized Poisson Probability Models. *Biostatistics*, 259–268. [https://doi.org/10.1007/978-94-009-4794-8\\_15](https://doi.org/10.1007/978-94-009-4794-8_15)
- [13] Shukla, K. K., Shanker, R., & Tiwari, M. K. (2021). A new one parameter discrete distribution and its applications. *Journal of Statistics and Management Systems*, 25(1), 269–283. <https://doi.org/10.1080/09720510.2021.1893475>
- [14] Tüzen, F., Erbaş, S., & Olmuş, H. (2018). A simulation study for count data models under varying degrees of outliers and zeros. *Communications in Statistics – Simulation and Computation*, 49(4), 1078–1088. <https://doi.org/10.1080/03610918.2018.1498886>
- [15] Wani, M. K., & Ahmad, P. B. (2023). Zero-inflated Poisson-Akash distribution for count data with excessive zeros. *Journal of the Korean Statistical Society*. <https://doi.org/10.1007/s42952-023-00216-5>
- [16] Kekalaki, E. (2014). On the distribution theory of over-dispersion. *Journal of Statistical Distributions and Applications*, 1(1). <https://doi.org/10.1186/s40488-014-0019-z>
- [17] Yau, K. K. W., Wang, K., & Lee, A. H. (2003). Zero-Inflated Negative Binomial Mixed Regression Modeling of Over-Dispersed Count Data with Extra Zeros. *Biometrical Journal*, 45(4), 437–452. <https://doi.org/10.1002/bimj.200390024>

[18] Young, D. S., Roemmele, E. S., & Shi, X. (2022). Zero-inflated modeling part II: Zero-inflated models for complex data structures. *Wiley Interdisciplinary Reviews: Computational Statistics*, 14(2), e1540. <https://doi.org/10.1002/wics.1540>

[19] Zhang, C., Tian, G.-L., & Ng, K.-W. (2016). Properties of the zero-and-one inflated Poisson distribution and likelihood-based inference methods. *Statistics and Its Interface*, 9(1), 11–32. <https://doi.org/10.4310/sii.2016.v9.n1.a2>



# STOCHASTIC OPTIMIZATION AND RELIABILITY ANALYSIS OF MUSHROOM PLANT

Shakuntla Singla<sup>1</sup>, Sonia<sup>2,\*</sup>, Poonam Panwar<sup>3</sup>

<sup>1,2\*</sup> Department of Mathematics and Humanities, MMEC, Maharishi Markandeshwar  
(Deemed to be University), Mullana, Ambala-133207, India

<sup>3</sup>MM Institute of Computer Technology and Business Management, Maharishi Markandeshwar  
(Deemed to be University), Mullana, Ambala-133207, India

<sup>1</sup>shakus25@gmail.com , <sup>2\*</sup>sonia\_garg99@yahoo.com , <sup>3</sup>rana.poonam1@gmail.com

## Abstract

*In the present paper the reliability model for availability analysis of mushroom plant is developed in three sub-units like water pump, winter cold standby unit A.C., and packing machine. We assume a doctor of mushroom and workers are available who examines and repairs the elements as when we need. A mathematical model of the system is developed by using all these considerations. MTSF, Availability, server of busy period and expected number of servers visit of mushroom plant are determined with the assistance of RPGT. Graphs and tables are draw to depict the behavior of various parameters such as MTSF, Availability, server of busy period and expected number of servers visits and the effect of various parameters of the plant is analyzed when repair and failure rate both are vary and also when one of them is constant*

**Keywords:** Availability, MTSF, RPGT, Straw.

## 1. Introduction

In Modern time, Production have vided variety from modest to complicate; So, mushroom manufacturer must have superior strategy of optimum accessibility for optimum grouping elements. In The era of competition, all mushroom manufacturer face challenges for assurance ideal manufacturing charges and nominal period to achieve implementation and reliability. Mushroom plants need extremely hard work for production [1]. For production of mushroom, we need storage rooms and in a storage room, wooden beds are required to put the bags of raw food of mushrooms on them. To prepare fertilize for mushroom we need wheat straw which is easily available everywhere. In alternate of wheat straw, we use rice straw, mustard straw, lentils straw and guar straw and other things which can be used in preparing of compost are bran, chicken beet, urea, gypsum and water as shown in figure 1. Bihar is top most state for producing mushroom. Mushroom contains components calcium, phosphorus, Potassium, iron and copper [2]. In winter season (sept. to march), two types of crops of mushroom are there. For increasing quality and production of mushroom Govt. has open mushroom centre at Solan collaborating with UNDP and the purpose of this centre is to provide technical knowledge about production and creating interest in farmers. The cost of this project is Rs. 1.26 crore with the following objectives: To make availability of quality spawn and compost, provide the latest production technologies for mushroom farming and provide marketing facilities for cultivation and distribution [3]. Govt. provide financial help to increase the production of mushroom in the form of subsidies and provide free training to farmers for production of mushroom. Nowadays, Mushroom production become popular in whole of the world [4].

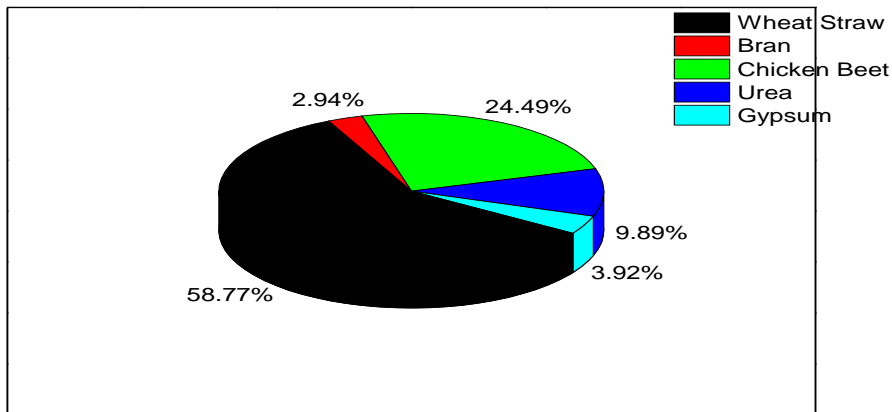





Figure 1. Compost for Mushroom

Approximately six days are required to prepare fertilizer. After this procedure production of mushroom take approximate 20 days. For our discussion, we take three units like water pump, winter cold standby unit and packing machine. A Doctor is required for examining whole activity [5]. In winter, demand of mushroom increases due to the benefits of mushroom for our health like decreases the risk of cancer, promote lower cholesterol, protect brain health, provide the source of vitamin D and support a healthy immune system [13][14]. Keeping in view the defective and maintenance charges are fixed while fluctuate other charges, their influence on grouping activity elements is shown by illustrating tables and charts, precede by discussions [6][7].

## 2. Assumptions and Notations

- Facility of doctor of mushroom is always available.
- Workers are available as required as we need.
- Repairs and failures are not dependent numerically [15].
- After Repairing, system is fully worked as the new one.
- The system is discussed in steady-state situations.

Table 1. Notations Used

Symbol	Represent
	Working state
	Regenerative state
	Failed state
$g_i$	Repair rate
$h_i$	Failure rate

### 3. System Description

The sub-system and their working are described as given below

- **Water Pump (A):** Water Pump for watering mushroom.
- **Winter cold stand by unit A.C. (B):** Production of mushroom required low temperature. For this purpose, mushroom produces in winter and in summer for this purpose uses of A.C. Button mushroom requires 20<sup>o</sup>-28<sup>o</sup>for vegetative growth 12-18<sup>o</sup>C for reproductive growth [16].
- **Packing Machine (M):** Seiler are used for packing mushroom.

### 4. Transition Diagrams

By taking all the described notations and assumptions [8-9], the Transition Diagram of the system is shown in Figure 2.

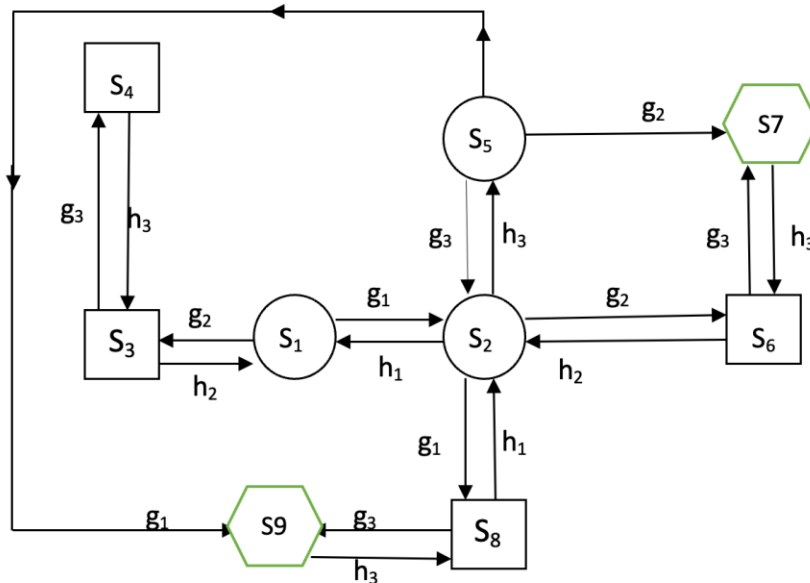


Figure 2. Transition Diagram

where,

- |                 |                |                 |                 |
|-----------------|----------------|-----------------|-----------------|
| $S_1 = ABB'$ ,  | $S_2 = AbB'$ , | $S_3 = aBB'$ ,  | $S_4 = aBB'M$ , |
| $S_5 = AbB'M$ , | $S_6 = abB'$ , | $S_7 = abB'M$ , | $S_8 = Abb'$ ,  |
| $S_9 = Abb'M$ , |                |                 |                 |

### 5. Model Description

A mushroom plant contains of following sub-units water pump(A), Winter (B) with cold standby unit B', Packing Machine (M). Implication order to repair the elements and system are  $M > A > B$ . In the start the sub-unit is in state  $S_1[ABB']$  where unit 'B', it's cold standby sub-unit, unit 'A' and server are in good operational condition, hence the framework works in full volume. The cold redundant sub-unit when decent is shown in (') which is prepared online directly with the assistance of a perfect switch over framework upon the disappointment of chief sub-unit 'B'. From stage  $S_1$  upon the disappointment of online unit 'B', disappointment rate of which is  $g_1$ , framework enters the stage  $S_2 [AbB']$ , here framework again works at full capacity as cold standby sub-unit is mode online. From

stage  $S_2$  upon repair of fizzled sub-unit, repair rate of which is  $h_1$ , framework again joins stage  $S_1$ . In stage  $S_1$ , if unit 'A' flops of which rate is  $g_2$ , upon its repair (repair rates  $h_2$ ) over the framework come again into the stage  $S_1$  while in stage  $S_3$  if it fails with failure rate  $g_3$ , framework enters the stage  $S_4$  [aBB'M] upon its repair (repair rate  $h_3$ ) framework re-enters the stage  $S_3$ . In stage  $S_2$  [AbB'] if online unit 'B' bombs at rates, the framework enters the stage  $S_8$  [Abb'], upon repair of unit 'B' at rate  $h_1$ . The scheme come again into stage  $S_2$  though in stage  $S_8$  if the M unit fails (whose disappointment rate is  $g_3$ ) structure joins the failed stage  $S_9$  [Abb'M] upon its repair, behavior of the structure rejoins the stage  $S_8$ , where it resumes repairing the fizzled sub-unit 'B'. Also, in stage  $S_2$  if unit 'A' fizzled at rate  $g_2$ , the framework takes stage  $S_6$  [abB'] upon repair of unit 'A' structure rejoins the stage  $S_2$  while in stage  $S_6$ . If associated fails at rate  $g_3$ , the framework takes the stage  $S_7$  [abB'M], upon its reparation as it is assumed top priority, the structure rejoins the stage  $S_6$ . In state  $S_2$  if the server failed the structure joins the stage  $S_5$  [AbB'M], here the structure continues to work of full volume, as the attendant is given top priority in repair, so upon its repair the structure rejoins stage  $S_2$ , in stage  $S_5$  if online sub-unit 'B' at rate  $g_1$ , the structure joins the stage  $S_9$  and if the sub-unit 'A' then the structure joins the stage  $S_7$ .

**Table 2.** Transition Probabilities

$q_{ij}(t)$	$p_{ij} = \int_0^{\infty} q_{ij}(t) dt$
$q_{1,2}(t) = g_1 e^{-(g_1 + g_2)t}$	$p_{1,2} = g_1 / (g_1 + g_2)$
$q_{1,3}(t) = g_2 e^{-(g_1 + g_2)t}$	$p_{1,3} = g_2 / (g_1 + g_2)$
$q_{2,1}(t) = h_1 e^{-(g_1 + g_2 + g_3 + h_1)t}$	$p_{2,1} = g_1 / (g_1 + g_2 + g_3 + h_1)$
$q_{2,5}(t) = g_3 e^{-(g_1 + g_2 + g_3 + h_1)t}$	$p_{2,5} = g_3 / (g_1 + g_2 + g_3 + h_1)$
$q_{2,6}(t) = g_2 e^{-(g_1 + g_2 + g_3 + h_1)t}$	$p_{2,6} = g_2 / (g_1 + g_2 + g_3 + h_1)$
$q_{2,8}(t) = g_1 e^{-(g_1 + g_2 + g_3 + h_1)t}$	$p_{2,8} = g_1 / (g_1 + g_2 + g_3 + h_1)$
$q_{3,1}(t) = h_2 e^{-(h_2 + g_3)t}$	$p_{3,1} = h_2 / (h_2 + g_3)$
$q_{3,4}(t) = g_3 e^{-(h_2 + g_3)t}$	$p_{3,4} = g_3 / (h_2 + g_3)$
$q_{4,3}(t) = h_3 e^{-h_3 t}$	$p_{4,3} = 1$
$q_{5,2}(t) = h_3 e^{-(g_1 + g_2 + h_3)t}$	$p_{5,2} = h_3 / (g_1 + g_2 + h_3)$
$q_{5,7}(t) = g_2 e^{-(g_1 + g_2 + h_3)t}$	$p_{5,7} = g_2 / (g_1 + g_2 + h_3)$
$q_{5,9}(t) = g_1 e^{-(g_1 + g_2 + h_3)t}$	$p_{5,9} = g_1 / (g_1 + g_2 + h_3)$
$q_{6,2}(t) = h_2 e^{-(g_3 + h_2)t}$	$p_{6,2} = h_2 / (g_3 + h_2)$
$q_{6,7}(t) = g_3 e^{-(g_3 + h_2)t}$	$p_{6,7} = g_3 / (g_3 + h_2)$
$q_{7,6}(t) = h_3 e^{-h_3 t}$	$p_{7,6} = 1$
$q_{8,2}(t) = h_1 e^{-(g_3 + h_1)t}$	$p_{8,2} = h_1 / (h_1 + g_3)$
$q_{8,9}(t) = g_3 e^{-(g_3 + h_1)t}$	$p_{8,9} = g_3 / (h_1 + g_3)$
$q_{9,8}(t) = h_3 e^{-h_3 t}$	$p_{9,8} = 1$

**Table 3.** Mean Sojourn Times

$R_i(t)$	$\mu_i = \int_0^{\infty} R_i(t) dt$
$R_1(t) = e^{-(g_1 + g_2)t}$	$\mu_1 = 1 / (g_1 + g_2)$
$R_2(t) = e^{-(g_1 + g_2 + g_3 + h_1)t}$	$\mu_2 = 1 / (g_1 + g_2 + g_3 + h_1)$
$R_3(t) = e^{-(h_2 + g_3)t}$	$\mu_3 = 1 / (h_2 + g_3)$
$R_4(t) = e^{-h_3 t}$	$\mu_4 = 1 / h_3$
$R_5(t) = e^{-(g_1 + g_2 + h_3)t}$	$\mu_5 = 1 / (g_1 + g_2 + h_3)$
$R_6(t) = e^{-(g_3 + h_2)t}$	$\mu_6 = 1 / (g_3 + h_2)$
$R_7(t) = e^{-h_3 t}$	$\mu_7 = 1 / h_3$
$R_8(t) = e^{-(h_1 + g_3)t}$	$\mu_8 = 1 / (h_1 + g_3)$
$R_9(t) = e^{-h_3 t}$	$\mu_9 = 1 / h_3$

### 5.1 Evaluation of Path Probabilities

Applying RPGT and use '1' as initial-stage of the structure, we detect all transition possibilities aspects of all accessible stages from initial stage 'ξ' = '1' [10] [11].

We will discover probabilities after state '1' to various vertices which are defined as follows:

$$V_{1,1} = 1 \text{ (Verified)}$$

$$V_{1,2} = (1, 2) / \{1 - (2, 5, 2)\} [1 - (2, 6, 2) / \{1 - (6, 7, 6)\}] [1 - (2, 8, 2) / \{1 - (8, 9, 8)\}]$$

$$= p_{1,2} / (1 - p_{2,5} p_{5,2}) [1 - \{(p_{2,6} p_{6,2} / (1 - p_{6,7} p_{7,6}))\}] [1 - \{(p_{2,8} p_{8,2} / (1 - p_{8,9} p_{9,8}))\}]$$

$$V_{1,3} = (1, 3) / \{1 - (3, 4, 3)\}$$

$$= p_{1,3} / (1 - p_{3,4} p_{4,3})$$

$$V_{1,4} = (1, 3, 4) / \{1 - (3, 4, 3)\}$$

$$= p_{1,3} p_{3,4} / (1 - p_{3,4} p_{4,3})$$

$$V_{1,5} = (1, 2, 5) / \{1 - (2, 5, 2)\} [1 - (2, 6, 2) / \{1 - (6, 7, 6)\}] [1 - (2, 8, 2) / \{1 - (8, 9, 8)\}]$$

$$= p_{1,2} p_{2,5} / (1 - p_{2,5} p_{5,2}) [1 - \{(p_{2,6} p_{6,2} / (1 - p_{6,7} p_{7,6}))\}] [1 - \{(p_{2,8} p_{8,2} / (1 - p_{8,9} p_{9,8}))\}]$$

$$V_{1,6} = (1, 2, 6) / \{1 - (2, 5, 2)\} [1 - (2, 6, 2) / \{1 - (6, 7, 6)\}] [1 - (2, 8, 2) / \{1 - (8, 9, 8)\}] \{1 - (6, 7, 6)\}$$

$$= (1, 2, 5, 7, 6) / \{1 - (2, 5, 2)\} [1 - (2, 6, 2) / \{1 - (6, 7, 6)\}] [1 - (2, 8, 2) / \{1 - (8, 9, 8)\}] \{1 - (6, 7, 6)\}$$

$$= p_{1,2} p_{2,6} / (1 - p_{2,5} p_{5,2}) [1 - \{(p_{2,6} p_{6,2} / (1 - p_{6,7} p_{7,6}))\}] [1 - \{(p_{2,8} p_{8,2} / (1 - p_{8,9} p_{9,8}))\}] (1 - p_{6,7} p_{7,6})$$

$$= p_{1,2} p_{2,5} p_{5,7} p_{7,6} / (1 - p_{2,5} p_{5,2}) [1 - \{(p_{2,6} p_{6,2} / (1 - p_{6,7} p_{7,6}))\}] [1 - \{(p_{2,8} p_{8,2} / (1 - p_{8,9} p_{9,8}))\}] (1 - p_{6,7} p_{7,6})$$

$$V_{1,7} = (1, 2, 5, 7) / \{1 - (2, 5, 2)\} [1 - (2, 6, 2) / \{1 - (6, 7, 6)\}] [1 - (2, 8, 2) / \{1 - (8, 9, 8)\}] \{1 - (8, 6, 8)\}$$

$$= (1, 2, 6, 7) / \{1 - (2, 5, 2)\} [1 - (2, 6, 2) / \{1 - (6, 7, 6)\}] [1 - (1, 7, 1) / \{1 - (7, 8, 7)\}] \{1 - (5, 6, 5)\}$$

$$= p_{1,2} p_{2,5} p_{5,7} / (1 - p_{2,5} p_{5,2}) [1 - \{(p_{2,6} p_{6,2} / (1 - p_{6,7} p_{7,6}))\}] [1 - \{(p_{2,8} p_{8,2} / (1 - p_{8,9} p_{9,8}))\}] (1 - p_{7,6} p_{6,7})$$

$$= p_{1,2} p_{2,6} p_{6,7} / (1 - p_{2,5} p_{5,2}) [1 - \{(p_{2,6} p_{6,2} / (1 - p_{6,7} p_{7,6}))\}] [1 - \{(p_{2,8} p_{8,2} / (1 - p_{8,9} p_{9,8}))\}]$$

$$(1 - p_{6,7} p_{7,6})$$

$$V_{1,8} = (1, 2, 8) / \{1 - (2, 5, 2)\} [1 - (2, 6, 2) / \{1 - (6, 7, 6)\}] [1 - (2, 8, 2) / \{1 - (8, 9, 8)\}] \{1 - (8, 9, 8)\}$$

$$= (1, 2, 8, 9, 8) / \{1 - (2, 5, 2)\} [1 - (2, 6, 2) / \{1 - (6, 7, 6)\}] [1 - (2, 8, 2) / \{1 - (8, 9, 8)\}] \{1 - (8, 9, 8)\}$$

$$= p_{1,2} p_{2,8} / (1 - p_{2,5} p_{5,2}) [1 - \{(p_{2,6} p_{6,2} / (1 - p_{6,7} p_{7,6}))\}] [1 - \{(p_{2,8} p_{8,2} / (1 - p_{8,9} p_{9,8}))\}] (1 - p_{8,9} p_{9,8})$$

$$= p_{1,2} p_{2,5} p_{5,9} p_{9,8} / (1 - p_{2,5} p_{5,2}) [1 - \{(p_{2,6} p_{6,2} / (1 - p_{6,7} p_{7,6}))\}] [1 - \{(p_{2,8} p_{8,2} / (1 - p_{8,9} p_{9,8}))\}] (1 - p_{8,9} p_{9,8})$$

$$V_{1,9} = (1, 2, 5, 9) / \{1 - (2, 5, 2)\} [1 - (2, 6, 2) / \{1 - (6, 7, 6)\}] [1 - (2, 8, 2) / \{1 - (8, 9, 8)\}] \{1 - (9, 8, 7)\}$$

$$= (1, 2, 8, 9) / \{1 - (2, 5, 2)\} [1 - (2, 6, 2) / \{1 - (6, 7, 6)\}] [1 - (2, 8, 2) / \{1 - (8, 9, 8)\}] \{1 - (9, 8, 9)\}$$

$$= p_{1,2} p_{2,5} p_{5,9} / (1 - p_{2,5} p_{5,2}) [1 - \{(p_{2,6} p_{6,2} / (1 - p_{6,7} p_{7,6}))\}] [1 - \{(p_{2,8} p_{8,2} / (1 - p_{8,9} p_{9,8}))\}] (1 - p_{9,8} p_{8,9})$$

$$= p_{1,2} p_{2,8} p_{8,9} / (1 - p_{2,5} p_{5,2}) [1 - \{(p_{2,6} p_{6,2} / (1 - p_{6,7} p_{7,6}))\}] [1 - \{(p_{2,8} p_{8,2} / (1 - p_{8,9} p_{9,8}))\}]$$

$$(1 - p_{9,8} p_{8,9})$$

Transition stage possibilities from base stage '2' are

$$V_{2,1} = (2, 1) / \{1 - (1, 3, 1)\} / \{1 - (3, 4, 3)\}$$

$$= p_{2,1} / \{(1 - p_{1,3} p_{3,1}) / (1 - p_{3,4} p_{4,3})\}$$

$$V_{2,2} = 1$$

$$V_{2,3} = (2, 1, 3) / \{1 - (1, 3, 1)\} / \{1 - (3, 4, 3)\} \{1 - (3, 4, 3)\}$$

$$= p_{2,1} p_{1,3} / \{(1 - p_{1,3} p_{3,1}) / (1 - p_{3,4} p_{4,3})\} (1 - p_{3,4} p_{4,3})$$

$$V_{2,4} = (2, 1, 3, 4) / \{1 - (1, 3, 1)\} / \{1 - (3, 4, 3)\} \{1 - (3, 4, 3)\}$$

$$= p_{2,1} p_{1,3} p_{3,4} / \{(1 - p_{1,3} p_{3,1}) / (1 - p_{3,4} p_{4,3})\} (1 - p_{3,4} p_{4,3})$$

$$V_{2,5} = (2, 5)$$

$$= p_{2,5}$$

$$V_{2,6} = (2, 6) / \{1 - (6, 7, 6)\} + (2, 5, 7, 6)$$

$$= p_{2,6} / (1 - p_{6,7} p_{7,6}) + p_{2,5} p_{5,7} p_{7,6}$$

$$V_{2,7} = (2, 5, 7) / \{1 - (7, 6, 7)\} / \{1 - (6, 2, 6)\} + (2, 6, 7)$$

$$= p_{2,5} p_{5,7} / \{(1 - p_{6,7} p_{7,6}) / (1 - p_{6,2} p_{2,6})\} + p_{2,6} p_{6,7}$$

$$V_{2,8} = (2, 8) / \{1 - (8, 9, 8)\} + (2, 5, 9, 8)$$

$$= p_{2,8} / (1 - p_{8,9} p_{9,8}) + p_{2,5} p_{5,9} p_{9,8}$$

$$V_{2,9} = (2, 8, 9) / \{1 - (8, 9, 8)\} + (2, 5, 9)$$

$$= p_{2,8} p_{8,9} / (1 - p_{8,9} p_{9,8}) + p_{2,5} p_{5,9}$$

## 6. Modeling System Parameters by using RPGT

### 6.1. Mean time to system failure ( $T_0$ )

Regenerative working stages [12], where the framework can transit (base stage '2'), earlier incoming into failed stage are: 'i' = 1, 2, 5 attractive 'ξ' = '1' [12]

$$T_0 = (V_{1,1} \mu_1 + V_{1,2} \mu_2 + V_{1,5} \mu_5) / \{1 - (1, 2, 1)\}$$

### 6.2 Availability of the system ( $A_0$ )

Regenerative stages, where the framework is accessible are 'i' = 1, 2, 5 attractive 'ξ' = '1' whole fraction of time for which the framework is accessible is assumed by

$$A_0 = (V_{2,1} \mu_1 + V_{2,2} \mu_2 + V_{2,5} \mu_5) / Z_1$$

$$\therefore Z = V_{1,1} \mu_1 + V_{1,2} \mu_2 + V_{1,3} \mu_3 + V_{1,4} \mu_4 + V_{1,5} \mu_5 + V_{1,6} \mu_6 + V_{1,7} \mu_7 + V_{1,8} \mu_8 + V_{1,9} \mu_9$$

$$\therefore Z_1 = V_{2,1} \mu_1 + V_{2,2} \mu_2 + V_{2,3} \mu_3 + V_{2,4} \mu_4 + V_{2,5} \mu_5 + V_{2,6} \mu_6 + V_{2,7} \mu_7 + V_{2,8} \mu_8 + V_{2,9} \mu_9$$

### 6.3 Server of busy period ( $B_0$ )

Regenerative stages where repairman is busy are  $2 \leq j \leq 9$ , whole fraction of time for which server remains eventful is by equation:

$$B_0 = (V_{1,2} \mu_2 + V_{1,3} \mu_3 + V_{1,4} \mu_4 + V_{1,5} \mu_5 + V_{1,6} \mu_6 + V_{1,7} \mu_7 + V_{1,8} \mu_8 + V_{1,9} \mu_9) / D$$

$$= 1 - (\mu_1 / D)$$

### 6.4 Expected number of server visit's ( $V_0$ )

Regenerative stages, where repair man do this job are  $j = 2, 5$  number of visit by repair man is given by:

$$V_0 = (V_{1,2} + V_{1,5}) / D$$

## 7. Behavior Analysis (Particular Cases: - $h_i = h$ ; $g_i = g$ )

### 7.1 Mean Time to System Failure (MTSF)

By taking values of repair and failure rates as  $g$ 's and  $h$ 's, Value of MTSF is calculated by RPGT

**Table 4.** MTSF ( $T_0$ )

	$h = .55$	$h = .65$	$h = .75$
$g = .15$	5.32	5.25	5.05
$g = .25$	4.49	4.42	4.37
$g = .35$	3.53	3.49	3.42

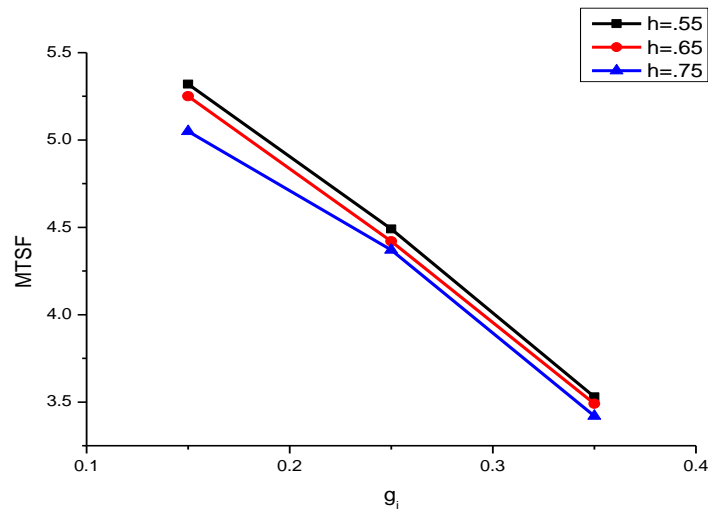


Figure 3. MTSF

From the above Figure 3 and Table 4 demonstrations the performance of MTSF Vs Repair rate of the sub-unit of the framework for various values of the disappointment rate. From the above Figure 3 one can determine that MTSF is increasing which must be so once the repair rate amassed and decreases when the disappointment rate rises which should be so in practical situations.

### 7.2 Availability of the system ( $A_0$ ):

Table 5: Availability of the system

	h = .55	h = .65	h = .75
g = .15	.84	.88	.93
g = .25	.72	.75	.79
g = .35	.62	.67	.72

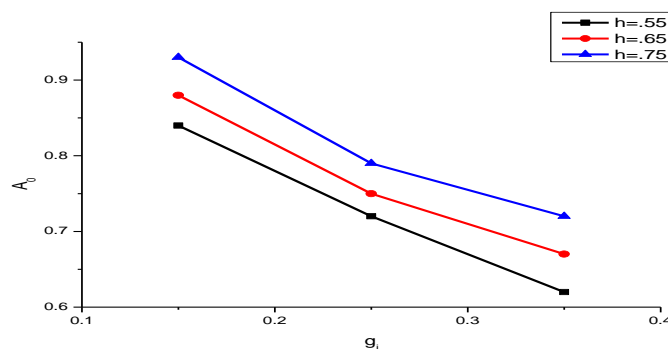


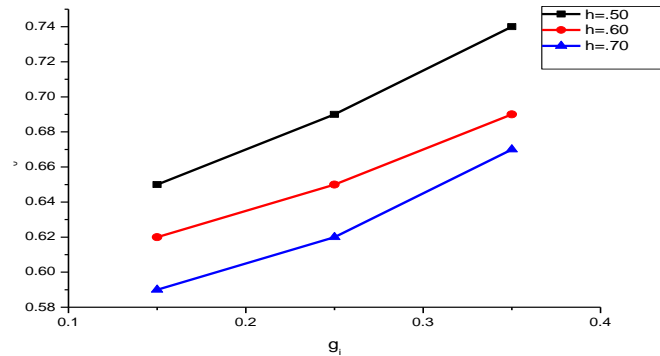
Figure 4: Availability of the system

The above Table 5 shows that the Availability is increasing when the repair rate is increasing and decrease with the rise in disappointment rate, which ought to be actually.

7.3 Server of the busy period ( $B_0$ ):

**Table 6:** Server of the busy period

	h = .50	h = .60	h = .70
g = .15	.65	.62	.59
g = .25	.69	.65	.62
g = .35	.74	.69	.67



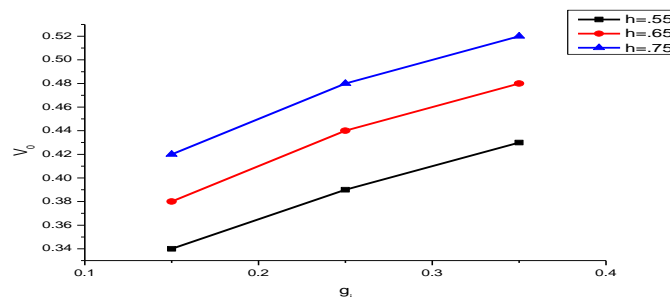
**Figure 5:** Server of the busy period

It can be concluded from the above Figure 5 that the values of server of busy period shows the expected trend for various values of disappointment rate, as server of busy period decreases with the rise in the values of repair rate.

7.4 Expected number of server visits ( $V_0$ ):

**Table 7:** Expected number of server visits

	h = .55	h = .65	h = .75
g = .15	.34	.38	.42
g = .25	.39	.44	.48
g = .35	.43	.48	.52



**Figure 6:** Expected number of server visits

It can be concluded from the above Figure 6 and Table 7 that the values of Expected number of server visits demonstrates the expected trend for various values of disappointment rate, as Expected number of server visits increases with the rise in the values of repair rate.

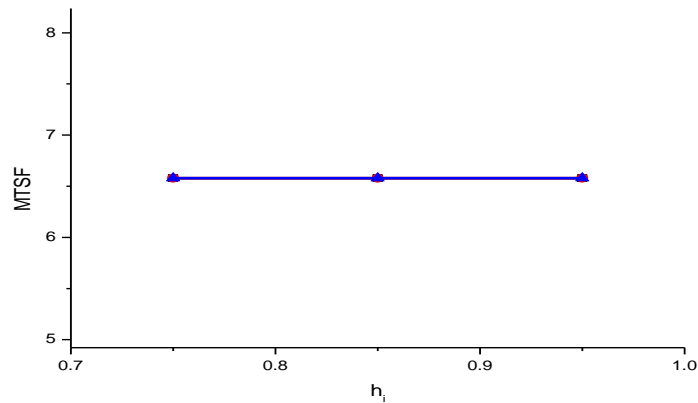


## 8. Effect of Repair Rates on System (Keeping Failure Rates Fixed)

### 8.1 Effect on MTSF (T0) parameters

**Table 8:** MTSF

$h_i$	$h_1$	$h_2$	$h_3$
0.75	6.58	6.58	6.58
0.85	6.58	6.58	6.58
0.95	6.58	6.58	6.58



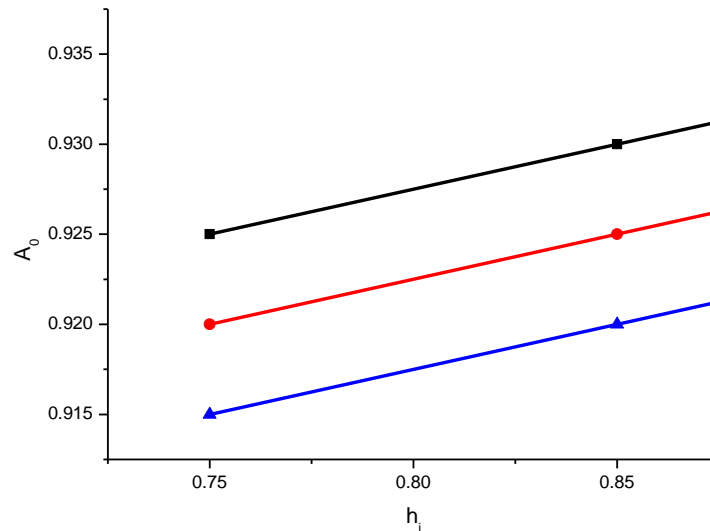
**Figure 7:** MTSF

From the above Table 8 in affecting in rows, from the 1<sup>st</sup> row it is understood that, MTSF is similar for secure and server. From the subsequent row it is determined that MTSF is constant when repair rate of server is increase. On associating the columns, it is experimental that MTSF constant at higher rates as growing repair rate of server. From the Figure 7, it is determined that MTSF is constant in repair rates.

### 8.2 Effect on Availability of the system ( $A_0$ )

**Table 9:** Availability of the system

$h_i$	$h_1$	$h_2$	$h_3$
0.75	0.925	0.920	0.915
0.85	0.930	0.925	0.920
0.95	0.935	0.930	0.925



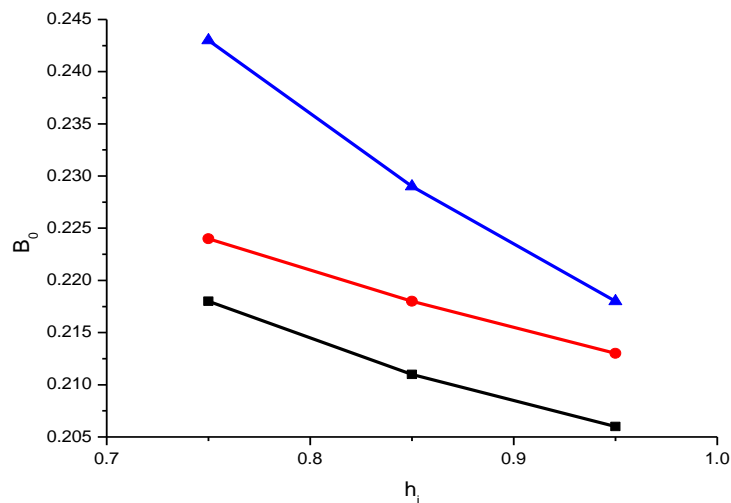
**Figure 8:** Availability of the system

From the Figure 8 and Table 9, it is realized there is not much implication change in value of Availability of the system parallel to rise in repair rates of sub-units and server. However, from the Figure 8 and Table 9 it is determined that to have extreme value of Availability of the system repair rate of server must be supreme.

### 8.3 Effect on Server of the busy period ( $B_0$ )

**Table 10.** Server of the busy period

$h_i$	$h_1$	$h_2$	$h_3$
0.75	0.218	0.224	0.243
0.85	0.211	0.218	0.229
0.95	0.206	0.213	0.218



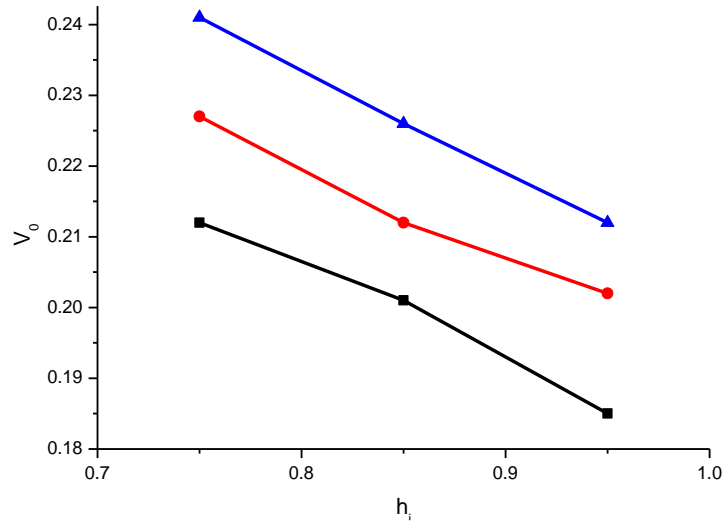
**Figure 9.** Server of the busy period

Observing in columns of Table 10, one sees that Server of the busy period reductions with the rise in repair rates which is applied but it decreases less, increasing the comparative repair rate of sub-unit. Same is the opinion while examining the values in rows.

### 8.4 Effect on Expected number of server visits ( $V_0$ )

**Table 11.** *Expected number of server visits*

$h_i$	$h_1$	$h_2$	$h_3$
0.75	0.212	0.227	0.241
0.85	0.201	0.212	0.226
0.95	0.185	0.202	0.212



**Figure 10.** *Expected number of server visits*

From the Figure 11 and Table 10 it is seen that cost of Expected number of server visits is optimal repair rate of sub-unit 'is 0.95 and associating the rates it is seen there is no significant, hence to keep assessment of Expected number of server visits lowest for minimum cost sub-unit need more care in terms of maintenance facilities.

## 9. Effect of Change of Failure Rates (Keeping Repair Rate Fixed)

### 9.1 MTSF

**Table 12.** *MTSF*

$g_i$	$g^1$	$g^2$	$g^3$
0.15	3.27	3.35	3.42
0.25	3.19	3.27	3.34
0.35	3.13	3.19	3.27

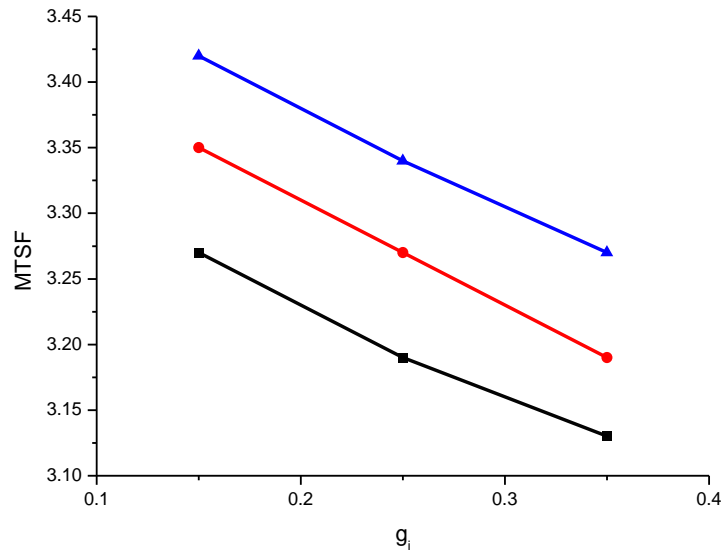


Figure 11. MTSF

For an ideal structure value of MTSF must be biggest possible from the Figure 11 and Table 12. It is determined that value of MTSF is supreme when disappointment rates of all sub-units and server are least and go as reducing as the disappointment rates of units rise. But value of MTSF decreases more quickly with rise in failure rate of first sub-unit over another sub-units, hence necessity be taken care of in terms of disappointment rate over another sub-units and server for greatest value of MTSF.

### 9.2 Availability of the system( $A_0$ )

Table 13. Availability of the system

$g^1$	$g^1$	$g^2$	$g^3$
0.15	0.927	0.948	0.963
0.25	0.907	0.927	0.945
0.35	0.886	0.906	0.927

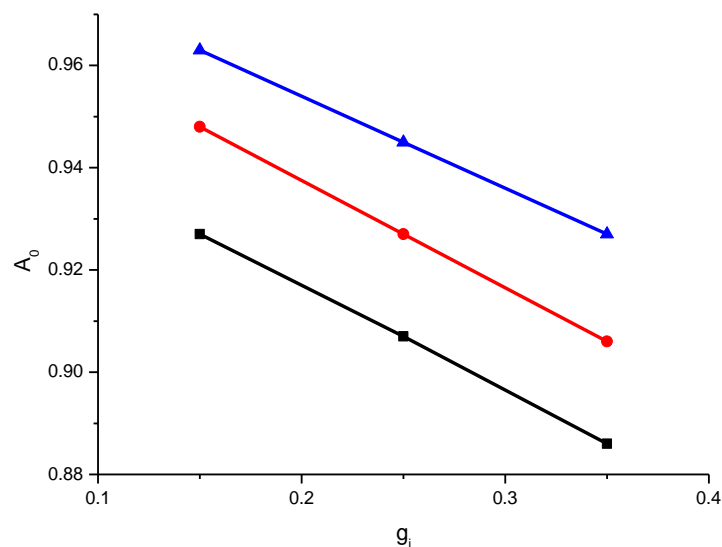


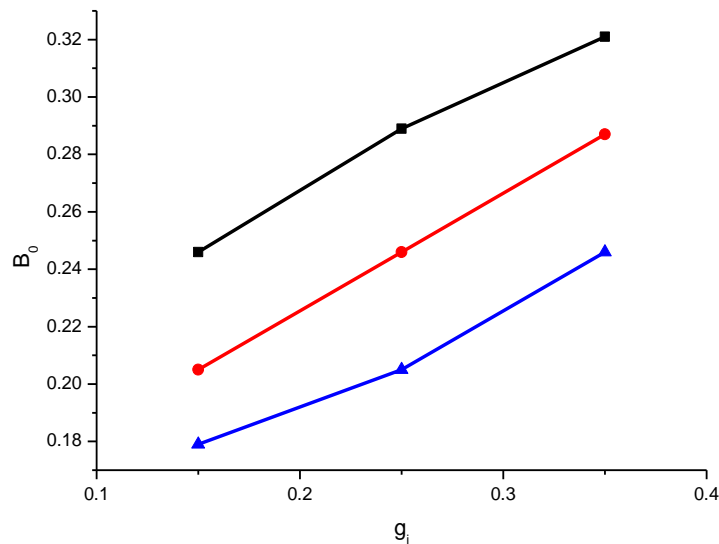
Figure 12. Availability of the system

An ideal structure value of Availability of the system ought to be supreme from the 1<sup>st</sup> row of above Table 13 and Figure 12, it is understood that Availability of the system is best when disappointment rate of sub-units and server are smallest on associating the columns Availability of the system decreases more quickly with the rise in disappointment of units, hence Availability of the system value of Availability of the system biggest, it is optional that that first sub-unit needs more care for upkeep facilities.

### 9.3 Server of the busy period ( $B_0$ )

**Table 14.** *Server of the busy period*

$g^i$	$g^1$	$g^2$	$g^3$
0.15	0.246	0.205	0.179
0.25	0.289	0.246	0.205
0.35	0.321	0.287	0.246



**Figure 13.** *Server of the busy period*

To do study with esteem to value of Server of the busy period in the exceeding Table 14, it is decent to keep value of Server of the busy period minimum, on associating the columns, it is experimental that Server of the busy period have similar values for disappointment rate of server in assessment to units first and second unit.

### 9.4 Expected number of server visits ( $V_0$ )

**Table 15.** *Expected number of server visits*

$g^i$	$g^1$	$g^2$	$g^3$
0.15	0.312	0.299	0.275
0.25	0.319	0.312	0.301
0.35	0.325	0.321	0.312

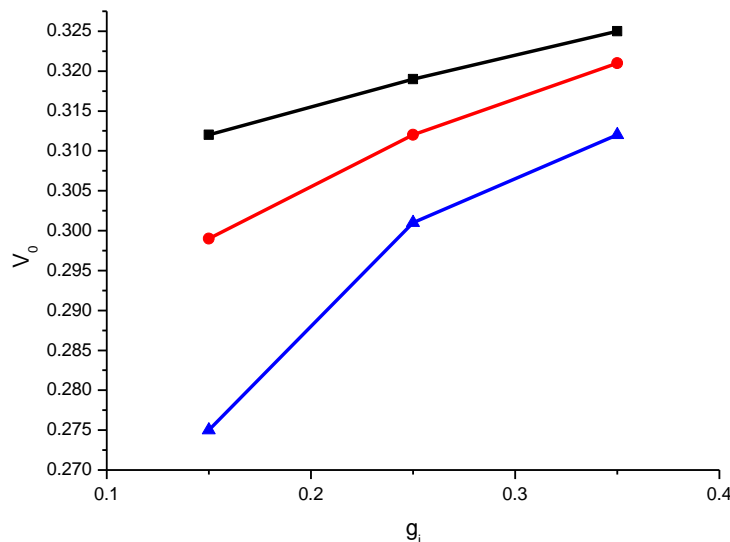


Figure 14. Expected number of server visits

A structure will be named Table 15 free if the Expected number of server visits are small foam the table and Figure 14, it is optional that for small value of Expected number of server visits, disappointment rates of sub-units and server to be kept smallest i.e., sub-units and server must be best in enterprise and quality, however value of Expected number of server visits rise proportional less in assessment to increasing disappointment rate of server. Thus, online sub-units need more upkeep the whole server. In all to keep reduced value of cost and that  $V_0$ , disappointment rates of sub-units and server to be kept small.

## 10. Results

- Value of  $MTSF(T_0)$  is decreased with increasing of repair and failure rate and  $T_0$  is fix when failure rate is fixed and  $T_0$  is increased when repair rate is fixed.
- Availability of the system ( $A_0$ ) is increased with increasing of failure rate and decreased with increasing of repair rate and when failure rate is fixing value of  $A_0$  is increased with the value of repair rate and in case when repair rate is fixing value of  $A_0$  is decreased with the value of rising failure rate.
- Server of busy period ( $B_0$ ) is increased with increased of repair rate and decreased with the increasing of failure rate and when failure rate is fixed value of  $B_0$  is decreased with the increasing of repair rate and in case if repair rate is fixed then value of  $B_0$  is increased with the increasing of failure rate.
- Expected number of server visit ( $V_0$ ) is increased with the increasing of repair and failure rate and when failure rate is fixed the value of  $V_0$  is decreased with increasing of repair rate and in case if repair rate is fixed then value of  $V_0$  is fixed with the increasing of failure rate.

## References

- [1] Bayer, C. (2012). *Strahlentherapie und Onkologie*. Strahlenther Onkol, 188, 616-627.
- [2] Basavanagoud, B., & Patil, S. (2016). *Upper Bounds for the Modified Second Multiplicative Zagreb Index of Graph Operations*. Bulletin of Mathematical Sciences and Applications Vol, 17, 11.
- [3] 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.

- [4] Rohilla, N., & Rai, M. (2021, December). *Advance machine learning techniques used for detecting and classification of disease in plants: A review*. In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) (pp. 490-494). IEEE.
- [5] Panwar, P., & Garg, A. (2013). *Analysis of reliability and cost tradeoffs in architecture based software applications using a genetic algorithm*. International Journal of Computer Applications, 72(1).
- [6] Panwar, P., & Lal, A. K. (2014). *Predicting total number of failures in a software using NHPP software reliability growth models*. In Proceedings of the Third International Conference on Soft Computing for Problem Solving: SocProS 2013, Volume 2 (pp. 715-727). Springer India.
- [7] Gong, Q., Li, J., Jiang, Z., & Wang, Y. (2024). *A hierarchical integration scheduling method for flexible job shop with green lot splitting*. Engineering Applications of Artificial Intelligence, 129, 107595.
- [8] Barak, M. S., Garg, R., & Kumar, A. (2021). *Reliability measures analysis of a milk plant using RPGT*. Life Cycle Reliability and Safety Engineering, 10(3), 295-302.
- [9] Kaur, R., & Panwar, P. (2015). *Study of Perfect and Imperfect Debugging NHPP SRGMs used for Prediction of Faults in a Software*. IJCSC, available online at [www.csjournalss.com](http://www.csjournalss.com), 6(1), 73-78.
- [10] Bhardwaj, R. K., Sonker, P., & Singh, R. (2022). *A semi-Markov model of a system working under uncertainty*. In System Assurances (pp. 175-187). Academic Press.
- [11] Shakuntla, S., Lal, A. K., Bhatia, S. S., & Singh, J. (2011). *Reliability analysis of polytube industry using supplementary variable technique*. Applied Mathematics and Computation, 218(8), 3981-3992.
- [12] Naithani, A., Parashar, B., Bhatia, P., & 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 Modeling and Optimization, 15(2), 499-509.
- [13] 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.
- [14] Niwas, R., & Garg, H. (2018). *An approach for analyzing the reliability and profit of an industrial system based on the cost free warranty policy*. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 40(5), 265.
- [15] Singla, S., Kaur, H., Gupta, D., & Kaur, J. (2023, September). *No Idle Constraint In Flow Shop Scheduling With Transportation Time, Weightage of Jobs And Job Block Criteria*. In 2023 IEEE 2nd International Conference on Industrial Electronics: Developments & Applications (ICIDEA)(pp. 450-454). IEEE.
- [16] Su, Y., & Landis, C. M. (2007). *Continuum thermodynamics of ferroelectric domain evolution: Theory, finite element implementation, and application to domain wall pinning*. Journal of the Mechanics and Physics of Solids, 55(2), 280-305.

# NUMERICAL INVESTIGATION OF RETRIAL QUEUEING INVENTORY SYSTEM WITH A CONSTANT RETRIAL RATE, WORKING VACATION, FLUSH OUT, COLLISION AND IMPATIENT CUSTOMERS

G. AYYAPPAN, N. ARULMOZHI



Department of Mathematics,  
Puducherry Technological University,  
Puducherry, India.

ayyappan@ptuniv.edu.in, arulmozhisathya@gmail.com,

## Abstract

*The retrial queueing inventory system with working vacation, flush out, balking, breakdown, and repair, as well as a constant retrial rate and orbital client collision are all examined in this study. We made the assumption that customers arrive through a Markovian arrival process and that they would get phase-type services from the server. The inventory is replenished using a  $(s, S)$  and  $(s, Q)$  strategy, and it is expected that the replenishment time will follow an exponential distribution. If there are zero inventory items, no customers in the orbit, or both, the server will go into working vacation mode. When a customer retries an orbit while the server is serving arriving customers, the orbital customer may collide with the arriving customer during that retry, in which case both of them will be shifted back into orbit; otherwise, the orbital customer may avoid colliding with the arriving customer and may rejoin the orbit for another retry. The number of customers in the orbit and the inventory level may be found in the steady state. A cost analysis is produced along with the establishment of various important performance measures. Moreover, some numerical examples are provided to clarify our mathematical notion.*

**Keywords:** Markovian arrival process, PH-distribution, working vacation, collision of orbital customers, flush out.

**AMS Subject Classification (2010):** 60K25, 68M30, 90B22.

## 1. INTRODUCTION

Retrial queues occur when initial consumers identify all servers and/or waiting space full. They may choose to try again after a random length of time or abandon the system permanently. RQ models have been thoroughly researched in a significant number of papers. Artalejo et al. [3] introduced the concept of retrial requests for inventory. They assumed that demand points are Poisson processes, whereas lead and retrial time points are exponential. They thought that the orbit's size is limitless. Manuel et al. [8] proposed a retrial inventory system that includes a service facility. They assumed clients come according to a Markovian arrival process (MAP), that service time for each client follows a phase-type distribution (PH), that lead time, lifetime of each item, and retrial times follow an exponential distribution.

Customers arrive at the single server retrial queueing-inventory system under consideration in this study using a Markovian Arrival Process, also known as the flexible point process. The MAP tries to accomplish significant generalisation of the Poisson process while keeping it tractable. Many real-world applications do not require a renewal procedure before arriving. As a result, the most useful tool for simulating renewal and non-renewal appearance situations is



the *MAP*. We can have realistic arrival patterns in this model because of the *MAP*, which also accounts for correlations and dependencies between arrivals. Furthermore, the continuous-time case is necessary, even though the *MAP* is defined for both discrete and continuous periods. See Chakravarthy [5] and Neuts [10] for further details on the *MAP* and its properties.

The notion of server vacation was first presented in the retry inventory system by Sivakumar [17]. For lead, inter-trial, inter-demand, and server vacation durations, he made the assumption that the distributions would be exponential. He also believed that these incidents are unrelated to one another. He instituted a programme of repeated vacations. A two-commodity substitutable retrial inventory system with a shared ordering strategy was examined by Sivakumar [15]. Sivakumar [16] examined a system of perishable inventory that had requests for retrials. The exponentially distributed lead periods for orders, the finite source of requests, the exponentially distributed life durations for stored objects, and the exponentially distributed inter-retrial intervals have all been assumed by the author. A two-commodity stochastic inventory technique with a complement item was proposed by Jeganathan et al. [11] in the context of a traditional retrial facility. When the primary item is out of supply, each new client will immediately enter an orbit of infinite capacity.

A  $M/M/1$  retrial queue under  $(s, S)$  policy with a storage system was examined by Shajin and Krishnamoorthy [14]. The authors use the assumption that when the server is inactive, the external arrivals immediately enter an orbit and that the time between two successive retrials has an exponential distribution. Only the client at the head of the orbit is allowed to reach the server. In contrast to the traditional method of employing just one vendor, Chakravarthy and Hayat [6] established the idea of multiple vendors responsible for replacing inventories. This way, replenishment happens via two vendors. The authors used the *MAM* to analyse the model in steady-state under the assumptions of a two-vendor system, where the lead times are exponentially distributed with a parameter that depends on the vendor, the demands occur according to a *MAP*, and the service times are PH. There are also interesting numerical examples given, such as a comparison of the systems with one and two vendors.

A queueing inventory model in which a new customer comes and waits for service when the server is unavailable due to vacation was examined by Y Zhang et al. [19]. The model included the server's multiple vacations and dissatisfied clients. They were able to extract some significant performance metrics and find the matrix geometric solution of the steady-state probability by using the truncated approximation approach. Using numerical analysis, the impact of the probability and impatience rate on a few performance metrics was examined. Using the genetic algorithm, the authors calculated the best possible policy and cost and arrived at the ideal service rate. Ayyappan et al. [4] studied the notions of working breakdown, collision, vacation, and reneging in a non-preemptive priority retrial queueing system with immediate feedback. They applied the supplementary variable technique to their model and also provided particular cases.

Service interruptions were originally implemented in an inventory model by Krishnamoorthy et al. [7]. They also believed that orders are processed instantly and that there is no limit to the amount of disruptions that can happen during a single service. Ushakumari [18] examined a  $(s, S)$  inventory system with recurrent demands for unfulfilled requests from the orbit and a random lead time. In their paper [1], Amirthakodi and Sivakumar spoke about retrial inventory queueing with a single server and customer feedback, where the orbit size is finite. The retrial queueing model with exponential service time, Poisson arrival, and delayed feedback was examined by Melikov et al. [9]. They used both  $(s, S)$  and  $(s, Q)$  replenishment policies for their study. In their analysis of an  $M/M/1/N$  queueing system with reverse balking, Kumar et al. [13] incorporate the idea of reverse reneging. Customers' input is used by Kumar and Som [?] in an  $M/M/1/N$  queueing system with reverse balking, reverse reneging, and retention of renegeed customers. They calculate the system size stationary probability.

## 2. MODEL DESCRIPTION

- We examine a single-server retrial queueing inventory model in which customers arrive at the system as represented by *MAP*, with  $D_0$  and  $D_1$  matrices as its dimension  $m$ . The service times, denoted as  $(\gamma, U)$  of order  $n$ , are assumed to follow the PH-distribution with  $U^0 + Ue = 0$ .
- If the server is available, he serves the customer right away upon their arrival. If not, the customer must enter the orbit of infinite. Every customer retries from the orbit at a constant rate, despite the size of the orbit. The inter-retrial times follow an exponential distribution with parameter  $\delta$ .
- If the orbit is empty, the inventory is zero, or both, then the server goes on vacation after serving the customer. Additionally, the vacation periods are expected to follow a  $\eta$ -parameter exponential distribution. In the event that a customer arrives during vacation time, the server will start charging the customer less for services than usual. Additionally, it is expected that the service times throughout the vacation period follow the PH distribution, denoted as  $(\gamma, \theta U)$ , with  $0 < \theta < 1$ . If the server examines the customer who is waiting in the system after completing this vacation, he will begin a normal busy period. Otherwise, he is dormant.
- The incoming customer may enter the orbit for a retry with probability  $q_1$  or balk the system with probability  $p_1$  during the service delivery, repair, and no inventory items, ensuring that  $p_1 + q_1 = 1$ .
- When a customer retries an orbit while the server is servicing incoming customers, there is a chance that the orbital customer and the incoming customer will collide and be shifted to the orbit with a probability of  $q_2$ ; if not, the orbital customer may not collide and will rejoin the orbit for a subsequent retry with a probability of  $p_2$ , such that  $p_2 + q_2 = 1$ .
- During regular busy periods, the server may get breakdown. As a result, the customer getting service at the moment must enter the orbit of limitless capacity. The server goes into idle mode when the repair operation is completed. The breakdown times are exponentially distributed with parameter  $\psi$ , whereas the repair times are PH-distributed with rate  $(\alpha, T)$ .
- All the customers in the orbit are flushed out periodically and the flush out times follow exponential distribution with parameter  $\sigma$ . The schematic picture of this model is provided in Figure 1.
- $\otimes$  - Kronecker product of two matrices of different dimensions.  $\oplus$  - Kronecker sum of two matrices of different dimensions.  $e$  - Column vector has a suitable size with each of its entries as 1.  $\mathbf{0}$  - It denotes zero matrices in the suitable order.

## 3. ANALYSIS

In the following section, we establish the queueing-inventory system's transition rate matrix. Assume that  $N(t), J(t), I(t), R(t), S(t), A(t)$  describe the total customers in the orbit, status of server, stock level, repair phases, service phases, arrival phases, respectively.

$$J(t) = \begin{cases} 0, & \text{server is idle in normal service mode,} \\ 1, & \text{server is busy in normal service mode,} \\ 2, & \text{server is idle in WV mode,} \\ 3, & \text{server is busy in WV mode,} \\ 4, & \text{server is repair mode.} \end{cases}$$

Consider  $X(t) = \{N(t), J(t), I(t), R(t), S(t), A(t)\}$  is a CTMC with state space

$$\Phi = \phi(0) \bigcup_{i=1}^{\infty} \phi(i). \tag{1}$$

where

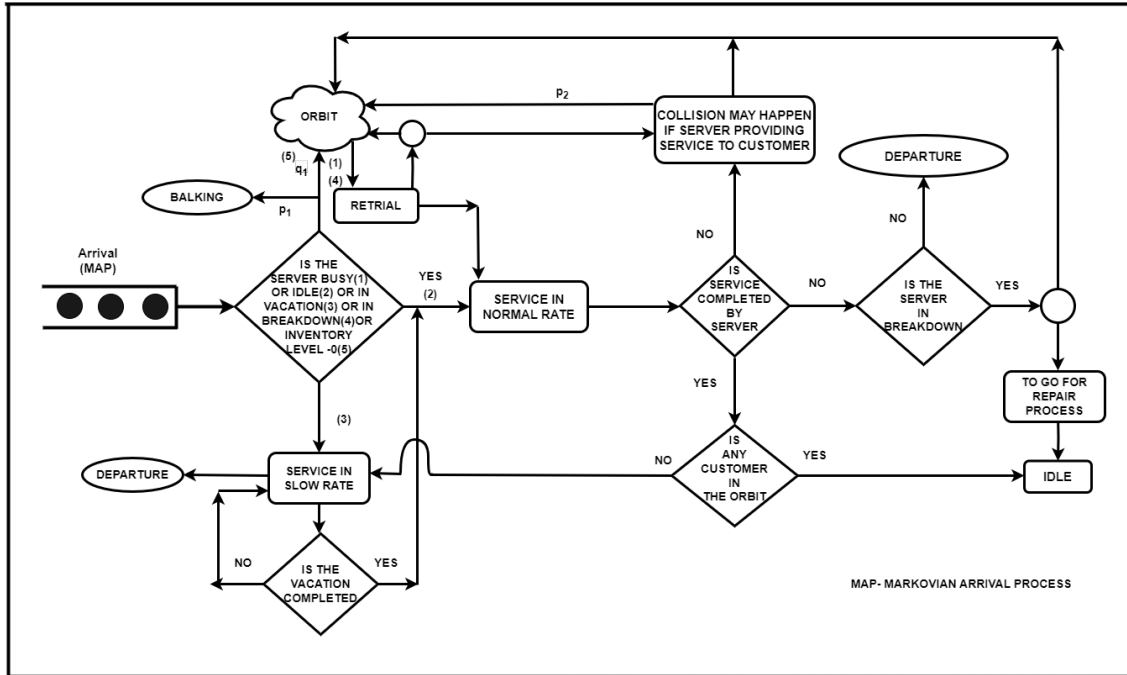


Figure 1: Schematic representation

$$\begin{aligned} \phi(0) = & \{(0, 0, u_1, u_4) : 0 \leq u_1 \leq S, 1 \leq u_4 \leq m\} \\ & \cup \{(0, 1, u_1, u_3, u_4) : 1 \leq u_1 \leq S, 1 \leq u_3 \leq n, 1 \leq u_4 \leq m\} \\ & \cup \{(0, 2, u_1, u_4) : 0 \leq u_1 \leq S, 1 \leq u_4 \leq m\} \\ & \cup \{(0, 3, u_1, u_3, u_4) : 1 \leq u_1 \leq S, 1 \leq u_3 \leq n, 1 \leq u_4 \leq m\} \\ & \cup \{(0, 4, u_1, u_2, u_4) : 1 \leq u_1 \leq S, 1 \leq u_2 \leq l, 1 \leq u_4 \leq m\} \end{aligned}$$

and for  $i \geq 1$ ,

$$\begin{aligned} \phi(i) = & \{(i, 0, u_1, u_4) : 0 \leq u_1 \leq S, 1 \leq u_4 \leq m\} \\ & \cup \{(i, 1, u_1, u_3, u_4) : 1 \leq u_1 \leq S, 1 \leq u_3 \leq n, 1 \leq u_4 \leq m\} \\ & \cup \{(i, 2, u_1, u_4) : 0 \leq u_1 \leq S, 1 \leq u_4 \leq m\} \\ & \cup \{(i, 3, u_1, u_3, u_4) : 1 \leq u_1 \leq S, 1 \leq u_3 \leq n, 1 \leq u_4 \leq m\} \\ & \cup \{(i, 4, u_1, u_2, u_4) : 1 \leq u_1 \leq S, 1 \leq u_2 \leq l, 1 \leq u_4 \leq m\} \end{aligned}$$

### 3.1. Construction of the QBD process for Model 1

The generator matrix of the Markov chain under  $(s, S)$  policy is given by:

$$Q = \begin{bmatrix} A_{00} & A_{01} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \dots \\ A_{10} & F_1 & F_0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \dots \\ A & F_2 & F_1 & F_0 & \mathbf{0} & \mathbf{0} & \dots & \dots \\ A & \mathbf{0} & F_2 & F_1 & F_0 & \mathbf{0} & \dots & \dots \\ A & \mathbf{0} & \mathbf{0} & F_2 & F_1 & F_0 & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \dots & \dots \end{bmatrix}$$

The entries in the block matrices of  $Q$  are defined as follows,

$$A_{00} = \begin{bmatrix} A_{00}^{11} & A_{00}^{12} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_{00}^{22} & A_{00}^{23} & \mathbf{0} & \mathbf{0} \\ A_{00}^{31} & \mathbf{0} & A_{00}^{33} & A_{00}^{34} & \mathbf{0} \\ \mathbf{0} & A_{00}^{42} & A_{00}^{43} & A_{00}^{44} & \mathbf{0} \\ A_{00}^{51} & \mathbf{0} & \mathbf{0} & \mathbf{0} & A_{00}^{55} \end{bmatrix},$$

where

$$A_{00}^{11} = \begin{bmatrix} C_1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_2 \\ \mathbf{0} & C_3 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_2 \\ \mathbf{0} & \mathbf{0} & C_3 & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_3 & \mathbf{0} & \dots & \mathbf{0} & C_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_4 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_4 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_4 \end{bmatrix},$$

$$A_{00}^{22} = \begin{bmatrix} C_5 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_6 \\ \mathbf{0} & C_5 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_6 \\ \mathbf{0} & \mathbf{0} & C_5 & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_6 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_5 & \mathbf{0} & \dots & \mathbf{0} & C_6 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_7 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_7 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_7 \end{bmatrix},$$

where  $C_1 = (D_0 + p_1 D_1) - \beta I_m$ ,  $C_2 = \beta I_m$ ,  $C_3 = D_0 - \beta I_m$ ,  $C_4 = D_0$ ,  
 $C_5 = U \oplus (D_0 + p_1 D_1) - (\psi + \beta) I_{nm}$ ,  $C_6 = \beta I_{nm}$ ,  $C_7 = U \oplus (D_0 + p_1 D_1) - \psi I_{nm}$ .  
 $A_{00}^{23} = I_S \otimes U^0 \otimes I_m$ ,  $A_{00}^{31} = I_{S+1} \otimes \eta I_m$ ,

$$A_{00}^{12} = \begin{bmatrix} \mathbf{0} \\ I_S \otimes \gamma \otimes D_1 \end{bmatrix},$$

$$A_{00}^{33} = \begin{bmatrix} C_8 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_9 \\ \mathbf{0} & C_{10} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_9 \\ \mathbf{0} & \mathbf{0} & C_{10} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_9 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_{10} & \mathbf{0} & \dots & \mathbf{0} & C_9 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{11} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{11} \end{bmatrix},$$

where  $C_8 = (D_0 + p_1 D_1) - (\eta + \beta) I_m$ ,  $C_9 = \beta I_m$ ,  $C_{10} = D_0 - (\eta + \beta) I_m$ ,  $C_{11} = D_0 - \eta I_m$ .  
 $A_{00}^{42} = I_{S+1} \otimes \eta I_{nm}$ ,  $A_{00}^{43} = I_S \otimes \theta U^0 \otimes I_m$ ,

$$A_{00}^{34} = \begin{bmatrix} \mathbf{0} \\ I_S \otimes \gamma \otimes D_1 \end{bmatrix}, A_{00}^{44} = \begin{bmatrix} C_{12} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{13} \\ \mathbf{0} & C_{12} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{13} \\ \mathbf{0} & \mathbf{0} & C_{12} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{13} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_{12} & \mathbf{0} & \dots & \mathbf{0} & C_{13} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{14} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_{14} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{14} \end{bmatrix},$$

where  $C_{12} = \theta U \oplus (D_0 + p_1 D_1) - (\eta + \beta) I_{nm}$ ,  $C_{13} = \beta I_{nm}$ ,  $C_{14} = \theta U \oplus (D_0 + p_1 D_1) - \eta I_{nm}$ .

$$A_{00}^{51} = [\mathbf{0} \quad I_S \otimes T^0 \otimes I_m], A_{00}^{55} = \begin{bmatrix} C_{15} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{16} \\ \mathbf{0} & C_{15} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{16} \\ \mathbf{0} & \mathbf{0} & C_{15} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{16} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_{15} & \mathbf{0} & \dots & \mathbf{0} & C_{16} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{17} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_{17} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{17} \end{bmatrix},$$

where  $C_{15} = T \oplus (D_0 + p_1 D_1) - \beta I_{lm}$ ,  $C_{16} = \beta I_{lm}$ ,  $C_{17} = T \oplus (D_0 + p_1 D_1)$ .

$$A_{01} = \begin{bmatrix} A_{01}^{11} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_{01}^{22} & \mathbf{0} & \mathbf{0} & A_{01}^{25} \\ \mathbf{0} & \mathbf{0} & A_{01}^{33} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & A_{01}^{44} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & A_{01}^{55} \end{bmatrix},$$

$$A_{01}^{11} = \begin{bmatrix} q_1 D_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, A_{01}^{22} = I_S \otimes I_n \otimes q_1 D_1, A_{01}^{25} = I_S \otimes e_n \alpha \otimes \psi I_m, A_{01}^{33} = \begin{bmatrix} q_1 D_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$A_{01}^{44} = I_S \otimes I_n \otimes q_1 D_1, A_{01}^{55} = I_S \otimes I_l \otimes q_1 D_1,$$

$$A_{10} = \begin{bmatrix} A_{10}^{11} & A_{10}^{12} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A_{10}^{21} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A_{10}^{33} & A_{10}^{34} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A_{10}^{43} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & A_{10}^{55} \end{bmatrix},$$

where

$$A_{10}^{11} = I_{S+1} \otimes \sigma I_m, A_{10}^{12} = \begin{bmatrix} \mathbf{0} \\ I_S \otimes \delta \gamma \otimes I_m \end{bmatrix}, A_{10}^{21} = [\mathbf{0} \quad I_S \otimes e_n \otimes \sigma I_m], A_{10}^{33} = I_{S+1} \otimes \sigma I_m,$$

$$A_{10}^{34} = \begin{bmatrix} \mathbf{0} \\ I_S \otimes \delta \gamma \otimes I_m \end{bmatrix}, A_{10}^{43} = [\mathbf{0} \quad I_S \otimes e_n \otimes \sigma I_m], A_{10}^{55} = I_S \otimes \sigma I_m.$$

$$F_1 = \begin{bmatrix} F_1^{11} & F_1^{12} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ F_1^{21} & F_1^{22} & F_1^{23} & \mathbf{0} & \mathbf{0} \\ F_1^{31} & \mathbf{0} & F_1^{33} & F_1^{34} & \mathbf{0} \\ \mathbf{0} & F_1^{42} & F_1^{43} & F_1^{44} & \mathbf{0} \\ F_1^{51} & \mathbf{0} & \mathbf{0} & \mathbf{0} & F_1^{55} \end{bmatrix},$$

where

$$F_1^{11} = \begin{bmatrix} C_{18} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{19} \\ \mathbf{0} & C_{20} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{19} \\ \mathbf{0} & \mathbf{0} & C_{20} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{19} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_{20} & \mathbf{0} & \dots & \mathbf{0} & C_{19} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{21} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_{21} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{21} \end{bmatrix},$$

where  $C_{18} = (D_0 + p_1 D_1) - (\sigma + \beta)I_m$ ,  $C_{19} = \beta I_m$ ,  $C_{20} = D_0 - (\delta + \sigma + \beta)I_m$ ,

$C_{21} = D_0 - (\delta + \sigma)I_m$ .  $F_1^{12} = \begin{bmatrix} \mathbf{0} \\ I_S \otimes \gamma \otimes D_1 \end{bmatrix}$ ,  $F_1^{21} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{S-1} \otimes U^0 \otimes I_m & \mathbf{0} \end{bmatrix}$ ,

$$F_1^{23} = \begin{bmatrix} U^0 \otimes I_m & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, F_1^{22} = \begin{bmatrix} C_{22} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{23} \\ \mathbf{0} & C_{22} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{23} \\ \mathbf{0} & \mathbf{0} & C_{22} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{23} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_{22} & \mathbf{0} & \dots & \mathbf{0} & C_{23} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{24} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_{24} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{24} \end{bmatrix},$$

where  $C_{22} = U \oplus (D_0 + p_1 D_1) + [(q_2 \delta - \delta) - (\psi + \sigma + \beta)]I_{nm}$ ,  $C_{23} = \beta I_{nm}$ ,

$C_{24} = U \oplus (D_0 + p_1 D_1) + [(q_2 \delta - \delta) - (\psi + \sigma)]I_{nm}$ ,  $F_1^{31} = I_{S+1} \otimes \eta I_m$ ,

$$F_1^{34} = \begin{bmatrix} \mathbf{0} \\ I_S \otimes \gamma \otimes D_1 \end{bmatrix}, F_1^{33} = \begin{bmatrix} C_{25} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{26} \\ \mathbf{0} & C_{27} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{26} \\ \mathbf{0} & \mathbf{0} & C_{27} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{26} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_{27} & \mathbf{0} & \dots & \mathbf{0} & C_{26} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{28} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_{28} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{28} \end{bmatrix},$$

where  $C_{25} = (D_0 + p_1 D_1) - (\sigma + \eta + \beta)I_m$ ,  $C_{26} = \beta I_m$ ,  $C_{27} = D_0 - (\sigma + \delta + \eta + \beta)I_m$ ,

$C_{28} = D_0 - (\sigma + \delta + \eta)I_m$ ,  $F_1^{42} = I_{S+1} \otimes \eta I_{nm}$ ,  $F_1^{43} = [I_S \otimes \theta U^0 \otimes I_m \quad \mathbf{0}]$ ,

$$F_1^{44} = \begin{bmatrix} C_{29} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{30} \\ \mathbf{0} & C_{29} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{30} \\ \mathbf{0} & \mathbf{0} & C_{29} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{30} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_{29} & \mathbf{0} & \dots & \mathbf{0} & C_{30} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{31} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_{31} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{31} \end{bmatrix},$$

where  $C_{29} = \theta U \oplus (D_0 + p_1 D_1) + [(q_2 \delta - \delta) - (\sigma + \eta + \beta)]I_{nm}$ ,  $C_{30} = \beta I_{nm}$ ,

$C_{31} = \theta U \oplus (D_0 + p_1 D_1) + [(q_2 \delta - \delta) - (\sigma + \eta)]I_{nm}$ .

$$F_1^{51} = [\mathbf{0} \quad I_S \otimes T^0 \otimes I_m], F_1^{55} = \begin{bmatrix} C_{32} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{33} \\ \mathbf{0} & C_{32} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{33} \\ \mathbf{0} & \mathbf{0} & C_{32} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{33} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_{32} & \mathbf{0} & \dots & \mathbf{0} & C_{33} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{34} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_{34} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{34} \end{bmatrix},$$

where  $C_{32} = T \oplus (D_0 + p_1 D_1) - (\sigma + \beta) I_m$ ,  $C_{33} = \beta I_m$ ,  $C_{34} = T \oplus (D_0 + p_1 D_1) - \sigma I_m$ .

$$F_0 = \begin{bmatrix} F_0^{11} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ F_0^{21} & F_0^{22} & \mathbf{0} & \mathbf{0} & F_0^{25} \\ \mathbf{0} & \mathbf{0} & F_0^{33} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & F_0^{43} & F_0^{44} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & F_0^{55} \end{bmatrix},$$

$$F_0^{11} = \begin{bmatrix} q_1 D_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, F_0^{21} = [\mathbf{0} \quad I_S \otimes e_n \otimes p_2 \delta I_m], F_0^{22} = I_S \otimes I_n \otimes q_1 D_1, F_0^{25} = I_S \otimes e_n \alpha \otimes \psi I_m,$$

$$F_0^{33} = \begin{bmatrix} q_1 D_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, F_0^{43} = [\mathbf{0} \quad I_S \otimes e_n \otimes p_2 \delta I_m], F_0^{44} = I_S \otimes I_n \otimes q_1 D_1, F_0^{55} = I_S \otimes I_l \otimes q_1 D_1.$$

$$F_2 = \begin{bmatrix} \mathbf{0} & F_2^{12} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & F_2^{34} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where  $F_2^{12} = \begin{bmatrix} \mathbf{0} \\ I_S \otimes \delta \gamma \otimes I_m \end{bmatrix}, F_2^{34} = \begin{bmatrix} \mathbf{0} \\ I_S \otimes \delta \gamma \otimes I_m \end{bmatrix},$

$$A = \begin{bmatrix} A^{11} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A^{21} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A^{33} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A^{43} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & A^{55} \end{bmatrix},$$

where  $A^{11} = I_{S+1} \otimes \sigma I_m, A^{21} = [\mathbf{0} \quad I_S \otimes e_n \otimes \sigma I_m], A^{33} = I_{S+1} \otimes \sigma I_m,$   
 $A^{43} = [\mathbf{0} \quad I_S \otimes e_n \otimes \sigma I_m], A^{53} = [\mathbf{0} \quad I_S \otimes e_l \otimes \sigma I_m]. A^{55} = I_S \otimes \sigma I_m,$

### Stability condition for Model I

To discuss the stability condition, we first consider the generator matrix  $F = F_0 + F_1 + F_2$ . If  $\chi = (\chi_0, \chi_1, \chi_2, \chi_3, \chi_4) = (\chi_{00}, \chi_{01}, \dots, \chi_{0s}, \chi_{0s+1}, \dots, \chi_{0S}, \chi_{11}, \chi_{12}, \dots, \chi_{1s}, \chi_{1s+1}, \dots, \chi_{1S}, \chi_{20}, \chi_{21}, \dots, \chi_{2s}, \chi_{2s+1}, \dots, \chi_{2S}, \chi_{31}, \chi_{32}, \dots, \chi_{3s}, \chi_{3s+1}, \dots, \chi_{3S}, \chi_{41}, \chi_{42}, \dots, \chi_{4s}, \chi_{4s+1}, \dots, \chi_{4S})$ .

The vector  $\chi$  represents the invariant vector of matrix  $F$ . Consequently, we have the relations  $\chi F = 0$  and  $\chi e = 1$ . For the Markov process with a QBD structure to exhibit stability, our model must satisfy the condition  $\chi F_0 e < \chi F_2 e$ . This condition is both necessary and sufficient for the stability of the queueing model under study and reduces to the inequality  $\lambda < \mu$ .

### 3.2. QBD process for Model II

In accordance with the assumptions outlined in the "Model Description" section, we will now examine Model II, while solely modifying the ordering policy from  $(s, S)$  to  $(s, Q)$ . The generator matrix of the process for the  $(s, Q)$  policy takes on the following form:

$$\tilde{Q} = \begin{bmatrix} \tilde{A}_{00} & A_{01} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \dots \\ A_{10} & \tilde{F}_1 & F_0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \dots \\ A & F_2 & \tilde{F}_1 & F_0 & \mathbf{0} & \mathbf{0} & \dots & \dots \\ A & \mathbf{0} & F_2 & \tilde{F}_1 & F_0 & \mathbf{0} & \dots & \dots \\ A & \mathbf{0} & \mathbf{0} & F_2 & \tilde{F}_1 & F_0 & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \dots & \dots \end{bmatrix}$$

The entries in the block matrices of  $\tilde{Q}$  are defined as follows,

$$\tilde{A}_{00} = \begin{bmatrix} \tilde{A}_{00}^{11} & \tilde{A}_{00}^{12} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \tilde{A}_{00}^{22} & \tilde{A}_{00}^{23} & \mathbf{0} & \mathbf{0} \\ \tilde{A}_{00}^{31} & \mathbf{0} & \tilde{A}_{00}^{33} & \tilde{A}_{00}^{34} & \mathbf{0} \\ \mathbf{0} & \tilde{A}_{00}^{42} & \tilde{A}_{00}^{43} & \tilde{A}_{00}^{44} & \mathbf{0} \\ \tilde{A}_{00}^{51} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \tilde{A}_{00}^{55} \end{bmatrix},$$

$$\tilde{A}_{00}^{11} = \begin{bmatrix} C_1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_2 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_3 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_2 & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & C_3 & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_3 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_4 & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_4 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_4 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_4 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_4 \end{bmatrix},$$

$$\tilde{A}_{00}^{12} = \begin{bmatrix} \mathbf{0} \\ I_S \otimes \gamma \otimes D_1 \end{bmatrix}, \tilde{A}_{00}^{23} = I_S \otimes U^0 \otimes I_m, \tilde{A}_{00}^{31} = I_{S+1} \otimes \eta I_m,$$

$$\tilde{A}_{00}^{22} = \begin{bmatrix} C_5 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_6 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_5 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_6 & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & C_5 & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_5 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_6 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_7 & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_7 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_7 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_7 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_7 \end{bmatrix},$$







$$\tilde{F}_1^{44} = \begin{bmatrix} C_{29} & 0 & 0 & \dots & 0 & 0 & \dots & C_{30} & 0 & \dots & 0 & 0 \\ 0 & C_{29} & 0 & \dots & 0 & 0 & \dots & 0 & C_{30} & \dots & 0 & 0 \\ 0 & 0 & C_{29} & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{29} & 0 & \dots & 0 & 0 & \dots & 0 & C_{30} \\ 0 & 0 & 0 & \dots & 0 & C_{31} & \dots & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{31} & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{31} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & C_{31} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{31} \end{bmatrix},$$

$$\tilde{F}_1^{51} = [0 \quad I_S \otimes T^0 \otimes I_m],$$

$$\tilde{F}_1^{55} = \begin{bmatrix} C_{32} & 0 & 0 & \dots & 0 & 0 & \dots & C_{33} & 0 & \dots & 0 & 0 \\ 0 & C_{32} & 0 & \dots & 0 & 0 & \dots & 0 & C_{33} & \dots & 0 & 0 \\ 0 & 0 & C_{32} & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{32} & 0 & \dots & 0 & 0 & \dots & 0 & C_{33} \\ 0 & 0 & 0 & \dots & 0 & C_{34} & \dots & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{34} & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{34} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & C_{34} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{34} \end{bmatrix},$$

### Stability condition for Model II

To discuss the stability condition, we first consider the generator matrix  $F = F_0 + \tilde{F}_1 + F_2$ . If  $\chi = (\chi_0, \chi_1, \chi_2, \chi_3, \chi_4) = (\chi_{00}, \chi_{01}, \dots, \chi_{0s}, \chi_{0s+1}, \dots, \chi_{0Q}, \dots, \chi_{0S}, \chi_{11}, \chi_{12}, \dots, \chi_{1s}, \chi_{1s+1}, \dots, \chi_{1Q}, \dots, \chi_{1S}, \chi_{20}, \chi_{21}, \dots, \chi_{2s}, \chi_{2s+1}, \dots, \chi_{2Q}, \dots, \chi_{2S}, \chi_{31}, \chi_{32}, \dots, \chi_{3s}, \chi_{3s+1}, \dots, \chi_{3Q}, \dots, \chi_{3S}, \chi_{41}, \chi_{42}, \dots, \chi_{4s}, \chi_{4s+1}, \dots, \chi_{4Q}, \dots, \chi_{4S})$ . Considering the QBD structure of the Markov process, stability exists in our model if it satisfies the condition  $\chi F_0 e < \chi F_2 e$ . This condition is both necessary and sufficient for the stability of this queueing model under study, and it reduces to  $\lambda < \mu$ .

### 3.3. The stationary probability vector

Let  $X$  be the stationary probability vector of the infinitesimal generator  $Q$  of the process  $\{X(t); t \geq 0\}$ . The subdivision of  $X = (x_0, x_1, x_2, \dots)$ , where  $x_0$  is of dimension  $2(S+1)m + 2Snm$  and  $x_1, x_2, \dots$  are of dimension  $2(S+1)m + 2Snm + Slm$ . As  $X$  is a vector satisfying the relation  $XQ = 0$  and  $Xe = 1$ . The probability vector  $X$  follows a matrix geometric structure under the steady state is

$$x_j = x_1 R^{j-1}, \quad j \geq 2 \tag{2}$$

where  $R$  is the quadratic equation's lowest non-negative solution  $R^2F_2 + RF_1 + F_0 = 0$  and the vector  $x_0, x_1$  are obtained with the help of succeeding equations:

$$x_0A_{00} + x_1A_{10} + \sum_{i=2}^{\infty} x_iA_i = 0, \tag{3}$$

$$x_0A_{01} + x_1[F_1 + RF_2] = 0, \tag{4}$$

subject to a condition normalization

$$x_0e_{2(s+1)m+2Snm} + x_1[I - R]^{-1}e_{2(s+1)m+2Snm+Slm} = 1. \tag{5}$$

The rate matrix  $R$  can be computed with the help of the following iteration formula which has been suggested by Neuts [10]  $R(n+1) = -F_0F_1^{-1} - R^2(n)F_2F_1^{-1}$  for  $n \geq 0$  where  $R(0) = 0$ . Since  $F_1^{-1}$  and  $(F_0 + R_2F_2)$  are positive, the rate matrix  $R$  will converge and so the entries of  $R$  will increase monotonically in the successive iterations. Iteration may be terminated when the condition  $\max_{i,j}[R_{ij}(n+1) - R_{ij}(n)] < \epsilon$  is attained. Here,  $\epsilon$  denotes the degree of accuracy and  $R(n)$  indicates the value of the rate matrix at the  $n$ -th iteration.

#### 4. SYSTEM CHARACTERISTICS

- Probability that the server is idle in regular process  
 $P_{INM} = \sum_{i=0}^{\infty} \sum_{u_1=0}^S \sum_{u_4=1}^m x_{i0u_1u_4}$ .
- Probability that the server is idle in working vacation process  
 $P_{I WV} = \sum_{i=0}^{\infty} \sum_{u_1=0}^S \sum_{u_4=1}^m x_{i2u_1u_4}$ .
- Probability that the server is busy in regular process  
 $P_{BNM} = \sum_{i=0}^{\infty} \sum_{u_1=1}^S \sum_{u_3=1}^n \sum_{u_4=1}^m x_{i1u_1u_3u_4}$ .
- Probability that the server is busy in working vacation  
 $P_{BWV} = \sum_{i=0}^{\infty} \sum_{u_1=1}^S \sum_{u_3=1}^n \sum_{u_4=1}^m x_{i3u_1u_3u_4}$ .
- Probability that the server is breakdown  
 $P_{BD} = \sum_{i=0}^{\infty} \sum_{u_1=1}^S \sum_{u_2=1}^l \sum_{u_4=1}^m x_{i4u_1u_2u_4}$ .
- Expected number of customers in the orbit  
 $E_{orbit} = \sum_{i=1}^{\infty} ix_i e$ .
- Probability that the server is busy  
 $P_{Busy} = P_{BNM} + P_{BWV}$ .
- Expected number of customers in the system  
 $E_{system} = E_{orbit} + P_{Busy}$ .
- Expected number of items in the inventory level  
 $E_{IL} = \sum_{i=0}^{\infty} \sum_{u_1=1}^S \sum_{u_4=1}^m u_1 x_{i0u_1u_4} + \sum_{i=0}^{\infty} \sum_{u_1=1}^S \sum_{u_3=1}^n \sum_{u_4=1}^m u_1 x_{i1u_1u_3u_4}$   
 $+ \sum_{i=0}^{\infty} \sum_{u_1=1}^S \sum_{u_4=1}^m u_1 x_{i2u_1u_4} + \sum_{i=0}^{\infty} \sum_{u_1=1}^S \sum_{u_3=1}^n \sum_{u_4=1}^m u_1 x_{i3u_1u_2u_3u_4}$   
 $+ \sum_{i=0}^{\infty} \sum_{u_1=1}^S \sum_{u_2=1}^l \sum_{u_4=1}^m u_1 x_{i4u_1u_2u_4}$ .
- Expected reorder rate  
 $E_R = \sum_{i=0}^{\infty} \sum_{u_3=1}^n \sum_{u_4=1}^m x_{i1(s+1)u_3u_4} (U^0 \otimes I_m) e + \sum_{i=0}^{\infty} \sum_{u_3=1}^n \sum_{u_4=1}^m x_{i3(s+1)u_3u_4} (\theta U^0 \otimes I_m) e$ .
- The effective retrial rate  
 $\Delta = \delta \sum_{i=1}^{\infty} \sum_{u_1=1}^S \sum_{u_4=1}^m x_{i0u_1u_4} + \delta \sum_{i=1}^{\infty} \sum_{u_1=1}^S \sum_{u_4=1}^m x_{i2u_1u_4}$ .

#### 5. COST ANALYSIS

The total cost for our model is given below, with the cost elements (per unit time) related to various system measures.

$$TC = c_w E_{system} + c_h E_{IL} + c_s E_R$$

where

- $TC$ : Total cost (per unit time)

- $c_h$ : The inventory holding cost (per unit time)
- $c_w$ : Waiting cost of a customer in the system (per unit time)
- $c_s$ : Setup cost (per order)

## 6. NUMERICAL IMPLEMENTATION

To compute numerical outcomes, we have employed diverse MAP demonstrations for the incoming arrival in a manner that ensures their mean values are 1, as recommended by [5].

- **Erlang arrival (ERA):**

$$D_0 = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix} D_1 = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}$$

- **Exponential arrival (EXA):**

$$D_0 = [-1] D_1 = [1]$$

- **Hyper exponential arrival (HEXA):**

$$D_0 = \begin{bmatrix} -1.90 & 0 \\ 0 & -0.19 \end{bmatrix} D_1 = \begin{bmatrix} 1.710 & 0.190 \\ 0.171 & 0.019 \end{bmatrix}$$

Consider the following PH-distributions for the service and repair progression:

- **Erlang service (ERS):**

$$\gamma = [1, 0] \quad U = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}$$

- **Erlang repair (ERR):**

$$\alpha = [1, 0] \quad T = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}$$

- **Exponential service (EXS):**

$$\gamma = [1] \quad U = [-1]$$

- **Exponential repair (EXR):**

$$\alpha = [1] \quad T = [-1]$$

- **Hyper exponential service (HEXS):**

$$\gamma = [0.8, 0.2] \quad U = \begin{bmatrix} -2.8 & 0 \\ 0 & -0.28 \end{bmatrix}$$

- **Hyper exponential repair (HEXR):**

$$\alpha = [0.8, 0.2] \quad T = \begin{bmatrix} -2.8 & 0 \\ 0 & -0.28 \end{bmatrix}$$

### Illustration 1

For this both policies, it was assumed that values of all parameters of the QIS were fixed except the service rate  $\mu$ :  $\lambda = 1$ ,  $\eta = 3$ ,  $\theta = 0.6$ ,  $\tau = 2$ ,  $\beta = 2$ ,  $\psi = 1$ ,  $\delta = 3$ ,  $\sigma = 0.5$ ,  $p_1 = p_2 = 0.6$ ,  $q_1 = q_2 = 0.4$ ,  $s = 5$ ,  $S = 15$ .

Here, we compare and analyse the two policy  $(s, S)$  and  $(s, Q)$  as follows in tables 1-6:

- First, we observe that both  $E_{system}$  and  $E_{orbit}$  in Table 1-6 under varying service rate  $\mu$ , it is gradually decreases as  $\mu$  increase for both  $(s, S)$  and  $(s, Q)$  but the notable is  $(s, S)$  policy give the minimum for both  $E_{system}$  and  $E_{orbit}$ .
- Observe the service times,  $E_{system}$  and  $E_{orbit}$  are decreases highly in HEXS and slowly decrease in ERS than all other service times. Likewise, from the view point of arrival times,  $E_{system}$  and  $E_{orbit}$  are decreases highly for HEXA compared to other arrival times.

**Table 1:** Service rate ( $\mu$ ) vs  $E_{system}$  and  $E_{orbit}$  - ERA

$\mu$	ERS		EXS		HEXS	
	$E_{system}$	$E_{orbit}$	$E_{system}$	$E_{orbit}$	$E_{system}$	$E_{orbit}$
15	0.081396697	0.047355675	0.116583261	0.046932851	0.060209747	0.031898478
16	0.075864324	0.043882684	0.109224350	0.043648505	0.057304727	0.030565025
17	0.071064407	0.040901610	0.102739862	0.04079181	0.054674483	0.029325579
18	0.066854402	0.038311241	0.096982359	0.038284564	0.052279029	0.028171817
19	0.063128015	0.036037241	0.091835886	0.036066509	0.050086560	0.027096253
20	0.059803954	0.034023538	0.087207964	0.034090466	0.048071247	0.02609211
21	0.056818744	0.032226894	0.083023918	0.032318974	0.046211764	0.025153247
22	0.054121959	0.030613354	0.079222779	0.030721913	0.044490267	0.024274106
23	0.051672939	0.029155827	0.075754267	0.029274789	0.042891663	0.023449656
24	0.049438484	0.027832404	0.072576547	0.027957481	0.041403065	0.022675351

**Table 2:** Service rate ( $\mu$ ) vs  $E_{system}$  and  $E_{orbit}$  - EXA

$\mu$	ERS		EXS		HEXS	
	$E_{system}$	$E_{orbit}$	$E_{system}$	$E_{orbit}$	$E_{system}$	$E_{orbit}$
15	0.093658859	0.057831180	0.125620027	0.057370051	0.077226462	0.047628434
16	0.087616380	0.053656465	0.117884279	0.053393412	0.073004243	0.044783511
17	0.082319352	0.050040468	0.111041640	0.049917853	0.069231262	0.042257589
18	0.077636049	0.046878488	0.104946594	0.046856002	0.065837924	0.039999559
19	0.073464370	0.044090388	0.099483485	0.044139417	0.062768568	0.037968788
20	0.069723889	0.041613800	0.094559255	0.041713775	0.059978093	0.036132537
21	0.066350351	0.039399467	0.090098203	0.039535447	0.057429527	0.034464111
22	0.063291771	0.037407970	0.086038129	0.037569008	0.055092251	0.032941515
23	0.060505619	0.035607397	0.082327459	0.035785419	0.052940673	0.031546455
24	0.057956750	0.033971637	0.078923079	0.034160655	0.050953225	0.030263586

**Table 3:** Service rate ( $\mu$ ) vs  $E_{system}$  and  $E_{orbit}$  - HEXA

$\mu$	ERS		EXS		HEXS	
	$E_{system}$	$E_{orbit}$	$E_{system}$	$E_{orbit}$	$E_{system}$	$E_{orbit}$
15	0.130072755	0.085272901	0.140741030	0.072324015	0.085907558	0.047067013
16	0.118644770	0.076620377	0.131673874	0.066854199	0.080218556	0.043713552
17	0.109278270	0.069644502	0.123722961	0.062135811	0.075394726	0.040903528
18	0.101432497	0.063889862	0.116692238	0.058026255	0.071233209	0.03850289
19	0.094745436	0.059054418	0.110429349	0.054416549	0.06759258	0.036419856
20	0.088964620	0.054929433	0.104814044	0.051222021	0.064370838	0.034589272
21	0.083908013	0.051365737	0.099750091	0.048375934	0.061492291	0.032963458
22	0.079440679	0.048253720	0.095159543	0.045824997	0.058899399	0.031506611
23	0.075460231	0.045510930	0.090978563	0.043526164	0.056547500	0.030191241
24	0.071887408	0.043074081	0.087154360	0.041444300	0.054401305	0.02899583

**Table 4:** Service rate ( $\mu$ ) vs  $E_{system}$  and  $E_{orbit}$  - ERA

$\mu$	ERS		EXS		HEXS	
	$E_{system}$	$E_{orbit}$	$E_{system}$	$E_{orbit}$	$E_{system}$	$E_{orbit}$
15	0.082004602	0.047355519	0.116584109	0.046933442	0.060824563	0.031913956
16	0.076429926	0.043882824	0.109225231	0.043649085	0.057874696	0.030578216
17	0.071593705	0.040901916	0.102740765	0.040792375	0.055206267	0.029336964
18	0.067352119	0.038311641	0.096983274	0.038285111	0.052777843	0.028181749
19	0.063597952	0.036037694	0.091836807	0.036067038	0.050556568	0.027104999
20	0.060249224	0.034024019	0.087208889	0.034090977	0.048515832	0.026099873
21	0.057241938	0.032227388	0.083024844	0.032319470	0.046633719	0.025160188
22	0.054525258	0.030613851	0.079223704	0.030722393	0.044891928	0.024280350
23	0.052058203	0.029156322	0.075755191	0.029275254	0.043275004	0.023455306
24	0.049807313	0.027832894	0.072577470	0.027957933	0.041769774	0.022680491

**Table 5:** Service rate ( $\mu$ ) vs  $E_{system}$  and  $E_{orbit}$  - EXA

$\mu$	ERS		EXS		HEXS	
	$E_{system}$	$E_{orbit}$	$E_{system}$	$E_{orbit}$	$E_{system}$	$E_{orbit}$
15	0.094342828	0.057831783	0.125622511	0.057371230	0.077912574	0.047638686
16	0.088262387	0.053657353	0.117886912	0.053394652	0.073653585	0.044793161
17	0.082931496	0.050041572	0.111044412	0.049919149	0.069847634	0.042266700
18	0.078217783	0.046879757	0.104949496	0.046857350	0.066424553	0.040008185
19	0.074018636	0.044091787	0.099486509	0.044140813	0.063328223	0.037976979
20	0.070253213	0.041615303	0.094562394	0.041715216	0.060513165	0.036140334
21	0.066856921	0.039401053	0.090101450	0.039536929	0.057942099	0.034471552
22	0.063777495	0.037409625	0.086041476	0.037570530	0.055584149	0.032948633
23	0.060972173	0.035609109	0.082330902	0.035786976	0.053413507	0.031553280
24	0.058405612	0.033973398	0.078926612	0.034162246	0.051408424	0.030270142

**Table 6:** Service rate ( $\mu$ ) vs  $E_{system}$  and  $E_{orbit}$  - HEXA

$\mu$	ERS		EXS		HEXS	
	$E_{system}$	$E_{orbit}$	$E_{system}$	$E_{orbit}$	$E_{system}$	$E_{orbit}$
15	0.131245994	0.085194321	0.14075261	0.072332391	0.087390312	0.047182656
16	0.119735103	0.076569695	0.131688802	0.066864727	0.081577421	0.043823642
17	0.110296386	0.069612267	0.123740628	0.062148038	0.076649406	0.041007718
18	0.102387591	0.063870343	0.116712188	0.058039850	0.072399139	0.03860134
19	0.09564524	0.059043966	0.110451236	0.054431263	0.068682030	0.036512935
20	0.089815602	0.054925634	0.104837597	0.051237669	0.065393700	0.034677418
21	0.084715599	0.051366937	0.099775098	0.048392369	0.062456649	0.033047113
22	0.080209444	0.048258756	0.095185831	0.045842105	0.059811936	0.031586194
23	0.07619406	0.045518963	0.091005993	0.043543856	0.057413806	0.030267136
24	0.072589624	0.043084494	0.087182817	0.041462502	0.055226098	0.029068385

### Illustration 2

We picture the consequences of the breakdown rate  $\psi$  against the  $P_{busy}$ . Fix  $\lambda = 1$ ,  $\mu = 15$ ,  $\theta = 0.6$ ,  $\eta = 3$ ,  $\tau = 5$ ,  $\beta = 2$ ,  $\delta = 3$ ,  $\sigma = 0.5$ ,  $p_1 = p_2 = 0.6$ ,  $q_1 = q_2 = 0.4$ ,  $s = 5$ ,  $S = 15$ , these values satisfy the condition for stability. From the figures 2 - 4: we can explore that while increasing the server's breakdown rate ( $\psi$ ),  $P_{busy}$  decreases for all feasible provisions of incoming arrival and service patterns. As increase in breakdown rate indicates that customers will frequently be unable to access the server, which is decreases of  $P_{busy}$  is higher for HEXA and lower for ERA. Like wise, it is higher for ERS and lower for HEXS.



Illustration 3

To investigate the impact of the  $TC$  on both the service ( $\mu$ ) and repair ( $\tau$ ) rates in the Figures 5-13. Fix  $\lambda = 1, \sigma = 0.2, \theta = 0.6, \beta = 3, \delta = 3, p_1 = p_2 = 0.6, q_1 = q_2 = 0.4, s = 5, S = 15, C_H = 70, C_I = 110, C_R = 120$ , such that the system leftovers stable.

From the viewpoint of Figures 5-13, we maximize both the service and repair rates for all possible groups of arrival and service times, we notice that the  $TC$  decreases. Consider the service times,  $TC$  decreases exceedingly for  $ERS$  and decreases moderately for  $EXS$ . Therefore,  $TC$  decreases slowly for  $ERA$  and rapidly for  $HEXA$ .

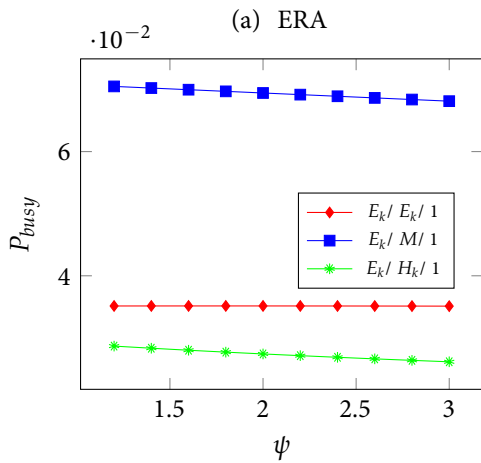


Figure 2: Breakdown rate vs.  $P_{busy}$

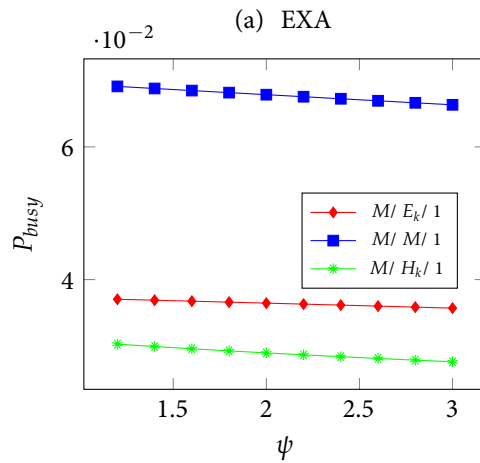


Figure 3: Breakdown rate vs.  $P_{busy}$

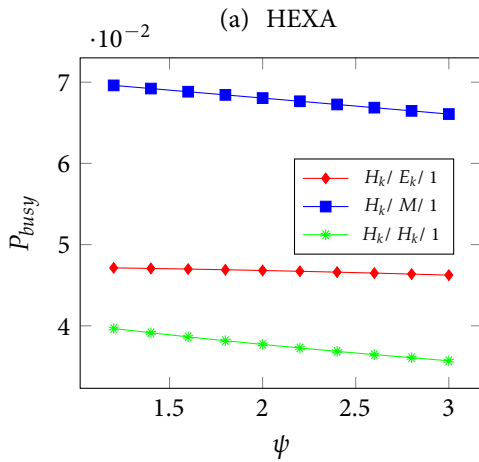


Figure 4: Breakdown rate vs.  $P_{busy}$

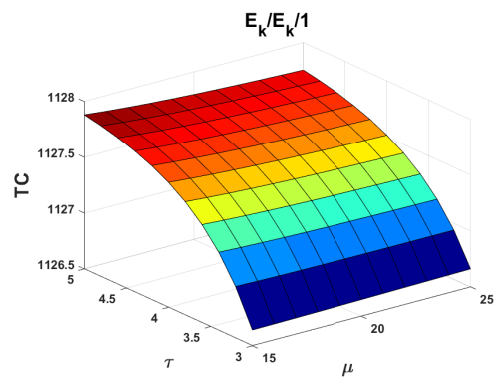


Figure 5: Service and repair rates vs.  $TC$

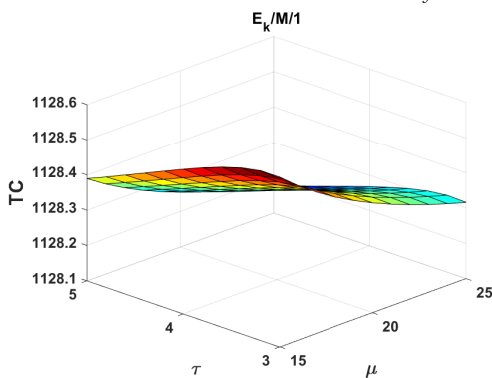


Figure 6: Service and repair rates vs.  $TC$

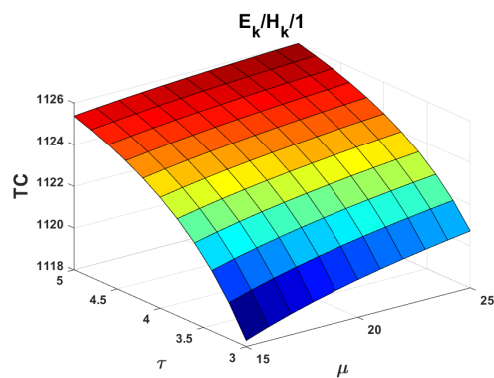


Figure 7: Service and repair rates vs.  $TC$

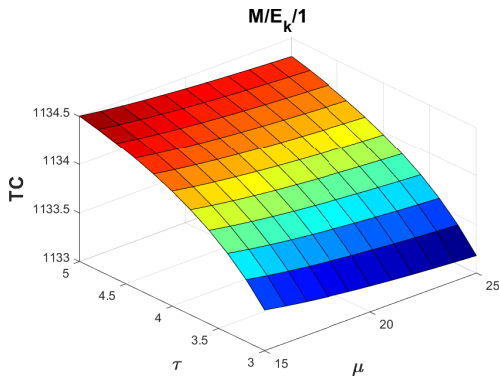


Figure 8: Service and repair rates vs. TC

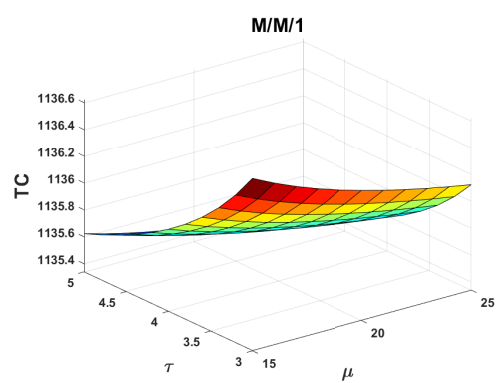


Figure 9: Service and repair rates vs. TC

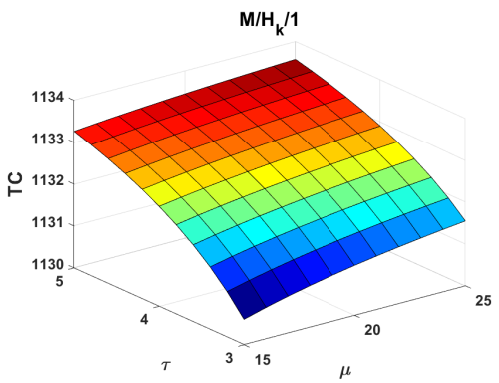


Figure 10: Service and repair rates vs. TC

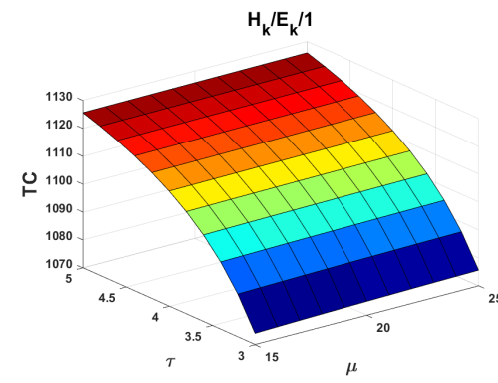


Figure 11: Service and repair rates vs. TC

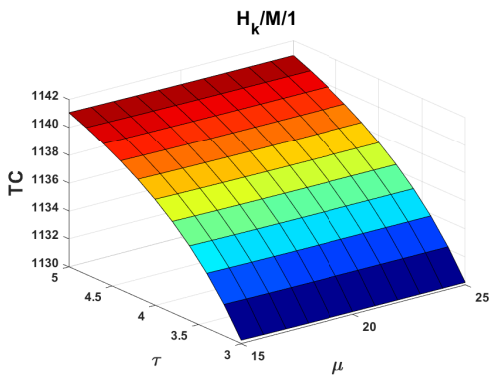


Figure 12: Service and repair rates vs. TC

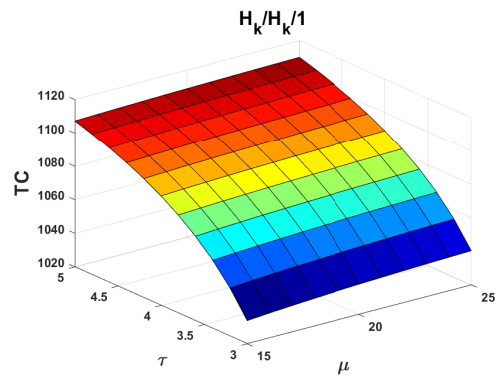


Figure 13: Service and repair rates vs. TC

## 7. CONCLUSION

A retrial inventory model with MAP arrivals, PH-distributed service, working vacations, collision of orbital customers, flush out, balking, breakdown and repair has been investigated. The peculiarity of this model is that the server can offer service even in the vacation period and the system is always stable because of the flush out of the system. We have considered MAP for arrivals and would like to extend our models by considering BMAP for arrivals which is best suited for modelling arrivals which come in batches.

## REFERENCES

- [1] Amirthakodi, M. and Sivakumar, B. (2014). An inventory system with service facility and finite orbit size for feedback customers, *OPSEARCH*, 52(2): 225-255.
- [2] Artalejo, J.R. and Gomez-Corral, A. *Retrial Queueing Systems: A Computational Approach*, Springer: London, UK, 2008.
- [3] Artalejo, J.R., Krishnamoorthy, A. and Lopez-Herrero, M.J. (2006). Numerical analysis of (s, S) inventory systems with repeated attempts, *Annals of Operations Research*, 141: 67-83.
- [4] Ayyappan, G., Udayageetha, J. and Somasundaram, B. (2020). Analysis of non-preemptive priority retrial queueing system with two-way communication, Bernoulli vacation, collisions, working breakdown, immediate feedback and reneging, *International Journal of Mathematics in Operational Research*, 16(4):480-498.
- [5] Chakravarthy, S.R. Markovian arrival processes, In *Wiley Encyclopedia of Operations Research and Management Science*, Wiley: Hoboken, NJ, USA, 2010.
- [6] Chakravarthy, S.R. and Hayat, K. (2020). Queueing-Inventory Models for a Two-Vendor System with Positive Service Times, *Queueing Models and Service Management*, 3(1): 1-35.
- [7] Krishnamoorthy, A., Nair, S.S. and Narayanan, V.C. (2010). An inventory model with server interruptions, In: *Proceedings of the 5th International Conference on Queueing Theory and Network Applications*, 1(1): 132-139.
- [8] Manuel, P., Sivakumar, B., & Arivarganan, G. (2008). A perishable inventory system with service facilities and retrial customers, *Computers & Industrial Engineering*, 54(3): 484- 501.
- [9] Melikov, A., Aliyeva, S., Nair, S.S., Kumar, B.K. (2022). Retrial Queueing- Inventory Systems with Delayed Feedback and Instantaneous Damaging of Items, *Axioms*, 11(5):241.
- [10] Neuts, M. F. (1981). *Matrix-geometric solutions in stochastic models: An algorithmic approach*, The Johns Hopkins University Press, Baltimore and London.
- [11] Nithya, M., Sugapriya, C., Selvakumar, S., Jeganathan, K., and Harikrishnan, T., (2022). A Markovian two-commodity queueing-inventory system with complimentary item and classical retrial facility, *Ural Mathematical Journal*, 8(1):90-116.
- [12] Jain, N. K, Kumar, R. and Som, B.K. (2014). An M/M/1/N queueing system with reverse balking, *American Journal of Operational Research*, 4(2):17-20.
- [13] Kumar, R. and Som, B.K. (2015). An M/M/1/N Queueing System with Reverse Balking, Reverse Reneging, and Retention of Reneged Customers, *Indian Journal of Industrial and Applied Mathematics*, 6(1):73-79.
- [14] Shajin, D. and Krishnamoorthy, A. (2020). Stochastic decomposition in retrial queueing-inventory system, *RAIRO Oper. Res.*, 54(1):81-99.
- [15] Sivakumar, B. (2008). Two-commodity inventory system with retrial demand, *European Journal of Operational Research*, 187(1): 70-83.
- [16] Sivakumar, B. (2009). A perishable inventory system with retrial demands and a finite population, *Journal of Computational and Applied Mathematics*, 224(1):29-38.
- [17] Sivakumar, B. (2011). An inventory system with retrial demands and multiple server vacation, *Quality Technology and Quantitative Management*, 8(2):125-146.
- [18] Ushakumari, P.V. (2006). On (s, S) inventory system with random lead time and repeated demands, *Journal of Applied Mathematics and Stochastic Analysis*, 2006:1-22.
- [19] Zhang, Y., Yue, D., Sun, L. and Zuo, J. (2022). Analysis of the Queueing-Inventory System with Impatient Customers and Mixed Sales, *Discrete Dynamics in Nature and Society, Hindawi*, 2022(1):1-12.

# ON $\varphi$ -CONHARMONICALLY FLAT LORENTZIAN PARA-KENMOTSU MANIFOLDS

I. V. VENKATESWARA RAO<sup>1</sup>, S. SUNITHA DEVI<sup>2</sup> AND K. L. SAI PRASAD<sup>3,\*</sup>

•  
Department of Mathematics

<sup>1</sup> P. B. Siddhartha College of Arts and Science, Vijayawada, Andhra Pradesh, India

<sup>2</sup> KL University, Vijayawada, Andhra Pradesh, 520002, INDIA

<sup>3,\*</sup> Gayatri Vidya Parishad College of Engineering for Women, Visakhapatnam, 530 048, INDIA

venkat\_inturi@rediffmail.com <sup>1</sup> sunithamallakula@yahoo.com <sup>2</sup> klsprasad@yahoo.com <sup>3,\*</sup>

## Abstract

*The present paper deals with a class of Lorentzian almost paracontact metric manifolds namely Lorentzian para-Kenmotsu (briefly LP-Kenmotsu) manifolds. We study and have shown that a quasi-conformally flat Lorentzian para-Kenmotsu manifold is locally isomorphic with a unit sphere  $S^n(1)$ . Further it is shown that an LP-Kenmotsu manifold which is  $\varphi$ -conharmonically flat is an  $\eta$ -Einstein manifold with the zero scalar curvature. At the end, we have shown that a  $\varphi$ -projectively flat LP-Kenmotsu manifold is an Einstein manifold with the scalar curvature  $r = n(n - 1)$ .*

**Keywords:** Lorentzian para-Kenmotsu manifold, Weyl-projective curvature tensor, conformal curvature tensor, Einstein manifold.

**2010 Mathematics Subject Classification:** 53C07, 53C05, 53C15

## I. INTRODUCTION

In 1989, K. Matsumoto [3] introduced the notion of Lorentzian paracontact and in particular, Lorentzian para-Sasakian (briefly LP-Sasakian) manifolds. Later, these manifolds have been widely studied by many geometers such as Matsumoto and Mihai [5], Mihai and Rosca [6], Mihai, Shaikh and De [7], Venkatesha and Bagewadi [16], Venkatesha, Pradeep Kumar and Bagewadi [17, 18] and obtained several results of these manifolds.

In 1995, Sinha and Sai Prasad [14] defined a class of almost paracontact metric manifolds namely para-Kenmotsu (briefly P-Kenmotsu) and Special Para-Kenmotsu (briefly SP-Kenmotsu) manifolds in similar to P-Sasakian and SP-Sasakian manifolds. In 2018, Abdul Haseeb and Rajendra Prasad defined a class of Lorentzian almost paracontact metric manifolds namely Lorentzian para-Kenmotsu (briefly LP-Kenmotsu) manifolds [1]. As an extension, Rajendra Prasad *et al.*, [10] have studied  $\varphi$ -semisymmetric LP-Kenmotsu manifolds with a quarter-symmetric non-metric connection admitting Ricci solitons.

On the other hand, In 1970, Pokhariyal and Mishra [9] introduced new tensor fields, called the Weyl-projective curvature tensor  $P(X, Y)Z$  of type (1, 3) and the tensor field  $E$  on a Riemannian manifold. Further many geometers have studied the properties of these tensor fields [2, 4, 8, 11, 12, 13, 15] as they play an important role in the theory of projective transformations of connections.

The projective curvature tensor  $P(X, Y)Z$ , with respect to the Riemannian connection on a Riemannian manifold  $(M_n, g)$ , is given by:

$$P(X, Y)Z = R(X, Y)Z + \frac{1}{n-1}[g(X, Z)QY - g(Y, Z)QX],$$

where  $QX = (n-1)X$ , and the Riemannian Christoffel curvature tensor  $R$  of type (1, 3) is given by:

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]}Z. \tag{1}$$

Here  $\nabla$  is said to be the Levi-Civita connection.

In the present work, we study a class of LP-Kenmotsu manifolds and it is organized as follows. Section 2 is equipped with some prerequisites about Lorentzian para-Kenmotsu manifolds. In section 3, we study the quasi-conformally flat Lorentzian para-Kenmotsu manifolds. Sections 4 and 5 respectively deals with  $\varphi$ -conharmonically flat and  $\varphi$ -projectively flat LP-Kenmotsu manifolds.

## II. PRELIMINARIES

An  $n$ -dimensional differentiable manifold  $M_n$  admitting a (1, 1) tensor field  $\phi$ , contravariant vector field  $\xi$ , a 1-form  $\eta$  and the Lorentzian metric  $g(X, Y)$  satisfying

$$\eta(\xi) = -1, \tag{2}$$

$$\phi^2 X = X + \eta(X)\xi, \tag{3}$$

$$g(\phi X, \phi Y) = g(X, Y) + \eta(X)\eta(Y), \tag{4}$$

$$g(X, \xi) = \eta(X), \tag{5}$$

$$\phi\xi = 0, \eta(\phi X) = 0, \text{rank } \phi = n - 1; \tag{6}$$

is called Lorentzian almost paracontact manifold [3].

In a Lorentzian almost paracontact manifold, we have

$$\Phi(X, Y) = \Phi(Y, X) \text{ where } \Phi(X, Y) = g(\phi X, Y). \tag{7}$$

A Lorentzian almost paracontact manifold  $M_n$  is called Lorentzian para-Kenmotsu manifold if [1]

$$(\nabla_X \phi)Y = -g(\phi X, Y)\xi - \eta(Y)\phi X, \tag{8}$$

for any vector fields  $X$  and  $Y$  on  $M_n$ , and  $\nabla$  is the operator of covariant differentiation with respect to the Lorentzian metric  $g$ .

It can be easily seen that in a LP-Kenmotsu manifold  $M_n$ , the following relations hold [1]:

$$\nabla_X \xi = -\phi^2 X = -X - \eta(X)\xi, \tag{9}$$

$$(\nabla_X \eta)Y = -g(X, Y)\xi - \eta(X)\eta(Y), \tag{10}$$

for any vector fields  $X$  and  $Y$  on  $M_n$ .

Also, in an LP-Kenmotsu manifold, the following relations hold [1]:

$$g(R(X, Y)Z, \xi) = \eta(R(X, Y)Z) = g(Y, Z)\eta(X) - g(X, Z)\eta(Y) \tag{11}$$

$$R(\xi, X)Y = g(X, Y)\xi - \eta(Y)X, \tag{12}$$

$$R(X, Y)\xi = \eta(Y)X - \eta(X)Y, \tag{13}$$

$$S(X, \xi) = (n - 1)\eta(X), \tag{14}$$

$$S(\phi X, \phi Y) = S(X, Y) + (n - 1)\eta(X)\eta(Y), \tag{15}$$

$$S(X, Y) = ag(X, Y) + b\eta(X)\eta(Y); \tag{16}$$

for any vector fields  $X, Y$  and  $Z$ , where  $R$  is the Riemannian curvature tensor and  $S$  is the Ricci tensor of  $M_n$ .

### III. LP-KENMOTSU MANIFOLDS WITH $\tilde{C}(X, Y)Z = 0$

The quasi-conformal curvature tensor  $\tilde{C}$  is defined as

$$\begin{aligned} \tilde{C}(X, Y)Z = & aR(X, Y)Z + b\{S(Y, Z)X - S(X, Z)Y + g(Y, Z)QX \\ & - g(X, Z)QY\} - \frac{r}{n} \left( \frac{a}{n-1} + 2b \right) \{g(Y, Z)X - g(X, Z)Y\} \end{aligned} \tag{17}$$

where  $a, b$  are constants such that  $ab \neq 0$  and

$$S(Y, Z) = g(QY, Z).$$

From (17), we get

$$\begin{aligned} R(X, Y)Z = & -\frac{b}{a}\{S(Y, Z)X - S(X, Z)Y + g(Y, Z)QX \\ & - g(X, Z)QY\} + \frac{r}{n} \left( \frac{a}{n-1} + 2b \right) \{g(Y, Z)X - g(X, Z)Y\}. \end{aligned} \tag{18}$$

Taking  $Z = \xi$  in (18) and on using (5), (13), (14), we get

$$\eta(Y)X - \eta(X)Y = -\frac{b}{a}\{\eta(Y)QX - \eta(X)QY\} \left\{ \frac{r}{an} \left( \frac{a}{n-1} + 2b \right) - \frac{b}{a}(n-1) \right\} \{\eta(Y)X - \eta(X)Y\}. \tag{19}$$

Taking  $Y = \xi$  and applying (2) we have

$$\begin{aligned} QX = & \left\{ \frac{r}{bn} \left( \frac{a}{n-1} + 2b \right) - (n-1) - \frac{a}{b} \right\} X \\ & + \left\{ \frac{r}{bn} \left( \frac{a}{n-1} + 2b \right) - \frac{a}{b} - 2(n-1) \right\} \eta(X)\xi. \end{aligned} \tag{20}$$

Contracting (20), we get after a few steps

$$r = n(n-1). \tag{21}$$

Using (21) in (20), we get

$$QX = (n-1)X. \tag{22}$$

Finally, using (22), we find from (18)

$$R(X, Y)Z = g(Y, Z)X - g(X, Z)Y.$$

Thus, we state

**Theorem 3.1:** A quasi-conformally flat LP-Kenmotsu manifold is locally isometric with a unit sphere  $S^n(1)$ .

#### IV. LP-KENMOTSU MANIFOLDS WITH $\varphi$ -CONHARMONICALLY FLAT CURVATURE TENSOR

The conharmonic curvature tensor  $K$  is defined as

$$K(X, Y)Z = R(X, Y)Z - \frac{1}{n-2} [S(Y, Z)X - S(X, Z)Y + g(Y, Z)SX - g(X, Z)SY].$$

A differentiable manifold  $(M_n, g), n > 3$ , satisfying the condition

$$\varphi^2 K(\varphi X, \varphi Y)\varphi Z = 0 \tag{23}$$

is called  $\varphi$ -conharmonically flat.

In this section, we study LP-Kenmotsu manifolds with the condition (23).

**Theorem 4.1:** Let  $M_n$  be an  $n$ -dimensional,  $(n > 3)$ ,  $\varphi$ -conharmonically flat LP-Kenmotsu manifold. Then  $M_n$  is an  $\eta$ -Einstein manifold with the zero-scalar curvature.

**Proof:** Assume that  $(M_n, g), n > 3$ , is a  $\varphi$ -conformally flat LP-Kenmotsu manifold. It can be easily seen that  $\varphi^2 K(\varphi X, \varphi Y)\varphi Z = 0$  holds if and only if

$$g(K(\varphi X, \varphi Y)\varphi Z, \varphi W) = 0,$$

for any  $X, Y, Z, W \in \chi(M_n)$ .

$$g(R(\varphi X, \varphi Y)\varphi Z, \varphi W) = \frac{1}{n-2} [g(\varphi Y, \varphi Z)S(\varphi X, \varphi W) - g(\varphi X, \varphi Z)S(\varphi Y, \varphi W) + g(\varphi X, \varphi W)S(\varphi Y, \varphi Z) - g(\varphi Y, \varphi W)S(\varphi X, \varphi Z)]. \tag{24}$$

We suppose that  $\{e_1, \dots, e_{n-1}, \xi\}$  is a local orthonormal basis of vector fields in  $M_n$ . By using the fact that  $\{\varphi e_1, \dots, \varphi e_{2n}, \xi\}$  is also a local orthonormal basis, if we put  $X=W=e_i$  in (23) and sum up with respect to  $i$ , then

$$\sum_{i=1}^{n-1} g(R(\varphi e_i, \varphi Y)\varphi Z, \varphi e_i) = \frac{1}{n-2} \sum_{i=1}^{n-1} [g(\varphi Y, \varphi Z)S(\varphi e_i, \varphi e_i) - g(\varphi e_i, \varphi Z)S(\varphi Y, \varphi e_i) + g(\varphi e_i, \varphi e_i)S(\varphi Y, \varphi Z) - g(\varphi Y, \varphi e_i)S(\varphi e_i, \varphi Z)], \tag{25}$$

where

$$\sum_{i=1}^{n-1} g(R(\varphi e_i, \varphi Y)\varphi Z, \varphi e_i) = S(\varphi Y, \varphi Z) + g(\varphi Y, \varphi Z), \tag{26}$$

$$\sum_{i=1}^{n-1} S(\varphi e_i, \varphi e_i) = r + n - 1, \tag{27}$$

$$\sum_{i=1}^{n-1} g(\varphi e_i, \varphi Z)S(\varphi Y, \varphi e_i) = S(\varphi Y, \varphi Z), \tag{28}$$

$$\sum_{i=1}^{n-1} g(\varphi e_i, \varphi e_i) = n + 1. \tag{29}$$

So, by the use of (26)-(29) the equation (25) turns into

$$-S(\varphi Y, \varphi Z) = (r+1)g(\varphi Y, \varphi Z). \tag{30}$$

Then by using (4) and (15), from equation (30) we get

$$S(Y, Z) = -(r+1)g(Y, Z) - (n+r)\eta(Y)\eta(Z), \tag{31}$$

which gives us, from (16),  $M_n$  is an  $\eta$ -Einstein manifold. Hence on contracting (31) we obtain  $nr=0$ , which implies the scalar curvature  $r=0$ , which proves the theorem.

V. LP-KENMOTSU MANIFOLDS WITH  $\varphi$ -PROJECTIVELY FLAT CURVATURE TENSOR

A differentiable manifold  $(M_n, g), n > 3$ , satisfying the condition

$$\varphi^2 P(\varphi X, \varphi Y)\varphi Z = 0 \tag{32}$$

is called  $\varphi$ -projectively flat, where  $P(X, Y)Z$  is the Weyl-projective curvature tensor of  $(M_n, g)$ .

**Theorem 5.1:** Let  $M_n$  be an  $n$ -dimensional,  $(n > 3)$ ,  $\varphi$ -projectively flat LP-Kenmotsu manifold. Then  $M_n$  is an Einstein manifold with the scalar curvature  $r = n(n-1)$ .

**Proof:** It can be easily seen that  $\varphi^2 P(\varphi X, \varphi Y)\varphi Z = 0$  holds if and

$$g(P(\varphi X, \varphi Y)\varphi Z, \varphi W) = 0,$$

for any  $X, Y, Z, W \in \chi(M_n)$ .

$$g(R(\varphi X, \varphi Y)\varphi Z, \varphi W) = \frac{1}{n-2} [g(\varphi Y, \varphi Z)S(\varphi X, \varphi W) - g(\varphi X, \varphi Z)S(\varphi Y, \varphi W)]. \tag{33}$$

By choosing  $\{e_1, \dots, e_{n-1}, \xi\}$  as a local orthonormal basis of vector fields in  $M_n$  and using the fact that  $\{\varphi e_1, \dots, \varphi e_{n-1}, \xi\}$  as a local orthonormal basis, on putting  $X=W=e_i$  in (33) and summing up with respect to  $i$ , we have

$$\sum_{i=1}^{n-1} g(R(\varphi e_i, \varphi Y)\varphi Z, \varphi e_i) = \frac{1}{n-2} \sum_{i=1}^{n-1} [g(\varphi Y, \varphi Z)S(\varphi e_i, \varphi e_i) - g(\varphi e_i, \varphi Z)S(\varphi Y, \varphi e_i)]. \tag{34}$$

Therefore, by using (26)-(29) into (34) we get

$$nS(\varphi Y, \varphi Z) = rg(\varphi Y, \varphi Z).$$

Hence by virtue of (4) and (15) we obtain

$$S(Y, Z) = \frac{r}{n}g(Y, Z) + \left(\frac{r}{n} - (n-1)\right)\eta(Y)\eta(Z). \tag{35}$$

Therefore from (35), by contraction, we obtain

$$r = n(n-1). \tag{36}$$

Then by substituting (36) into (35) we get

$$S(Y, Z) = (n-1)g(Y, Z),$$

which implies  $M_n$  is an Einstein manifold with the scalar curvature  $r = n(n-1)$ .

This completes the proof of the theorem.

**Acknowledgements:** The authors acknowledge Dr. A. Kameswara Rao, Assistant Professor of G.V.P. College of Engineering for Women for his valuable suggestions in preparation of the manuscript.

**Conflicts of interest:** The authors declare that there is no conflict of interests regarding the publication of this paper.


REFERENCES

[1] Abdul Haseeb and Rajendra Prasad. Certain results on Lorentzian para-Kenmotsu manifolds. Bulletin of Parana's Mathematical Society, (2018), 201-220.



- [2] U. C. De and Avijit Sarkar . On a type of P-Sasakian manifolds. Math. Reports, 11(2009), no. 2, 139-144.
- [3] K. Matsumoto. On Lorentzian Paracontact manifolds. Bulletin of the Yamagata University Natural Science, 12 (2)(1989), 151-156.
- [4] K. Matsumoto, S. Ianus and I. Mihai. On P-Sasakian manifolds which admit certain tensor fields. Publ. Math. Debrecen, 33(1986), 61-65.
- [5] K. Matsumoto and I. Mihai. On a certain transformation in a Lorentzian para-Sasakian manifold. Tensor, N.S., 47(1988), 189-197.
- [6] I. Mihai and R. Rosca. On Lorentzian P-Sasakian manifolds. Classical Analysis, World Scientific Publ., Singapore,(1992), 155-169.
- [7] I. Mihai, A. A. Shaikh, U.C. De. On Lorentzian para-Sasakian manifolds. Rendiconti del Seminario Matematico de Messina, Serie II (1999), 75-77.
- [8] G. P. Pokhariyal. Study of a new curvature tensor in a Sasakian manifold. Tensor (N.S.), 36(1982), 222-225.
- [9] G. P. Pokhariyal and R. S. Mishra. The curvature tensors and their relativistic significance. Yokohama Math. J., 18(1970), 105-108.
- [10] Rajendra Prasad, Abdul Haseeb and Umesh Kumar Gautam. On  $\phi$ -semisymmetric LP Kenmotsu manifolds with a QSNM-connection admitting Ricci solitons. Kragujevac Journal of Mathematics, 45(5)(2021), 815-827.
- [11] K. L. Sai Prasad and T. Satyanarayana. Some curvature properties on a Special paracontact Kenmotsu manifold with respect to Semi-symmetric connection. Turkish Journal of Analysis and Number Theory, 3(2015), no. 4, 94-96.
- [12] K. L. Sai Prasad, S. Sunitha Devi and G. V. S. R. Deekshitulu. On a class of Lorentzian para Kenmotsu manifolds admitting the Weyl-projective curvature tensor of type (1,3). Italian Journal of Pure and Applied Mathematics, 45(2021), 990-1001.
- [13] K. L. Sai Prasad, S. Sunitha Devi and G. V. S. R. Deekshitulu. On a class of Lorentzian paracontact metric manifolds. Italian Journal of Pure and Applied Mathematics, 49(2023), 514-527.
- [14] B. B. Sinha and K. L. Sai Prasad. A class of almost paracontact metric Manifold. Bulletin of the Calcutta Mathematical Society, 87(1995), 307-312.
- [15] S. Sunitha Devi, K. L. Sai Prasad and T. Satyanarayana. Certain curvature conditions on Lorentzian para-Kenmotsu Manifolds. Reliability Theory & Applications, 17(2022), no. 2(68), 413-421.
- [16] Venkatesha and C. S. Bagewadi, On concircular  $\mathcal{C}\mathcal{E}$ -recurrent LP-Sasakian manifolds. Differ. Geom. Dyn. Syst., 10(2008), 312-319.
- [17] Venkatesha, K. T. Pradeep Kumar and C. S. Bagewadi. On Lorentzian para-Sasakian manifolds satisfying  $W_2$  curvature tensor. IOSR J. of Mathematics, 9(2014), no. 6, 124-127.
- [18] Venkatesha, K. T. Pradeep Kumar and C. S. Bagewadi. On quarter-symmetric metric connection in a Lorentzian para-Sasakian manifold. Ajerbaijan Journal of Mathematics, 5(2015), no. 1, 3-12.

# DATA ANALYSIS AND CLASSICAL ESTIMATION METHODS OF THE BOUNDED POWER LOMAX DISTRIBUTION

Amal S. Hassan <sup>1</sup> , Asma M. Khalil <sup>2</sup> and Heba F. Nagy <sup>3,\*</sup> 

1 Faculty of Graduate Studies for Statistical Research, Cairo University, 12613, Giza, Egypt:  
amal52\_soliman@cu.edu.eg

2 Faculty of Graduate Studies for Statistical Research, Cairo University, 12613, Giza, Egypt:  
12422020463929@pg.cu.edu.eg

3 Faculty of Graduate Studies for Statistical Research, Cairo University, 12613, Giza, Egypt:  
heba\_nagy\_84@cu.edu.eg

\* **Correspondence:** Email: [heba\\_nagy\\_84@cu.edu.eg](mailto:heba_nagy_84@cu.edu.eg)

## Abstract

*In this work, a novel bounded three-parameter power Lomax distribution termed the unit power Lomax (UPLoD) is presented. The UPLoD is capable of handling data with left and right skewed shapes according to its probability density function. Additionally, according to the hazard rate function, the distribution may be used to analyse data containing J-shaped hazard rates. It is possible to determine some of the distribution's mathematical characteristics like moments, probability-weighted moments, incomplete moments, residual and reversed residual life, quantile function, stress strength model, and entropy (Rényi, Havrda and Charvát, Tsallis, and Arimoto) measures. The Cramér–von Mises, weighted least squares, maximum likelihood, Anderson–Darling, maximum product of spacing, and least squares approaches are among the conventional estimating techniques that are taken into account. The performance of the resulting estimates is compared using a Monte Carlo simulation based on some precision metrics. An actual data application is presented using water capacity data, and data about the Susquehanna River's maximum flood levels to show the importance of the new distribution compared to several other known distributions.*

**Keywords:** Unit Power Lomax distribution, Entropy measures, Parameter estimation, Goodness-of-fit test.

## 1 Introduction

The Lomax distribution (LoD), sometimes referred to as the Pareto II distribution, was first presented by Lomax [1] to model business failure data, but it has since been widely used in a wide range of applications. Harris [2] utilized the LoD for data on wealth and income, In the case of severely tailed data, Bryson [3] suggested employing it in place of the exponential distribution. Atkinson and Harrison [4] used it to model data on business failure. It was utilized in the biological sciences and even for modeling the distribution of server computer file sizes, as mentioned by Holland *et al.* [5]. The LoD has been used to model a variety of data that many writers have explored.

Rady *et al.* [6] presented the power Lomax distribution (PLoD) as a generalization of the LoD that includes an additional shape parameter. The PLoD has been used in many applications and fields, like

those pertaining to biological sciences, engineering sciences, medical research, econometrics, and life testing. The PLoD's probability density function (PDF) is provided by:

$$f(x; \omega) = \kappa \tau^{-1} \eta x^{\kappa-1} \left(1 + \tau^{-1} x^{\kappa}\right)^{-\eta-1}; x > 0, \quad (1)$$

where  $\omega = (\kappa, \eta, \tau)$  is the set of parameters,  $\kappa > 0$  and  $\eta > 0$  are shape parameters and  $\tau > 0$  is the scale parameter. The following defines the PLoD's cumulative distribution function (CDF)

$$F(x; \omega) = 1 - \left(1 + \frac{x^{\kappa}}{\tau}\right)^{-\eta}; x > 0. \quad (2)$$

On the other hand, statistics professionals have recently become interested in the creation and the development of novel probability distributions that can provide models that fit datasets ranging from zero to one. To model proportions, percentages, and probabilities, bounded distributions are required. The study of datasets on  $(0, 1)$  regard to parametric or semi-parametric regression models is crucial in applied fields. Additionally, unit distributions add more flexibility over the course of the unit interval without changing the core distribution's properties. When modeling proportions that are typically seen in industry, medical applications, and risk analysis, unit distributions are an essential tool.

Here are a few of the most significant unit distributions with varying numbers of parameters. The log Lindley distribution (Gómez-Déniz *et al.* [7]), unit-Gompertz distribution (UGoD) (Mazucheli *et al.* [8]), unit Lindley distribution (ULD) (Mazucheli *et al.* [9]), unit modified Burr-III distribution (Haq *et al.* [10]), unit generalized half normal distribution (Korkmaz [11]), unit-Weibull distribution (UWD) (Mazucheli *et al.* [12]), unit Gamma/Gompertz distribution (UG/GD) (Bantan *et al.* [13]), unit log logistic distribution (ULLD) Ribeiro-Reis [14]), unit Burr-XII distribution (UBXIID) (Korkmaz and Chesneau [15]), unit half-logistic geometric distribution (Ramadan *et al.* [16]), unit power-skew-normal distribution (Martínez-Flórez *et al.* [17]), unit exponentiated half logistic distribution (Hassan *et al.* [18]), unit Teissier distribution (Krishna *et al.* [19]), unit Xgamma distribution (Hashmi *et al.* [20]), unit-exponentiated Pareto distribution (UEPD) (Haj Ahmad *et al.* [21]), unit exponentiated Lomax (Fayomi *et al.* [22]), Kumaraswamy unit-Gompertz distribution (Akata *et al.* [23]), unit inverse exponentiated Weibull distribution (Hassan and Alharbi [24]) and unit-power Burr X distribution (Fayomi *et al.* [25]).

The main goal of this work is to present a new and adaptable probabilistic model for the PLoD with a domain  $(0,1)$ . This model refers to the unit PLoD (UPLoD) that can be used to evaluate a wide range of data sets with values ranging from zero to one. The UPLoD is presented in light of the following details:

- a) To offer a new distribution that is specified on  $(0,1)$  to compete with the current bounded distributions.
- b) There are several possible forms for the density function: symmetric, unimodal, reversed J-shaped, left- and right-skewed. Furthermore, J-shaped and rising hazard rate function (HF) plots of the UPLoD are possible.
- c) Statistical characteristics are given, including quantile function, stress strength (SS) reliability model, moments, incomplete moments (IMs), probability-weighted moments (PWMs), residual and inverted residual lives, and entropy measures.
- d) The performance of parameter estimate for the UPLoD is assessed and compared using six traditional estimation techniques: least squares (LS), weighted LS (WLS), maximum likelihood (ML), maximum product spacing (MPS), Anderson-Darling (AD), and Cramer-von Mises (CvM).
- e) To evaluate the validity of different estimates, simulation research is conducted. The Susquehanna River's maximum flood level and water capacity data are used to evaluate the UPLoD's usefulness to a number of alternative models.

The structure of the article is as follows: A new bounded distribution is shown in Section 2. In Section 3, the statistical characteristics of the UPLoD are covered. The model parameter estimators based on ML, LS, CvM, WLS, AD, and MPS methods are derived in Section 4. In order to make sense of the findings in Section 5, a simulation study is conducted. In Section 6, two real data sets are used to demonstrate the UPLoD's utility. Section 7 presents the conclusions.

## 2 Unit Power Lomax Distribution

In this section, a new lifetime model called the UPLoD is introduced and investigated. The UPLoD is obtained by using the exponential function transformation in the form  $Y = e^{-X}$  where  $X$  has the PLoD with density function (1), hence the UPLoD's PDF is provided by:

$$f(y; \omega) = \frac{\kappa\eta}{\tau y} (-\ln y)^{\kappa-1} \left(1 + \tau^{-1} (-\ln y)^\kappa\right)^{-\eta-1}; \quad 0 < y < 1; \kappa, \eta, \tau > 0, \quad (3)$$

where  $\omega = (\kappa, \eta, \tau)$  is the set of parameters, where  $\kappa$  and  $\eta$  are shape parameters, while  $\tau$  is scale parameter. The CDF of the UPLoD is provided as follows:

$$F(y; \omega) = \left(1 + \tau^{-1} (-\ln y)^\kappa\right)^{-\eta}, \quad 0 < y < 1. \quad (4)$$

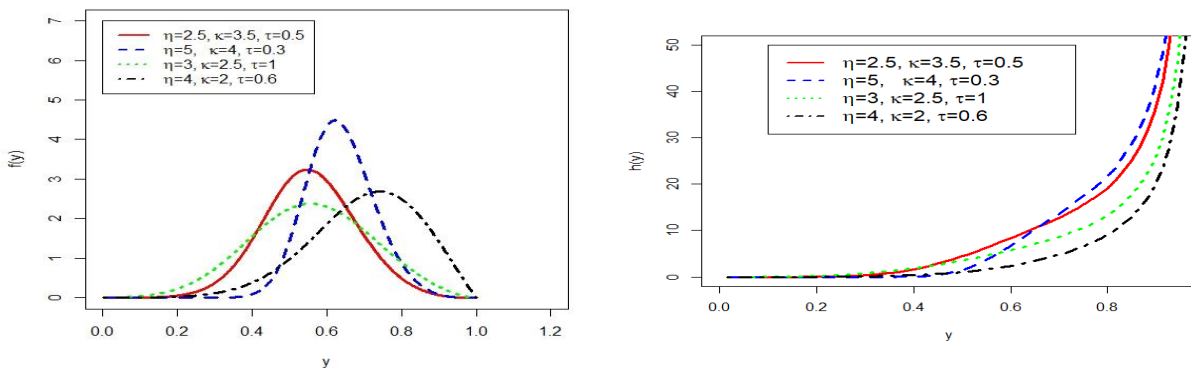
The survival and HF of the UPLoD, for  $0 < y < 1$ , are given, in that order, by

$$S(y; \omega) = 1 - \left(1 + \tau^{-1} (-\ln y)^\kappa\right)^{-\eta},$$

and,

$$h(y; \omega) = \frac{\eta\kappa(-\ln y)^{\kappa-1} \left(1 + \tau^{-1} (-\ln y)^\kappa\right)^{-\eta-1}}{\tau y \left[1 - \left(1 + \tau^{-1} (-\ln y)^\kappa\right)^{-\eta}\right]}.$$

Figure 1 represents the PDF plots of the UPLoD for selected parameter values. It shows that the UPLoD exhibits symmetric, unimodal, reversed J-shaped, left-skewed and right-skewed shapes. Also, the HF plots of UPLoD for some values of parameters are increasing and J-shaped.



**Figure 1:** Plots of the PDF and HF for the UPLoD

The quantile function of a random variable  $Y$  has the UPLoD is obtained. The quantile function of the UPLoD, say  $y = Q(p) = F^{-1}(p)$ , where  $p \sim$  uniform  $(0,1)$  can be obtained by inverting CDF (4) as follows:

$$p = \left[1 + \tau^{-1} (-\ln y)^\kappa\right]^{-\eta}.$$

Then, the quantile function of the UPLoD takes the following form

$$Q(p) = e^{-\tau^{1/\kappa} \left( (p)^{-1/\eta} - 1 \right)^{1/\kappa}} \quad (5)$$

The first quantile ( $Q_1$ ) is obtained by setting  $p = 0.25$  in (5); the median ( $Q_2$ ) is obtained by setting  $p = 0.5$  in (5); and the third quantile ( $Q_3$ ) is obtained by setting  $p = 0.75$  in (5).

### 3 Some Statistical Properties

In this section, some statistical properties of the UPLoD, including, the  $r^{th}$  moment, PWMs, IMs, and moments of residual, some entropy measures, and SS reliability are derived.

#### 3.1 Moments & Some Measures

Ordinary moments can be used to gain a large number of a UPLoD's significant properties and attributes. It is simple to extract the UPLoD's  $r^{th}$  moment from PDF (3) in the manner that follows

$$\begin{aligned} E(Y^r) &= \int_0^1 \frac{\eta\kappa}{\tau y} y^r (-\ln y)^{\kappa-1} \left( 1 + \tau^{-1} (-\ln y)^\kappa \right)^{-\eta-1} dy \\ &= \sum_{h=0}^{\infty} A_h(\omega) B\left(\frac{h}{\kappa} + 1, \eta - \frac{h}{\kappa}\right); \eta > \frac{h}{\kappa}, \end{aligned}$$

where  $A_h(\omega) = \frac{(-r)^h \eta \tau^{h/\kappa}}{h!}$  and  $B(.,.)$  is the Beta function. Certain numerical values of mean ( $\mu_1'$ ), variance ( $\sigma^2$ ), skewness ( $\alpha_3$ ), kurtosis ( $\alpha_4$ ) and coefficient of variation (CV) are mentioned in Table 1 for some selected parameter values.

**Table 1:** Some moments measures of the UPLoD

$\tau$	$\eta$	$\kappa$	$\mu_1'$	$\sigma^2$	CV	$\alpha_3$	$\alpha_4$
1	0.7	7	0.337	0.011	0.031	-0.113	3.317
	0.9	9	0.361	0.006	0.015	-0.049	3.709
	1.1	11	0.373	0.003	0.009	0.093	3.853
2	0.7	7	0.302	0.010	0.034	0.009	3.280
	0.9	9	0.333	0.006	0.016	0.022	3.684
	1.1	11	0.350	0.003	0.0095	0.137	3.859
3	0.7	7	0.283	0.009	0.035	0.084	3.285
	0.9	9	0.317	0.005	0.017	0.066	3.679
	1.1	11	0.337	0.003	0.0097	0.164	3.866

Table 1 indicates that when the values of  $\eta$  and  $\kappa$  rises while the value of  $\tau$  remain constant, the values of  $\sigma^2$  and CV fall and the values of the other measures increase. It is therefore possible to draw the conclusion that, as the value of  $\tau$  rises for predetermined values of  $\eta$  and  $\kappa$ , then the values of the mean and variance decrease, while the values of other measures increase. Thus, it can be claimed that the distribution is skewed to right and left, according to the values of skewness. Finally, according to values of  $\alpha_4$  in Table 1, the UPLoD is leptokurtic.

#### i. The probability-Weighted Moments

It was Greenwood *et al.* [26] who proposed the PWM. Estimators of parameters and the quantile function of the generalized distributions expressible in inverse form are derived using the PWM. Given two positive integers,  $s$  and  $r$ , the PWM of a random variable  $Y$  is defined as

$$v_{s,r} = \int_{-\infty}^{\infty} y^s [F(y)]^r f(y) dy. \quad (6)$$

Using PDF (3) and CDF (4) in (6), the PWM of the UPLoD is derived as follows

$$\begin{aligned} v_{s,r} &= \frac{\eta\kappa}{\tau} \int_0^1 y^{s-1} (-\ln y)^{\kappa-1} \left(1 + \tau^{-1} (-\ln y)^\kappa\right)^{-\eta(r+1)-1} dy \\ &= \sum_{h=0}^{\infty} \frac{(-1)^h s^h (\tau)^{\frac{h}{\kappa}} \eta}{h!} B\left(\frac{h}{\kappa} + 1, \eta(r+1) - \frac{h}{\kappa}\right). \end{aligned}$$

### ii. The incomplete Moments

Understanding a distribution's form as well as its mean is vital for the solutions to many important economic concerns. This is made clear throughout the study of econometrics (for instance, asymmetric error terms cannot continue to be produced by the widely held spherical distributions). The  $r^{\text{th}}$  IM of the UPLoD is obtained as follows by utilizing PDF (3)

$$\begin{aligned} \phi_r(x) &= \frac{\eta\kappa}{\tau} \int_0^x y^{r-1} (-\ln y)^{\kappa-1} \left(1 + \tau^{-1} (-\ln y)^\kappa\right)^{-\eta-1} dy \\ &= \sum_{h=0}^{\infty} A_h(\omega) B\left(\frac{h}{\kappa} + 1, \eta - \frac{h}{\kappa}, \left(1 + \frac{(-\ln x)^\kappa}{\tau}\right)^{-1}\right), \end{aligned}$$

where  $B(\cdot, \cdot, x)$  is incomplete Beta function.

### iii. Residual and Reversed Residual Life's

Residual life and reversed residual life random variables are often used in risk analysis. As a result, Balkema and De Haan [27] looked into some related statistical functions, including the survival function, mean, and variance. The residual life is defined as the interval between time  $t$  and the time of failure for the conditional random variable. The  $r^{\text{th}}$  moment of the residual life, say  $I_r(t)$  is defined as follows:

$$I_r(t) = \frac{1}{S(t)} \sum_{n=0}^r \int_t^{\infty} (y-t)^n f(y) dy = \frac{1}{S(t)} \sum_{n=0}^r \binom{r}{n} (-t)^{r-n} \int_t^{\infty} y^n f(y) dy. \quad (7)$$

Additionally, by combining PDF (3) into (7), the  $r^{\text{th}}$  moment of residual life of the UPLoD can be obtained as follows:

$$I_r(t) = \frac{1}{S(t; \omega)} \sum_{n=0}^r \binom{r}{n} (-t)^{r-n} \int_t^{\infty} \frac{\eta\kappa}{\tau} y^{n-1} (-\ln y)^{\kappa-1} \left(1 + \tau^{-1} (-\ln y)^\kappa\right)^{-\eta-1} dy.$$

After some manipulation,  $I_r(t)$  takes the following form

$$I_r(t) = \frac{1}{S(t; \omega)} \sum_{n=0}^r l_{h,n} B\left(\frac{h}{\kappa} + 1, \eta - \frac{h}{\kappa}, \left(1 + \tau + (-\ln t)^{-\kappa}\right)^{-1}\right),$$

where  $l_{h,n} = (-t)^{r-n} \binom{r}{n} \sum_{j=0}^{\infty} \frac{(-1)^h n^h \tau^{\frac{h}{\kappa}} \eta}{h!}$ .

Further, the  $r^{\text{th}}$  moment of reversed residual life of the UPLoD is derived as follows:

$$\begin{aligned} \varepsilon_r(t) &= \frac{1}{F(t)} \sum_{n=0}^r \int_t^{\infty} (t-y)^r f(y) dy \\ &= \frac{1}{F(t; \omega)} \sum_{n=0}^r (-1)^n \binom{r}{n} t^{r-n} \int_0^t \frac{\eta \kappa}{\tau} y^{n-1} (-\ln y)^{\kappa-1} \left(1 + \tau^{-1} (-\ln y)^\kappa\right)^{-\eta-1} dy, \end{aligned}$$

which is the incomplete beta function, and takes the following form

$$\varepsilon_r(t) = \frac{1}{F(t; \omega)} \sum_{n=0}^r l'_{h,n} B\left(\frac{h}{\kappa} + 1, \eta - \frac{h}{\kappa}, \left(1 + \tau^{-1} (-\ln x)^\kappa\right)^{-1}\right),$$

where  $l'_{h,n} = (-1)^n t^{r-n} \binom{r}{n} \sum_{h=0}^{\infty} \frac{(-1)^h n^h \tau^{\frac{h}{\kappa}} \eta}{h!}$ .

### 3.2 Some Entropy measures

In research on reliability and risk assessment, entropy measures are essential. It has been used in many biological applications and in the physical and medical domains. Entropy measures how much the uncertainty associated with a random variable  $Y$ 's distribution fluctuates. Some entropy measures of the UPLoD as Rényi, Havrda and Charvat, Tsallis and Arimoto are obtained here.

The Rényi entropy, of order  $\gamma > 0$  and  $\gamma \neq 1$ , for the UPLoD is defined by:

$$R_\gamma = \frac{1}{1-\gamma} \log \left[ \int_{-\infty}^{\infty} f(y)^\gamma dy \right]. \tag{8}$$

The Rényi entropy of the UPLoD is obtained by using PDF (3) in (8) as follows:

$$R_\gamma = (1-\gamma)^{-1} \log \left[ \left(\frac{\eta \kappa}{\tau}\right)^\gamma \int_0^1 (-\ln y)^{\gamma(\kappa-1)} y^{-\gamma} \left(1 + \tau^{-1} (-\ln y)^\kappa\right)^{-\gamma(\eta+1)} dy \right].$$

Hence, the Rényi entropy of the UPLoD takes the following form

$$R_\gamma = (1-\gamma)^{-1} \log \left[ \sum_{h=0}^{\infty} \delta_h(\omega, \gamma) B\left(\frac{h-\gamma+1}{\kappa} + \gamma, \gamma \eta - \frac{(h-\gamma+1)}{\kappa}\right) \right],$$

where  $\delta_h(\omega, \gamma) = \frac{\kappa^{\gamma-1} \eta^\gamma (\gamma-1)^h \tau^{\kappa^{-1}(h-\gamma+1)}}{h!}$ .

Shannon's entropy was extended by Havrda and Charvát [28]. Havrda and Charvat (HC) of the UPLoD is obtained from PDF (3) as follows

$$HC_\gamma = \frac{1}{2^{1-\gamma} - 1} \left[ \int_0^1 \left(\frac{\eta \kappa}{\tau y}\right)^\gamma (-\ln y)^{\gamma(\kappa-1)} \left(1 + \tau^{-1} (-\ln y)^\kappa\right)^{-\gamma(\eta+1)} dy - 1 \right].$$

Using the same procedure in Rényi entropy, then, the UPLoD's HC entropy has the following structure.

$$HC_\gamma = \frac{1}{2^{1-\gamma} - 1} \left[ \left[ \sum_{h=0}^{\infty} \delta_h(\omega, \gamma) B\left(\gamma + \frac{h-\gamma+1}{\kappa}, \gamma \eta - \frac{(h-\gamma+1)}{\kappa}\right) \right]^\gamma - 1 \right].$$

Tsallis [29] proposed an extension of Shannon's entropy. Tsallis entropy of the UPLoD is acquired as follows from PDF (3)

$$T_\gamma = (\gamma - 1)^{-1} \left[ 1 - \left( \frac{\eta\kappa}{\tau} \right)^\gamma \int_0^1 (-\ln y)^{\gamma(\kappa-1)} y^{-\gamma} \left( 1 + \tau^{-1} (-\ln y)^\kappa \right)^{-\gamma(\eta+1)} dy \right].$$

By the similar way used above, hence the UPLoD's Tsallis entropy has the following structure

$$T_\gamma = \frac{1}{\gamma-1} \left[ 1 - \sum_{h=0}^{\infty} \delta_h(\omega, \gamma) B \left( \gamma + \frac{h-\gamma+1}{\kappa}, \gamma\eta - \frac{(h-\gamma+1)}{\kappa} \right) \right].$$

An alternative entropy metric with comparable qualities to the Shannon entropy measure was proposed by Arimoto [30] and named the Arimoto's entropy. The following is how to derive Arimoto's entropy of the UPLoD from PDF (3).

$$A_\gamma = \frac{\gamma}{1-\gamma} \left[ \left( \int_0^1 \left( \frac{\eta\kappa}{\tau y} \right)^\gamma (-\ln y)^{\gamma(\kappa-1)} \left( 1 + \tau^{-1} (-\ln y)^\kappa \right)^{-\gamma(\eta+1)} dy \right)^{1/\gamma} - 1 \right].$$

In the same vein used above, the UPLoD's Arimoto's entropy has the following structure

$$A_\gamma = \frac{\gamma}{1-\gamma} \left[ \left[ \sum_{h=0}^{\infty} \delta_h(\omega, \gamma) B \left( \gamma + \frac{h-\gamma+1}{\kappa}, \gamma\eta - \frac{(h-\gamma+1)}{\kappa} \right) \right]^{1/\gamma} - 1 \right].$$

Certain numerical values of some entropy measures of the UPLoD are mentioned in Table 2 for some predetermined values of the parameters.

**Table 2:** A selection of entropy measures of the UPLoD

$\gamma$	$\tau$	$\eta$	$\kappa$	$R_\gamma$	$T_\gamma$	$HC_\gamma$	$A_\gamma$
0.3	0.5	0.5	1	-0.1550	-0.1469	0.2675	0.1301
		2	2	-0.1168	-0.1122	0.2043	0.1023
		5	3	-0.5033	-0.4242	0.7724	0.2961
		7	4	-0.6937	-0.5495	1.0006	0.3436
		9	5	-0.8420	-0.6362	1.1584	0.3685
		11	6	-0.9676	-0.7029	1.2799	0.3837
0.5		0.5	1	-0.3397	-0.3125	0.5334	0.2880
		2	2	-0.1780	-0.1703	0.2908	0.1631
		5	3	-0.6263	-0.5377	0.9179	0.4654
		7	4	-0.8295	-0.6790	1.1591	0.5637
		9	5	-0.9916	-0.7818	1.3347	0.6290
		11	6	-1.1300	-0.8633	1.4737	0.6770
0.8		0.5	1	-1.1101	-0.9955	1.5380	0.9694
		2	2	-0.2502	-0.2440	0.3770	0.2425
		5	3	-0.7354	-0.6839	1.0566	0.6718
		7	4	-0.9504	-0.8655	1.3372	0.8459
		9	5	-1.1235	-1.0062	1.5546	0.9795
		11	6	-1.2709	-1.1222	1.7339	1.0888

Table 2 indicates that when the values of  $\gamma$  increases while the value of  $\eta, \kappa, \tau$ , remain constant the values of the  $T_\gamma$  and  $R_\gamma$  decrease and values of the other measures increase. Consequently, it can be deduced that when the value of  $\eta$  and  $\kappa$  grow for constant values of  $\tau$  and  $\gamma$ , then the values of the  $T_\gamma$  and  $R_\gamma$  fall and values of the other measures rise.



### 3.3 Stress-Strength Reliability

The SS model, say  $R = P [ Y < X ]$ , where  $X$  is the strength and  $Y$  is the stress of the system, is widely used in several fields such as engineering, statistics, and biostatistics. A few real-world examples include buildings, the deterioration of rocket engines, the aging of concrete pressure vessels, and the fatigue failure of aircraft structures. For more applications and examples (see [31-33]). Assume that  $X$  be the system's strength and  $Y$  stress, where  $X$  and  $Y$  are independent random variables having UPLoD  $(\kappa, \eta_1, \tau)$  and UPLoD  $(\kappa, \eta_2, \tau)$ , respectively, then the SS reliability is given as follows

$$R = \int_{y=0}^1 \int_{x=0}^y \frac{\kappa\eta_2}{\tau x} (-\ln x)^{\kappa-1} \left(1 + \tau^{-1}(-\ln x)^\kappa\right)^{-\eta_2-1} \frac{\kappa\eta_1}{\tau y} (-\ln y)^{\kappa-1} \left(1 + \tau^{-1}(-\ln y)^\kappa\right)^{-\eta_1-1} dx dy.$$

Hence, the stress strength of the UPLoD takes the following form

$$R = \frac{\eta_1}{(\eta_1 + \eta_2)}.$$

## 4 Parameter Estimation

In this section, six different estimation methods for estimating model parameters are presented. These methods are ML, LS, WLS, CvM, AD and MPS.

### 4.1 Maximum Likelihood Estimator

The estimation of the UPLoD parameters is deemed using the ML method. Let  $y_1, y_2, \dots, y_m$  be a random sample of size  $m$  from the UPLoD, the log-likelihood function, pointed by  $\ln M^*$ , is given by

$$\begin{aligned} \ln M^* &= m \ln(\eta) + m \ln(\kappa) - m \ln(\tau) - \sum_{r=1}^m \ln y_r + (\kappa - 1) \sum_{r=1}^m \ln(-\ln y_r) \\ &\quad - (\eta + 1) \sum_{r=1}^m \ln \left(1 + \tau^{-1}(-\ln y_r)^\kappa\right). \end{aligned}$$

An alternative to previous equation, we obtain the ML equations as below:

$$\begin{aligned} \frac{\partial \ln M^*}{\partial \eta} &= \frac{m}{\eta} - \sum_{r=1}^m \left(1 + \frac{1}{\tau}(-\ln y_r)^\kappa\right), \\ \frac{\partial \ln M^*}{\partial \tau} &= \frac{-m}{\tau} + \sum_{r=1}^m \frac{(\eta + 1)}{\tau^2(-\ln y_r)^{-\kappa} + \tau}, \end{aligned}$$

and

$$\frac{\partial \ln M^*}{\partial \kappa} = \frac{m}{\kappa} + \sum_{r=1}^m \ln(-\ln y_r) - \sum_{r=1}^m \frac{(\eta + 1) \ln(-\ln y_r)}{\left(1 + \tau(-\ln y_r)^{-\kappa}\right)}.$$

By numerically solving  $\partial \ln M^* / \partial \eta = 0$ ,  $\partial \ln M^* / \partial \tau = 0$ , and  $\partial \ln M^* / \partial \kappa = 0$ , based on optimization algorithm as optim using R program, the ML estimates (MLEs) of  $\eta, \tau$ , and  $\kappa$  are produced.

### 4.2 Least Squares & Weighted Least Squares

Let  $y_1, y_2, \dots, y_m$  be a random sample of size  $m$  from the UPLoD. Suppose that  $y_{(1)} < y_{(2)} < \dots < y_{(m)}$  denotes the corresponding ordered sample. Minimizing the sum squares error yields the LS and WLS estimators of the unknown parameters of the UPLoD.

$$l^*(\omega) = \sum_{r=1}^m g_r \left[ \left( 1 + \frac{1}{\tau} (-\ln y_{(r)})^\kappa \right)^{-\eta} - \frac{r}{m+1} \right]^2. \quad (9)$$

Alternatively, the LS estimates (LSEs) of  $\eta, \tau, \kappa$ , can be obtained by setting  $g_r = 1$  in (9). Similarly, the WLS estimates (WLSEs) of unknown parameters are obtained from (9) by putting  $g_r = \frac{(m+1)^2(m+2)}{r(m-r+1)}$ . These estimates can also be obtained by solving the following non-linear equations using an optimization algorithm

$$\begin{aligned} \frac{\partial l^*(\omega)}{\partial \tau} &= \sum_{r=1}^m g_r \left[ \left( (\tau^{-1}(-\ln y_{(r)})^\kappa + 1)^{-\eta} - \frac{r}{m+1} \right) \mathfrak{F}_1(y_{(r)} | \omega) \right] = 0, \\ \frac{\partial l^*(\omega)}{\partial \eta} &= \sum_{r=1}^m g_r \left[ \left( (\tau^{-1}(-\ln y_{(r)})^\kappa + 1)^{-\eta} - \frac{r}{m+1} \right) \mathfrak{F}_2(y_{(r)} | \omega) \right] = 0, \end{aligned}$$

and,

$$\frac{\partial l^*(\omega)}{\partial \kappa} = \sum_{r=1}^m g_r \left[ \left( (\tau^{-1}(-\ln y_{(r)})^\kappa + 1)^{-\eta} - \frac{r}{m+1} \right) \mathfrak{F}_3(y_{(r)} | \omega) \right] = 0,$$

where,

$$\begin{aligned} \mathfrak{F}_1(y_{(r)} | \omega) &= \frac{\partial \left[ \left( (\tau^{-1}(-\ln y_{(r)})^\kappa + 1)^{-\eta} \right) \right]}{\partial \tau} = \eta \tau^{-2} (-\ln y_{(r)})^\kappa \left( 1 + \tau^{-1}(-\ln y_{(r)})^\kappa \right)^{-\eta-1}, \\ \mathfrak{F}_2(y_{(r)} | \omega) &= \frac{\partial \left[ \left( (\tau^{-1}(-\ln y_{(r)})^\kappa + 1)^{-\eta} \right) \right]}{\partial \eta} = - \left( 1 + \tau^{-1}(-\ln y_{(r)})^\kappa \right)^{-\eta} \ln \left( 1 + \tau^{-1}(-\ln y_{(r)})^\kappa \right), \\ \mathfrak{F}_3(y_{(r)} | \omega) &= \frac{\partial \left[ \left( (\tau^{-1}(-\ln y_{(r)})^\kappa + 1)^{-\eta} \right) \right]}{\partial \kappa} \\ &= -\eta \tau^{-1} (-\ln y_{(r)})^\kappa \ln(-\ln y_{(r)}) \left( 1 + \tau^{-1}(-\ln y_{(r)})^\kappa \right)^{-\eta-1}. \end{aligned} \quad (10)$$

### 4.3 Maximum Product Spacing

The ML approach can be replaced by the MPS method, which approaches the Kullback-Leibler information metric. Although ML estimation is the most popular and extensively used approach, it does not work well in some situations involving big samples and complex continuous distributions. The spacing between the values of the CDF at consecutive data points is the foundation of the MPS approach. The MPS has been used in several applications such as pure mathematics, statistics, hydrology, econometrics, magnetic resonance imaging and others. Let  $Y_{(1)} < Y_{(2)} < \dots < Y_{(m)}$  be the ordered statistics from the distribution with sample size  $m$ , and  $y_{(1)} < y_{(2)} < \dots < y_{(m)}$  be the ordered observed values. Cheng and Amin [34] introduced the MPS method serving as an alternative to the ML method.

Let  $y_{(1)} < y_{(2)} < \dots < y_{(m)}$  are ordered random samples from the UPLoD having CDF(4). The uniform spacings can be defined as follows, based on a size  $m$  random sample from the UPLoD.

$$D_r(\omega) = F(y_{(r)} | \omega) - F(y_{(r-1)} | \omega), \quad r = 1, 2, \dots, m+1,$$

where,  $F(y_{(0)}|\omega)=0$ ,  $F(y_{(m+1)}|\omega)=1$  and  $\sum_{r=1}^{m+1} D_r(\omega)=1$ .

The MPS estimate (MPSE) for the UPLoD is given by maximizing the geometric mean of the spacings

$$S^*(\omega) = \frac{1}{1+m} \sum_{r=1}^{m+1} \ln D_r(\omega) = \frac{1}{1+m} \sum_{r=1}^{m+1} \ln \left[ \left( 1 + \tau^{-1} (-\ln y_{(r)})^\kappa \right)^{-\eta} - \left( 1 + \tau^{-1} (-\ln y_{(r-1)})^\kappa \right)^{-\eta} \right].$$

The MPSE of  $\eta, \tau$ , and  $\kappa$  are obtained by solving the following non-linear equations technique

$$\frac{\partial S^*(\omega)}{\partial \tau} = \frac{1}{1+m} \sum_{r=1}^{m+1} \left[ \frac{\mathfrak{J}_1(y_{(r)}|\omega) - \mathfrak{J}_1(y_{(r-1)}|\omega)}{D_r(\omega)} \right] = 0,$$

$$\frac{\partial S^*(\omega)}{\partial \eta} = \frac{1}{1+m} \sum_{r=1}^{m+1} \left[ \frac{\mathfrak{J}_2(y_{(r)}|\omega) - \mathfrak{J}_2(y_{(r-1)}|\omega)}{D_r(\omega)} \right] = 0,$$

and,

$$\frac{\partial S^*(\omega)}{\partial \kappa} = \frac{1}{1+m} \sum_{r=1}^{m+1} \left[ \frac{\mathfrak{J}_3(y_{(r)}|\omega) - \mathfrak{J}_3(y_{(r-1)}|\omega)}{D_r(\omega)} \right] = 0,$$

where  $\mathfrak{J}_k(y_{(r)}|\omega)$ ,  $k=1, 2$  and  $3$  are given in Equation (10). Also,  $\mathfrak{J}_k(y_{(r-1)}|\omega)$ ,  $k=1, 2$  and  $3$  are given in (10) by replacing  $(r)$  with  $(r-1)$ .

#### 4.4 Cramer-von Mises & Anderson-Dalring Estimators

The CvM estimates (CvMEs) and AD estimates (ADEs) of set parameters of  $\omega = (\eta, \kappa, \tau)$ , are obtained by minimizing the following functions:

$$C(\omega) = \frac{1}{12m} + \sum_{r=1}^m \left[ \left( 1 + \tau^{-1} (-\ln y_{(r)})^\kappa \right)^{-\eta} - \frac{2r-1}{2m} \right]^2,$$

and,

$$A^\bullet(\omega) = -m - \sum_{r=1}^m \frac{2r-1}{m} \left[ \ln \left( 1 + \tau^{-1} (-\ln y_{(r)})^\kappa \right)^{-\eta} + \ln \left[ 1 - \left( 1 + \tau^{-1} (-\ln y_{(m+1-r)})^\kappa \right)^{-\eta} \right] \right].$$

with respect to  $\tau, \eta$  and  $\kappa$ .

### 5 Numerical Study

In this section, a numerical analysis was conducted to assess and compare the performance of the estimates with regard to their relative absolute biases (RAB), chosen parameter values and mean squared errors (MSEs) for different sample sizes. The following steps provide a description of the numerical techniques:

**Step 1:** A random sample is created from the UPLoD by using the inverse transformation (5) with sample sizes  $m = (50, 75, 100, 125, 150, \text{ and } 175)$ .

**Step 2:** Some parameter values are selected as, **Set1:**  $(\eta = 2, \kappa = 0.8, \tau = 0.05)$ , **Set2:**  $(\eta = 2.5, \kappa = 0.8, \tau = 0.05)$ , **Set3:**  $(\eta = 3, \kappa = 0.67, \tau = 0.05)$ , **Set4:**  $(\eta = 3.5, \kappa = 0.67, \tau = 0.05)$ , **Set5:**  $(\eta = 3, \kappa = 0.8, \tau = 0.05)$ , and **Set 6:**  $(\eta = 3.5, \kappa = 0.8, \tau = 0.05)$ .

**Step 3:** Obtain the parameter estimates of  $\eta, \kappa, \tau$  using the provided estimation methods for the selected sample sizes.

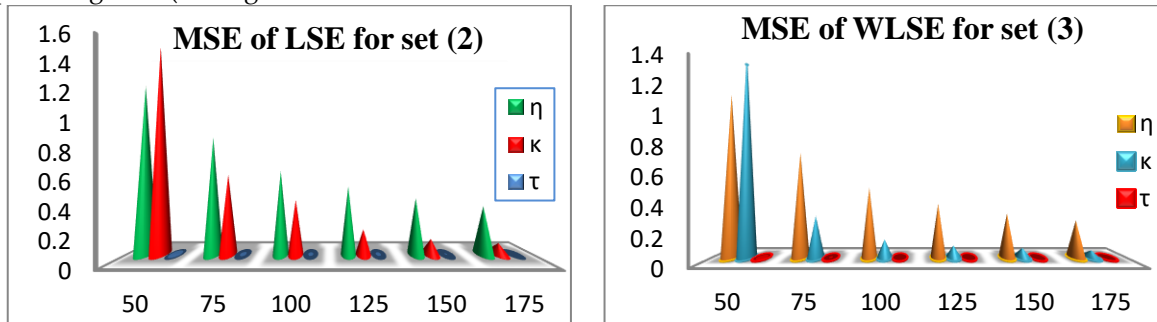
**Step 4:** Steps 1 through 4 are repeated 1000 times for each sample size and the chosen parameter values. Then, the MSEs and RABs of different estimates of  $\eta, \kappa, \tau$  are computed. The MSEs and RABs have the following formulas

$$RAB(\omega) = \frac{1}{1000} \sum_{k=1}^{1000} \left| \frac{\hat{\omega}_k - \omega_k}{\omega_k} \right|, \quad MSE(\omega) = \frac{1}{1000} \sum_{k=1}^{1000} (\hat{\omega}_k - \omega_k)^2.$$

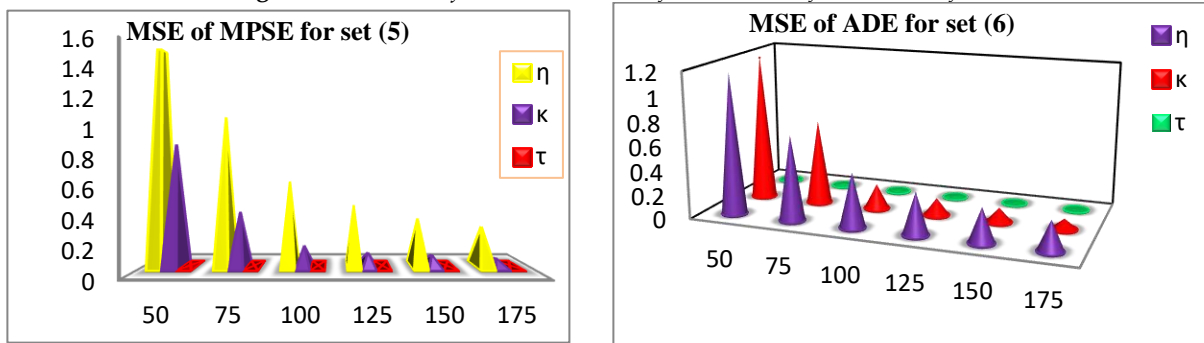
**Step 5:** The numerical results of the simulation study are listed in Tables (3–8).

The findings obtained regarding the behavior of the estimated parameters from the UPLoD are as follows:

1. The RABs of all estimates decrease with increasing sample sizes based on different estimation techniques (see Tables (3–8)).
2. The MSEs for the  $\eta$  estimate increase as value of  $\eta$  increases and the MSEs for the  $\kappa$  estimate decrease as the value of  $\eta$  increases, for all estimation methods (see Tables 3, 7 and 8).
3. From Tables 5 and 7 it is observed that the RAB of  $\kappa$  estimate generally increase and the RAB of  $\eta$  estimate generally decrease when  $\kappa$  increases for all estimation methods.
4. From Tables 4 and 8 it is observed that the RAB of  $\eta, \kappa$  estimates generally constant when  $\eta$  increases for all estimation methods.
5. The RAB of  $\tau$  estimate generally decreases when  $\kappa$  and  $\eta$  increase.
6. For all selected sets of parameters, the MSEs of all estimates based on various approaches decrease as sample size grows (see Figures 2 and 3).



**Figure 2:** The MSEs of the LSE and WLSE for the UPLoD for all values of  $m$



**Figure 3:** The MSEs of ADE and MPSE for the UPLoD for all values of  $m$

7. It can be seen from Figure 4 that the MSE of  $\kappa$  estimates from AD and WLS methods gets the least value followed by the ML method compared to other methods for set 3 and set 4.

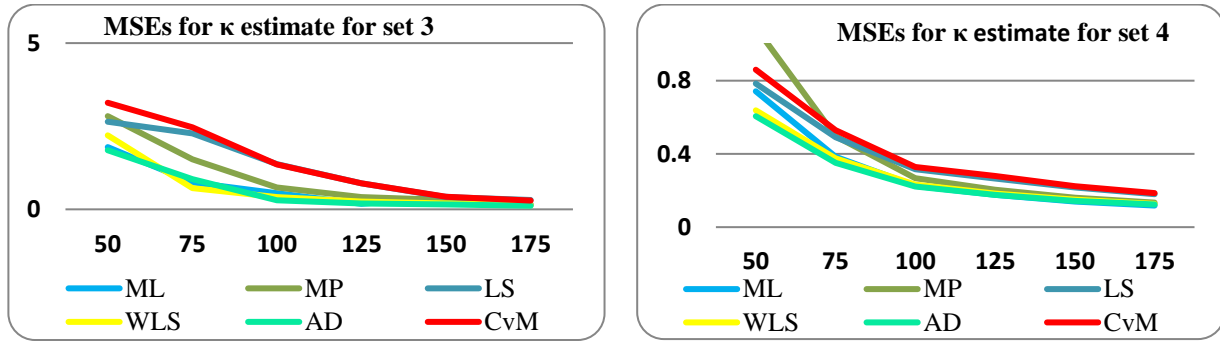


Figure 4: The MSEs for different  $\kappa$  estimates of the UPLoD for all  $m$  values

8. For sets 2 and 6, the MSE of  $\eta$  estimates based on the AD and WLS techniques yields the lowest values, followed by the ML approach in comparison to other methods (see Figure (5)).

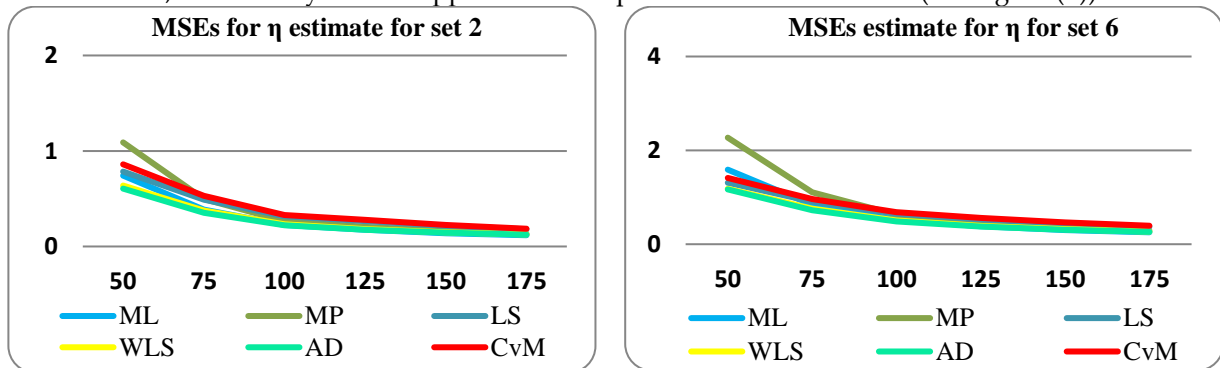


Figure 5: The MSEs of different  $\eta$  estimates of the UPLoD for all values of  $m$

Table 3: MSEs and RABs of different estimates for  $\eta = 2, \kappa = 0.8, \tau = 0.05$

$m$		ML		MP		LS		WLS		AD		CvM	
		MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB
50	$\eta$	0.425	0.074	0.586	0.129	0.369	0.046	0.0339	0.004	0.333	0.059	0.402	0.073
	$\kappa$	1.872	0.436	2.803	0.607	2.635	0.476	2.226	0.443	1.777	0.408	3.206	0.545
	$\tau$	0.001	1.276	0.002	1.303	0.001	1.362	0.002	1.449	0.001	1.36	0.001	1.226
75	$\eta$	0.227	0.049	0.291	0.085	0.259	0.045	0.216	0.039	0.212	0.048	0.272	0.64
	$\kappa$	0.81	0.252	1.507	0.363	2.286	0.433	0.641	0.216	0.911	0.228	2.466	0.419
	$\tau$	0.0009	1.067	0.0009	1.025	0.001	1.188	0.001	1.014	0.001	0.944	0.001	1.15
100	$\eta$	0.144	0.034	0.171	0.06	0.184	0.035	0.146	0.028	0.14	0.035	0.197	0.049
	$\kappa$	0.481	0.172	0.658	0.243	1.363	0.287	0.368	0.151	0.268	0.125	1.349	0.294
	$\tau$	0.0007	0.845	0.0007	0.827	0.0011	1.093	0.0008	0.905	0.0007	0.792	0.0011	1.031
125	$\eta$	0.109	0.029	0.125	0.049	0.153	0.031	0.114	0.024	0.109	0.03	0.159	0.043
	$\kappa$	0.171	0.108	0.367	0.169	0.779	0.214	0.23	0.115	0.179	0.095	0.781	0.223
	$\tau$	0.0005	0.673	0.0005	0.671	0.0009	0.971	0.0006	0.742	0.0005	0.654	0.0009	0.923
150	$\eta$	0.087	0.021	0.099	0.039	0.127	0.025	0.093	0.019	0.089	0.023	0.133	0.034
	$\kappa$	0.38	0.112	0.286	0.145	0.368	0.159	0.172	0.097	0.149	0.084	0.376	0.166
	$\tau$	0.0003	0.552	0.0004	0.544	0.0009	0.947	0.0004	0.617	0.0003	0.526	0.0008	0.905
175	$\eta$	0.073	0.019	0.083	0.033	0.109	0.023	0.079	0.018	0.078	0.021	0.114	0.031
	$\kappa$	0.13	0.077	0.153	0.108	0.281	0.117	0.123	0.068	0.107	0.061	0.264	0.12
	$\tau$	0.0003	0.455	0.0003	0.439	0.0006	0.728	0.0003	0.492	0.0003	0.432	0.0006	0.713

**Table 4:** MSEs and RABs of different estimates for  $\eta = 2.5, \kappa = 0.8, \tau = 0.05$

$m$		ML		MP		LS		WLS		AD		CvM	
		MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB
50	$\eta$	0.742	0.079	1.09	0.14	0.784	0.064	0.638	0.052	0.606	0.066	0.86	0.095
	$\kappa$	1.883	0.45	2.249	0.563	1.941	0.419	1.297	0.369	1.209	0.362	2.36	0.47
	$\tau$	0.026	1.082	0.031	1.189	0.03	1.205	0.027	1.127	0.027	1.139	0.03	1.151
75	$\eta$	0.385	0.053	0.497	0.09	0.491	0.054	0.376	0.042	0.352	0.049	0.529	0.074
	$\kappa$	0.495	0.215	1.342	0.356	1.43	0.326	0.738	0.217	0.695	0.211	1.421	0.336
	$\tau$	0.018	0.084	0.018	0.807	0.018	0.84	0.017	0.794	0.017	0.763	0.019	0.826
100	$\eta$	0.226	0.034	0.267	0.06	0.317	0.038	0.232	0.029	0.222	0.036	0.328	0.052
	$\kappa$	0.479	0.172	0.716	0.246	1.309	0.279	0.368	0.151	0.268	0.125	1.242	0.282
	$\tau$	0.011	0.621	0.012	0.632	0.015	0.738	0.012	0.623	0.011	0.577	0.015	0.715
125	$\eta$	0.176	0.031	0.204	0.052	0.266	0.035	0.183	0.025	0.176	0.031	0.28	0.046
	$\kappa$	0.279	0.113	0.235	0.148	0.461	0.179	0.174	0.103	0.141	0.087	0.49	0.189
	$\tau$	0.007	0.447	0.008	0.485	0.015	0.722	0.009	0.524	0.008	0.444	0.015	0.702
150	$\eta$	0.139	0.022	0.161	0.041	0.216	0.027	0.15	0.019	0.144	0.024	0.224	0.036
	$\kappa$	0.194	0.097	0.189	0.127	0.264	0.14	0.133	0.088	0.118	0.077	0.274	0.147
	$\tau$	0.006	0.386	0.006	0.389	0.011	0.619	0.007	0.44	0.006	0.388	0.011	0.584
175	$\eta$	0.118	0.019	0.134	0.035	0.18	0.025	0.127	0.018	0.124	0.022	0.186	0.033
	$\kappa$	0.097	0.07	0.115	0.096	0.195	0.102	0.096	0.062	0.086	0.056	0.199	0.107
	$\tau$	0.005	0.311	0.005	0.318	0.008	0.469	0.005	0.327	0.004	0.293	0.008	0.456

**Table 5:** MSEs and RABs of different estimates for  $\eta = 3, \kappa = 0.67, \tau = 0.05$

$m$		ML		MP		LS		WLS		AD		CvM	
		MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB
50	$\eta$	1.187	0.089	1.661	0.154	0.959	0.051	0.873	0.048	0.851	0.064	1.065	0.082
	$\kappa$	0.726	0.325	0.905	0.427	1.201	0.416	1.566	0.431	0.549	0.288	1.306	0.443
	$\tau$	0.002	1.393	0.002	1.355	0.002	1.497	0.002	1.509	0.001	1.386	0.002	1.438
75	$\eta$	0.607	0.059	0.802	0.103	0.671	0.051	0.552	0.043	0.536	0.05	0.712	0.071
	$\kappa$	0.328	0.174	0.402	0.237	0.699	0.279	0.301	0.169	0.237	0.152	0.797	0.304
	$\tau$	0.0009	1.009	0.0009	0.902	0.001	1.154	0.0009	1.075	0.0008	0.938	0.001	1.088
100	$\eta$	0.379	0.042	0.463	0.074	0.474	0.039	0.367	0.031	0.358	0.039	0.501	0.054
	$\kappa$	0.136	0.112	0.152	0.141	0.267	0.161	0.131	0.104	0.104	0.089	0.285	0.173
	$\tau$	0.0006	0.776	0.0006	0.691	0.0009	0.995	0.0007	0.842	0.0006	0.757	0.0009	0.943
125	$\eta$	0.282	0.035	0.334	0.06	0.394	0.036	0.287	0.027	0.275	0.032	0.413	0.047
	$\kappa$	0.09	0.08	0.099	0.104	0.195	0.131	0.086	0.079	0.076	0.069	0.202	0.139
	$\tau$	0.0005	0.62	0.0004	0.569	0.0009	0.962	0.0006	0.723	0.0005	0.626	0.0009	0.936
150	$\eta$	0.221	0.026	0.26	0.047	0.319	0.028	0.234	0.021	0.225	0.025	0.336	0.038
	$\kappa$	0.073	0.072	0.082	0.092	0.119	0.104	0.069	0.069	0.063	0.062	0.123	0.109
	$\tau$	0.0003	0.494	0.0003	0.462	0.0007	0.831	0.0004	0.579	0.0003	0.511	0.0006	0.795
175	$\eta$	0.186	0.022	0.216	0.041	0.279	0.027	0.198	0.019	0.194	0.023	0.288	0.034
	$\kappa$	0.049	0.053	0.057	0.069	0.088	0.074	0.051	0.048	0.047	0.045	0.089	0.079
	$\tau$	0.0003	0.428	0.0003	0.388	0.0005	0.676	0.0003	0.478	0.0003	0.437	0.0005	0.647

**Table 6:** MSEs and RABs of different estimates for  $\eta = 3.5, \kappa = 0.67, \tau = 0.05$

<i>m</i>		ML		MP		LS		WLS		AD		CvM	
		MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB
50	$\eta$	1.589	0.089	2.271	0.153	1.31	0.052	1.182	0.048	1.168	0.068	1.412	0.079
	$\kappa$	0.729	0.325	0.904	0.426	1.623	0.464	1.493	0.418	0.549	0.288	1.674	0.487
	$\tau$	0.002	1.393	0.002	1.353	0.002	1.498	0.002	1.509	0.001	1.386	0.002	1.438
75	$\eta$	0.84	0.06	1.106	0.103	0.909	0.051	0.761	0.043	0.725	0.052	0.966	0.07
	$\kappa$	0.328	0.174	0.403	0.237	0.62	0.27	0.301	0.168	0.237	0.152	0.775	0.302
	$\tau$	0.0009	1.009	0.0008	0.903	0.001	1.153	0.0009	1.075	0.0009	0.957	0.001	1.088
100	$\eta$	0.516	0.042	0.629	0.074	0.647	0.039	0.506	0.031	0.482	0.038	0.686	0.055
	$\kappa$	0.136	0.112	0.152	0.141	0.424	0.179	0.131	0.104	0.104	0.089	0.385	0.187
	$\tau$	0.0006	0.775	0.0006	0.692	0.0009	0.996	0.0007	0.842	0.0006	0.757	0.0009	0.943
125	$\eta$	0.383	0.035	0.454	0.06	0.528	0.036	0.389	0.027	0.376	0.033	0.56	0.047
	$\kappa$	0.09	0.08	0.099	0.104	0.195	0.131	0.086	0.079	0.076	0.069	0.202	0.139
	$\tau$	0.0005	0.62	0.0005	0.588	0.0009	0.963	0.0006	0.723	0.0005	0.626	0.0009	0.936
150	$\eta$	0.302	0.026	0.354	0.047	0.436	0.028	0.319	0.021	0.307	0.025	0.462	0.038
	$\kappa$	0.073	0.072	0.082	0.092	0.119	0.104	0.069	0.069	0.063	0.062	0.123	0.109
	$\tau$	0.0003	0.494	0.0003	0.462	0.0007	0.831	0.0004	0.579	0.0003	0.511	0.0006	0.795
175	$\eta$	0.254	0.023	0.293	0.041	0.374	0.026	0.269	0.019	0.264	0.023	0.394	0.034
	$\kappa$	0.049	0.053	0.057	0.069	0.087	0.074	0.051	0.048	0.047	0.045	0.089	0.079
	$\tau$	0.0003	0.428	0.0003	0.388	0.0005	0.676	0.0003	0.478	0.0003	0.437	0.0005	0.647

**Table 7:** MSEs and RABs of different estimates for  $\eta = 3, \kappa = 0.8, \tau = 0.05$

<i>m</i>		ML		MP		LS		WLS		AD		CvM	
		MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB
50	$\eta$	1.074	0.079	1.617	0.141	1.122	0.064	0.914	0.052	0.861	0.066	1.24	0.095
	$\kappa$	1.885	0.451	2.241	0.563	1.9	0.411	1.296	0.369	0.899	0.453	2.247	0.471
	$\tau$	0.026	1.082	0.031	1.189	0.03	1.204	0.027	1.127	0.098	1.046	0.0301	1.151
75	$\eta$	0.549	0.052	0.707	0.089	0.712	0.054	0.546	0.042	0.499	0.049	0.747	0.073
	$\kappa$	0.495	0.215	1.345	0.356	1.429	0.326	0.739	0.217	0.695	0.211	1.438	0.339
	$\tau$	0.018	0.84	0.018	0.807	0.018	0.84	0.017	0.794	0.017	0.763	0.019	0.837
100	$\eta$	0.337	0.036	0.402	0.063	0.471	0.039	0.344	0.029	0.329	0.036	0.498	0.054
	$\kappa$	0.282	0.148	0.416	0.209	0.857	0.24	0.374	0.146	0.205	0.113	0.77	0.241
	$\tau$	0.011	0.615	0.01	0.579	0.014	0.701	0.011	0.604	0.01	0.55	0.014	0.678
125	$\eta$	0.253	0.031	0.293	0.051	0.384	0.035	0.264	0.025	0.253	0.031	0.404	0.046
	$\kappa$	0.279	0.113	0.235	0.148	0.461	0.179	0.174	0.103	0.141	0.087	0.491	0.189
	$\tau$	0.007	0.447	0.009	0.504	0.015	0.722	0.009	0.524	0.008	0.444	0.015	0.603
150	$\eta$	0.201	0.002	0.231	0.041	0.31	0.027	0.216	0.019	0.208	0.024	0.328	0.037
	$\kappa$	0.194	0.097	0.188	0.128	0.264	0.14	0.133	0.088	0.118	0.077	0.274	0.147
	$\tau$	0.006	0.386	0.006	0.39	0.011	0.619	0.007	0.44	0.006	0.388	0.011	0.584
175	$\eta$	0.169	0.019	0.193	0.035	0.259	0.025	0.183	0.018	0.179	0.022	0.269	0.033
	$\kappa$	0.097	0.07	0.115	0.096	0.195	0.102	0.096	0.062	0.086	0.056	0.199	0.107
	$\tau$	0.005	0.311	0.005	0.318	0.008	0.469	0.004	0.327	0.004	0.293	0.008	0.456

**Table 8:** MSEs and RABs of different estimates for  $\eta = 3.5, \kappa = 0.8, \tau = 0.05$

<i>m</i>		ML		MP		LS		WLS		AD		CvM	
		MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB	MSE	RAB
50	$\eta$	1.462	0.079	2.179	0.141	1.521	0.064	1.233	0.051	1.164	0.065	1.698	0.095
	$\kappa$	1.888	0.451	2.242	0.563	1.958	0.42	1.297	0.369	1.215	0.362	2.245	0.471
	$\tau$	0.026	1.082	0.029	1.171	0.03	1.205	0.027	1.127	0.027	1.139	0.03	1.151
75	$\eta$	0.745	0.052	0.969	0.089	0.984	0.055	0.739	0.042	0.696	0.05	1.028	0.073
	$\kappa$	0.495	0.215	1.347	0.356	1.432	0.326	0.733	0.216	0.695	0.221	1.433	0.339
	$\tau$	0.018	0.84	0.018	0.807	0.018	0.84	0.017	0.794	0.017	0.762	0.0195	0.837
100	$\eta$	0.459	0.036	0.547	0.063	0.638	0.039	0.468	0.029	0.447	0.036	0.681	0.054
	$\kappa$	0.282	0.148	0.414	0.209	0.858	0.241	0.375	0.146	0.205	0.113	0.769	0.241
	$\tau$	0.011	0.616	0.01	0.58	0.014	0.701	0.011	0.604	0.01	0.55	0.014	0.678
125	$\eta$	0.344	0.031	0.399	0.052	0.517	0.034	0.359	0.025	0.346	0.031	0.538	0.046
	$\kappa$	0.279	0.113	0.237	0.148	0.461	0.179	0.174	0.103	0.141	0.087	0.49	0.189
	$\tau$	0.007	0.447	0.008	0.485	0.015	0.722	0.009	0.524	0.008	0.444	0.014	0.602
150	$\eta$	0.274	0.022	0.315	0.041	0.415	0.027	0.294	0.019	0.282	0.024	0.438	0.036
	$\kappa$	0.194	0.097	0.188	0.182	0.264	0.14	0.133	0.088	0.118	0.077	0.274	0.147
	$\tau$	0.006	0.386	0.006	0.39	0.011	0.619	0.007	0.44	0.006	0.388	0.011	0.584
175	$\eta$	0.231	0.019	0.262	0.035	0.354	0.025	0.248	0.018	0.244	0.022	0.368	0.033
	$\kappa$	0.097	0.07	0.115	0.096	0.195	0.102	0.096	0.062	0.087	0.056	0.199	0.107
	$\tau$	0.005	0.311	0.005	0.318	0.008	0.469	0.005	0.327	0.004	0.293	0.008	0.456

### 6 Applications to Real Data

In this section, a data analysis is provided in order to examine the goodness-of-fit of the UPLoD when compared to some other models, namely UWD, ULD, UGoD, ULLD, UBXIID, UG/GD, Kumaraswamy distribution (KumD) (Kumaraswamy [35]) and Toppe-Leone distribution (TLD) (Nadarajah and Kotz [36]).

#### i. First Data Set

The first real data set represents 20 observations of comprised water capacity month-wise from the Shasta reservoir in California in the month of February from 1991-2010. Hashmi *et al.* [20] provided the dataset. The following are the data details

0.0833	0.0833	0.1167	0.1167	0.1167	0.15	0.1833	0.2167	0.2167
0.25	0.25	0.25	0.25	0.2833	0.3167	0.35	0.3833	0.4167
0.4167	0.45	0.4833	0.4833	0.7167	0.7167	0.75	0.75	0.85
0.9167								

Some of the data's values can be summarized as follows:  $Q_1 = 0.208, Q_2 = 0.300, Q_3 = 0.483, \text{mean} = 0.377, \alpha_3 = 0.765, \text{ and } \alpha_4 = 2.421$ . The MLEs and standard errors (SEs) for all models are given in Table 9. The measures of fit statistic using the maximized log-likelihood ( $-2\log L$ ), Akaike information criterion ( $E_1$ ), Bayesian information criterion ( $E_2$ ), the correct Akaike information criterion ( $E_3$ ), Hannan-Quinn information criterion ( $E_4$ ), the Kolmogorov Smirnov (KS) statistic values along with P-value, CvM test (CvMT) and AD test (ADT) are calculated in Table 9. The model with minimum values for  $-2\log L, E_1, E_2, E_3$  and  $E_4$  can be selected as the model that best fits the data.



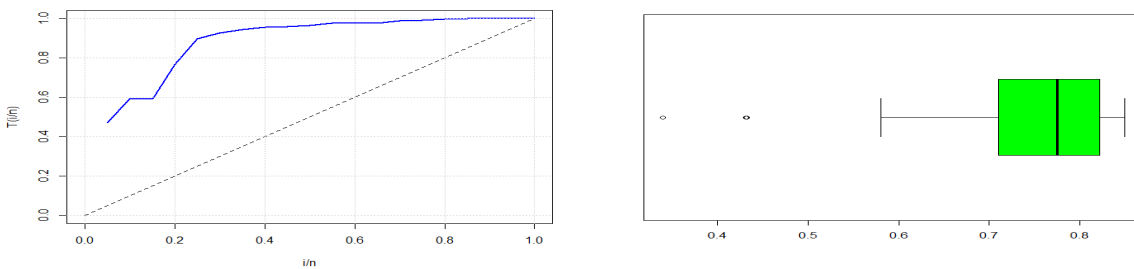
**Table 9:** MLEs and SEs of all model parameters for the first data

	Models					
	UBXIID	UG/GD	UPLoD	UWD	KumD	TLD
$\kappa$	1.765	28.097	3.833	1.57	—	—
SE	(0.254)	(8.125)	(0.427)	(0.248)	—	—
$\eta$	5.856	0.168	0.589	4.207	4.489	—
SE	(1.549)	(0.069)	(0.165)	(1.12)	(2.041)	—
$\tau$	—	63.538	0.003	—	6.347	0.867
SE	—	(67.518)	(0.002)	—	(1.558)	(1.938)

**Table 10:** The statistical measures for the first data

Measures	Models					
	UBXIID	UG/GD	UPLoD	UWD	KumD	TLD
-2log L	-11.744	-15.366	-16.272	-10.957	-13.475	-11.587
$E_1$	-19.488	-24.732	-26.543	-17.914	-22.949	-21.175
$E_2$	-17.497	-21.744	-23.556	-15.922	-20.958	-20.179
$E_3$	-18.783	-24.149	-25.043	-17.208	-22.244	-20.953
$E_4$	-19.099	-24.149	-25.96	-17.525	-22.561	-20.981
KS	0.225	0.192	0.1504	0.242	0.221	0.255
P-value	0.224	0.399	0.701	0.1638	0.245	0.124
CvMT	0.295	0.174	0.105	0.332	0.241	0.313
ADT	1.701	1.081	0.726	1.874	1.425	1.786

The results show that the UPLoD provides a significantly more suited compared to the other five models. The left panel of Figure 6 shows that the box plots is left-skewed. Also, the right panel of Figure 6 shows that the total time on test (TTT) plot is concave; that is, TTT plot was obtained and compared the hazard line, which is an increasing function. Figure 7 shows the probability-probability (PP) plots, also referred to as "parametric plots," and the CDF line empirically (red) utilizing the projected CDF line (black) of the UPLoD of monthly water capacity from the Shasta reservoir in California for the month of February from 1991 to 2010 to illustrate the empirical findings reported in Table 10.



**Figure 6:** Boxplot and TTT plots of the UPLoD for the first data

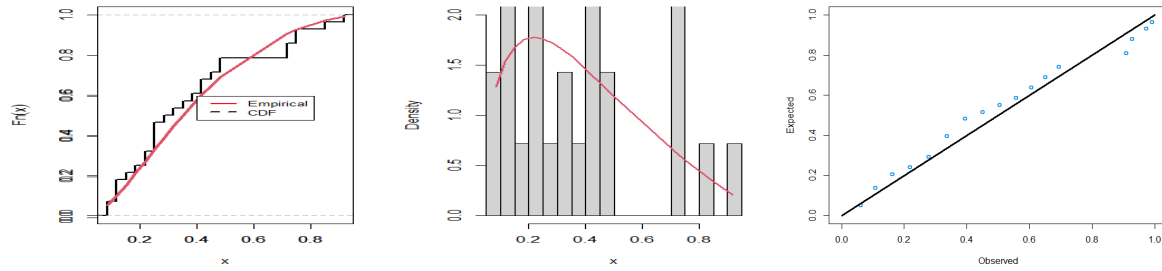


Figure 7: The CDF plot with line empirically, fitted PDF and PP plots for the first data

ii. Second Data Set

The second data set represents 20 observations of the maximum flood level (in millions of cubic feet per second) for the Susquehanna River at Harrisburg, Pennsylvania. The data set was taken from Mazucheli *et al.* [8]. The data are as below:

0.26	0.27	0.3	0.32	0.32	0.34	0.38	0.38	0.39	0.4
0.41	0.42	0.42	0.42	0.45	0.48	0.49	0.61	0.65	0.74

Here's a summary of some of the data's values:  $Q_1 = 0.335$ ,  $Q_2 = 0.405$ ,  $Q_3 = 0.458$ ,  $\text{mean} = 0.423$ ,  $\alpha_3 = 1.07$ , and  $\alpha_4 = 3.66$ . The MLEs and SEs for all models are given in Table 11. The measures of fit are calculated in Table 12. The model with minimum values for the proposed measures can be chosen as the best model to fit the data.

Table 11: MLEs and SEs of all model parameters for the second data

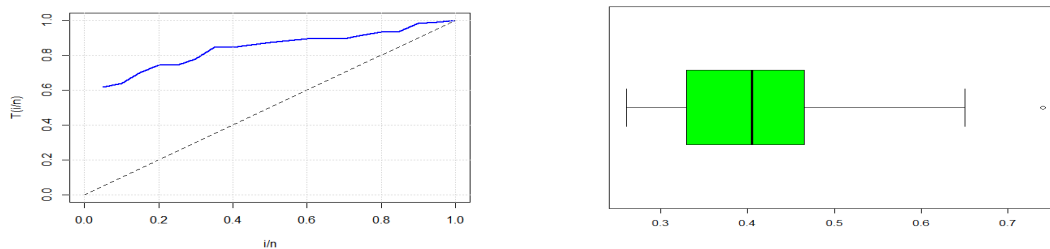
	Models					
	UBXIID	ULLD	UG/GD	UPLoD	KumD	TLD
$\kappa$	1.646	5.274	4.021	3.967	12.005	—
SE	(0.37)	(1.023)	(1.134)	(0.741)	(5.474)	—
$\eta$	4.848	—	2.019	27.109	3.377	—
SE	(0.918)	—	(1.859)	(70.70)	(0.604)	—
$\tau$	—	0.894	76.349	25.757	—	2.241
SE	—	(0.064)	(48.889)	(68.84)	—	(0.501)

Table 12: The statistical measures for the second data

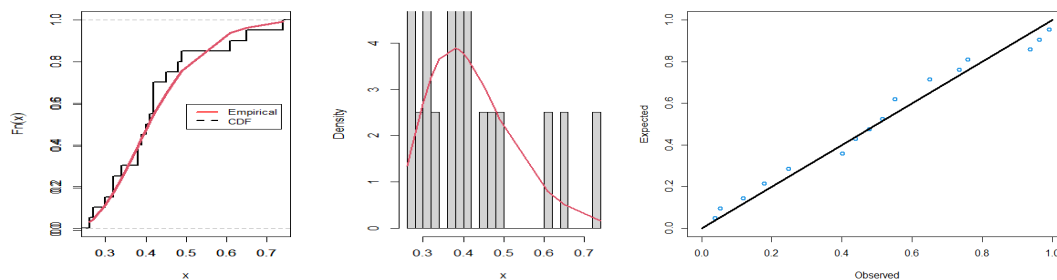
Measures	Models					
	UBXIID	ULLD	UG/GD	UPLoD	KumD	TLD
-2log L	-14.747	-13.853	-15.162	-16.100	-12.973	-7.381
$E_1$	-25.494	-23.707	-24.324	-26.200	-21.241	-12.763
$E_2$	-23.102	-21.715	-21.337	-23.213	-21.241	-11.767
$E_3$	-24.608	-23.007	-22.824	-24.700	-21.132	-12.541
$E_4$	-25.105	-23.318	-23.741	-25.617	-21.558	-12.568
KS	0.1882	0.1604	0.204	0.147	0.2175	0.3409
P-value	0.4777	0.6823	0.374	0.781	0.3005	0.0191
CvMT	0.0930	0.1197	0.061	0.059	0.1673	0.1195
ADT	0.5765	0.7323	0.29	0.353	0.9747	0.7111

The results show that the UPLoD provides a significantly more suited compared to the other five models. The left panel of Figure 8 shows that the box plots is right skewed. Also, the right panel of Figure 8 shows that the TTT plot is concave. Figure 9 illustrates the empirical finding given in Table 12 by showing

the PP plots, and the CDF line empirically (red) utilizing the projected CDF line (black), for the UPLoD of the maximum flood level cubic feet per for the Susquehanna River.



**Figure 8:** Boxplot and TTT plot of the UPLoD for the second data



**Figure 9:** The CDF plot with line empirically, fitted PDF and PP plots employing the second data

## 7 Conclusion

We offer a new bounded distribution in this study, which we term the unit power Lomax distribution, as an alternative to several new bounded distributions. The UPLoD captures several kinds of density and hazard functions. Moments, incomplete moments, PWM, residual and inverted residual lives, quantile function, and entropy measurements are some of the mathematical characteristics of the proposed UPLoD. Some metrics of entropy have also been determined. The unknown parameters of the proposed UPLoD are estimated using the ML, LS, CvM, WLS, AD, and MPS techniques. The asymptotic behaviour of the parameter estimates for the UPLoD was investigated using a simulated study. The findings of the simulation indicate that the WLS and AD approaches are better than the others. Two real-data examples demonstrate that the UPLoD outperforms all competitors in fitting this type of data set.

**Funding:** No funding was received for this work.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest.

## References

- [1] Lomax, K.S. (1954), Business failures: Another example of the analysis of failure data. *Journal of the American statistical association*, **49**(268): 847–852
- [2] Harris, C.M. (1968), The Pareto distribution as a queue service discipline. *Operations Research*, **16**(2): 307–313
- [3] Bryson, M.C. (1974), Heavy-tailed distributions: properties and tests. *Technometrics*, **16**(1): 61–68
- [4] Atkinson, A.B. and Harrison, A.J. (1978), Distribution of personal wealth in Britain. *The Economic Journal*, **88**(351): 581–583
- [5] Holland, O., Golaup, A., and Aghvami, A.H. (2006), Traffic characteristics of aggregated module downloads for mobile terminal reconfiguration. *IEE Proceedings-Communications*, **153**(5): 683–690

- [6] Rady, E.-H.A., Hassanein, W.A., and Elhaddad, T.A. (2016), The power Lomax distribution with an application to bladder cancer data. *SpringerPlus*, **5**: 1–22
- [7] Gómez-Déniz, E., Sordo, M.A., and Calderín-Ojeda, E. (2014), The Log–Lindley distribution as an alternative to the beta regression model with applications in insurance. *Insurance: Mathematics and Economics*, **54**: 49–57
- [8] Mazucheli, J., Menezes, A.F., and Dey, S. (2019), Unit-Gompertz distribution with applications. *Statistica*, **79**(1): 25–43
- [9] Mazucheli, J., Menezes, A.F.B., and Chakraborty, S. (2019), On the one parameter unit-Lindley distribution and its associated regression model for proportion data. *Journal of Applied Statistics*, **46**(4): 700–714
- [10] Haq, M.A.u., Hashmi, S., Aidi, K., Ramos, P.L., and Louzada, F. (2020), Unit modified Burr-III distribution: Estimation, characterizations and validation test. *Annals of Data Science*: 1–26
- [11] Korkmaz, M.Ç. (2020), The unit generalized half normal distribution: A new bounded distribution with inference and application. *UPB Scientific Bulletin, Series A: Applied Mathematics and Physics open access*, **82**(2): 133–140
- [12] Mazucheli, J., Menezes, A.F.B., Fernandes, L.B., De Oliveira, R.P., and Ghitany, M.E. (2020), The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modeling of quantiles conditional on covariates. *Journal of Applied Statistics*, **47**(6): 954–974
- [13] Bantan, R.A.R., Jamal, F., Chesneau, C., and Elgarhy, M. (2021), Theory and applications of the unit gamma/Gompertz distribution. *Mathematics*, **9**(16): 1850. <https://doi.org/10.3390/math9161850>
- [14] Ribeiro-Reis, L.D. (2021), Unit log-logistic distribution and unit log-logistic regression model. *Journal of the Indian Society for Probability and Statistics*, **22**(2): 375–388
- [15] Korkmaz, M.Ç. and Chesneau, C. (2021), On the unit Burr-XII distribution with the quantile regression modeling and applications. *Computational and Applied Mathematics*, **40**(1): 29. <https://doi.org/10.1007/s40314-021-01418-5>
- [16] Ramadan, A.T., Tolba, A.H., and El-Desouky, B.S. (2022), A unit half-logistic geometric distribution and its application in insurance. *Axioms*, **11**(12): 676. <https://doi.org/10.3390/axioms11120676>
- [17] Martínez-Flórez, G., Azevedo-Farias, R.B., and Tovar-Falón, R. (2022), New class of unit-power-skew-normal distribution and its associated regression model for bounded responses. *Mathematics*, **10**(17): 3035
- [18] Hassan, A.S., Fayomi, A., Algarni, A., and Almetwally, E.M. (2022), Bayesian and non-Bayesian inference for unit-exponentiated half-logistic distribution with data analysis. *Applied Sciences*, **12**(21): 11253. <https://doi.org/10.3390/app122111253>
- [19] Krishna, A., Maya, R., Chesneau, C., and Irshad, M.R. (2022), The unit Teissier distribution and its applications. *Mathematical and Computational Applications*, **27**(1): 12. <https://doi.org/10.3390/mca27010012>
- [20] Hashmi, S., Ahsan-ul-Haq, M., Zafar, J., and Khaleel, M.A. (2022), Unit Xgamma distribution: its properties, estimation and application: Unit-Xgamma distribution. *Proceedings of the Pakistan Academy of Sciences: A. Physical and Computational Sciences*, **59**(1): 15–28
- [21] Haj Ahmad, H., Almetwally, E.M., Elgarhy, M., and Ramadan, D.A. (2023), On unit exponential pareto distribution for modeling the recovery rate of COVID-19. *Processes*, **11**(1): 232. <https://doi.org/10.3390/pr11010232>
- [22] Fayomi, A., Hassan, A.S., and Almetwally, E.M. (2023), Inference and quantile regression for the unit-exponentiated Lomax distribution. *Plos one*, **18**(7): e0288635. <https://doi.org/10.1371/journal.pone.0288635>
- [23] Akata, I.U., Opone, F.C., and Osagiede, F.E.U. (2023), The Kumaraswamy unit-Gompertz distribution and its application to lifetime datasets. *Earthline Journal of Mathematical Sciences*, **11**(1): 1–22

- [24] Hassan, A.S. and Alharbi, R.S. (2023), Different estimation methods for the unit inverse exponentiated Weibull distribution. *Communications for Statistical Applications and Methods*, **30**(2): 191–213
- [25] Fayomi, A., Hassan, A.S., Baaqeel, H., and Almetwally, E.M. (2023), Bayesian inference and data analysis of the unit-power Burr X distribution. *Axioms*, **12**(3): 297. <https://doi.org/10.3390/axioms12030297>
- [26] Greenwood, J.A., Landwehr, J.M., Matalas, N.C., and Wallis, J.R. (1979), Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, **15**(5): 1049–1054
- [27] Balkema, A.A. and De Haan, L. (1974), Residual life time at great age. *The Annals of Probability*, **2**(5): 792–804
- [28] Havrda, J. and Charvát, F. (1967), Quantification method of classification processes. Concept of structural-entropy. *Kybernetika*, **3**(1): 30–35
- [29] Tsallis, C. (1988), Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, **52**(1): 479–487
- [30] Arimoto, S. (1971), Information-theoretical considerations on estimation problems. *Information and Control*, **19**(3): 181–194
- [31] Hassan, A.S., Elshaarawy, R., and Nagy, H.F. (2023), Reliability analysis of exponentiated exponential distribution for neoteric and ranked sampling designs with applications. *Statistics, Optimization & Information Computing*, **11**(3): 580–594
- [32] Hassan, A.S., Ismail, D.M., and Nagy, H.F. (2023), Analysis of a non-identical component-strengths system based on lower record data. *Reliability: Theory & Applications*, **18**(2 (73)): 513–528
- [33] Hassan, A.S., Elshaarawy, R., and Nagy, H.F. (2024), Estimation Study of Multicomponent Stress-Strength Reliability Using Advanced Sampling Approach. *Gazi University Journal of Science*, **37**(1): 465–481
- [34] Cheng, R.C.H. and Amin, N.A.K. (1979), Maximum product-of-spacings estimation with applications to the lognormal distribution. *University of Wales IST: Cardiff, UK, Mathematical Report*, **79-1**
- [35] Kumaraswamy, P. (1980), A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, **46**(1-2): 79–88
- [36] Nadarajah, S. and Kotz, S. (2003), Moments of some J-shaped distributions. *Journal of Applied Statistics*, **30**(3): 311–317

# LIMIT CYCLES OF LENGTH TWO IN THE RIKKER MODEL AND THEIR APPLICATION IN FISHING

<sup>1</sup> GURAMI TSITSIASHVILI, <sup>2</sup> TATYANA SHATILINA, <sup>1,3</sup> MARINA OSIPOVA,  
<sup>1</sup> TATYANA RADCHENKOVA



<sup>1</sup> Russia, 690041, Vladiv ostok, Radio street 7, IAM FEB RAS,

<sup>2</sup>VNIRO (TINRO), 690091, Vladiv ostok, lane. Shevchenko, 4

<sup>3</sup> Russia, 690002, Vladiv ostok, FEFU Campus 10 Ajax Bay, Russian Island

guram@iam.dv o.ru, tatyana.shatilina@tinr o-center.ru, mao1975@list.ru, tarad@y andex.ru

## Abstract

*The paper investigates the limit cycle of length two in the Rikker model. It is established that the dependence of the ratio of the maximum value of the cycle to the minimum depends monotonously and almost linearly on the growth coefficient of the Rikker model. Models of the parity shift of the limit cycle of length two are constructed, which is provided by a simultaneous sharp decrease/increase in the growth coefficient. On the example of the Amur salmon in 1994 It is shown that a decrease in the growth coefficient, leading to a shift in the parity of the cycle of length two, is accompanied by a low temperature during the life cycle of pink salmon, when the pink salmon population is in a state of spawn and when the young are rolling.*

**Keywords:** the limit cycle of length two, the parity shift of the cycle and its conditions, the growth coefficient of the Rikker model.

## 1. INTRODUCTION

In chaos theory, and specifically in population dynamics, the Rikker model is a population growth model [1]. In ichthyology research, it aroused the interest of mathematicians in determining the length of limit cycles depending on the value of the growth coefficient (Malthusian parameter) [2] - [5]. Despite numerous studies of the Rikker model, the analysis of limit cycles of length two may lead to new results.

In this paper, the ratio of the maximum value of the limit cycle of length two to the minimum is investigated. With the help of computational experiments, it is shown that this ratio depends monotonously and almost linearly on the growth coefficient for almost the entire range of values. This result may be applied to solving an important applied problem formulated by A.A. Goryainov, an employee of the Tinro Center [6], [7] on changing the parity of the limit cycle of length two in the Rikker model. If the maximum value of a cycle of length two is taken in even (odd) years, then a change in parity is understood as such a change in the cycle in which the maximum value begins to be taken in odd (even) years.

The task of changing the parity is important for predicting significant changes in populations, the dynamics of which is subject to the Rikker model. Such parity changes occur very rarely and are the result of a significant influence of external (hydro meteorological) conditions on population dynamics. A feature of the method for solving the problem of parity change is the consideration of the Rikker model at an extreme value of the growth coefficient, which makes it possible to determine the conditions for parity shift from hydro meteorological data. It is

shown that the very shift of the cycle of length two at the minimum point of the cycle leads to a significant decrease in catches. On the contrary, at the maximum point of the cycle, a very high growth coefficient is required for its implementation. These circumstances allow us to point out an analogy between the parity shift (number/catch of pink salmon) and the failure of the technical system.

## 2. METHODS

Consider the Rikker model  $x_{n+1} = a_n x_n \exp(-bx_n)$ ,  $n \geq 0$ . Using the standard substitution  $y_n = bx_n$ , we arrive at a recurrent sequence

$$y_{n+1} = f(y_n) = \alpha_n y_n \exp(-y_n), \quad n \geq 0. \tag{1}$$

Here  $\alpha_n = a_n/b$  is the growth coefficient and  $c_n = y_{n+1}/y_n$  is the return coefficient. It follows from the formula (1) that the equality  $\alpha_n = c_n \exp(y_n)$  is fulfilled, linking the growth coefficient  $\alpha_n$  with the return coefficient  $c_n$  and with  $y_n$ .

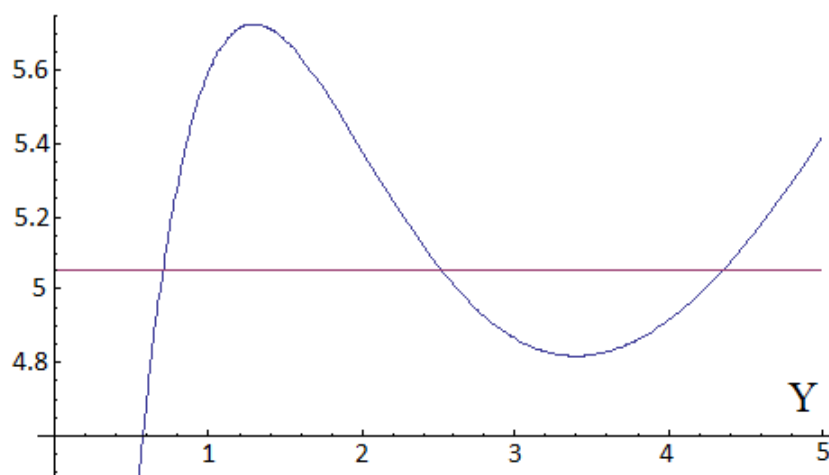
Let's focus on the case when  $\alpha_n \equiv \alpha$ . In this case, the following classification of stable limit modes in the Rikker model [1] - [4] is known. For  $0 < \alpha < \beta_0 = 1$ , the sequence  $y_n$ ,  $n \geq 0$ , has a stable rest point  $Y_1 = 0$ . At  $\beta_0 < \alpha < \beta_1 \approx e^2 \approx 7,39$  the sequence  $y_n$ ,  $n \geq 0$ , has a stable rest point  $Y_2 = Y_2(\alpha) > 0$ . At  $\beta_1 < \alpha < \beta_2 \approx 12.49$  the sequence  $y_n$ ,  $n \geq 0$ , has a stable limit cycle of length two. At  $\beta_2 < \alpha < \beta_3 \approx 14.68$  the sequence  $y_n$ ,  $n \geq 0$ , has a stable limit cycle of length four, etc.

Let's calculate the components of the limit cycle of length two  $Y$ ,  $f(Y)$ , defined for  $Y > 0$  by the relations

$$Y = f(f(Y)) \Rightarrow 1 = \alpha^2 \exp(-Y(1 + \alpha e^{-Y})) \Rightarrow \varphi(Y) = \psi(Y), \tag{2}$$

$$\varphi(Y) = 2 \ln \alpha, \quad \psi(Y) = Y(1 + \alpha e^{-Y}).$$

Numerical calculations of the roots of the equation (2) show that for  $\beta_1 < \alpha < \beta_2$  this equation has three roots, because the function  $\psi(Y)$  has both a minimum and a maximum. Moreover, the minimum root  $Y_{min}$  and the maximum root  $Y_{max}$  are related by the relations  $Y_{max} = f(Y_{min})$ ,  $Y_{min} = f(Y_{max})$ . The root  $Y_{mid} = \ln \alpha$ , contained between the minimum and maximum roots, corresponds to the unstable rest point of the sequence  $y_n$ ,  $n \geq 0$ , (see Fig. 1).



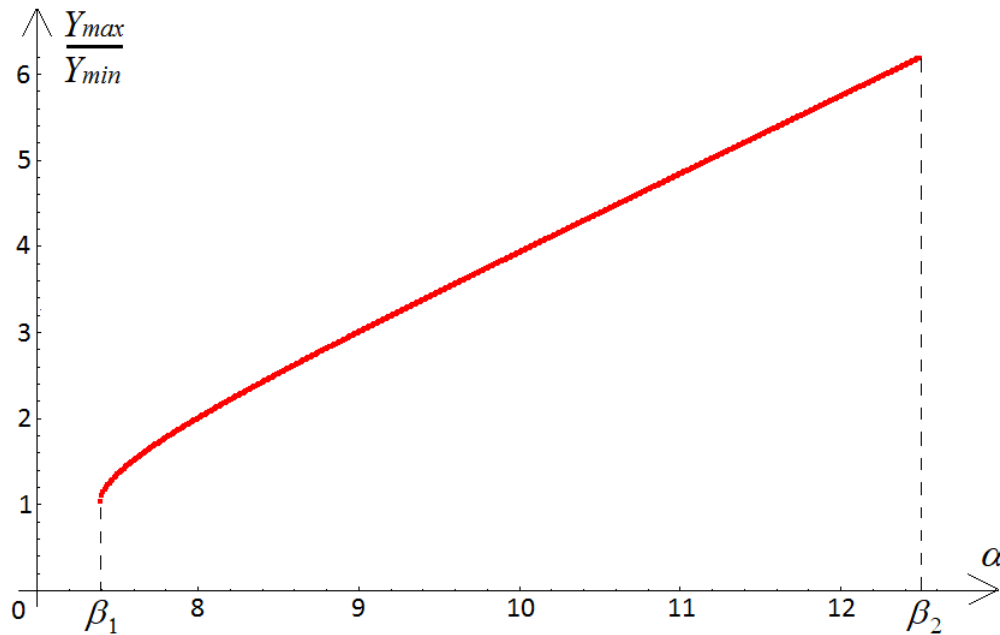
**Figure 1.** Graphs of functions  $\varphi(Y)$  (red line),  $\psi(Y)$  (blue line) at  $\alpha = \beta_2$ .

The results of calculating the roots of the equation (2) are presented in Table 1.

**Table 1.** The values of the roots  $Y_{min}$ ,  $Y_{mid}$ ,  $Y_{max}$  and the ratio  $Y_{max}/Y_{min}$  depending on the growth coefficient  $\alpha$ ,  $\beta_1 < \alpha < \beta_2$ .

$\alpha$	$Y_{max}/Y_{min}$	$Y_{min}$	$Y_{mid}$	$Y_{max}$
7.39	1.02807	1.97244	2.00013	2.02781
8	2	1.38629	2.07944	2.77259
9	3	1.09861	2.19722	3.29584
10	3.92745	0.934596	2.30259	3.67057
11	4.83672	0.821659	2.3979	3.97413
12	5.74133	0.737215	2.48491	4.2326
12.15	5.87697	0.726287	2.49733	4.26837
12.3	6.01263	0.715737	2.5096	4.30346
12.45	6.14831	0.705543	2.52172	4.3379
12.49	6.1845	0.702882	2.52493	4.34697

From the table 1 it can be seen that the ratio  $Y_{max}/Y_{min}$  increases with the growth coefficient  $\alpha$  increases from a value close to one at  $\alpha \approx \beta_1$  to a significantly larger unit ( $\approx 6.1845$ ) value at  $\alpha \approx \beta_2$ . This is shown in more detail in Fig. 2.

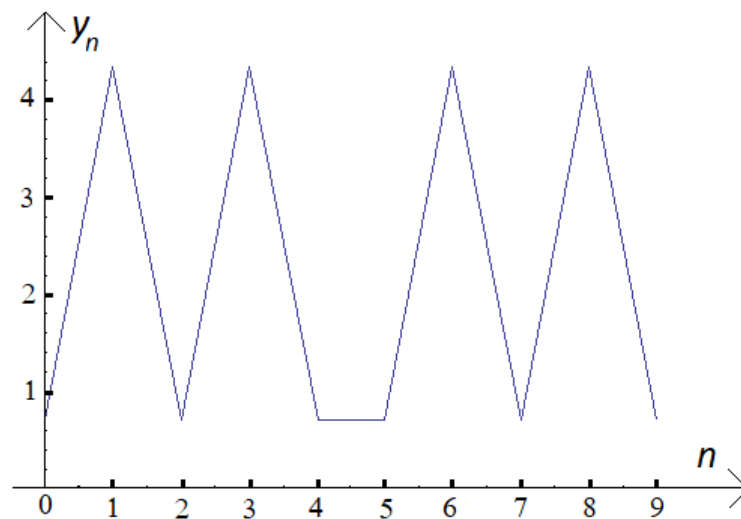


**Figure 2.** Graph of the dependence of  $Y_{max}/Y_{min}$  on the growth coefficient of  $\alpha$  at  $\beta_1 < \alpha < \beta_2$ .

As a result, it has been empirically established that on the segment  $\beta_1 < \alpha < \beta_2$  the ratio  $Y_{max}/Y_{min}$  depends on the growth coefficient  $\alpha$  almost linearly (except for a small tail on the left). Moreover, this fact cannot be established analytically.

Using these estimates and formulas (1), (2), we give one example of the parity shift of a stable cycle at the minimum point (see Fig. 3). Let  $\alpha_n = \beta_2$ ,  $0 \leq n \leq 9$ ,  $n \neq 4$ ,  $\alpha_4 = \beta_2^* = \exp(0.702882) \approx 2.01956$ , and the sequence  $y_n$ ,  $0 \leq n \leq 9$ ,  $n \neq 4$  coincides with a stable cycle of length two  $y_0 = y_2 = y_4 = y_5 = y_7 = y_9 = 0, 702882$ ,  $y_1 = y_3 = y_6 = y_8 = 4, 34697$ . However, at  $n = 4$ , there is a shift in the parity of the stable cycle.

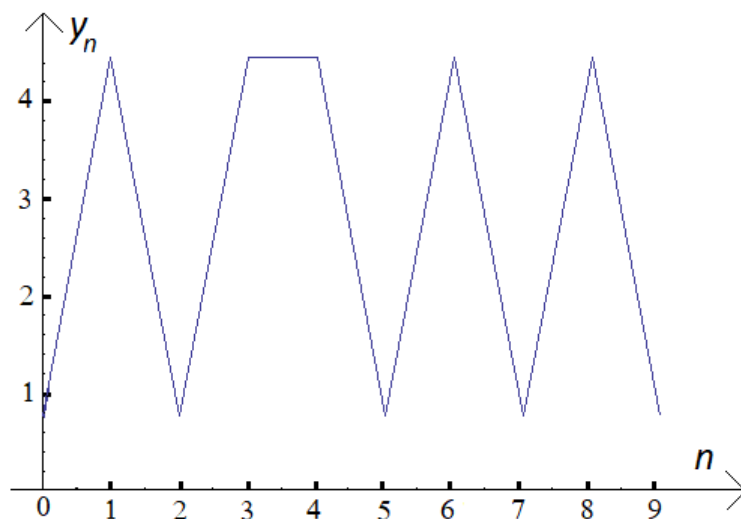




**Figure 3.** The graph of the parity shift in the Rikker model at the minimum point.

This calculation gives some idealized example of the parity shift of a cycle of length two. To achieve such a shift, it is necessary at the moment  $n = 4$  to significantly reduce the growth coefficient, namely, by  $\beta_2 / \beta_2^* \approx 0.6, 184$  times.

In turn, the graph of the parity shift of the limit cycle of length two at the maximum point is possible at the point  $n = 3$  at  $\alpha_3 = \beta_2^{**} = \exp(4, 34697) \approx 77, 244$ . Let  $\alpha_n = \beta_2, 0 \leq n \leq 9, n \neq 3, \alpha_3 = \beta_2^{**}$ , and the sequence  $y_n, 0 \leq n \leq 9, n \neq 3$  coincides with a stable cycle of length two:  $y_0 = y_2 = y_4 = y_6 = y_8 = 0.702882, y_1 = y_3 = y_5 = y_7 = y_9 = 4.34697$ . However, at  $n = 3$ , a steady cycle parity shift occurs. To achieve such a shift, we need to greatly increase the growth coefficient, namely, by  $\beta_2^{**} / \beta_2 \approx 0.6, 184$  times.



**Figure 4.** The graph of the parity shift in the Rikker model at the maximum point.

Thus, in the idealized parity shift model of the limit cycle of length two, it is necessary to significantly (6,184 times) simultaneously reduce/increase the growth coefficient. Therefore, a more realistic procedure is to reduce the growth coefficient and the corresponding shift of the limit cycle of length two at the minimum point.

### 3. RESULTS

Let's focus on the meteorological conditions that ensure a parity shift at the minimum point of the limit cycle for the dynamics of the Amur pink salmon population. Similar studies have previously been conducted for the Seaside pink salmon [8]. Detailed ichthyology studies show that negative meteorological effects leading to a parity shift are possible in two periods of the pink salmon life cycle. The first period is incubation: in January, when the pink salmon population is in a state of caviar. The second period is in June-July during the decline of young pink salmon. Table 2 shows data on the air temperature in Nikolae vsk on Amur in the first period and in the second period on the water temperature in the Tatar Strait. These data show that the parity shift of the length two cycle occurs at sufficiently low temperatures during the specified periods of the Amur pink salmon life cycle.

**Table 2.** Temperature data for 1994.

Month	HMS Air Temperature Nikolae vsk-on-Amur	Water Temperature in Tatar Strait
January	-6.4	0.5
February	3.3	0.3
March	0.1	0.3
April	-0.8	0.1
May	-2.2	-0.6
June	-1.5	-0.9
July	-0.9	0.1
August	1.2	0.9
September	0.5	0.2
October	0.9	0.2
November	-0.8	0.1
December	-2.8	0.5

### 4. DISCUSSION

The problem of shifting the parity of a cycle of length two on the one hand is of serious theoretical and practical interest. When solving it, we have to limit ourselves to cycles of length two in order to compress the initial biological and hydrometeorological information. Of course, this approach to analyzing the source information allows for a certain approximation. But such an approximation can be justified by setting the problem of shifting the parity of a cycle of length two. Moreover, when solving this problem, it is necessary to analyze in detail the life cycle of the Amur pink salmon and identify critical moments in it.

### 5. CONCLUSION

In conclusion, it should be noted that the idealized model of the parity shift of a cycle of length two is not always implemented in practice. The parity shift can occur in states where the ratio of the maximum cycle value to the minimum value is close to one. Nevertheless, even in this case, there is a decrease in the growth coefficient caused by adverse meteorological conditions.

### REFERENCES

- [1] Rikker, U. E. (1979) Methods of assessment and interpretation of biological indicators of fish populations. Moscow: Food Industry, 408 p. (In Russian).
- [2] Shapiro, A. P., Luppov, S. P. (1983) Recurrent equations in the theory of population biology. Moscow: Nauka, 132 p. (In Russian).

- [3] Lasunsky , A. V. (2012) On the period of solutions of a discrete periodic logistic equation Proceedings of the KarSC RAS, vol. 5. Ser. Mathematical Modeling and Information Technology , No. 3, pp. 44-48. (In Russian).
- [4] Elaydi, S. N., Luis, R., Oliveira, H. (2011) Towards a theory of periodic difference equations and its application to population dynamics. Dynamics, Games and Science. Springer Proc. Math. Vol. 1, pp. 287-321.
- [5] Shlufman, K. V., Neverova, G. P. , Frisman, E. Ya. (2017) Dynamic modes of the Riker model with a periodically changing Malthusian parameter . Nonlinear dynamics. Vol. 13. No. 3, pp. 363-380. (In Russian).
- [6] Goryainov, A. A., Shatilina, T. A. (2003) Dynamics of the Asian pink salmon and climatic changes over the Asia-Pacific region in the twentieth century. Marine Biology. Vol. 29, No. 6, pp. 429-435. (In Russian).
- [7] Goryainov, A. A., Krupyanko, N. I., Shatilina, T. A. (2013) Comparative analysis of catch dynamics of the Primorye and Amur pink salmon Bulletin No. 8 of the study of Pacific salmon in the Far East. Vladivostok: TINRO Center, pp. 106-118. (In Russian).
- [8] Lysenko, A. V., Shatilina, T. A., Gaiko, L. A. (2021) The influence of hydrometeorological conditions on the dynamics of catch (abundance) of the Primorsky pink salmon *ONCORHYNCHUS GORBUSCHA* (SALMONIDAE) based on retrospective data (Sea of Japan, Tatar Strait). Questions of Ichthyology, Vol. 61. No. 2, pp. 206-218. (In Russian).

# INFERENCE ON THE INVERSE POWER BURR-HATKE DISTRIBUTION UNDER TYPE II CENSORING

Pavitra Kumari

•

Department of Statistics, Central University of Haryana, India

[pavitra@cuh.ac.in](mailto:pavitra@cuh.ac.in)

Vinay Kumar

•

Department of Mathematics and Statistics, Chaudhary Charan Singh Haryana Agricultural  
University, India

[vinatstats@hau.ac.in](mailto:vinatstats@hau.ac.in)

## Abstract

*There are many real-life situations, where data require probability distribution function which have decreasing or upside-down bathtub (UBT) shaped failure rate function. The inverse power burr hatke distribution consists both decreasing and UBT shaped failure rate functions. Here, we address the different estimation methods of the parameter and reliability characteristics of the inverse Pareto distribution from both classical and Bayesian approaches. We consider classical estimation procedures to estimate the unknown parameter of inverse power burr-hatke distribution, such as maximum likelihood. Also, we consider Bayesian estimation using squared error loss function based joint priors. The Monte Carlo simulations are performed to compare the performances of the obtained estimators in mean square error sense. Finally, the flexibility of the proposed distribution is illustrated empirically using one real-life datasets. The analyzed data shows that the introduced distribution provides a superior fit than some important competing distributions such as the Weibull, inverse Pareto and Burr-Hatke distributions.*

**Keywords:** Burr-Hatke Distribution, Inverse Power Burr- Hatke Distribution, Type II censoring, Bayesian estimation, Lindley's Approximation technique.

## I. Introduction

Statistical distributions can be used to model many real-life scenarios, such as reliability, actuarial science, survival analysis and lifetime data. Different lifetime distributions have been introduced in the statistical literature to provide greater flexibility in modelling data in these applied sciences. One of the important features of generalized distributions is their capability for providing superior fit for various life-time data encountered in the applied fields. Hence, the statisticians have been interested in constructing new families of distributions to model such data.

Recently, several new distributions and regression models to provide inferences on these distributions have been developed for modeling health and biomedical data, among other fields. Some distributions and classes of distributions developed include exponentiated Burr XII Poisson distribution by da Silva et al. [1], Weibull Burr XII (WBXII) distribution by Afify et al. [2], odd log logistic Topp-Leone G family of distributions by Alizadeh et al. [3], Burr-Hatke exponential (BHE)

distribution by Abouelmagd [4] and Yadav et al. [5], odd generalized gamma-G family of distributions by Nasir et al. [6], Chen-G family of distributions by Anzagra et al. [7], inverse-power Burr-Hatke distribution by Afify et al. [8], harmonic mixture Weibull-G family of distributions by Zamanah et al. [9], harmonic mixture G family of distributions by Kharazmi et al. [10] and Alshenawy R. [11] studied Progressive Type-II Censoring Schemes of Extended Odd Weibull Exponential Distribution with Applications in Medicine and Engineering. Ahmed et. al. [12] studied Bayesian and Classical Inference under Type-II Censored Samples of the Extended Inverse Gompertz Distribution with Engineering Applications. Hassan [13] studied Statistical Inference of Chen distribution Based on Two Progressive Type-II Censoring Schemes.

Burr Hatke model provides only a decreasing hazard rate (HR) shape; hence, its use will be limited to modelling the data that exhibits only increasing failure rate. IPBH model can accommodate right-skewed shape, symmetrical shape, reversed J shape and left-skewed shape densities. Its hazard rate (HR) can be an increasing shape, a unimodal shape, or a decreasing shape. IPBH distribution provides more accuracy and flexibility in fitting engineering and medicine data. The IPBH distribution was constructed using the inverse-power (IP) transformation. The aim of this article is to develop the classical and Bayesian estimation procedures for the parameters of the IPBH.

The rest of the article is organized as follows: IPBH is discussed in Section 2. Also, mathematical formulation is given for type II censoring with failure and censoring time distributions in this section. Section 3 deals with the maximum likelihood estimation and asymptotic confidence intervals of the parameters. Section 4 describes asymptotic confidence interval. Sections 5 describe the formulation of Bayes estimation procedure using Markov chain Monte Carlo (MCMC) methods under SELF loss function using gamma informative priors. Section 6 deals with a Monte Carlo simulation study to explore the properties of various estimates developed in this article. Real life dataset is analyzed for illustration purposes in Section 7. Finally, conclusive remarks are given in section 8. Also, it is essential to mention that the statistical software R 3.5.2, [R Core Team (2018)] is used for computation purposes throughout the article.

## II. The Model

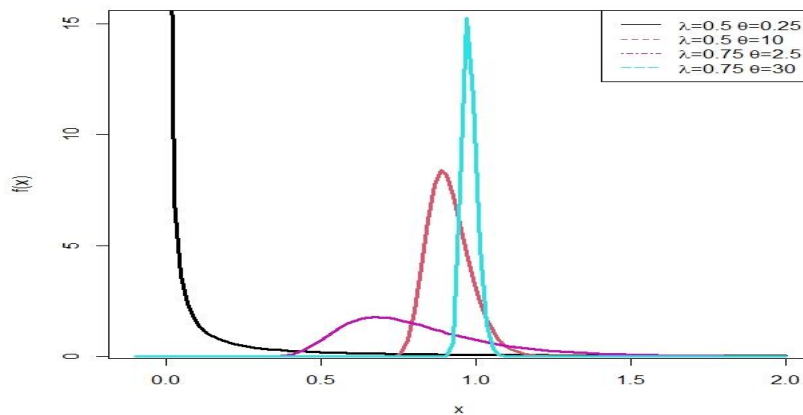
If a random variable X follows IPBH with parameter  $(\lambda, \theta)$  the cdf is given by:

$$F(x; \lambda, \theta) = \frac{\exp(-\lambda x^{-\theta})}{x^{-\theta+1}}, \quad \lambda, \theta > 0 \tag{2.1}$$

Therefore, the corresponding probability density function is given by

$$f(x; \theta, \lambda) = \frac{\theta \exp(-\lambda x^{-\theta})[\lambda + (1 + \lambda)]x^{-\theta}}{x(x^\theta + 1)^2}, \quad \lambda, \theta > 0 \tag{2.2}$$

Where  $\theta$  and  $\lambda$  are shape parameters, respectively.



**Figure 1.** Possible density shapes of the IPBH distribution for several values of  $\lambda$  and  $\theta$ .

The survival function (SF) and HR function of the IPBH distribution take the following forms, respectively:

$$S(x: \lambda, \theta) = 1 - \frac{x^\theta \exp(-\lambda x^\theta)}{x^{-\theta} + 1} \tag{2.3}$$

$$h(x: \lambda, \theta) = \frac{\theta [\lambda + (\lambda + 1)x^{-\theta}]}{x(x^\theta + 1)[(x^\theta + 1) \exp(-\lambda x^{-\theta}) - x^\theta]} \tag{2.4}$$

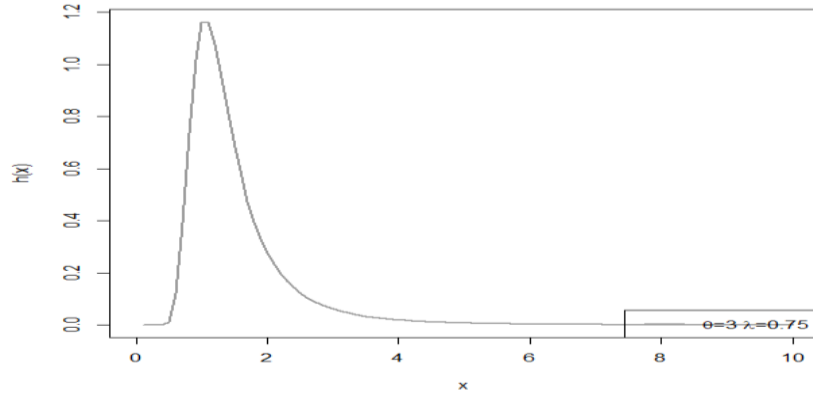


Figure 2. Possible failure rate shape of the IPBH distribution for values of  $\lambda = 0.75$  and  $\theta = 3$

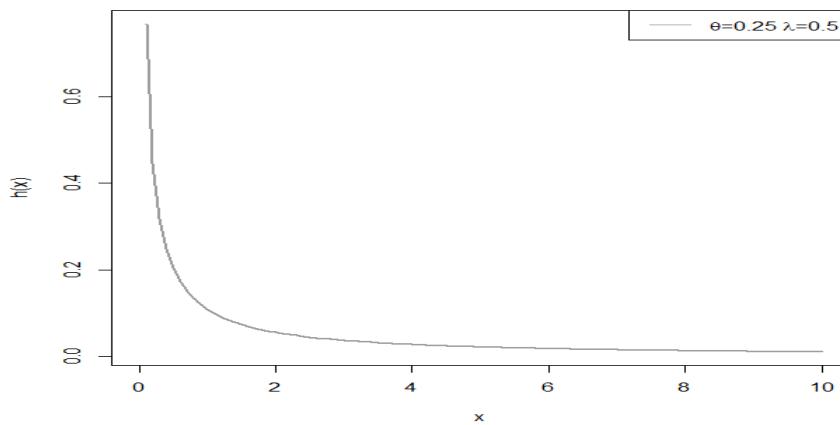


Figure 3. Possible failure rate shape of the IPBH distribution for values of  $\lambda = 0.5$  and  $\theta = 0.25$

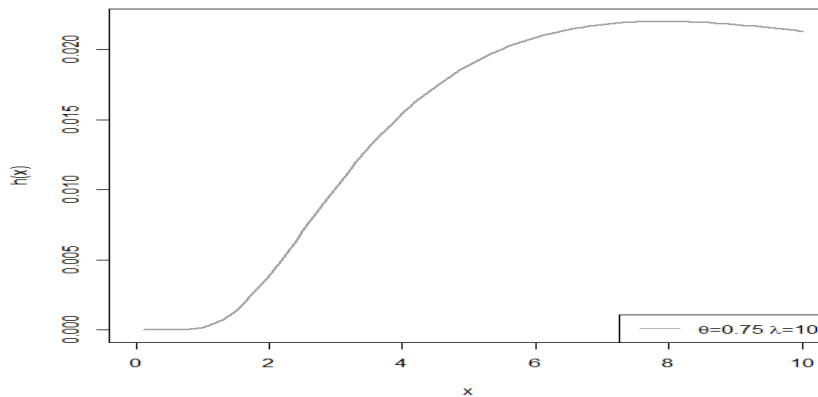


Figure 4. Possible failure rate shape of the IPBH distribution for values of  $\lambda = 10$  and  $\theta = 0.75$

### III. Maximum Likelihood Estimation

In the literature, Several censoring schemes have been discussed. Even though, Type-I and Type-II censoring schemes are most popular censoring. Consider a life test where  $n$  independent units taken from a IPBH distribution are placed under observation and failure time of each unit is recorded. Suppose that the test is terminated when  $r$ th, ( $1 \leq r \leq n$ ),  $r$  is prefixed unit fails. These observed failure times, say  $(x_1, x_2, \dots, x_r)$  is a Type-II censored sample of size  $r$ . In this censoring scheme  $n-r$  units remain unobserved and survive beyond the time of termination. In Type-II censoring the time of termination is a random variable and the likelihood function based on  $(x_1, x_2, \dots, x_r)$  is given by Cohen [14].

$$L(\lambda, \theta | \underline{x}) = \frac{n}{(n-r)} \prod_{i=0}^r f(x_i) [1 - F(x_{(r)})]^{n-r} \tag{2.5}$$

Assume that  $n$  independent observed values taken of IPBH distribution as presented in (2) are put on a test. Using the Type-II censoring, we obtained the ordered  $r$  failures. If the ordered  $r$  failures are then  $(x_1, x_2, \dots, x_r)$  the likelihood function of  $(\lambda, \theta)$  under Type-II censored data drawn of an IPBD distribution, is obtained as follows:

$$L(\lambda, \theta | \underline{x}) = \frac{n}{(n-r)} \prod_{i=0}^r f(x_i) [1 - F(x_{(r)})]^{n-r}$$

$$L(\lambda, \theta | \underline{x}) = r \log(\theta) - \lambda \sum_{i=1}^r x_i^{-\theta} + \sum_{i=1}^r \log[\lambda + (\lambda + 1)x_i^\theta] - 2 \sum_{i=1}^r \log(x_i^\theta + 1) - \sum_{i=1}^r \log(x_i) - \eta$$

MLEs of  $\lambda$  and  $\theta$  is a solution of equation (2.5) accomplished by addressing the first partial derivatives of the total log-likelihood to be zero. So, we consider the equation as follows,

$$\frac{d \log L}{d \lambda} = \sum_{i=1}^r \frac{x_i^\theta + 1}{\lambda + (\lambda + 1)x_i^\theta} + \sum_{i=1}^r x_i^\theta + \frac{(n-r)}{(x_{(r)}^\theta + 1)e^{\lambda x_{(r)}^{-\theta}} - x_{(r)}^\theta}$$

$$\frac{d \log L}{d \theta} = -\lambda \sum_{i=1}^r x_i^{-\theta} \log(x_i) + \sum_{i=1}^r \frac{(\lambda + 1)x_i^\theta \log(x_i)}{\lambda + (\lambda + 1)x_i^\theta} - 2 \sum_{i=1}^r \frac{x_i^\theta \log(x_i)}{x_i^\theta + 1} + \eta_1(x)$$

The closed form solutions to the nonlinear Equations are difficult to reach and a numerical method must be applied to solve these simultaneous equation for obtaining the MLE of  $\lambda$  and  $\theta$ .

### IV. Asymptotic Confidence Intervals

The maximum likelihood estimators of the unknown parameters are not in closed form, it is not easy to drive the exact distributions of the MLEs. Thus, we use the asymptotic distribution of MLEs for the constructions of asymptotic confidence intervals of the parameters based on observed Fisher information matrix. Let  $\hat{\alpha} = (\hat{\lambda}, \hat{\theta})$ , be the MLE of  $\alpha = (\lambda, \theta)$ . The observed Fisher information matrix is given by:

$$I(\alpha) = \begin{bmatrix} \frac{\partial \ln L(\theta, \lambda)}{\partial \lambda^2} & \frac{\partial \ln L(\theta, \lambda)}{\partial \lambda \partial \theta} \\ \frac{\partial \ln L(\theta, \lambda)}{\partial \theta \partial \lambda} & \frac{\partial \ln L(\theta, \lambda)}{\partial \theta^2} \end{bmatrix}$$

$$\frac{\partial \ln L(\theta, \lambda)}{\partial \lambda^2} = \sum_{i=1}^r \frac{(x_i^\theta + 1)^2}{((\lambda + (\lambda + 1)x_i^\theta))^2} + \frac{(n-r)(x_{(r)}^\theta + 1)x_{(r)}^\theta e^{\lambda x_{(r)}^{-\theta}}}{(x_{(r)}^\theta - (x_{(r)}^\theta + 1)e^{\lambda x_{(r)}^{-\theta}})^2}$$

$$\frac{\partial \ln L(\theta, \lambda)}{\partial \theta \partial \lambda} = \sum_{i=1}^r \left( \frac{x_i^{-\theta} \log(x_i)}{\lambda + (\lambda + 1)x_i^\theta} - \frac{(\lambda + 1)x_i^\theta (x_i^\theta + 1) \log(x_i)}{(\lambda + (\lambda + 1)x_i^\theta)^2} \right) - \sum_{i=1}^r x_i^{-\theta} (-\log(x_i))$$

Thus, the observed variance-covariance matrix becomes  $I^{-1}(\hat{\alpha})$ . The asymptotic distribution of MLE

$\hat{\alpha}$  is a bivariate normal distribution as  $\hat{\alpha}N(0, I^{-}(\hat{\alpha}))$ . Consequently, two sided equal tailed  $100(1-\eta)$  asymptotic confidence intervals for the parameters  $\lambda$  and  $\theta$  are given by  $\left[ \hat{\lambda} + Z_{\frac{\eta}{2}} \sqrt{var(\hat{\lambda})} \right]$  and  $\left[ \hat{\theta} + Z_{\frac{\eta}{2}} \sqrt{var(\hat{\theta})} \right]$  respectively. Here,  $Var(\hat{\lambda})$  and  $Var(\hat{\theta})$  are diagonal elements of the observed variance-covariance matrix  $I^{-}(\hat{\alpha})$  and  $Z_{\frac{\eta}{2}}$  is the upper  $\left( \frac{Z_{\eta}}{2} \right)^{th}$  percentile of the standard normal distribution.

### V. The Bayesian Estimation

In this section, we discuss the Bayes estimators of the unknown parameters of the model in (2) under square error loss function (SELF). In order to select the best decision in decision theory, an appropriate loss function must be specified. SELF is generally used for this purpose. The use of the SELF is well justified when over estimation and under estimation of equal magnitude has the same consequences. When the true loss is not symmetric with respect to over estimation and under estimation, asymmetric loss functions are used to represent the consequences of different errors. If all parameters of the model are unknown, a joint conjugate prior for the parameters does not exist. In such conditions there are numerous ways to choose the priors. Hence, we choose to consider the piecewise independent priors. The proposed priors for the parameters  $\lambda$  and  $\theta$  may be taken as:

$$\begin{aligned} g_1(\lambda) &= \lambda^{a_1-1} e^{-\lambda b_1}, & a_1, b_1 > 0 \\ g_2(\theta) &= \lambda^{a_1-1} e^{-\lambda b_1}, & a_2, b_2 > 0 \end{aligned}$$

Thus, the joint prior distribution of  $\lambda$  and  $\theta$  can be written as:

$$g(\lambda, \theta) = \lambda^{a_1-1} \theta^{a_1-1} e^{-(\lambda b_1 + \theta b_1)} \tag{4.1}$$

Now we derive the Bayes estimators for the unknown parameters  $\lambda$  and  $\theta$  under squared error loss function. If  $\mu$  is the parameter to be estimated by an estimator  $\hat{\mu}$  then the squared error loss function is defined as  $L_s(\mu, \hat{\mu}) = (\mu - \hat{\mu})^2$ . The joint posterior distribution of  $\lambda$  and  $\theta$  after simplification is:

$$\Pi(\lambda, \theta | \underline{x}) = \frac{\frac{n}{(n-r)} \lambda^{a_1-1} \theta^{a_2-1} e^{(\lambda b_1 + \theta b_2)} \prod_{i=0}^r f(x_i) (1 - F(x))^{n-r}}{\int_0^\infty \int_0^\infty \frac{n}{(n-r)} \lambda^{a_1-1} \theta^{a_2-1} e^{(\lambda b_1 + \theta b_2)} \prod_{i=0}^r f(x_i) (1 - F(x))^{n-r} \partial \lambda \partial \theta} \tag{4.2}$$

Therefore, the Bayes estimator of any function of  $\lambda$  and  $\theta$ , say  $\alpha(\hat{\lambda}, \hat{\theta})$  under squared error loss function is.

#### I. Subsection One

##### Lindley's Approximation

It is difficult to compute Eq. (4.2) analytically. Lindley's [15] approximation is used to compute the ratio of integrals of the form Eq. (4.3). Based on Lindley's approximation, the approximate Bayes estimator of  $\lambda$  under the squared error loss function is:

$$\hat{\lambda}_{lindley} = \hat{\lambda} + \frac{1}{2} [\mu_1 (2\rho_1 \sigma_{11} + 2\rho_2 \sigma_{21} + \sigma_{11}^2 L_{111} + 2\sigma_{12} \sigma_{21} L_{111} + \sigma_{11} \sigma_{22} L_{211} + \sigma_{12} \sigma_{22} L_{222})] \tag{4.4}$$

$$\hat{\theta}_{lindley} = \hat{\theta} + \frac{1}{2} [\mu_2 (2\rho_2 \sigma_{22} + 2\rho_1 \sigma_{21} + \sigma_{22}^2 L_{222} + 2\sigma_{12} \sigma_{11} L_{111} + 3\sigma_{12} \sigma_{22} L_{122})] \tag{4.5}$$

Here  $L(\lambda, \theta)$  is the log-likelihood and  $\rho(\lambda, \theta)$  is the log of prior distribution  $\pi(\lambda, \theta)$ ,  $\hat{\lambda}$  and  $\hat{\theta}$  are the MLEs of  $\lambda$  and  $\theta$  respectively.

### VI. Simulation Study

This section deals with a Monte Carlo simulation study. Here, we compare various estimators developed in the previous sections with the help of Monte Carlo simulation study. Six different sample sizes  $n = 50, 60, 70, 80$  and  $90$  are considered in the simulation study. Following combination



of the true values of the parameters  $(\lambda, \theta) = (0.5, 1)$  and  $(\lambda, \theta) = (1.5, 1)$  are taken. In each case the ML and Bayes estimates of the unknown parameters are computed. The whole process is simulated 1000 times. Tables 1–2 report the simulation results including Average Estimate (AE), MSE of the IPBH parameters.

**Table 1:** Bayes estimate of the parameter  $\lambda$  and  $\theta$  when  $\theta = 1$  and  $\lambda = 0.5$

n	r	Prior1		Prior2		Prior1		Prior2	
		$\hat{\lambda}$				$\hat{\theta}$			
		AE	MSE	AE	MSE	AE	MSE	AE	MSE
50	46	0.5332	0.018	0.5714	0.0152	1.0862	0.0082	1.0778	0.2459
50	48	0.5321	0.0137	0.5318	0.0142	1.0571	0.0715	1.0754	0.0821
60	56	0.5263	0.0124	0.5268	0.0122	1.0655	0.00615	1.0553	0.00567
60	58	0.5195	0.0102	0.5257	0.0099	1.0525	0.0516	1.0529	0.0588
70	66	0.5173	0.0089	0.5224	0.0091	1.0491	0.0506	1.0551	0.0511
70	68	0.5171	0.0084	0.5223	0.0092	1.0468	0.0485	1.0492	0.0492
80	76	0.5168	0.0071	0.5152	0.007	1.0423	0.0447	1.0468	0.0429
80	78	0.5156	0.0061	0.5187	0.0074	1.0387	0.0394	1.0271	0.0366
90	86	0.5078	0.0054	0.5162	0.0063	1.0311	0.0343	1.0327	0.0364
90	88	0.5115	0.0049	0.511	0.0053	1.0296	0.0316	1.0329	0.0349

**Table 2:** Bayes estimate of the parameter  $\lambda$  and  $\theta$  when  $\theta = 1$  and  $\lambda = 1.5$

n	r	Prior1		Prior2		Prior1		Prior2	
		$\hat{\lambda}$				$\hat{\theta}$			
		AE	MSE	AE	MSE	AE	MSE	AE	MSE
50	46	1.7311	0.464	1.7088	0.394	1.0504	0.0528	1.0498	0.0492
50	48	1.6536	0.2447	1.6543	0.2456	1.0369	0.0407	1.0439	0.0423
60	56	1.6038	0.1654	1.6382	0.1643	1.0295	0.0313	1.0453	0.0338
60	58	1.5855	0.1197	1.5290	0.1250	1.0244	0.0286	1.0306	0.0298
70	66	1.5841	0.1195	1.5771	0.1181	1.0258	0.0248	1.0221	0.0258
70	68	1.5723	0.0956	0.1181	0.0922	1.0243	0.0244	1.0202	0.0231
80	76	1.5636	0.0958	1.5771	0.1127	1.0143	0.0199	1.0285	0.0237
80	78	1.573	0.0856	1.5639	0.0827	1.0228	0.0198	1.0191	0.002
90	86	1.5614	0.0821	1.5587	0.0792	1.0186	0.0182	1.0162	0.0198
90	88	1.5534	0.0712	1.5489	0.0710	1.0199	0.0171	1.0276	0.0175

## VII. Real-Life Applications

In this section, we illustrate estimation procedures discussed in the previous sections with the help of one real datasets. Here, we consider a real dataset namely the strengths of glass fibres The Data I, respectively are given below:

**Data set:**

This dataset consists of 63 observations which are generated to simulate the strengths of glass fibres [18].The 63 observations of the dataset are as follows: “1.014, 1.081, 1.082, 1.185, 1.223, 1.248, 1.267, 1.271, 1.272, 1.275, 1.276, 1.278, 1.286, 1.288, 1.292, 1.304, 1.306, 1.355, 1.361, 1.364, 1.379, 1.409, 1.426, 1.459, 1.460, 1.476, 1.481, 1.484, 1.501, 1.506, 1.524, 1.526, 1.535, 1.541, 1.568, 1.579, 1.581, 1.591, 1.593,

1.602, 1.666, 1.670, 1.684, 1.691, 1.704, 1.731, 1.735, 1.747, 1.748, 1.757, 1.800, 1.806, 1.867, 1.876, 1.878, 1.910, 1.916, 1.972, 2.012, 2.456, 2.592, 3.197, and 4.121”.

We calculate MLEs of the unknown parameters together with some useful measure of goodness-of fit tests for one dataset, namely, the negative log likelihood function  $-\ln L$ , the Akaike information criterion denoted by  $AIC = 2k - 2\ln L$ , proposed by Akaike [16] and Bayesian information criterion denoted by  $BIC = k\ln(n) - 2\ln L$ , proposed by Schwarz [17], where  $k$  is the number of parameters in the model,  $n$  is the number of observations in the given datasets,  $L$  is the maximized value of the likelihood function for the estimated model and Kolmogorov-Smirnov (K-S) statistic with its  $p$ -value. The best distribution corresponds to the lowest  $-\ln L$ , AIC, BIC and K-S statistic and the highest  $p$  values. The K-S statistic with its  $p$ -value is obtained using  $ks$  test function in statistical software R. The results of the MLEs and measures of goodness-of-fit tests are reported in Tables 3 and 4, respectively. These results show that IPBH distribution is the best choice for the considered datasets. However, for Data I, according to K-S test IPBH is better than the BH.

**Table 3:** Data Summary for the Data Set

Min	1 <sup>st</sup> Qu.	Median	mean	3 <sup>rd</sup> Qu.	Max
1.014	1.305	1.526	1.616	1.741	4.121

**Table 4:** Goodness of Fit criterions on the data set

Distribution	Estimates	$-\log L$	AIC	BIC	K-S (stat)	P-value
IPBD	$\hat{\theta} = 5.7408$ $\hat{\lambda} = 6.03415$	15.403	26.8066	39.0929	0.08507	0.7197
BR	$\hat{\theta} = 0.2325$ 0	113.364	222.7286	235.0148	0.77052	< 0.001
Weibull	$\hat{\theta} = 3.0521$ $\hat{\lambda} = 1.7873$	43.254	96.7345	101.12	0.2051	0.009
Exponential	$\hat{\lambda} = 0.6189$	93.222	187.432	190.523	0.4721	< 0.001

### VIII. Conclusion

This article deals with the classical and Bayesian estimation procedures for parameters of inverse power Burr-Hatke distribution using second type censoring. The maximum likelihood estimators and corresponding asymptotic confidence intervals based on observed Fisher information matrix of the unknown parameters were derived. The Bayes estimates of the parameters under square error loss function were approximated using Lindley’s approximation. The performance of these estimators was examined by extensive Monte Carlo simulation study, which indicated that the MLEs can be obtained easily and quickly with satisfactory estimates. For more efficient estimators, Bayes estimation method with available prior information or convenient non-informative priors in the absence of prior information is recommended.

### References

[1] Da Silva, R. V., Gomes-Silva, F., Ramos, M. W. A., & Cordeiro, G. M. (2015). The exponentiated Burr XII Poisson distribution with application to lifetime data. *International Journal of Statistics and Probability*, 4(4): 112-131.

[2] Afify, A. Z., Cordeiro, G. M., Ortega, E. M., Yousof, H. M. and Butt, N. S. (2018). The four

parameter Burr XII distribution: properties, regression model, and applications. *Communications in Statistics-Theory and Methods*, 47(11): 2605–2624.

[3] Alizadeh, M., Lak, F., Rasekhi, M., Ramires, T. G., Yousof, H. M., and Altun, E. (2018). The odd log-logistic Topp–Leone G family of distributions: heteroscedastic regression models and applications. *Computational Statistics*, 33(3): 1217-1244.

[4] Abouelmagd, T. H. M. (2018). The Logarithmic Burr-Hatke Exponential Distribution for Modeling Reliability and Medical Data. *International Journal of Statistics and Probability*, 7(5): 73-85.

[5] Yadav, A. S., Altun, E. and Yousof, H. M. (2021). Burr–Hatke Exponential Distribution: A Decreasing Failure Rate Model, Statistical Inference and Applications. *Annals of Data Science*, 8(2): 241–260.

[6] Nasir, M. A., Tahir, M. H., Chesneau, C., Jamal, F., and Shah, M. A. A. (2020). The odds generalized gamma-G family of distributions: Properties, regressions and applications. *Statistica*, 80(1): 3-38.

[7] Anzagra, L., Sarpong, S., and Nasiru, S. (2020). Chen-G class of distributions. *Cogent Mathematics and Statistics*, 7(2):1721.

[8] Afify, A. Z., Aljohani, H. M., Alghamdi, A. S., Gemeay, A. M. and Sarg, A. M. (2021). A new two-parameter burr-hatke distribution: properties and bayesian and non-bayesian inference with applications. *Journal of Mathematics*. 2021: 16.

[9] Zamanah, E., Nasiru, S., and Luguterah, A. (2022). Harmonic Mixture Weibull-G Family of Distributions: Properties, Regression and Applications to Medical Data. *Computational and Mathematical Methods*, 2022: 24.

[10] Khaazmi, O., Nik, A. S., Hamedani, G. G., and Altun, E. (2022). Harmonic Mixture-G Family of Distributions: Survival Regression, Simulation by Likelihood, Bootstrap and Bayesian Discussion with MCMC Algorithm. *Austrian Journal of Statistics*, 51(2): 1-27.

[11] Altun, E., Alizadeh, M., & Yousof, H. M. (2022). The Odd Log-Logistic Weibull-G Family of Distributions with Regression and Financial Risk Models. *Journal of the Operations Research Society of China*, 10(1): 133-158.

[12] Ahmed, E., Hassan M. A. and Ahmed, Z. A. (2021). Bayesian and Classical Inference under Type-II Censored Samples of the Extended Inverse Gompertz Distribution with Engineering Applications. *Entropy*, 23: 1578.

[13] Hassan, M. A. (2021). Statistical inference of chen distribution based on two progressive type-ii censoring schemes. *Computers, Materials & Continua*, 66: 2797–2814.

[14] Cohen, A. C. (1965). Maximum Likelihood Estimation in the Weibull Distribution Based on Complete and Censored Samples. *Technometrics*, 7: 579–588.

[15] Lindley, D. V. (1980). Approximate Bayesian methods (with discussions), *Trabajos de Estadística*. 31: 232–245.

[16] Akaike, H. A. (1974). New look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19: 716-723.

[17] Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6: 461-464.

## A COMPARATIVE STUDY OF INVENTORY MODELLING: DETERMINISTIC OVER STOCHASTIC APPROACH

Lalji Kumar, Pratima Singh Ghoshi, Shreyashi Saxena, Kajal Sharma



Dr. Harisingh Gour Vishwavidyalaya, Sagar, Madhya Pradesh, India – 470003

*laljikumar07@gmail.com, pratimasingh04625@gmail.com, shreyashisaxena09@gmail.com,  
kajal100136@gmail.com*

### Abstract

*This research study provides a comprehensive comparison of two critical approaches to inventory modelling- deterministic and stochastic. The deterministic model employs traditional optimization techniques to optimize complex systems, while the stochastic model leverages Particle Swarm Optimization (PSO) simulations to tackle the challenges posed by uncertain dynamics. This approach enables us to develop effective strategies for optimizing complex systems. After conducting sensitivity analyses, it was found that the deterministic model oversimplifies demand dynamics, whereas the stochastic model more adeptly captures market uncertainties. As a result, this study suggests that businesses adopt stochastic approaches to inventory management to better engage in adaptive decision-making, contingency planning, optimal resource allocation, risk mitigation, and realistic performance metrics. The research provides valuable insights for businesses seeking to navigate the complexities of modern supply chains.*

**Keywords:** Inventory, Deterioration, Stochastic optimization, Risk analysis, Particle swarm optimization (PSO).

**MSC Classification:** 90B05, 90B30, 90B50, 91B70, 93E20

### I. Introduction

Inventory modelling plays an integral role in contemporary supply chain management. A comprehensive understanding of the relationship between inventory dynamics and market uncertainties is essential, prompting a comprehensive exploration of deterministic and stochastic approaches. This research delves into the core of this dichotomy, aiming to provide invaluable insights for businesses grappling with the challenges of unpredictable market conditions. In the global marketplace, businesses encounter continuous volatility and uncertainty. The traditional deterministic approach to inventory modelling, relying on fixed parameters and constant demand assumptions, has limitations in capturing the fluidity of real-world markets. A sudden surge in market trends or an unforeseen external event, such as a pandemic, can disrupt this equilibrium, leaving the inventory misaligned with actual demand. This mismatch results in potential revenue loss due to stockouts or excessive holding costs and underscores the urgency for a more adaptive and resilient approach. On the other hand, the stochastic paradigm acknowledges the inherent variability in market dynamics. In a world where demand fluctuations are normal, businesses cannot afford to disregard the impact of uncertainty on inventory management. For example, a manufacturing company that utilizes stochastic modelling may adjust its production levels dynamically based on probabilistic demand forecasts. This approach allows for real-time

responsiveness to market shifts, minimizing the risks associated with stockouts, excess inventory, and subsequent financial repercussions.

To address this problem, a meticulous methodology has been crafted, leveraging both classical optimization techniques for deterministic modelling and Particle Swarm Optimization (PSO) simulations for the stochastic scenario. The deterministic approach involves traditional optimization algorithms, aiming to find the optimal solution based on fixed parameters. While this method is widely used, it often must account for the inherent uncertainties in dynamic markets. In contrast, the stochastic approach utilizes PSO, a nature-inspired optimization algorithm that mimics the social behavior of particles to search for optimal solutions in a multidimensional space. In the context of inventory modelling, PSO enables the exploration of diverse demand scenarios, considering the stochastic nature of the objective function. By simulating multiple scenarios, PSO provides a more realistic representation of the potential outcomes in uncertain market conditions. The application of PSO in stochastic modelling is particularly relevant when dealing with complex and dynamic objective functions influenced by stochastic variables, such as fluctuating demand patterns. This approach allows the model to adapt and evolve as market conditions change, providing decision-makers with a versatile tool for strategic inventory planning.

This study holds significant importance as it has the potential to revolutionize the way businesses approach inventory management. It presents a paradigm shift from rigid and deterministic methods to adaptable and stochastic techniques. Given supply chains' growing interconnectedness and susceptibility to global disruptions, the need for a responsive and agile inventory modelling framework is increasingly crucial. From a managerial perspective, this study empowers decision-makers to make informed decisions amidst uncertainty. By highlighting the limitations of deterministic models and the benefits of stochastic approaches, it encourages managers to adopt adaptive strategies that align with the ever-evolving nature of modern markets. For example, a retail manager equipped with insights from stochastic modelling can proactively adjust inventory levels based on probabilistic demand forecasts, thereby minimizing the impact of unforeseen events such as stockouts or excess inventory. Furthermore, the study contributes to academic discourse by comparing deterministic and stochastic inventory modelling comprehensively. By contrasting classical optimization techniques with advanced optimization algorithms such as PSO, it provides a holistic understanding of the strengths and weaknesses of each approach. This nuanced understanding is essential for researchers and academicians seeking to advance the theoretical foundations of inventory management. The study's real-world applicability extends beyond conventional industries to emerging sectors like e-commerce, where demand patterns are subject to rapid and unpredictable changes. By highlighting the adaptability and effectiveness of stochastic modelling, the study provides a roadmap for businesses navigating the complexities of a digital economy.

Altogether, this research aims to redefine the contours of inventory management, transcending the limitations of deterministic paradigms. By combining real-world context, meticulous methodologies, and a profound understanding of the problem at hand, this study aims to guide businesses in navigating the uncertainty of modern supply chains.

## II. Literature Review

The recent literature encompasses diverse studies, highlighting the ongoing debate between stochastic and deterministic approaches in various operations research and management domains. The work on crude oil price forecasting ([17]) emphasizes the advantages of their stochastic pruning DE-DL method and shows superior results compared to deterministic counterparts. Using a two-stage stochastic programming model, [6] explores the ability of Industrial Symbiosis networks to

withstand fluctuations in demand, revealing their resilience. Adopting a stochastic perspective [2], this article demonstrates the efficacy of their modified particle swarm optimization algorithm within a rolling horizon framework for contributing to aggregate production planning under uncertainty. A two-stage stochastic programming model is presented for disaster preparedness [9], which considers uncertainties in emergency demand and road network congestion. Proposing a two-stage stochastic programming model, [14] advocates for smaller initial networks to adapt to future uncertainties in district cooling network design. A study on forecasting intermittent demand ([20]) uses genetic algorithms and particle swarm optimization, highlighting the stochastic nature of these optimization methods.

Modelling the hot deformation ([16]) of multiphase steels requires advanced numerical models for deterministic and stochastic approaches. A stochastic inventory model that incorporates quadratic price-sensitive demand ([12]). The effects of different probability distributions are compared. Genetic Algorithms is advocated for optimizing inventory management, favoring their efficiency over traditional deterministic systems ([4]). In conversion processes, [11] explore homogeneous and heterogeneous scenarios, acknowledging the stochastic nature of optimal conversion timing, quantity, cost, and time considerations and optimizes using a metaheuristic algorithm. The collective findings suggest a growing preference for stochastic approaches, recognizing their ability to capture and address uncertainties inherent in real-world operational scenarios. The contributions to the comparative study between stochastic and deterministic approaches in inventory control ([4]), particularly in a pharmaceutical distribution setting. Their conceptual model, rooted in modern control theory, addresses practical supply chain constraints. The dynamic mathematical model considers multiple products, variable lead time, deterministic and stochastic demand, and various ordering policies. Objective functions maximize planned versus realized inventory levels and minimize stock-out situations. Real-life data validate the model, providing a comprehensive solution to pharmaceutical supply chain inventory challenges. Exploring the intricate relationship between inventory and demand, [19] proposes a logistic growth model for inventory-dependent demand rates. The study begins with a deterministic optimal control problem, optimizing the present value of total net profit over an infinite horizon. It then extends to the stochastic version, solving the associated Hamilton-Jacobi-Bellman equation and demonstrating optimal inventory levels in a stochastic context. A study ([3]) investigates how prices and production are jointly determined over multiple periods in the face of non-stationary stochastic demand. Their study considers limited production capacity and discretionary sales, comparing partial planning or delayed strategies. The analysis, incorporating deterministic approximations, provides insights into the effectiveness of delayed production versus delayed pricing, with heuristics achieving a high percentage of the corresponding optimal strategy.

A deterministic inventory model is presented for a single item with two storage facilities ([10]). The model addresses linearly time-dependent demand over a fixed and finite time horizon. The model, applicable to scenarios like food grain production, offers a general solution through the gradient method, highlighting its versatility in products with periodic production and linearly increasing demand. Tackling the inventory control problem of nonstationary stochastic demand by incorporating a certainty-equivalent mixed integer linear programming model using the (R, S) policy ([18]). The study provides numerical examples and demonstrates the model's application through a piecewise linear approximation to handle non-linear cost functions. The focus is on inventory planning in closed-loop supply chains ([8]), specifically in equipment-intensive service industries. The planning approach is tactical, which is concerned with short-term decisions rather than long-term strategy. Their mixed-integer programming model, addressing conflicting business objectives, is accompanied by a metaheuristic approach to solution. Experimental evaluations demonstrate the model's effectiveness, emphasizing the impact of cost weightings on different planning strategies. A stochastic inventory model is presented ([13]), considering price-dependent demand, probabilistic

lead time, and allowances for shortages in a finite time horizon. The study emphasizes the financial implications of advance payment on unit prices, deriving an expected average profit expression. Numerical examples and solution techniques, such as the generalized reduced gradient technique and stochastic search genetic algorithm, show the model's applicability and benefits.

Inspired by these studies, the current research endeavors to conduct a comprehensive comparative analysis between deterministic and stochastic approaches in inventory management. The aim is to derive insights into each approach's trade-offs, advantages, and practical implications, contributing to the ongoing discourse in the field.

### III. Notations

The following notations are used in subsequent discussions, in accordance with usual tradition:

#### Parameters

$I(t)$	:	Instantaneous inventory level
$a$	:	Demand potential
$b$	:	Time dependent parameter
$c$	:	Time sensitive parameter
$\theta$	:	Constant deterioration rate per unit per unit of time
$Q$	:	Stock replenishment quantity
$h_c$	:	Per unit holding cost
$C_s$	:	Per unit shortage cost
$C_0$	:	Per unit purchasing cost

#### Decision Variable

$t'$	:	Stock ending time
$T$	:	Inventory cycle time
$p$	:	Per unit price
$\pi(T, t', p)$	:	Total profit per cycle

### IV. Model Formulation with Deterministic Approach

Consider the initial stock size at time  $t = 0$  is  $Q$ . As the business begins, the stock experiences depletion over time. The demand is price and time-dependent, expressed as  $D(p, t) = a - bt - cp$ , where  $a, b, c > 0$ . Here,  $a$  represents the base demand,  $b$  is time dependent parameter, and  $c$  reflects the price sensitivity parameter of demand. After the time  $t'$  stock will be end and then shortage begins. It is assumed that the shortage is fully backlogged during stock out time till the time  $T$ .

Based on this condition, the rate of the declining of the inventory level ( $I(t)$ ), due to demand and per unit deterioration rate  $\theta$ , is given as:

$$\frac{dI}{dt} = \begin{cases} -(a - bt - cp) - \theta I(t), & 0 \leq t \leq t' \\ -(a - bt - cp) & t' \leq t \leq T \end{cases} \quad (1)$$

with the conditions  $I(t') = 0, I(0) = Q$ .

Solving the differential equation we have,

$$I(t) = \begin{cases} \frac{1}{\theta} [e^{\theta t'}(a - bt' - c_p) - e^{\theta t}(a - bt - c_p)] & 0 \leq t \leq t' \\ \left( at' - \frac{bt'^2}{2} - c_p t' \right) - \left( at - \frac{bt^2}{2} - c_p t \right) & t' \leq t \leq T \end{cases} \quad (2)$$

Based on these inventory equations, there are several costs associated with the profit function. The cost incurred in holding the products with per unit holding cost  $h_c$  is given as:

$$THC = h_c \int_0^{t'} I(t) dt \quad (3)$$

Or,

$$THC = \frac{1}{2\theta^3} \left[ h \left( -2(a - c_p)\theta(1 - e^{t'\theta} + t'\theta) + b(-2 + t'^2\theta^2 + e^{t'\theta}(2 - 2t'\theta)) \right) \right] \quad (4)$$

Total shortage cost with per unit shortage cost  $c_s$ , during the stock-out period is given as:

$$TSC = c_s \int_{t'}^T I(t) dt \quad (5)$$

Or,

$$TSC = c_s \left( -\frac{aT^2}{2} + \frac{1}{2}c_p T^2 + \frac{bT^3}{6} + aTt' - c_p Tt' - \frac{at'^2}{2} + \frac{1}{2}c_p t'^2 - \frac{1}{2}bTt'^2 + \frac{bt'^3}{3} \right) \quad (6)$$

Total purchasing cost with per unit purchasing cost  $C_0$ , is given as:

$$TPC = C_0 Q, \quad \text{where } Q = I(0) \quad (7)$$

Or,

$$TPC = \frac{1}{2\theta^2} \left( c \theta \left( -b + (-a + c_p)\theta + e^{t'\theta}(b + (a - c_p - bt')\theta) \right) \right) \quad (8)$$

Total cost incurred in terms of deterioration is as follows:

$$TDC = c_d \theta \int_0^{t'} I(t) dt \quad (9)$$

Or,

$$TDC = \frac{1}{2\theta^3} \left( c_d \left( -2(a - c_p)\theta(1 - e^{t'\theta} + t'\theta) + b(-2 + t'^2\theta^2 + 2e^{t'\theta}(1 - t'\theta)) \right) \right) \quad (10)$$

Total revenue generated during the selling period is as follows:

$$TRV = p \int_0^T D(t, p) dt \quad (11)$$

Or,

$$TRV = P \left( (a - c_p)T - \frac{bT^2}{2} \right) \quad (12)$$

Combining these costs, we have formulated the profit function given as:

$$\pi(t', T, p) = TRV - TDC - TPC - TSC - THC \quad (13)$$

Or,



$$\begin{aligned} \pi(t', T, p) = & P \left( (a - c p)T - \frac{b T^2}{2} \right) \\ & - \frac{1}{2\theta^3} \left( c_d \left( -2(a - c p)\theta(1 - e^{t'\theta} + t'\theta) + b(-2 + t'^2\theta^2 + 2e^{t'\theta}(1 - t'\theta)) \right) \right) \\ & - \frac{1}{2\theta^2} \left( c \theta \left( -b + (-a + c p)\theta + e^{t'\theta}(b + (a - c p - bt')\theta) \right) \right) \\ & - \frac{1}{2\theta^3} \left[ h_c \left( -2(a - c p)\theta(1 - e^{t'\theta} + t'\theta) + b(-2 + t'^2\theta^2 + e^{t'\theta}(2 - 2t'\theta)) \right) \right] \\ & - \frac{1}{2\theta^3} \left[ h_c \left( -2(a - c p)\theta(1 - e^{t'\theta} + t'\theta) + b(-2 + t'^2\theta^2 + e^{t'\theta}(2 - 2t'\theta)) \right) \right] \end{aligned} \tag{14}$$

For the optimization of this model, we have utilized the classical optimization approach given in the following theorem.

**Theorem 1:** For the positive values parameters, the proposed profit function is concave with respect to the holding time  $t'$  and replenish time  $T$ .

**Proof:** Using the objective function, we have formulated the Hessian matrix, given as,

$$H = \begin{bmatrix} \frac{\partial^2 \pi}{\partial t'^2} & \frac{\partial^2 \pi}{\partial t' \partial T} \\ \frac{\partial^2 \pi}{\partial t' \partial T} & \frac{\partial^2 \pi}{\partial T^2} \end{bmatrix} \tag{15}$$

Where,

$$\begin{aligned} \frac{\partial^2 \pi}{\partial t'^2} = & \frac{1}{6\theta^3} \left( 3(a - c p)\theta \left( -2e^{\theta t'} h_c \theta^2 + \theta(2 c_s \theta - 2C_0 e^{\theta t'} \theta^2 - 2 c_d e^{\theta t'} \theta^2) \right) + \right. \\ & b \left( h_c \left( -6\theta^2 + 12 e^{\theta t'} \theta^2 + 6e^{\theta t'} \theta^2(-1 + \theta t') \right) - \theta(\theta^2(-8 c_s (T - t') + 2 c_s (T + t'))) + \right. \\ & \left. \left. c_d \left( 6 \theta^2 - 12 e^{\theta t'} \theta^2 + e^{\theta t'} \theta^2(6 - 6\theta t') \right) - 6c_0(2 e^{\theta t'} \theta^2 + e^{\theta t'} \theta^2(-1 + \theta t')) \right) \right); \end{aligned}$$

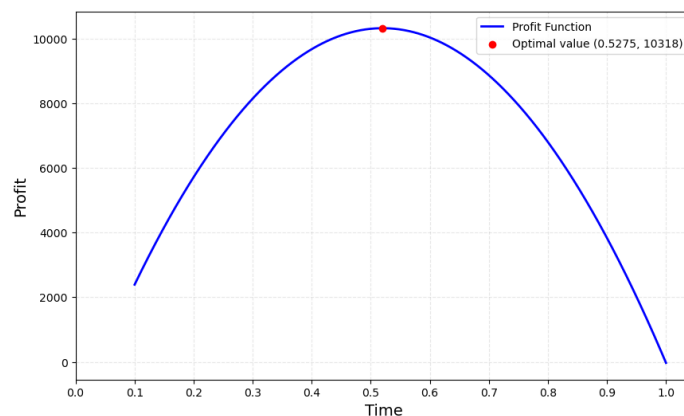
$$\frac{\partial^2 \pi}{\partial T \partial t'} = \frac{\partial^2 \pi}{\partial t' \partial T} = 0;$$

and,

$$\frac{\partial^2 \pi}{\partial T^2} = \frac{1}{6 \theta^3} (6 c_s (a - c p)\theta^3 - b \theta^3(6 p + 4c_s(T - t') + 2 c_s(T + 2t')))$$

From the evaluation, we have,  $\frac{\partial^2 \pi}{\partial t'^2}, \frac{\partial^2 \pi}{\partial T^2} > 0$ , and  $\frac{\partial^2 \pi}{\partial t'^2} \frac{\partial^2 \pi}{\partial T^2} - \frac{\partial^2 \pi}{\partial t' \partial T} < 0$ .

Thus, the objective function is concave with respect to replenish time  $T$  and  $t'$ . The figure 1 and 2, shows the concavity of the profit function plotted on the values provided in example 1.



**Figure 1:** Concavity of the profit function with respect to replenish time

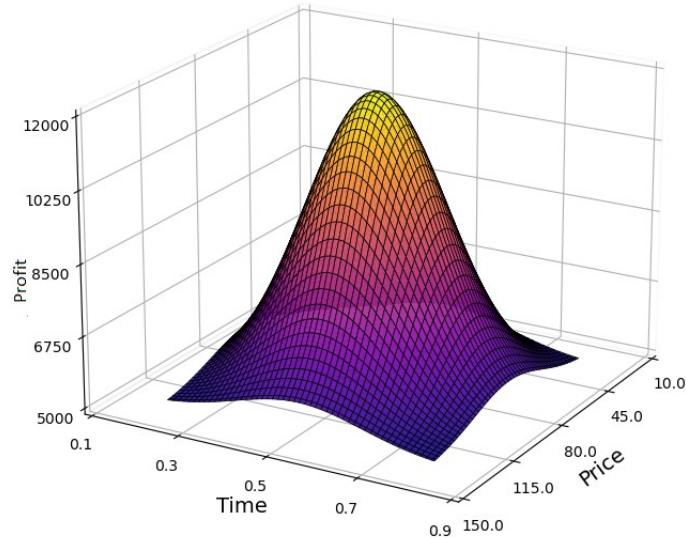


Figure 2: Concavity of the profit function with respect to replenish time and per unit selling price

**Theorem 2:** For the optimal value of  $p$ , the objective function (eq. 14) is concave with respect to the selling price.

**Proof:** From objective function eq. 14, we have

$$\frac{\partial \pi(\cdot)}{\partial p} = \frac{1}{2\theta^2} \left( T(2a - bT)\theta^2 + 2h_c(-1 + e^{t\theta} - t'\theta) - c_\theta \left( -2c_0(-1 + e^{t\theta}) + (4pT + c_s(T - t')^2)\theta + c_d(2 - 2e^{t\theta} + 2t'\theta) \right) \right) \quad (16)$$

Putting  $\frac{\partial \pi(\cdot)}{\partial p} = 0$ , we yield the optimal value of per unit selling price as,

$$p^* = -\frac{1}{4cT\theta^2} \left( 2c h_c - 2ce^{t\theta} h_c + 2cc_0\theta + 2cc_d\theta - 2cc_0e^{t\theta}\theta - 2cc_d e^{t\theta}\theta + 2h_c t'\theta - 2aT\theta^2 + bT^2\theta^2 + cc_s T^2\theta^2 + 2cc_d t'\theta^2 - 2cc_s T t'\theta^2 + c c_s t'^2\theta^2 \right)$$

Again differentiating eq. 14, we have,

$$\frac{\partial^2 \pi(\cdot)}{\partial p^2} = -2 c T \quad (17)$$

Thus, for  $c > 0$ , the proposed profit function is concave. The concavity of the profit function with respect to the price can also be seen in figure 3.

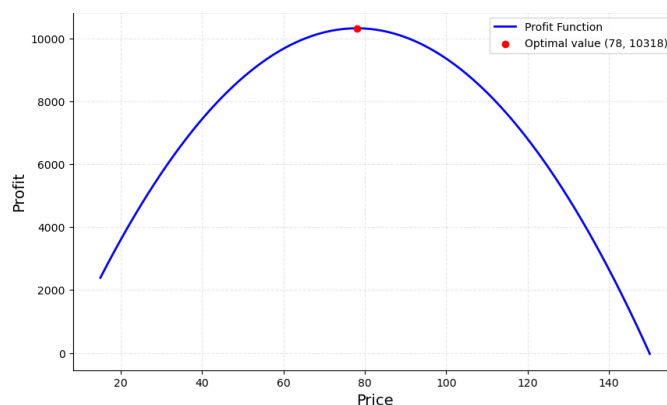


Figure 3: Concavity of the profit function with respect to per unit selling price

Following an in-depth exploration of the deterministic approach, the subsequent section delves into the stochastic approach. This approach intricately considers and integrates uncertain factors associated with demand, acknowledging the dynamic and unpredictable nature of variables. In

contrast to the deterministic approach, which assumes a fixed and known demand, the stochastic approach takes into account the inherent variability and unpredictability in demand, providing a more comprehensive and realistic perspective in decision-making processes.

### V. Model Formulation with Stochastic Approach

Let consider the general form where the demand  $D(p, t)$ , is the function of price and time, and  $f(p)$  is the probability distribution function over price. To be more specific, consider the pre-defined relation of the demand and uncertain factor, i.e.,  $D(p, t) = a - b t - c p + \epsilon$ , where  $b$  and  $c$  are time dependent and price sensitive parameters, and  $\epsilon$  can be determined with the specific distribution, such as uniform, normal etc. depends on the characteristic of the demand fluctuations.

For choosing specific  $\epsilon$ , follows the normal distribution with mean  $\mu$ , and standard deviation  $\sigma$ , the  $\epsilon \sim N(0, \sigma^2)$ , the pdf for the stochastic demand can be expressed as

$$f(p) = e^{-\frac{(D(p,t)-\mu)^2}{2\sigma^2}} \tag{18}$$

The uniform distribution is characterized by the constant probability density within a specific range. Consider the price range as  $[p_{max}, p_{min}]$ . In this case the probability distributon over this range will be as follows:

$$F(p) = \frac{1}{p_{max}-p_{min}} \tag{19}$$

Therefore, we can express the demand with uniform distribution as  $D(p, t) = a - b t - c p + \epsilon$  where  $\epsilon \sim U(p_{min}, p_{max})$ .

Taking these stochastic demand values and the inventory equation (1, 2), we have reworked for all the costs using these equations:

Expected holding cost with per unit holding cost  $h_c$  is as follows:

$$EHC = E \left( \int_{p_{min}}^{p_{max}} \left( h_c \left[ \int_0^{t'} I(t) dt \right] \right) f(p) dp \right) \tag{20}$$

Expected shortage cost with per unit shortage cost  $C_s$  is as follows:

$$ESC = E \left( \int_{p_{min}}^{p_{max}} \left( c_s \left[ \int_{t'}^T I(t) dt \right] \right) f(p) dp \right) \tag{21}$$

Expected shortage cost with per unit shortage cost  $C_s$  is as follows:

$$EPC = E \left( \int_{p_{min}}^{p_{max}} C_0 Q f(p) dp \right) \tag{22}$$

Expected deteriorating cost with per unit deterioration cost  $C_s$  is as follows:

$$EDC = E \left( \int_{p_{min}}^{p_{max}} \left( c_d \theta \int_0^{t'} I(t) dt \right) f(p) dp \right) \tag{23}$$

Expected revenue is as follows:

$$ERV = E \left( \int_{p_{min}}^{p_{max}} \left( p \int_0^T D(t, p) dt \right) f(p) dp \right) \tag{24}$$

Combining all the above cost, the profit function governs as:

$$\pi(t', T, p) = ERV - EHC - ESC - EPC - EDC \tag{25}$$

Or,

$$\begin{aligned} \pi(t', T, p) = & E \left( \int_{p_{min}}^{p_{max}} \left( p \int_0^T D(t, p) dt \right) f(p) dp \right) - E \left( \int_{p_{min}}^{p_{max}} \left( h_c \left[ \int_0^{t'} I(t) dt \right] \right) f(p) dp \right) - \\ & E \left( \int_{p_{min}}^{p_{max}} \left( c_s \left[ \int_{t'}^T I(t) dt \right] \right) f(p) dp \right) - E \left( \int_{p_{min}}^{p_{max}} c_0 Q f(p) dp \right) - E \left( \int_{p_{min}}^{p_{max}} \left( c_d \theta \int_0^{t'} I(t) dt \right) f(p) dp \right) \end{aligned} \quad (26)$$

subject to the conditions  $c_0 \leq p$ ,  $p_{min} \leq p^* \leq p_{max}$ , and  $t' < T$ .

As the objective function is probabilistic, we have utilized Particle Swarm Optimization (PSO) to maximize the profit function. PSO is advantageous in optimization processes and excels in navigating complex solution spaces by simulating the social behavior of particles. PSO facilitates swift convergence to optimal outcomes by continuously adapting individual positions guided by personal and global best solutions. This collaborative, swarm-based approach is particularly effective in tackling intricate profit optimization challenges, especially when confronted with uncertainties such as stochastic demand. The algorithm's capacity to balance exploration and exploitation makes it an adaptable and powerful tool, contributing to improved decision-making in scenarios where traditional optimization methods may fall short. Here is the algorithm (Algorithm 1), inspired by [11] that optimizes the profit function effectively.

---

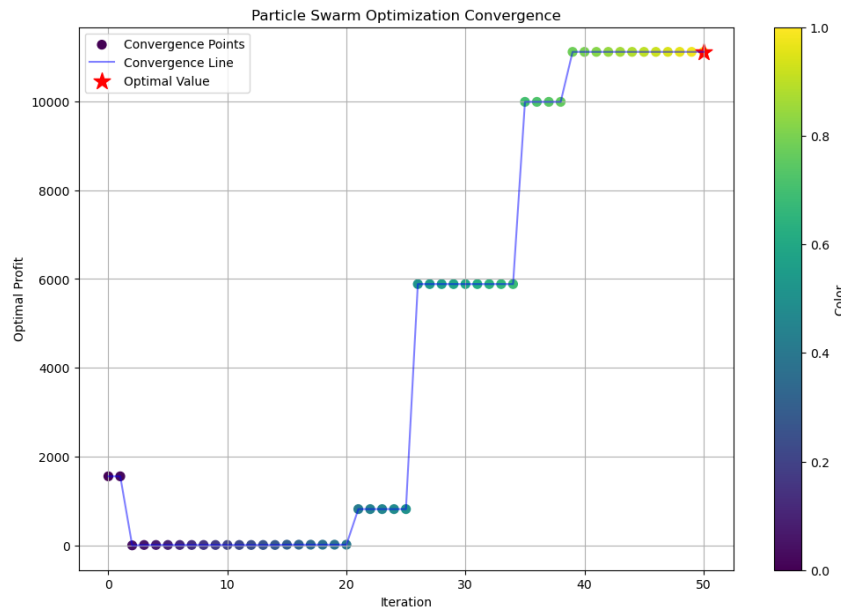
**Algorithm 1:** Algorithm to maximize the profit function using PSO.

**Input:** Parametric values, objective function, constraints.

**Output:** Global best values for  $t'$ ,  $T$ , and  $p$ .

1. Define module 1 taking argument  $(\mu, \sigma)$
  2. Evaluate the  $\epsilon$  using desired probability distribution from eq. 18
  3. Return  $\epsilon$
  4. Define module 2 taking the distribution function for desired probability distribution calling module 1, and objective function and return the value of objective function
  5. Initialize the parameters associated with PSO and identify the decision variables
  6. Initialize the maximum number of iterations for PSO
  7. Initialize the random position and velocities
  8. for  $i=1, 2, \dots$ , maximum iterations do
    9. Evaluate the fitness for each particle by calling module 2
    10. Identify the global best position for each particle
    11. Update the particle's position and velocity equations
    12. Check the convergence criteria for each iteration and find the global best value among them
    13. If maximum value found from existing value, replace the value and set new position
    14. Plot the iteration and function's value
  15. end for
  16. Return the optimal decision variables and corresponding maximum profit based on the best particle's position.
- 

Utilizing the above algorithm, the following plot (figure 4) illustrates the global convergence of the profit function with respect to their decision variables.



**Figure 4:** Iterative convergence of the probabilistic profit function

Now, in further section we will generalize the result difference for both the approaches, i.e., deterministic over stochastic and will find the superiority of the approaches through the numerical simulations.

## VI. Results

Effective inventory management is a critical aspect of supply chain optimization. Two primary approaches have been developed to address uncertainties: deterministic and stochastic. Each of these approaches offers distinct strategies for handling uncertainties in inventory management. In this study, we aim to explore the dynamics of both approaches by utilizing numerical formulations to visualize their impact on demand and profitability. The findings of this study will provide valuable insights for improving inventory management practices, which can ultimately contribute to the overall efficiency of the supply chain. Example 1 illustrates the deterministic method, which relies on known variables and minimizes uncertainties. In the following example, we compare the outcomes of a stochastic approach to a deterministic one. This exploration aims to understand better how different methods influence inventory management and decision-making for businesses seeking stability and precision in stock management.

**Example 1:** Consider a scenario where demand stability is critical and uncertainties are minimized through a deterministic approach. We assume a fixed potential demand of 150 units and factors like price sensitivity parameter ( $c$ ) = 0.1 and time sensitivity parameter ( $b$ ) = 0.1 to tackle the optimization process. We considered holding costs ( $h$ ) = \$2 per unit per unit of time, shortage costs ( $c_s$ ) is \$2 per unit of time, purchasing costs ( $c_0$ ) is 10 per unit, deteriorating rate ( $\theta$ ) = 0.001 per unit per unit of time and deterioration cost ( $c_d$ ) = \$0.1.

The analysis finds insightful metrics: the optimal replenishment time is  $t'$  is 1.69719 units, a shortage duration of 0.29951 units, a streamlined inventory cycle time ( $T$ ) = 1.9967 units, the optimal selling price per unit ( $p$ ) = \$78.01, the optimal profit is \$10318.

Demand Function (Mean = 0, Deviation =  $\pm 5$ )

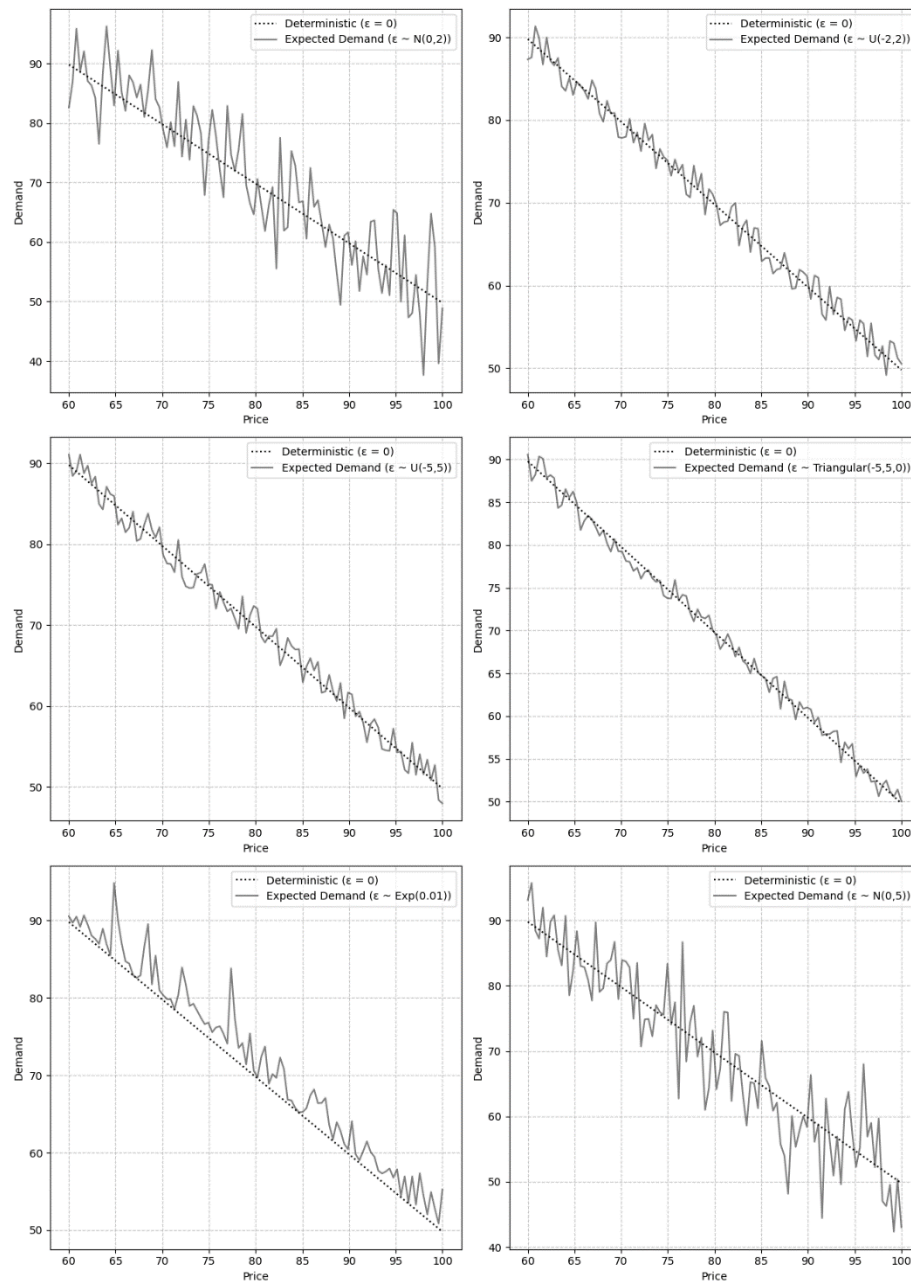


Figure 5: Demand under the uncertainty

The data set from example 1 has been utilized to formulate numerical results for the stochastic approach to compare results better. Using Figures 5 and 6, we have illustrated the disparities between deterministic and stochastic models and their implications for managerial decision-making in uncertain situations.

In the initial exploration, Figure 5 captures the essence of demand dynamics under deterministic and stochastic circumstances. The deterministic line, depicted by a dotted black line, represents a scenario where demand is predictable and follows a predefined pattern. In contrast, stochastic scenarios introduce variability, depicted through fluctuating demand graphs under various distributions. The numerical formulation of demand incorporates distributions such as normal, uniform, triangular, and exponential, simulating market conditions with different levels of unpredictability. This formulation allows us to visually notice how demand evolves when subject to

varying degrees of uncertainty. The result is a series of demand scenarios that reflect the potential variability inherent in real-world markets.

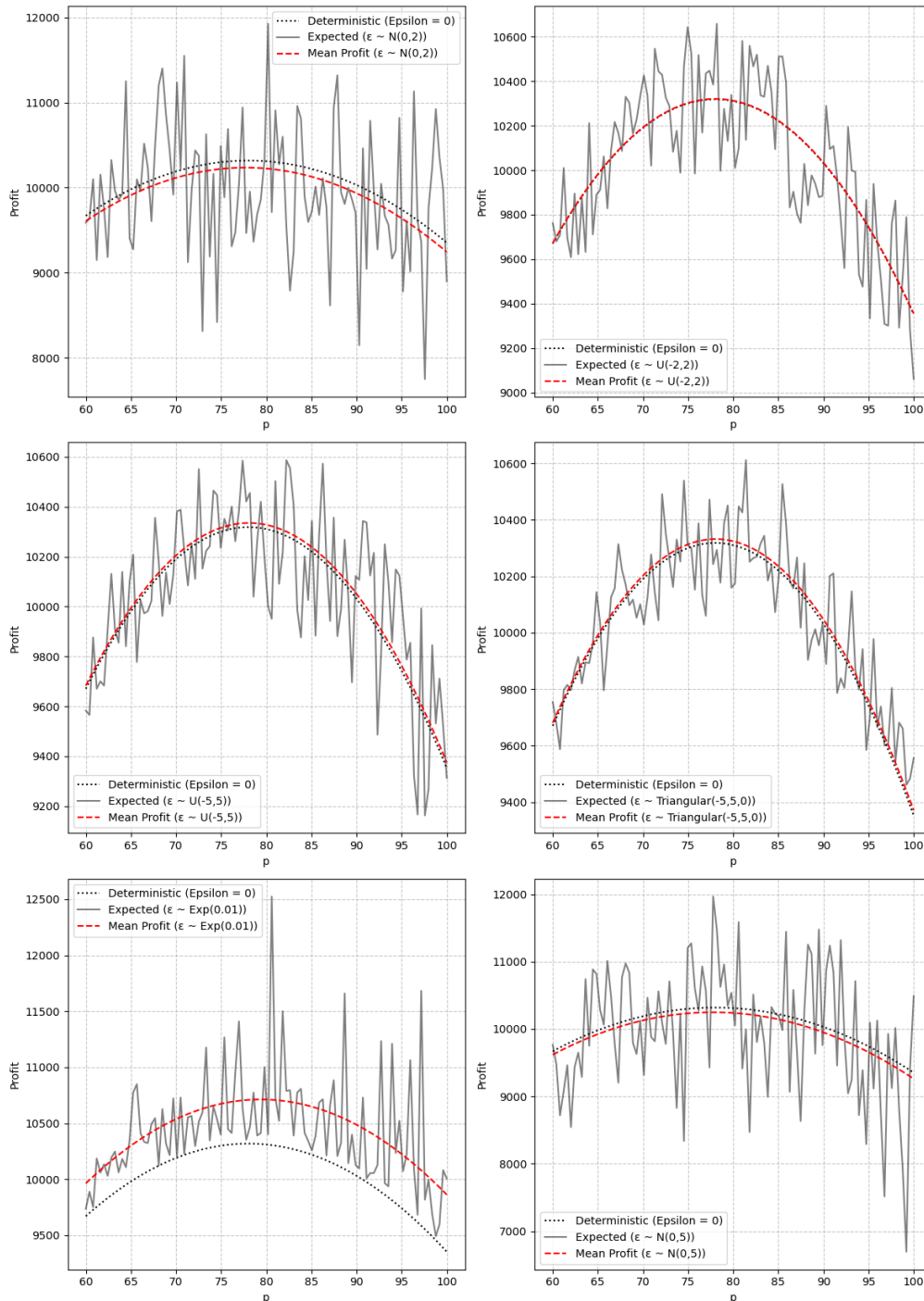


Figure 6: Profit under the uncertainty

The profit formulation integrates normal, uniform, triangular, and exponential distributions to simulate the impact of unpredictable market dynamics on profitability. The accompanying mean profit lines offer a glimpse into the expected profitability under stochastic conditions. Here, the interplay between deterministic and stochastic trends becomes apparent, illustrating how market uncertainties can significantly affect overall profitability.

The numerical exploration provides a foundation for understanding the managerial implications of deterministic and stochastic models in uncertain environments. The deterministic approach, while straightforward and easy to implement, may need to be revised when faced with the unpredictable nature of real-world markets. The illustrated figures provide evidence that deterministic models have the potential to oversimplify demand and profit scenarios, thereby leading to erroneous decisions. In contrast, the stochastic models offer a more nuanced perspective, acknowledging and embracing uncertainty. This acknowledgement is crucial for managerial decision-making in unpredictable environments. Managers armed with stochastic insights can anticipate a spectrum of possible outcomes and strategically plan for contingencies. The superiority of stochastic models logically unfolds through the comparison of deterministic and stochastic trends. In Figure 5, the deterministic line represents a singular path, unable to capture the diverse and fluctuating nature of market demand. The stochastic demand scenarios, on the other hand, reflect the inherent variability in market dynamics, allowing for a more comprehensive understanding.

Figure 6 reinforces this logic by illustrating the rigid nature of deterministic profit trends contrasted with the dynamic and adaptable nature of stochastic profitability. The mean profit lines in stochastic scenarios serve as beacons, guiding managers toward a more informed and resilient decision-making process. In uncertain environments, where market conditions are subject to change, the deterministic approach may lead to missed opportunities or unexpected challenges. Stochastic models, by accommodating variability, empower managers to make decisions that align with the complex reality of supply and demand fluctuations. The numerical exploration of deterministic and stochastic models in inventory management provides valuable insights for managerial decision-making. The visual representations in Figures 5 and 6 underscore the limitations of deterministic approaches in handling uncertainties compared to stochastic models' more adaptable and realistic nature. Managerial implications highlight the importance of embracing uncertainty and leveraging stochastic insights to navigate unpredictable market conditions effectively. The logical illustration of the superiority of stochastic models emphasizes their capacity to capture the dynamic nature of demand and profitability, offering a strategic advantage in decision-making.

As businesses operate in an increasingly complex and uncertain global landscape, adopting stochastic models becomes imperative for those seeking resilience, adaptability, and optimized decision outcomes. The numerical results presented here guide managers, encouraging them to explore and implement stochastic approaches in their quest for effective and agile inventory management strategies.

## VII. Conclusion

The study conducted a comparative study between deterministic and stochastic approaches in inventory modelling. The deterministic model was subjected to classical optimization techniques, while the stochastic optimizations were addressed using particle swarm optimization (PSO). The analysis presented above sheds light on the intricacies and implications of these approaches, unveiling valuable insights for inventory management strategies when dealing with uncertainty. The sensitivity analyses conducted on deterministic and stochastic models emphasize the significance of acknowledging uncertainty in inventory dynamics. The deterministic paradigm assumes that demand and other parameters remain constant, resulting in robust predictability. However, this approach needs to be more balanced with the complex nature of real-world markets and may lead to suboptimal decision-making. On the other hand, the stochastic model, which embraces variability in demand, offers a more realistic depiction of market dynamics. There are several insights into this study are given below:

1. Stochastic modelling enables managers to make adaptive decisions and respond to changing market conditions in real time. By contrast, deterministic methods may need to pay more attention to the dynamic nature of demand, putting businesses at a disadvantage.



2. Stochastic modelling captures a wide range of potential outcomes, making it an effective tool for robust contingency planning. Managers can anticipate and plan for uncertainties, reducing the impact of unexpected disruptions on inventory management.
3. Sensitivity analyses have demonstrated the superiority of the stochastic model in optimizing resource allocation. This helps managers efficiently use resources, minimize holding costs, and increase profitability.
4. By quantifying uncertainties, the stochastic model becomes a powerful tool for risk mitigation. Managers can use it to implement proactive risk management strategies, ensuring resilience in the face of unforeseen market fluctuations.
5. The stochastic model provides more realistic performance metrics, enabling managers to evaluate inventory management strategies against dynamic market conditions. This provides a comprehensive understanding of operational effectiveness.

The following are some potential avenues for extending this work, which may help to develop further and advance the research:

1. Integrate machine learning algorithms for a data-driven approach.
2. Incorporate multi-objective optimization techniques.
3. Incorporate real-time market feedback to enhance accuracy.
4. Explore cross-functional collaboration between inventory management and other business units.
5. Investigate the potential of leveraging blockchain for improved inventory visibility and risk management.

This research emphasizes the differences between deterministic and stochastic inventory modelling and offers practical suggestions for managerial decision-making. Businesses dealing with modern supply chains' complex and uncertain landscape would benefit from implementing stochastic approaches, particularly when combined with advanced optimization methods like PSO. The identified managerial implications and proposed future extensions pave the way for a more adaptive, resilient, and technologically advanced approach to inventory management.

#### Declarations

- The authors did not receive support from any organization for the submitted work
- No funding was received to assist with the preparation of this manuscript
- The authors declare no competing interests

#### Data Availability

All data supporting this study are available from the corresponding author upon reasonable request

#### References

- [1] Antic S., Djordjevic M.L., Lisec A., Dynamic discrete inventory control model with deterministic and stochastic demand in pharmaceutical distribution. *Applied Sciences*, 2022, v. 12, no. 3, p. 15-36.
- [2] Boujnah I., Tlili M., Korbaa O., A modified particle swarm optimization algorithm in a rolling horizon framework for the aggregate production planning problem: pharmaceutical industry case. *Annals of Operations Research*, 2024, p. 1-18.
- [3] Chan L. M., Simchi-Levi D., & Swann J., Pricing, Production, and Inventory Policies for Manufacturing with Stochastic Demand and Discretionary Sales. *Manufacturing & Service Operations Management*, 2006, v. 8, no. 2, p. 149-168.
- [4] Chaudhari R.H., Gor A.S., Narsingani F.J., Statistical model for inventory optimization using genetic approach. *AIP Conference Proceedings*, 2023, v. 2728, no. 1.
- [5] Darmawan A., Evaluating proactive and reactive strategies in supply chain network design

- with coordinated inventory control in the presence of disruptions. *Journal of Industrial and Production Engineering*, 2024, p. 1-17.
- [6] Daş G.S., Yeşilkaya M., Birgören B., A two-stage stochastic model for an industrial symbiosis network under uncertain demand. *Applied Mathematical Modelling*, 2024, v. 125, p. 444-462.
- [7] Datta A., Sarkar B., Dey B.K., Sangal I., Yang L., Fan S.K., Sardar S.K., Thangavelu L., The impact of sales effort on a dual-channel dynamical system under a price-sensitive stochastic demand. *Journal of Retailing and Consumer Services*, 2024, v. 76, 103561.
- [8] Desport P., Lardeux F., Lesaint D., Cairano-Gilfedder C. D., Liret A., Owusu G., A combinatorial optimisation approach for closed-loop supply chain inventory planning with deterministic demand. *European journal of industrial engineering*, 2017, v. 11, no. 3, p. 303-327.
- [9] Huatian G., Xiaoguang, Y., A two-stage stochastic programming for the integrated emergency mobility facility allocation and road network design under uncertainty. *Research Square*, 2024.
- [10] Kar S., Bhunia A. K., Maiti M., Deterministic inventory model with two levels of storage, a linear trend in demand and a fixed time horizon. *Computers & Operations Research*, 2001, v. 28, no. 13, p. 1315-1331.
- [11] Khedlekar U.K., Kumar L., Mathematical modelling for convertible items with rework using particle swarm optimization. *International Journal of Systems Science: Operations & Logistics*, 2024, v. 11, no. 1, 2306222.
- [12] Khedlekar U.K., Kumar L., Keswani M., A Stochastic Inventory Model with Price-Sensitive Demand, Restricted Shortage and Promotional Efforts. *Yugoslav Journal of Operations Research*. 2023, v. 33, no. 4, p. 613-642.
- [13] Maiti A.K., Maiti M.K., Maiti M., Inventory model with stochastic lead-time and price dependent demand incorporating advance payment. *Applied Mathematical Modelling*, 2009, v. 33, no. 5, p. 2433-2443.
- [14] Neri M., Guelpa E., Verda V., Two-stage stochastic programming for the design optimization of district cooling networks under demand and cost uncertainty. *Applied Thermal Engineering*, 2024, v. 236, 121594.
- [15] Olivares-Nadal A.V., Constructing decision rules for multiproduct newsvendors: An integrated estimation-and-optimization framework. *European Journal of Operational Research*, 2024.
- [16] Oprocha P., Czyżewska N., Klimczak K., Kusiak J., Morkisz P., Pietrzyk M., Szeliga D., A Comparative Study of Deterministic and Stochastic Models of Microstructure Evolution during Multi-Step Hot Deformation of Steels. *Materials*, 2023, v. 16, no. 9, 3316.
- [17] Purohit S.K., Panigrahi S., Novel deterministic and probabilistic forecasting methods for crude oil price employing optimized deep learning, statistical and hybrid models. *Information Sciences*, 2024, v. 658, 120021.
- [18] Tarim S.A., Kingsman B.G., Modelling and computing  $(R_n, S_n)$  policies for inventory systems with non-stationary stochastic demand. *European Journal of Operational Research*, 2006, v. 174, no. 1, p. 581-599.
- [19] Tsoularis A., Deterministic and stochastic optimal inventory control with logistic stock-dependent demand rate. *International Journal of Mathematics in Operational Research*, 2014, v. 6, no. 1, p. 41-69.
- [20] Yuna F., Erkeyman B., Yilmaz M., Inventory control model for intermittent demand: a comparison of metaheuristics. *Soft Computing*, 2023, v. 27, no. 10, p. 6487-6505.

# SEQUENTIAL TESTING PROCEDURE FOR THE PARAMETERS OF INVERSE DISTRIBUTION FAMILY

K. S. CHAUHAN<sup>1\*</sup>, A. SHARMA<sup>2</sup>



<sup>1,2</sup> Ram Lal Anand College, University of Delhi, New Delhi-110021

<sup>1\*</sup>kuldeepsinghchauhan.stat@rla.du.ac.in, <sup>2</sup> anurag.stats@rla.du.ac.in

## Abstract

*The sequential probability ratio test is a powerful statistical tool that is frequently employed for hypothesis testing, parameter estimation, and statistical inference. The aspect of robustness is of utmost importance when employing SPRTS in practical applications. Past studies have investigated the robustness of SPRTS for specific distributions. We have developed SPRTS for a family of inverse distributions that includes eleven distinct distributions. The primary objective of this study is to investigate and evaluate the robustness of SPRTS under various conditions and distributions, focusing on the parameters of the inverse distribution family. SPRTS efficacy is measured using OC and ASN functions. This study comprehensively covers the construction and rigorous evaluation of SPRTS, particularly in testing simple null hypotheses against simple alternative hypotheses. Additionally, we investigate the robustness of SPRTS under various factors, including the presence of other parameters and specified coefficients of variation. Conclusive results, graphic representations, tables, and acceptance and rejection regions add clarity to the findings.*

**Keywords:** Inverse Distributions Family, Sequential Probability Ratio Tests (SPRT), Operating Characteristics (OC), Average Sample Number (ASN).

## 1. INTRODUCTION

Sequential Probability Ratio Tests (SPRT) are innovative methodologies that prove highly effective for both hypothesis testing and parameter estimation in statistical inference. The foundational work by [21] introduced the concept of SPRT for analyzing simple null hypotheses against simple alternatives. To assess the effectiveness of SPRT, operating characteristic (OC) and average sample number (ASN) functions were developed as performance measures. Sequential probability ratio tests (SPRTS) have long been recognized as valuable tools for making efficient and prompt decisions in various statistical applications. These tests play a crucial role in scenarios where data are collected sequentially over time, and the goal is to make a conclusive determination about a specific hypothesis. Robustness, which ensures the validity and reliability of these tests under varying conditions, is an essential aspect to consider when employing SPRTS in real-world situations. Multiple studies have scrutinized the robustness of SPRTS in disparate scenarios, enhancing our understanding of their performance and versatility. For instance, [1] examined Wald's SPRT for Levy processes, while [3] explored the robustness of sequential testing procedures for generalized life distributions. Other research, such as that by [6] studied the robustness of sequential testing procedures for parameters of zero-truncated negative binomial, binomial and Poisson distributions. Previous works have also assessed the robustness of SPRTS in specific settings, such as [8], considered sequential life tests in the exponential case. [9] examined the robustness of sequential probability ratio tests in the presence of nuisance parameters.[11] evaluated exponential and Weibull test plans, whereas [12] concentrated on investigating the robustness of the SPRT for a negative binomial distribution in cases where the shape parameter

is not specified. Additionally, [13] investigated the robustness of the exponential SPRT when failures from a Weibull distribution were transformed using a known shape parameter. Other relevant research includes [17] discusses the performance analysis of the Sequential Probability Ratio Test (SPRT) under various conditions and [14] explored robustifying the SPRT for a discrete model under "contamination." In contrast, [15] analyzed the performance and robustness of an SPRT for non-identically distributed observations. The robustness of SPRTS has also been examined in the context of exponential life-testing procedures [18] and the scale parameter of gamma and exponential distributions [19]. [16] discusses the use of sequential probability ratio tests (SPRTS) for the statistical analysis of simulation outputs generated by computers. The type I and type II errors exponents of sequential probability ratio tests, when the actual distributions differ from the test distributions analyzed by [2]. In light of these studies, this research aims to investigate further and evaluate the robustness of sequential probability ratio tests under various conditions and distributions. In this study, we aim to extend the existing research and contribute to the robustness analysis of SPRTS for parameters of inverse distribution family suggested by [7]. Our focus will be on thoroughly examining the robustness of these tests using OC and ASN functions. We will develop and rigorously evaluate the SPRTS, with specific attention given to their robustness about the OC and ASN functions. Sections 3 and 5 will cover the essential elements of constructing and evaluating the SPRTS, including testing simple null hypotheses against simple alternatives, sequential analyses of composite hypotheses, and comprehensively examining their robustness. Section 4 shall analyze simple null hypotheses established on the parameter  $\gamma$ , taking into account the presence of the illustrious  $\delta$ . Furthermore, in Section 6, we shall investigate comparable hypotheses founded on the parameter  $\delta$ , factoring in the existence of  $\gamma$ . In Section 7, we will further investigate the robustness of the SPRTS in the presence of a specified coefficient of variation. Section 8 presents the regions of acceptance and rejection deduced for the null hypothesis  $H_0$  compared to the alternative hypothesis  $H_1$ . Finally, Section 9 will effectively explain the synthesized data and provide conclusive findings using a combination of tables and graphics.

Through this comprehensive analysis, we aim to gain valuable insights into the robustness, performance, and limitations of SPRTS in the inverse family of distributions.

## 2. INVERSE DISTRIBUTIONS FAMILY

Suppose a random variable ( $rv$ )  $x$  having p.d.f.

$$f(x; a^{-1}, \gamma, \delta, \theta) = \frac{\gamma^\delta g^{\delta-1}(x^{-1}; \theta) g'(x^{-1}; \theta)}{x^{2\Gamma(\delta)}} \exp(-\gamma g(x^{-1}; \theta)); \quad (1)$$

$$0 < x < a^{-1}, \quad \gamma > 0, \delta > 0.$$

Where,  $g(x^{-1}; \theta)$ , is a function of  $\theta$  and  $x$ . Moreover,  $g(x^{-1}; \theta)$  real-valued, Strict decreasing the function of  $x$  with  $g(\infty; \theta) = \infty$  and  $g'(x^{-1}; \theta)$  stances for the derivative of  $g(x; \theta)$  by  $x^{-1}$ .

the equation (1) shows that the above distribution can be converted in the following distributions as special cases: If  $g(x; \theta) = x^2, \delta = k + 1 (k \geq 0), (k = \frac{-1}{2})$  provide the inverse Half-normal distribution and ( $k = 0$ ) the inverse Rayleigh distribution. If  $g(x; \theta) = \log\left(1 + \frac{x^b}{v^b}\right), b > 0, v > 0, \delta = 1$ , provide the inverse log-logistic model. If  $g(x; \theta) = \log\left(1 + \frac{x^b}{v^b}\right), b > 0, v = 1, \delta > 1$ , provide the inverse Burr distribution. If  $g(x; \theta) = \log\left(1 + \frac{x^b}{v^b}\right), b = 1, v > 1, \delta > 1$ , provide the inverse Lomax distribution. If  $g(x; \theta) = \frac{x^2}{2}, \delta = \frac{h}{2} (h > 0)$ , it becomes inverse Chi-distribution. If  $g(x; \theta) = \log\left(\frac{x}{a}\right)$  and  $\delta = 1$ , obtain inverse Pareto distribution. If  $g(x; \theta) = x^r \exp(ax), r > 0, a > 0, \delta = 1$ , obtain inverse modified Weibull distribution. If  $g(x; \theta) = \mu x + \frac{vx^2}{2}, \gamma = \delta = 1$ , obtain inverse linear exponential distribution. If  $g(x; \theta) = \log x$ , obtain the inverse of the log-gamma distribution. If  $g(x; \theta) = x^p, p > 0, \delta > 0$ , obtained the inverse generalized gamma distribution.

### 3. SPRT FOR EVALUATING THE HYPOTHESES OF $\gamma$

Let a series  $X_1, X_2, \dots$  from (1), assume one needs to assess the simple hypotheses  $H_0 : \gamma = \gamma_0$  as opposed to  $H_1 : \gamma = \gamma_1 (> \gamma_0)$ . The analysis of SPRT on behalf of  $H_0$ , expressed in this manner

$$Z_i = \ln \left\{ \frac{f(X_i; a, \gamma_1, \delta, \underline{\theta})}{f(X_i; a, \gamma_0, \delta, \underline{\theta})} \right\} = \delta \ln \left( \frac{\gamma_1}{\gamma_0} \right) - g(x_i^{-1}; \underline{\theta}) (\gamma_1 - \gamma_0) \tag{2}$$

Admit  $H_0$  if  $\sum_{i=1}^n Z_i \leq \ln B$ , refuse  $H_0$  if  $\sum_{i=1}^n Z_i \geq \ln A$ , or else, carry on sampling using the value of  $(n + 1)^{\text{th}}$ . If  $\alpha$  and  $\beta$  belong to the interval  $(0, 1)$  and represent type I and type II errors sequentially, the work by [21] provides definitions for  $A$  and  $B$  that are specified as

$$A \cong \frac{(1 - \beta)}{\alpha}$$

and

$$B \cong \frac{\beta}{(1 - \alpha)}$$

Where  $0 < B < 1 < A$

The OC function is almost specified as

$$L(\gamma) \cong \frac{(A^{t_0} - 1)}{(A^{t_0} - B^{t_0})}$$

Where  $t_0$  is the non-zero result for equation

$$E(e^{t_0 z_i}) = 1 \tag{3}$$

Note 1: Use the statement that  $g(x^{-1}; \theta)$  follows gamma distribution Using (1) with (3), we find

$$\left( \frac{\gamma_1}{\gamma_0} \right)^{\delta t_0} \left\{ \frac{t_0 (\gamma_1 - \gamma_0) + \gamma}{\gamma} \right\}^{-\delta} = 1$$

or,

$$\gamma = \frac{t_0 (\gamma_1 - \gamma_0)}{\left( \frac{\gamma_1}{\gamma_0} \right)^{t_0} - 1} \tag{4}$$

To find the values of OC and ASN functions, evaluate (4) as

$$t_0 \ln \left( \frac{\gamma_1}{\gamma_0} \right) = \ln \left[ 1 + t_0 \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right) \right] \tag{5}$$

By utilizing the natural logarithm function of  $(1 + x)$ , which is defined for  $1 < x < 1$ , in (5).we can achieve the desired outcome from (6).

$$\left\{ \frac{1}{3} \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right)^3 \right\} t_0^2 - \left\{ \frac{1}{2} \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right)^2 \right\} t_0 + \left\{ \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right) - \ln \left( \frac{\gamma_1}{\gamma_0} \right) \right\} = 0 \tag{6}$$

Using (2), provides that

$$E(Z_i | \gamma) = \delta \left[ \ln \left( \frac{\gamma_1}{\gamma_0} \right) - \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right) \right] \tag{7}$$

Using (7), we get, the ASN function

$$E(N | \gamma) \cong \frac{L(\gamma) \ln B + \{1 - L(\gamma)\} \ln A}{\delta \left[ \ln \left( \frac{\gamma_1}{\gamma_0} \right) - \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right) \right]} \tag{8}$$

Using (8) the ASN function for  $H_0$  along with  $H_1$  specified as

$$E_0(N) \cong \frac{(1 - \alpha) \ln B + \alpha \ln A}{\delta \left[ \ln \left( \frac{\gamma_1}{\gamma_0} \right) - \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right) \right]}$$

and

$$E_1(N) \cong \frac{\beta \ln B + (1 - \beta) \ln A}{\delta \left[ \ln \left( \frac{\gamma_1}{\gamma_0} \right) - \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right) \right]}$$

#### 4. SPRT FOR EVALUATING THE HYPOTHESES OF $\gamma$ ALTHOUGH $\delta$ IS CHANGING

Using section (3), The maximum value of ASN gets on behalf of  $\gamma = \tilde{\gamma}$  where  $\tilde{\gamma}$  is getting from  $E(Z_i | \gamma) = 0$  and the maximum value is specified as

$$E_{\tilde{\gamma}}(N) \cong - \frac{(\ln A * \ln B)}{E(Z_i^2 | \tilde{\gamma})} \tag{9}$$

Also

$$\tilde{\gamma} = \left\{ \begin{array}{l} \gamma_1 - \gamma_0 \\ \ln \left( \frac{\gamma_1}{\gamma_0} \right) \end{array} \right\} \tag{10}$$

Also, using (7) we get

$$E(Z_i^2 | \tilde{\gamma}) = \delta \left[ \ln \left( \frac{\gamma_1}{\gamma_0} \right) - \left( \frac{\gamma_1 - \gamma_0}{\tilde{\gamma}} \right) \right]^2 + \frac{(\gamma_1 - \gamma_0)^2 \delta}{\tilde{\gamma}^2} \tag{11}$$

Utilizing (9) and (11), we find that

$$E_{\tilde{\gamma}}(N) \cong \frac{-(\ln A * \ln B)}{\left\{ \delta \ln \left( \frac{\gamma_1}{\gamma_0} \right) - \frac{(\gamma_1 - \gamma_0) \delta}{\tilde{\gamma}} \right\}^2 + \frac{(\gamma_1 - \gamma_0)^2 \delta}{\tilde{\gamma}^2}}$$

Assuming that there has been a modification to the parameter  $\delta$  and that (1) has transformed into  $f(x; a, \gamma, d, \theta)$ , this can be attained by replacing  $\delta$  with  $d$ . To analyze the robustness of SPRT, suggest  $t_0$  as the result of the equation

$$\int_0^{a^{-1}} \left\{ \frac{f(x_i; a, \gamma_1, \delta, \theta)}{f(x_i; a, \gamma_0, \delta, \theta)} \right\}^{t_0} f(x_i; a, \gamma, d, \theta) dx_i = 1 \tag{12}$$

We achieve from (12) and put  $\phi_1 = \left( \frac{\delta}{d} \right)$

$$\left( \frac{\gamma_1}{\gamma_0} \right)^{\delta t_0} \frac{\gamma^d}{\Gamma(d)} \int_0^{a^{-1}} \exp \left[ - \{ (\gamma_1 - \gamma_0) t_0 + \gamma \} g(x_i^{-1}; \theta) \right] \frac{g^{d-1}(x_i^{-1}; \theta) g'(x_i^{-1}; \theta)}{x_i^2} dx_i = 1,$$

or,

$$(\gamma_1 - \gamma_0) \frac{t_0}{\gamma} + 1 = \left( \frac{\gamma_1}{\gamma_0} \right)^{\frac{\delta t_0}{d}}$$

or,

$$\gamma = \frac{(\gamma_1 - \gamma_0) t_0}{\left( \frac{\gamma_1}{\gamma_0} \right)^{\phi_1 t_0} - 1} \tag{13}$$

To find the values of OC functions, evaluate (13) as

$$\phi_1 t_0 \ln \left( \frac{\gamma_1}{\gamma_0} \right) = \ln \left[ 1 + t_0 \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right) \right] \tag{14}$$

Equation (14), Solve as (5) and find the roots of  $t_0$  from (15)

$$\left\{ \frac{1}{3} \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right)^3 \right\} t_0^2 - \left\{ \frac{1}{2} \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right)^2 \right\} t_0 + \left\{ \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right) - \phi_1 \ln \left( \frac{\gamma_1}{\gamma_0} \right) \right\} = 0 \quad (15)$$

where  $\phi_1 = \left( \frac{\delta}{a} \right)$ . The ASN function coincides with (8)

$$E(Z_i | \gamma) = \phi_1 \left[ \ln \left( \frac{\gamma_1}{\gamma_0} \right) - \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right) \right] \quad (16)$$

### 5. SPRT FOR EVALUATING THE HYPOTHESES OF $\delta$

Suppose taking a sequence  $X_1, X_2, \dots$  from (1) are independently and identically distributed. To analyze the simple null hypotheses in contradiction of the simple alternative hypotheses when  $\gamma$  is identified.  $H_0 : \delta = \delta_0$  as opposed to  $H_1 : \delta = \delta_1 (> \delta_0)$ .

We suggest the resulting SPRT

$$Z_i = (\delta_1 - \delta_0) \ln \gamma + (\delta_1 - \delta_0) \ln \left\{ g \left( x_i^{-1}; \theta \right) \right\} + \ln \left( \frac{\Gamma(\delta_0)}{\Gamma(\delta_1)} \right) \quad (17)$$

Admit  $H_0$  on the  $n^{\text{th}}$  step, if

$$\sum_{i=1}^n \ln \left\{ g \left( x_i^{-1}; \theta \right) \right\} \leq \left\{ \ln B - n (\delta_1 - \delta_0) \ln \gamma - n \ln \left( \frac{\Gamma(\delta_0)}{\Gamma(\delta_1)} \right) \right\} / (\delta_1 - \delta_0) \quad (18)$$

Reject  $H_0$  if

$$\sum_{i=1}^n \ln \left\{ g \left( x_i^{-1}; \theta \right) \right\} \geq \left\{ \ln A - n (\delta_1 - \delta_0) \ln \gamma - n \ln \left( \frac{\Gamma(\delta_0)}{\Gamma(\delta_1)} \right) \right\} / (\delta_1 - \delta_0) \quad (19)$$

Then using the  $(n + 1)^{\text{th}}$  value carry on sampling if

$$\frac{\left\{ \ln B - n (\delta_1 - \delta_0) \ln \gamma - n \ln \left( \frac{\Gamma(\delta_0)}{\Gamma(\delta_1)} \right) \right\}}{(\delta_1 - \delta_0)} < \sum_{i=1}^n \ln \left\{ g \left( x_i^{-1}; \theta \right) \right\} < \frac{\left\{ \ln A - n (\delta_1 - \delta_0) \ln \gamma - n \ln \left( \frac{\Gamma(\delta_0)}{\Gamma(\delta_1)} \right) \right\}}{(\delta_1 - \delta_0)} \quad (20)$$

The OC function,  $A$  and  $B$  same as previously.

$$L(\delta) \cong \frac{(A^{t_0} - 1)}{(A^{t_0} - B^{t_0})} \quad (21)$$

Here  $t_0$  is the positive as well as negative but not zero

$$E \left\{ e^{t_0 Z_i} \right\} = 1. \quad (22)$$

Using Note 1 with (22), we get

$$\left\{ \frac{\Gamma(t_0(\delta_1 - \delta_0) + \delta)}{\Gamma(\delta)} \right\} = \left( \frac{\Gamma(\delta_1)}{\Gamma(\delta_0)} \right)^{t_0}. \quad (23)$$

Taking the logarithm of both sides of (23), with  $\ln(1 + x); -1 < x < 1$

$$\ln \Gamma(x) = \ln \sqrt{2\pi} - x + \left( x - \frac{1}{2} \right) \ln x \quad (24)$$

By using the equation (24) of approximation, we get

$$\frac{t_0^2}{6} \left( \frac{\delta_1 - \delta_0}{\delta} \right)^3 (\delta + 1) - \frac{t_0}{4} \left( \frac{\delta_1 - \delta_0}{\delta} \right)^2 (2\delta + 1) - \left( \delta_0 - \frac{1}{2} \right) \ln \delta_0 + \left( \delta_1 - \frac{1}{2} \right) \ln \delta_1 - \left( 1 + \ln \delta - \frac{1}{2\delta} \right) (\delta_1 - \delta_0) = 0 \tag{25}$$

Simplifying terms up to the third degree in  $t_0$ , we get the roots of  $t_0$  from (25).

$$E \left\{ \ln(g(X_i^{-1}; \theta)) \right\} = \frac{\gamma^\delta}{\Gamma(\delta)} \int_0^\infty (\ln x) x^{\delta-1} e^{-\gamma x} dx \tag{26}$$

We achieved, using [10], that

$$E \left\{ \ln(g(X_i^{-1}; \theta)) \right\} = \{ \psi(\delta) - \ln \gamma \}, \tag{27}$$

And  $\psi(\delta)$  is specified as

$$\psi(\delta) = \frac{d}{d(\delta)} \ln \Gamma(\delta)$$

Using (7) and (26), we find

$$E(Z_i | \delta) = [\ln \{ \Gamma(\delta_0) \} - \ln \{ \Gamma(\delta_1) \}] + (\delta_1 - \delta_0) \psi(\delta) \tag{28}$$

The ASN function for  $H_0$  and  $H_1$  using (22) and (27) are specified as

$$E_0(N) \cong \frac{(1 - \alpha) \ln B + \alpha \ln A}{\{ \ln(\Gamma(\delta_0)) - \ln(\Gamma(\delta_1)) \} + (\delta_1 - \delta_0) \psi(\delta)} \tag{29}$$

and

$$E_1(N) \cong \frac{\beta \ln B + (1 - \beta) \ln A}{\{ \ln(\Gamma(\delta_0)) - \ln(\Gamma(\delta_1)) \} + (\delta_1 - \delta_0) \psi(\delta)} \tag{30}$$

## 6. SPRT FOR EVALUATING THE HYPOTHESES OF $\delta$ ALTHOUGH $\gamma$ IS CHANGING

Using Section (5), The greatest value of ASN attained for  $\delta = \tilde{\delta}$ , where  $\tilde{\delta}$  is the result of  $E(Z_i | \delta) = 0$

$$\psi(\tilde{\delta}) = \frac{\{ \ln \Gamma(\delta_1) - \ln \Gamma(\delta_0) \}}{(\delta_1 - \delta_0)}$$

This gives the highest worth as

$$E_\delta(N) \cong - \frac{(\ln A * \ln B)}{E(Z_i^2 | \tilde{\delta})}$$

Using (17) and [10], we get

$$E(Z_i^2 | \tilde{\delta}) = \{ \ln(\Gamma(\delta_0) / \Gamma(\delta_1)) \}^2 + (\delta_1 - \delta_0)^2 \{ (\psi(\tilde{\delta}))^2 + \zeta(2, \tilde{\delta} - 1) \}$$

Where  $\zeta(z, q)$  is specified as

$$\zeta(z, q) = \sum_{n=0}^\infty \left( \frac{1}{(q+n)^2} \right)$$

Where  $t_0$  is the solution of the equation

$$\int_0^{a^{-1}} \left\{ \frac{f(x_i; a, \gamma, \delta_1, \theta)}{f(x_i; a, \gamma, \delta_0, \theta)} \right\}^{t_0} f(x_i; a, \eta, \delta, \theta) dx_i = 1. \tag{31}$$



We achieve this using (17) and (31),

$$\gamma^{(\delta_1 - \delta_0)t_0} \left\{ \frac{\Gamma(\delta_0)}{\Gamma(\delta_1)} \right\}^{t_0} \frac{\eta^\delta}{\Gamma(\delta)} \int_0^{a^{-1}} \frac{g^{(\delta_1 - \delta_0)h + \delta - 1}(x_i^{-1}; \theta) g'(x_i^{-1}; \theta) \exp(-\eta g(x_i^{-1}; \theta)) dx_i}{x_i^2} = 1$$

or,

$$\phi_2^{(\delta_1 - \delta_0)t_0} \left\{ \frac{\Gamma(\delta_0)}{\Gamma(\delta_1)} \right\}^{t_0} \frac{\Gamma((\delta_1 - \delta_0)t_0 + \delta)}{\Gamma(\delta)} = 1 \tag{32}$$

Where  $\phi_2 = \frac{\gamma}{\eta}$ .

By applying the logarithm function to both sides of the equation (32), and employing the approximation (24), the solutions for the variable  $t_0$  are obtained from the following equation,

$$\begin{aligned} \frac{t_0^2}{6} \left( \frac{\delta_1 - \delta_0}{\delta} \right)^3 (\delta + 1) - \frac{t_0}{4} \left( \frac{\delta_1 - \delta_0}{\delta} \right)^2 (2\delta + 1) - \left( \delta_0 - \frac{1}{2} \right) \ln \delta_0 + \left( \delta_1 - \frac{1}{2} \right) \ln \delta_1 \\ - (\delta_1 - \delta_0) \ln \phi_2 - \left( 1 + \ln \delta - \frac{1}{2\delta} \right) (\delta_1 - \delta_0) = 0 \end{aligned} \tag{33}$$

The ASN function coincides with (8),

$$E(Z_i | \delta) = \ln \left\{ \frac{\Gamma(\delta_0)}{\Gamma(\delta_1)} \right\} + (\delta_1 - \delta_0) \Gamma(\delta) + (\delta_1 - \delta_0) \ln \phi_2. \tag{34}$$

### 7. SPRT ROBUSTNESS FOR $\gamma$ WITH INDICATED COEFFICIENT OF VARIATION

If  $g(x; \theta) = \frac{x^2}{2}, \delta = \frac{h}{2} (h > 0)$  in (1), the values of  $\mu = \frac{h}{h-2}$ , for  $h > 2$  and  $\sigma^2 = \frac{2h^2}{(h-2)^2(h-4)}$ , for  $h > 4$ . Then, the coefficient of variation (CV)

$$C = \sqrt{\frac{2}{(h-4)}} \tag{35}$$

Assume that the value of the coefficient of variation alters from  $c$  to  $c^*$ , then  $\delta$  becomes

$$\delta^* = \frac{1}{C^{*2}} + 2 \tag{36}$$

The OC function is

$$\psi_1 t_0 \ln \left( \frac{\gamma_1}{\gamma_0} \right) = \ln \left[ 1 + t_0 \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right) \right] \tag{37}$$

Solve (37) as (5) up to the third degree in  $t_0$  and find the roots of  $t_0$  from (39)

$$\left\{ \frac{1}{3} \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right)^3 \right\} t_0^3 - \left\{ \frac{1}{2} \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right)^2 \right\} t_0^2 + \left\{ \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right) - \psi_1 \ln \left( \frac{\gamma_1}{\gamma_0} \right) \right\} t_0 = 0 \tag{38}$$

where  $\psi_1 = \left( \frac{\delta}{\delta^*} \right)$ .

The ASN function coincides with (8)

$$E(Z_i | \gamma) = \psi_1 \left[ \ln \left( \frac{\gamma_1}{\gamma_0} \right) - \left( \frac{\gamma_1 - \gamma_0}{\gamma} \right) \right] \tag{39}$$

### 8. ACCEPTANCE AND REJECTION REGION

we need to assess the simple hypotheses  $H_0 : \gamma = \gamma_0$  as opposed to  $H_1 : \gamma = \gamma_1 (> \gamma_0)$  having preassigned  $0 < \alpha$  and  $\beta < 1$  then  $Z_i$  is

$$Z_i = \delta \cdot \ln \left( \frac{\gamma_1}{\gamma_0} \right) - g \left( x_i^{-1}; \theta \right) (\gamma_1 - \gamma_0) \tag{40}$$

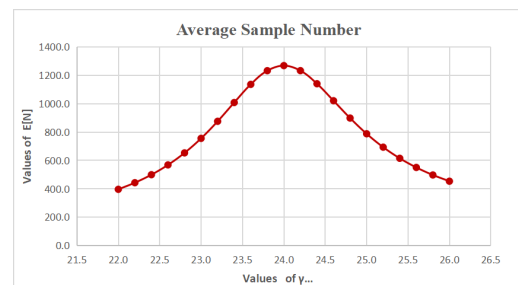
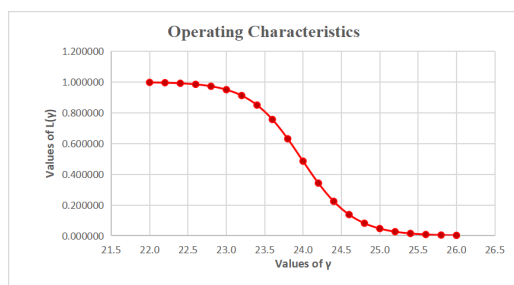
Define,  $Z(N) = \sum_{i=1}^n X_i$  and  $N =$  initial integer  $n(\geq 1)$ , so that the inequality is defined as  $Z(N) \leq c_1 + dn$  or  $Z(N) \geq c_2 + dn$  valid among the parameters.

$$c_1 = \frac{\ln B}{(\gamma_1 - \gamma_0)}, c_2 = \frac{\ln A}{(\gamma_1 - \gamma_0)} \text{ and } d = \frac{\delta \ln \left( \frac{\gamma_0}{\gamma_1} \right)}{(\gamma_1 - \gamma_0)}$$

### 9. RESULT AND DISCUSSION

**Table 1:**  $H_0 : \gamma_0 = 22, H_1 : \gamma_1 = 26 \quad H_0 : \delta_0 = 22, H_1 : \delta_1 = 26$

$\gamma$	$L(\gamma)$	$E[N]$	$\delta$	$L(\delta)$	$E[N]$
22.0	0.997848	396.3	22.0	0.997500	16.82
22.2	0.995846	442.9	22.2	0.995382	18.70
22.4	0.992101	499.5	22.4	0.991517	20.98
22.6	0.985191	568.6	22.6	0.984524	23.77
22.8	0.972657	653.2	22.8	0.972019	27.16
23.0	0.950427	755.6	23.0	0.950054	31.27
23.2	0.912296	875.9	23.2	0.912590	36.08
23.4	0.850178	1008.3	23.4	0.851663	41.38
23.6	0.756664	1136.4	23.6	0.759744	46.52
23.8	0.631008	1232.9	23.8	0.635534	50.40
24.0	0.485370	1268.9	24.0	0.490420	51.83
24.2	0.342685	1233.2	24.2	0.347054	50.32
24.4	0.224024	1140.9	24.4	0.227029	46.47
24.6	0.138008	1021.1	24.6	0.139693	41.46
24.8	0.081636	898.8	24.8	0.082409	36.35
25.0	0.047077	788.0	25.0	0.047344	31.74
25.2	0.026743	693.5	25.2	0.026777	27.81
25.4	0.015062	615.1	25.4	0.015011	24.55
25.6	0.008442	550.6	25.6	0.008374	21.88
25.8	0.004719	497.5	25.8	0.004660	19.69
26.0	0.002634	453.5	26.0	0.002590	17.87



**Figure 1:** OC and ASN Curve for section 3.

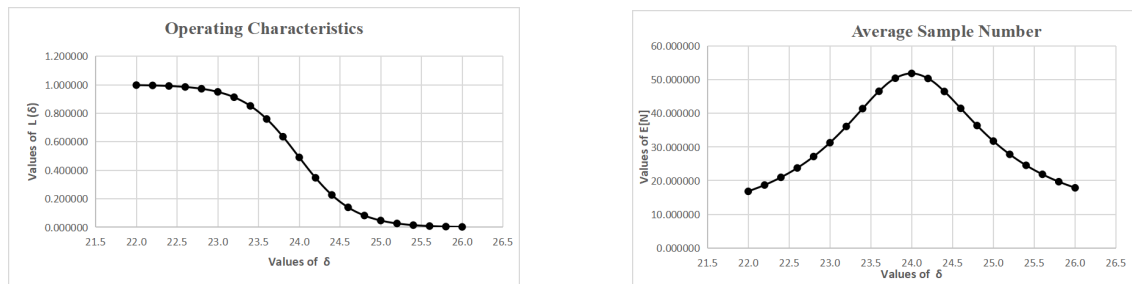


Figure 2: OC and ASN Curve for section 5.

I. The values denoted by the OC and ASN functions for sections 3 and 5 under  $\alpha = \beta = 0.05$ , corresponding to the parameters  $\gamma$  and  $\delta$  can be found in Table 1, while the visuals representing these values are illustrated in Figures 1 and 2. The table mentioned above and curves yield outcomes that are deemed acceptable.

Table 2: OC and ASN Functions for section 4, under  $\alpha = \beta = 0.05$ , where  $H_0 : \gamma_0 = 22, H_1 : \gamma_1 = 26$

$\gamma$	$\phi_1 = 0.95$		$\phi_1 = 0.98$		$\phi_1 = 1$		$\phi_1 = 1.02$		$\phi_1 = 1.05$	
	$L(\gamma)$	$E[N]$	$L(\gamma)$	$E[N]$	$L(\gamma)$	$E[N]$	$L(\gamma)$	$E[N]$	$L(\gamma)$	$E[N]$
22.0	0.999977	256.109	0.999593	325.533	0.997848	396.275	0.990174	501.269	0.925563	760.351
22.2	0.999949	275.377	0.999182	357.072	0.995846	442.881	0.981647	572.664	0.871228	878.478
22.4	0.999891	297.332	0.998388	394.382	0.992101	499.451	0.966278	659.846	0.787253	998.243
22.6	0.999773	322.565	0.996875	439.007	0.985191	568.593	0.939285	764.602	0.670219	1097.251
22.8	0.999539	351.837	0.994031	492.942	0.972657	653.184	0.893689	885.460	0.528271	1148.101
23.0	0.999084	386.143	0.988758	558.689	0.950427	755.616	0.821178	1013.745	0.382212	1134.421
23.2	0.998208	426.784	0.979124	639.166	0.912296	875.930	0.715853	1129.741	0.255118	1063.405
23.4	0.996548	475.461	0.961857	737.241	0.850178	1008.278	0.581040	1204.816	0.159600	959.524
23.6	0.993440	534.365	0.931752	854.390	0.756664	1136.436	0.433599	1215.202	0.095382	847.741
23.8	0.987698	606.223	0.881433	987.837	0.631008	1232.898	0.297492	1159.170	0.055358	743.688
24.0	0.977243	694.173	0.802653	1126.039	0.485370	1268.878	0.190055	1058.028	0.031567	653.709
24.2	0.958580	801.169	0.690763	1245.161	0.342685	1233.182	0.115203	939.908	0.017819	578.588
24.4	0.926211	928.388	0.551685	1313.657	0.224024	1140.910	0.067448	825.361	0.010001	516.709
24.6	0.872518	1071.916	0.404611	1309.361	0.138008	1021.076	0.038658	724.251	0.005595	465.802
24.8	0.789395	1217.829	0.273264	1236.136	0.081636	898.796	0.021884	639.054	0.003124	423.701
25.0	0.673200	1339.075	0.172445	1120.253	0.047077	788.025	0.012302	568.673	0.001742	388.586
25.2	0.531691	1402.143	0.103641	991.440	0.026743	693.530	0.006888	510.809	0.000971	359.005
25.4	0.385428	1386.471	0.060349	869.573	0.015062	615.103	0.003848	463.075	0.000540	333.831
25.6	0.257639	1300.211	0.034475	763.352	0.008442	550.600	0.002147	423.395	0.000301	312.195
25.8	0.161310	1173.099	0.019477	674.378	0.004719	497.491	0.001197	390.091	0.000167	293.426
26.0	0.096429	1035.896	0.010936	601.031	0.002634	453.477	0.000666	361.853	0.000093	277.004

II. Figure 3 illustrates the numerical values of the OC and ASN curves extracted from Table 2, corresponding to different  $\phi_1$  values. When  $\phi_1 < 1$  ( $\phi_1 > 1$ ), the OC curve shifts either towards the right or left direction, while the ASN curve shifts towards the upper right or lower left direction. Both curves demonstrate that the SPRT exhibits a high degree of sensitivity towards alterations in  $\delta$ .

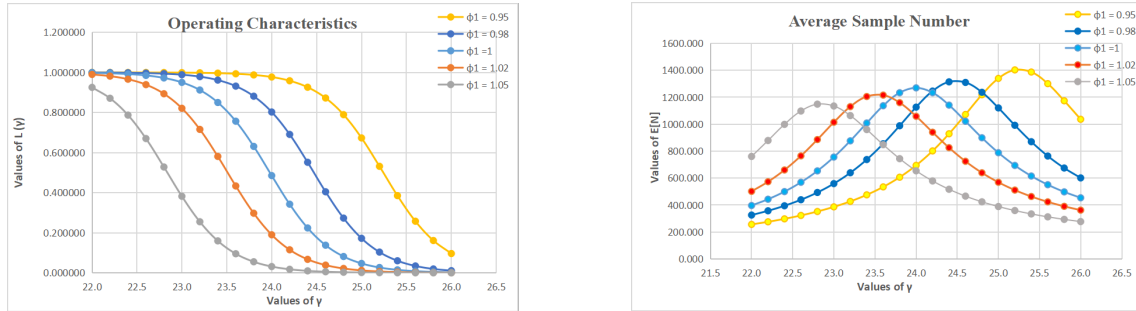


Figure 3: OC and ASN Curve for section 4.

Table 3: OC and ASN Functions for section 6, under  $\alpha = \beta = 0.05$  where  $H_0 : \delta_0 = 22, H_1 : \delta_1 = 26$

$\delta$	$\phi_2 = 0.95$		$\phi_2 = 0.99$		$\phi_2 = 1$		$\phi_2 = 1.02$		$\phi_2 = 1.05$	
	$L(\delta)$	$E[N]$	$L(\delta)$	$E[N]$	$L(\delta)$	$E[N]$	$L(\delta)$	$E[N]$	$L(\delta)$	$E[N]$
22.0	0.999949	12.805	0.998800	16.277	0.997500	16.821	0.989824	25.063	0.930141	38.018
22.2	0.999903	13.769	0.997769	17.854	0.995382	18.704	0.981471	28.633	0.879707	43.924
22.4	0.999816	14.867	0.995875	19.719	0.991517	20.984	0.966597	32.992	0.800964	49.912
22.6	0.999653	16.128	0.992418	21.950	0.984524	23.765	0.940663	38.230	0.689215	54.863
22.8	0.999350	17.592	0.986152	24.647	0.972019	27.161	0.896956	44.273	0.550252	57.405
23.0	0.998787	19.307	0.974920	27.934	0.950054	31.265	0.827248	50.687	0.403178	56.721
23.2	0.997751	21.339	0.955114	31.958	0.912590	36.081	0.725144	56.487	0.271799	53.170
23.4	0.995848	23.773	0.921109	36.862	0.851663	41.381	0.592699	60.241	0.171045	47.976
23.6	0.992379	26.718	0.865216	42.719	0.759744	46.521	0.445466	60.760	0.102434	42.387
23.8	0.986097	30.311	0.779487	49.392	0.635534	50.398	0.307359	57.959	0.059402	37.184
24.0	0.974842	34.709	0.660892	56.302	0.490420	51.829	0.196953	52.901	0.033783	32.685
24.2	0.955006	40.058	0.518186	62.258	0.347054	50.320	0.119421	46.995	0.018999	28.929
24.4	0.920963	46.419	0.372612	65.683	0.227029	46.466	0.069790	41.268	0.010619	25.835
24.6	0.865024	53.596	0.247071	65.468	0.139693	41.460	0.039871	36.213	0.005915	23.290
24.8	0.779245	60.891	0.153542	61.807	0.082409	36.355	0.022479	31.953	0.003290	21.185
25.0	0.660606	66.954	0.091170	56.013	0.047344	31.737	0.012580	28.434	0.001828	19.429
25.2	0.517881	70.107	0.052585	49.572	0.026777	27.807	0.007012	25.540	0.001015	17.950
25.4	0.372321	69.324	0.029808	43.479	0.015011	24.553	0.003901	23.154	0.000564	16.692
25.6	0.246826	65.011	0.016731	38.168	0.008374	21.882	0.002167	21.170	0.000313	15.610
25.8	0.153354	58.655	0.009340	33.719	0.004660	19.688	0.001204	19.505	0.000174	14.671
26.0	0.091038	51.795	0.005199	30.052	0.002590	17.872	0.000669	18.093	0.000097	13.850

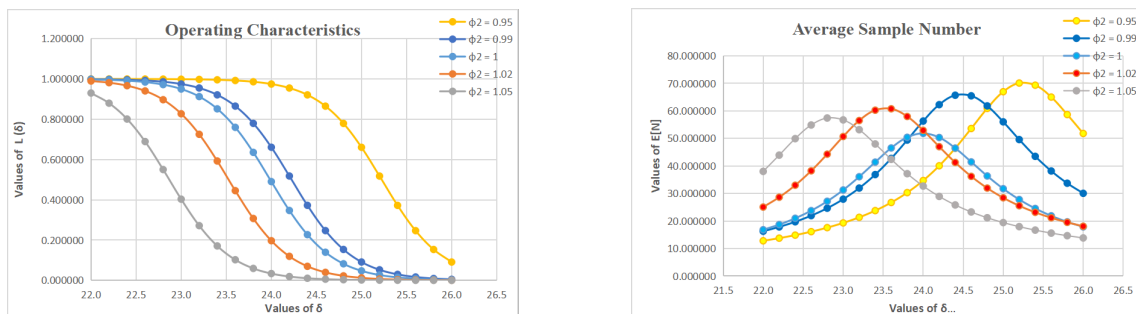


Figure 4: OC and ASN Curve for section 6.

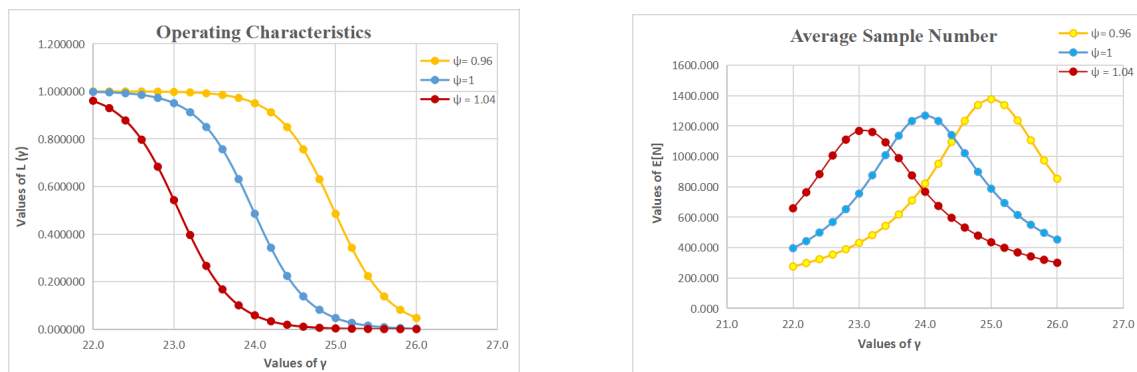
III. Figure 4 portrays the values of the operational characteristic (OC) and average sample num-

ber (ASN) curves derived from Table 3 across various magnitudes of  $\phi_2$ . When  $\phi_2 < 1$  ( $\phi_2 > 1$ ), the OC curve experiences a rightward (leftward) shift, while the ASN curve undergoes an upward rightward (downward leftward) shift. Both curves demonstrate the considerable sensitivity of the sequential probability ratio test (SPRT) to parameter  $\gamma$  alterations.

IV. Figure 5 illustrates the plotted values of the OC and ASN curves obtained from Table 4 while considering different values of ' $\psi$ '. When  $\psi < 1$  ( $\psi > 1$ ) is taken into account, the OC curve experiences a shift towards the right (left), while the ASN curve shifts upwards (downwards) towards the right. It is evident from both curves that the SPRT demonstrates a considerable level of sensitivity towards variations in ' $\psi$ '.

**Table 4:** OC and ASN Functions for section 7, under  $\alpha = \beta = 0.05$  where  $H_0 : \gamma_0 = 22, H_1 : \gamma_1 = 26$

$\gamma$	$\psi = 0.96$		$\psi = 1$		$\psi = 1.04$	
	$L(\gamma)$	$E[N]$	$L(\gamma)$	$E[N]$	$L(\gamma)$	$E[N]$
22.0	0.999936	275.746	0.997848	396.275	0.960659	658.849
22.2	0.999864	298.188	0.995846	442.881	0.929568	764.064
22.4	0.999719	324.055	0.992101	499.451	0.877735	883.456
22.6	0.999432	354.157	0.985191	568.593	0.796964	1006.162
22.8	0.998873	389.555	0.972657	653.184	0.683060	1110.363
23.0	0.997803	431.636	0.950427	755.616	0.542805	1168.026
23.2	0.995782	482.209	0.912296	875.930	0.396046	1160.562
23.4	0.992010	543.586	0.850178	1008.278	0.266293	1092.897
23.6	0.985068	618.594	0.756664	1136.436	0.167506	988.999
23.8	0.972493	710.346	0.631008	1232.898	0.100472	874.913
24.0	0.950216	821.423	0.485370	1268.878	0.058441	767.625
24.2	0.912031	951.852	0.342685	1233.182	0.033369	674.376
24.4	0.849857	1095.264	0.224024	1140.910	0.018849	596.352
24.6	0.756289	1234.025	0.138008	1021.076	0.010583	532.042
24.8	0.630591	1338.272	0.081636	898.796	0.005922	479.150
25.0	0.484941	1376.797	0.047077	788.025	0.003307	435.440
25.2	0.342284	1337.532	0.026743	693.530	0.001844	399.016
25.4	0.223688	1236.956	0.015062	615.103	0.001028	368.364
25.6	0.137752	1106.603	0.008442	550.600	0.000572	342.305
25.8	0.081457	973.726	0.004719	497.491	0.000318	319.929
26.0	0.046959	853.432	0.002634	453.477	0.000177	300.536



**Figure 5:** OC and ASN Curve for section 7.

V. The acceptance and rejection zones for the null hypothesis  $H_0$ , with  $H_0 : \gamma_0 = 22$  and the

alternative hypothesis  $H_0 : \gamma_0 = 26$ . Both the  $\alpha$  and  $\beta$  significance levels are set to 0.05, and the degrees of freedom  $\delta$  are set to 2. The values of the constants  $c_1$ ,  $c_2$ , and  $d$  are -287.0828, 287.0828, and -27.90466, respectively. As a result, if the observed value  $Z(n)$  is less than or equal to  $-27.90466N + 287.0828$ , we accept the null hypothesis  $H_0$ , and we accept the alternative hypothesis  $H_1$  if  $Z(n)$  is higher than or equal to  $-27.90466N - 287.0828$ . In the intermediate stages, the sampling procedure continues.

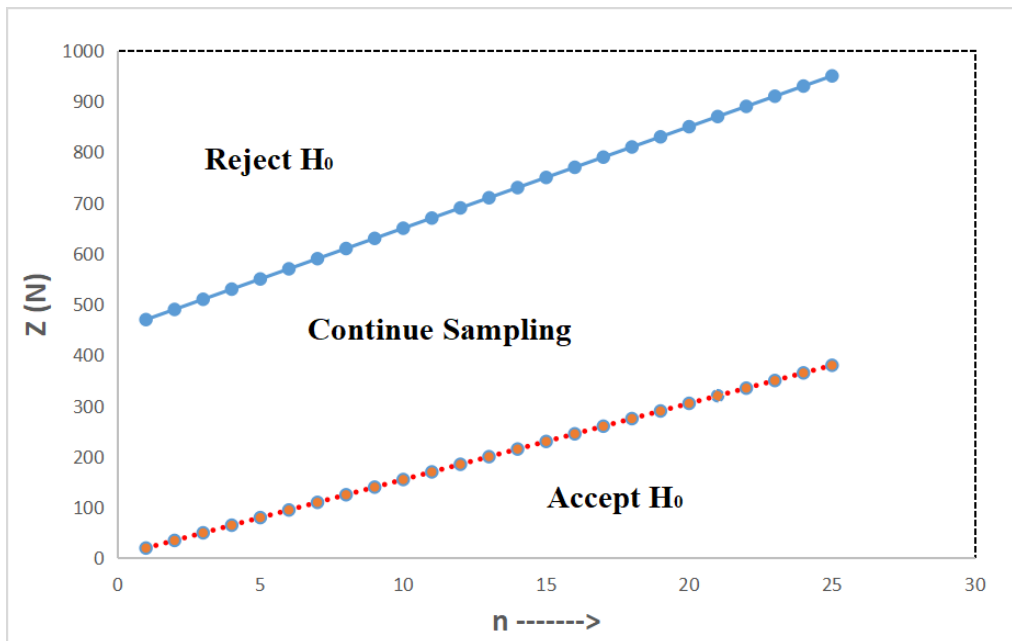


Figure 6: The Acceptance and Rejection zones for  $H_0$

## REFERENCES

- [1] Buonaguidi, B. and Muliere, P.(2013). On the Wald's Sequential Probability Ratio Test for Levy Processes. *Sequential Analysis*, 32(3), 267 - 287.
- [2] Boroumand, P. and Guillen i Fabregas, A. (2021). Error exponent sensitivity of sequential probability ratio testing. *IEEE International Symposium on Information Theory (ISIT)*, 184-189.
- [3] Chaturvedi, A., Tiwari, N. and Tomer, S.K. (2002). Robustness of the sequential testing procedures for the generalized life distributions. *Brazilian Journal of Probability and Statistics*, 16, 7-24.
- [4] Chaturvedi, A., Chauhan, K., and Alam, M. W. (2009). Estimation of the reliability function for a family of lifetime distributions under type I and type II censorings. *Journal of reliability and statistical studies*, 11-30.
- [5] Chaudhary, A. and Chauhan, K. (2009). Estimation and testing procedures for the reliability function of Weibull distribution under type I and type II censoring. *Journal of statistics sciences*, 1(2), 121-136.
- [6] Chaturvedi, A, Chauhan, K., Alam, M. W. (2013). Robustness of the sequential testing procedures for the parameters of zero-truncated negative binomial, binomial, and Poisson distributions. *Journal of the Indian Statistical Association*, 51(2), 313-328.
- [7] Chauhan, K. S. (2022). Estimation and testing procedures of  $P(Y < X)$  for the inverse distributions family under type-II censoring. *Reliability. Reliability: Theory Applications*, 17(3 (69)), 328-339.
- [8] Epstein, B. and Sobel, M. (1955). Sequential life test in exponential case. *The Annals of Mathematical Statistics*, 26, 82-93.

- [9] Eger, K. and Tsoy, E.B. (2008). Robustness of sequential probability ratio tests in case of nuisance parameters. *Third International Forum on Strategic Technologies*, 266-270.
- [10] Gradshteyn, I. S. and Ryzhik, I. M. (2007). *Tables of Integrals Series and Products*, Academic Press, New York.
- [11] Harter, H.L. and Moore, A.H. (1976). An Evaluation of Exponential and Weibull Test Plans. *IEEE Transactions on Reliability*, 25(2), 100-104.
- [12] Hubbard, D. J. and Allen, O.B. (1991). Robustness of the SPRT for a negative binomial to misspecification of the dispersion parameter. *Biometrics*, 47(2), 419-427.
- [13] Hauck, D. J. and Keats, J.B. (1997). Robustness of the exponential sequential probability ratio test (SPRT), when Weibull distributed failures are transformed using a known shape parameter. *Microelectronics Reliability*, 37(12), 1835-1840.
- [14] Kharin, A. Y. (2016). On Robustifying of the Sequential Probability Ratio Test for a Discrete Model under "Contaminations". *Austrian Journal of Statistics*, 31(4), 267-277.
- [15] Kharin, A. Y. and That Tu, T. (2018). Performance analysis and robustness evaluation of a sequential probability ratio test for non-identically distributed observations. *Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics Series* 54(2), 179-192.
- [16] Kleijnen, J. P. C. and Shi, W. (2021). Sequential probability ratio tests: conservative and robust. *SIMULATION*, 97(1), 33-43.
- [17] Liu, Y. and Li, X. R. (2013). Performance Analysis of Sequential Probability Ratio Test. *Sequential Analysis*, 32(4), 469-497.
- [18] Montage, E. R., and Singaurwalla, N. D. (1985). Robustness of the Sequential Exponential Life-Testing Procedures. *Jour. Amer. Statist. Assoc.*, 80, 715-719.
- [19] Pandit, P. V. and Gudaganavar, N. V. (2010). On Robustness of a Sequential Test for Scale Parameter of Gamma and Exponential Distributions. *Applied Mathematics-Journal of Chinese Universities Series B*, 01, 274-278.
- [20] Raghavachari, M. (1965). Operating Characteristic and Expected Sample Size of a Sequential Probability Ratio Test for the Simple Exponential Distribution. *Calcutta Statistical Association Bulletin*, 14, 65 - 73.
- [21] Wald, A. (1947). *Sequential Analysis*. John Wiley and Sons, New York.

# CONFIDENCE INTERVAL USING MAXIMUM LIKELIHOOD ESTIMATION FOR THE PARAMETERS OF POISSON TYPE RAYLEIGH CLASS MODEL

<sup>1</sup>Rajesh Singh, <sup>2</sup>Preeti A. Badge, <sup>3</sup>Pritee Singh

<sup>1,2</sup>Department of Statistics, S. G. B. Amravati University, Amravati, India

<sup>3</sup>Department of Statistics, Institute of Science, Nagpur, India.

[rsinghamt@hotmail.com](mailto:rsinghamt@hotmail.com)

[preetibadge10@gmail.com](mailto:preetibadge10@gmail.com)

[priteesingh25@gmail.com](mailto:priteesingh25@gmail.com)

## Abstract

*In this research paper, confidence interval using maximum likelihood estimation is obtained for Poisson type Rayleigh class for the parameters. The failure intensity function, mean time to failure function and likelihood function for the parameter is derived. Confidence interval has been obtained for the parameters using maximum likelihood estimation. To study the performance of proposed Confidence interval, average length and coverage probability are calculated by using Monte Carlo simulation technique. From the obtained intervals, it is concluded that Confidence interval for the parameters perform better for appropriate choice of execution time and certain values of parameters.*

**Keywords:** Rayleigh distribution, Software reliability growth model, Maximum likelihood estimation (MLE), Average length and coverage probability.

## 1. Introduction

Software reliability is the quality characteristic of operation system which can measure, predict and estimate quality of software system. In last several decades various model have been proposed to assess software reliability. Most of them are probabilistic models. Software modeling techniques can be divided into two categories: Prediction and estimation models. Estimation models determines the current software reliability by applying statistical inference techniques to failure data while the prediction models determines future software reliability based upon available software metrics and measures. The parameters involved in software reliability models can be estimated by using some basic procedures like maximum likelihood, least square estimation and Bayesian point estimation, etc. Among all software reliability models, software reliability growth models are very useful to assess software reliability. With the help of software reliability approach customer fulfill their requirements.

Most of the past research work in software reliability modeling has concentrated on the point estimation of the parameters. The uncertainty of the estimates by using interval estimation has not been fully discussed. The most commonly applied interval estimation technique is based on the central limit theorem assuming large sample size. In real world testing the number of software



failures observed is usually large. Maximum likelihood Estimation is most preferable because of its easy computation and it is suitable for large sample size. The Confidence interval provides two pairs of values based on sample (say an interval) within which the parameter will lie with certain probability

In this research paper Poisson type Rayleigh class model is considered according to Musa and Okumoto [10] classification scheme. The Rayleigh distribution is widely used for communications, physical sciences, medical imaging and engineering, applied statistics and clinical trials. Sinha and Howlader [16] have computed credible and HPD intervals of the parameters of Rayleigh distribution and also estimated credible interval using reliability function. Sinha [17] has estimated Bayesian interval for the parameters of Rayleigh distribution and its reliability function. Hirano [5] has described origin and properties of Rayleigh distribution. Merovci and Elbatal [9] have defined and studied Weibull Rayleigh distribution and its mathematical properties. Roy [13] proposed Discrete Rayleigh distribution for univariate and bivariate situations. Dey et al [2] have described different approaches for the estimation of two parameters Rayleigh distribution and also computed credible intervals. Rao et al [12] have proposed software reliability growth model of inverse Rayleigh distribution to assess the failure process of developed software and estimated the model parameters by maximum likelihood estimation. Fang and Yeh [3] have proposed a software reliability estimation that uses Stochastic differential equations i.e. SDEs with the fault detection function.

Rao and Cunha [1] have estimated credible intervals and confidence intervals through maximum likelihood estimators for lognormal distribution and also compared average length and coverage probability of the calculated interval. Jeske et al [6] have developed the confidence intervals of average failure rate on the basis of asymptotic theory. Fang and Yeh [4] have proposed confidence interval of the software fault detection process of software reliability growth models using stochastic differential equations. Saroj et al [14] have proposed transformed distribution called inverse Muth distribution and obtained asymptotic confidence interval for parameters of distribution in case of maximum likelihood estimation and maximum product spacing estimation (MPSE) is alternative method for MLE. Lalitha and Mishra [7] have obtained modified maximum likelihood estimate of the Rayleigh distribution using hyperbolic approximation. Lee et al [8] used Obha's model to build the SRGM with confidence intervals that can help the software developers to determine the optimal release time in practice. The Rayleigh distribution is widely used for communications, physical sciences, medical imaging and engineering, applied statistics and clinical trials. Seo et al [15] have obtained the exact confidence intervals for unknown parameters and predictive intervals for future upper record values by considering some pivotal quantities in the two parameter Rayleigh distribution. Yamada et al. [18] have proposed software reliability growth models incorporating the quantity of test-effort exhausted on software testing described by the Rayleigh curve and Rayleigh function used to estimate the detection rate of defects as a function of time during the software development process.

The frame of this paper is such that section 2 presents derivation of failure intensity and expected number of failures using Rayleigh distribution. Section 3 presents Likelihood function and construction of Confidence interval of parameters for Rayleigh distribution. Results are discussed in the section 4 while concluding remarks are provided in section 5.

## 2. Model Formulation and Evaluation

Considering that software system has experienced  $m_e$  failures at times  $t_i$ , where  $i = 1, 2, \dots, m_e$  during the execution time  $t_e$ . Let the parameter  $\gamma_0$  be the total number of failures and the second parameter is  $\gamma_1$ . Let 't' be the positive random variable having Rayleigh distribution then its probability density function in terms of 't' is given as

$$f(t) = \begin{cases} t\gamma_1^{-2} e^{-\frac{1}{2}\left[\frac{t}{\gamma_1}\right]^2} & , t > 0, \gamma_1 > 0 \\ 0 & , \text{Otherwise} \end{cases} \quad (1)$$

The failure intensity function of the above model can be given as:

$$\lambda(t) = \gamma_0 t \gamma_1^{-2} e^{-\frac{1}{2}[t\gamma_1^{-1}]^2} \quad , t > 0, \gamma_1 > 0, \gamma_0 > 0 \quad (2)$$

Also, the mean failures function i.e. expected number of failures at time  $t_e$  is given by:

$$\mu(t_e) = \gamma_0 \gamma_1^{-2} \int_0^{t_e} t \gamma_1^{-2} e^{-\frac{1}{2}[t\gamma_1^{-1}]^2} dx$$

After some algebraic simplification, the above equation can be given as:

$$\mu(t_e) = \gamma_0 \left[ 1 - e^{-\frac{1}{2}(t_e \gamma_1^{-1})^2} \right] \quad (3)$$

## 3. Likelihood function and Confidence Interval

Confidence interval is one of the estimation techniques to draw statistical inference. Confidence interval can construct with several different methods. The method of confidence interval constructed through maximum likelihood estimation is discussed here. Likelihood function is significant part of frequentist and Bayesian analysis. It can be used to compare probability of various parameter values. The likelihood function of  $(\gamma_0, \gamma_1)$  is obtained with the help of failure intensity (2) and expected number of failures (3) the likelihood function (see for details Musa et al [11]) and can be expressed as follows:

$$L(\gamma_0, \gamma_1) = \gamma_0^{m_e} \gamma_1^{-2m_e} \left[ \prod_{i=1}^{m_e} t_i \right] e^{-\frac{1}{2}T\gamma_1^{-2}} e^{-\gamma_0} \exp \left\{ \gamma_0 e^{-\frac{1}{2}\left(\frac{t_e}{\gamma_1}\right)^2} \right\} \quad (4)$$

$$\text{Where, } \sum_{i=1}^{m_e} t_i^2 = T$$

Maximum likelihood estimators for the parameters  $\gamma_0$  and  $\gamma_1$  are given by:

$$\hat{\gamma}_{m0} = m_e \left[ 1 - e^{-\frac{1}{2}\left(\frac{t_e}{\hat{\gamma}_{m1}}\right)^2} \right]^{-1} \quad (5)$$

$$\hat{\gamma}_{m1} = \frac{1}{2} \left\{ \frac{r}{m_e} - t_e^2 e^{-\frac{1}{2}(\frac{t_e}{\hat{\gamma}_1})^2} \left[ 1 - e^{-\frac{1}{2}(\frac{t_e}{\hat{\gamma}_1})^2} \right]^{-1} \right\} \quad (6)$$

To obtain confidence limits for the parameters  $\gamma_0$  and  $\gamma_1$  asymptotic variances of the maximum likelihood estimator of the parameters  $\gamma_0$  and  $\gamma_1$  are derived, which is the inverse of Fisher information matrix. The negative second order partial derivative of log likelihood function is obtained as follows:

$$\frac{\partial^2 \log L}{\partial \gamma_0^2} = -\frac{m_e}{\hat{\gamma}_0^2} \quad (7)$$

$$\frac{\partial^2 \log L}{\partial \gamma_1^2} = \frac{1}{\left[ \left( -\frac{2m_e}{\hat{\gamma}_1^2} + 3 \sum_{i=1}^m \frac{t_i}{\hat{\gamma}_1^4} + 3\hat{\gamma}_0 e^{-\frac{t_e^2}{2\hat{\gamma}_1}} \frac{t_e^2}{\hat{\gamma}_1^2} + \hat{\gamma}_0 e^{-\frac{t_e}{2\hat{\gamma}_1}} \frac{t_e^4}{\hat{\gamma}_1^6} \right) \right]} \quad (8)$$

Using equation (7) and (8) equations, the variance for the parameters  $\gamma_0$  and  $\gamma_1$  can be given as follows:

$$\text{Var}(\hat{\gamma}_0) = \frac{\hat{\gamma}_0^2}{m_e} \quad (9)$$

$$\text{Var}(\hat{\gamma}_1) = \frac{1}{\left[ \left( -\frac{2m_e}{\hat{\gamma}_1^2} + 3 \sum_{i=1}^m \frac{t_i}{\hat{\gamma}_1^4} + 3\hat{\gamma}_0 e^{-\frac{t_e^2}{2\hat{\gamma}_1}} \frac{t_e^2}{\hat{\gamma}_1^2} + \hat{\gamma}_0 e^{-\frac{t_e}{2\hat{\gamma}_1}} \frac{t_e^4}{\hat{\gamma}_1^6} \right) \right]} \quad (10)$$

Using equation (9) and (10) the 100(1- $\alpha$ ) % confidence interval for parameters  $\gamma_0$  i.e. ( $\tilde{\gamma}_{0L}, \tilde{\gamma}_{0U}$ ) and  $\gamma_1$  i.e. ( $\tilde{\gamma}_{1L}, \tilde{\gamma}_{1U}$ ) are given by

$$\tilde{\gamma}_{0L} = \hat{\gamma}_0 + Z_{\alpha/2} \sqrt{\frac{\hat{\gamma}_0}{m_e}} \quad (11)$$

$$\tilde{\gamma}_{0U} = \hat{\gamma}_0 - Z_{\alpha/2} \sqrt{\frac{\hat{\gamma}_0}{m_e}} \quad (12)$$

$$\tilde{\gamma}_{1L} = \hat{\gamma}_1 - Z_{\alpha/2} \sqrt{\frac{1}{\left[ \left( -\frac{2m_e}{\hat{\gamma}_1^2} + 3 \sum_{i=1}^m \frac{t_i}{\hat{\gamma}_1^4} + 3\hat{\gamma}_0 e^{-\frac{t_e^2}{2\hat{\gamma}_1}} \frac{t_e^2}{\hat{\gamma}_1^2} + \hat{\gamma}_0 e^{-\frac{t_e}{2\hat{\gamma}_1}} \frac{t_e^4}{\hat{\gamma}_1^6} \right) \right]}} \quad (13)$$

$$\tilde{\gamma}_{1U} = \hat{\gamma}_1 + Z_{\alpha/2} \sqrt{\frac{1}{\left[ \left( -\frac{2m_e}{\hat{\gamma}_1^2} + 3 \sum_{i=1}^m \frac{t_i}{\hat{\gamma}_1^4} + 3\hat{\gamma}_0 e^{-\frac{t_e^2}{2\hat{\gamma}_1}} \frac{t_e^2}{\hat{\gamma}_1^2} + \hat{\gamma}_0 e^{-\frac{t_e}{2\hat{\gamma}_1}} \frac{t_e^4}{\hat{\gamma}_1^6} \right) \right]}} \quad (14)$$

Where,  $Z_{\alpha/2}$  is the [100(1+ $\alpha$ )/2]<sup>th</sup> standard normal percentile.

Substituting the tabulated values of  $Z_{\alpha/2}$ , 95% confidence intervals ( $\tilde{\gamma}_{0L}, \tilde{\gamma}_{0U}$ ) and ( $\tilde{\gamma}_{1L}, \tilde{\gamma}_{1U}$ ) can be obtained.

#### 4. Results and Discussion

Here, two sided interval confidence interval at 95% confidence level using maximum likelihood estimation are obtained for the parameter  $\gamma_0$  and  $\gamma_1$ . A sample size  $m_e$  was generated up to execution time  $t_e$  and it was repeated 1000 times from the Rayleigh distribution by considering different values of  $\gamma_0$  and  $\gamma_1$  to study the performance of proposed confidence interval. 95% confidence intervals have been obtained by using Monte Carlo simulation technique. The values of average length and coverage probability have been obtained by assuming execution time  $t_e (= 5,6,7,8)$ , and parameters  $\gamma_0 (= 10(1)14)$ , and  $\gamma_1 (= 0.75(0.25)1.75)$ . Average length and coverage probability obtained for confidence interval has been summarized in the tables 1 to 8.

Tables 1 to 4 summarize average length and coverage probability for parameter  $\gamma_0$ . From these tables, it is seen that as the value of  $\gamma_0$  increases, average length calculated for parameter  $\gamma_0$  decreases and as value of  $\gamma_1$  increases, average length also increases. As the value of  $\gamma_0$  increases coverage probability also decreases. And coverage probability increases as  $\gamma_1$  increases. It was also found that as execution time increases average length also increases and coverage probability also increase.

From the tables 5 to 8 it is noticed that the average length calculated for parameter  $\gamma_1$  increases as the value of scale parameter  $\gamma_0$  and shape parameter  $\gamma_1$  increase. Coverage probability for  $\gamma_1$  parameter is increasing as value of  $\gamma_0$  increasing and coverage probability increases slightly as  $\gamma_1$  increases. It can also be observed that varying execution time, average length and coverage probability increases.

**Table 1:** Average length and coverage probability of 95% Confidence interval of  $\hat{\gamma}_{M0}$  calculated for different values of the parameters  $\gamma_0 = (10:1:14)$ ,  $\gamma_1 = (0.75:0.25:1.75)$  and  $t_e = 5$

$\gamma_1 \backslash \gamma_0$	10	11	12	13	14
0.75	8.58030 (0.994)	7.80447 (0.993)	6.94918 (0.992)	6.61624 (0.991)	6.36012 (0.991)
1	8.58031 (0.994)	8.01258 (0.993)	7.06641 (0.992)	6.92884 (0.992)	6.36018 (0.991)
1.25	8.67284 (0.994)	8.19769 (0.993)	7.19420 (0.992)	7.30324 (0.992)	6.53328 (0.991)
1.50	8.67286 (0.994)	8.38281 (0.993)	7.27498 (0.992)	7.58629 (0.992)	6.73742 (0.991)
1.75	8.76538 (0.995)	8.39524 (0.994)	7.60509 (0.993)	7.71259 (0.993)	6.79844 (0.991)

\*The values in the parenthesis is coverage probability.

**Table 2:** Average length and coverage probability of 95% Confidence interval of  $\hat{\gamma}_{M0}$  calculated for different values of the parameters  $\gamma_0 = (10:1:14), \gamma_1 = (0.75:0.25:1.75)$  and  $t=6$

$\gamma_1 \backslash \gamma_0$	10	11	12	13	14
0.75	8.58030 (0.994)	8.083131 (0.993)	7.55347 (0.993)	7.17032 (0.992)	6.38972 (0.991)
1	8.72548 (0.994)	8.382828 (0.993)	7.59294 (0.993)	7.28211 (0.992)	6.77412 (0.991)
1.25	8.76538 (0.995)	8.395234 (0.994)	7.68555 (0.993)	7.28413 (0.992)	6.93309 (0.992)
1.50	8.78338 (0.995)	8.395235 (0.994)	7.89537 (0.993)	7.36196 (0.992)	6.95633 (0.992)
1.75	8.86543 (0.995)	8.487783 (0.994)	8.092692 (0.993)	7.56139 (0.993)	7.30408 (0.992)

\*The values in the parenthesis is coverage probability.

**Table 3:** Average length and coverage probability of 95% Confidence interval of  $\hat{\gamma}_{M0}$  calculated for different values of the parameters  $\gamma_0 = (10:1:14), \gamma_1 = (0.75:0.25:1.75)$  and  $t=7$

$\gamma_1 \backslash \gamma_0$	10	11	12	13	14
0.75	8.65738 (0.994)	8.48777 (0.994)	8.21017 (0.993)	7.39136 (0.993)	7.26634 (0.992)
1	8.68386 (0.994)	8.55324 (0.994)	8.39537 (0.993)	8.10580 (0.993)	7.48448 (0.992)
1.25	8.72529 (0.994)	8.60538 (0.994)	8.47527 (0.994)	8.19886 (0.993)	7.48785 (0.992)
1.50	8.74578 (0.994)	8.65581 (0.994)	8.48778 (0.994)	8.28175 (0.993)	7.51383 (0.992)
1.75	8.76538 (0.995)	8.70539 (0.994)	8.70251 (0.994)	8.29022 (0.993)	7.81726 (0.993)

\*The values in the parenthesis is coverage probability.

**Table 4:** Average length and coverage probability of 95% Confidence interval of  $\hat{\gamma}_{M0}$  calculated for different values of the parameters  $\gamma_0 = (10:1:14), \gamma_1 = (0.75:0.25:1.75)$  and  $t=8$

$\gamma_1 \backslash \gamma_0$	10	11	12	13	14
0.75	8.75526 (0.995)	8.72563 (0.994)	8.48777 (0.993)	8.37329 (0.993)	7.70533 (0.992)
1	8.76530 (0.995)	8.72443 (0.994)	8.56786 (0.993)	8.38274 (0.993)	7.79001 (0.992)
1.25	8.77386 (0.995)	8.73284 (0.994)	8.67284 (0.993)	8.38875 (0.993)	7.79140 (0.992)
1.50	8.78628 (0.995)	8.76645 (0.995)	8.73713 (0.993)	8.58031 (0.993)	7.89511 (0.992)
1.75	8.79404 (0.996)	8.76852 (0.995)	8.74653 (0.993)	8.58157 (0.993)	8.080868 (0.993)

\*The values in the parenthesis is coverage probability.

**Table 5:** Average length and coverage probability of 95% Confidence interval of  $\hat{\gamma}_{M1}$  calculated for different values of the parameters  $\gamma_0 = (10:1:14), \gamma_1 = (0.75:0.25:1.75)$  and  $t=5$

$\gamma_1 \backslash \gamma_0$	10	11	12	13	14
0.75	0.452752 (0.992)	0.865715 (0.992)	1.027517 (0.993)	1.23690 (0.994)	1.61596 (0.994)
1	0.456343 (0.992)	0.866659 (0.992)	1.246347 (0.994)	1.59739 (0.994)	1.689033 (0.994)
1.25	0.558101 (0.992)	1.002341 (0.992)	1.319541 (0.994)	1.796597 (0.994)	1.705186 (0.994)
1.50	0.57393 (0.992)	1.02732 (0.993)	1.419113 (0.994)	1.923769 (0.994)	1.927146 (0.994)
1.75	0.606502 (0.992)	1.34733 (0.993)	1.426714 (0.994)	1.991414 (0.994)	2.031833 (0.995)

\*The values in the parenthesis is coverage probability.

**Table 6:** Average length and coverage probability of 95% Confidence interval of  $\hat{\gamma}_{M1}$  calculated for different values of the parameters  $\gamma_0 = (10:1:14), \gamma_1 = (0.75:0.25:1.75)$  and  $t=6$

$\gamma_1 \backslash \gamma_0$	10	11	12	13	14
0.75	0.45221 (0.992)	0.77637 (0.993)	0.85234 (0.993)	1.02542 (0.993)	1.19246 (0.994)
1	0.46395 (0.992)	0.77946 (0.993)	1.06090 (0.993)	1.22735 (0.994)	1.59603 (0.994)
1.25	0.53536 (0.992)	0.92815 (0.993)	1.07249 (0.993)	1.32042 (0.994)	1.71065 (0.994)
1.50	0.61473 (0.992)	1.03821 (0.993)	1.44207 (0.994)	1.46521 (0.994)	1.97013 (0.994)
1.75	0.72554 (0.993)	1.36680 (0.994)	1.47128 (0.994)	1.62992 (0.994)	2.22514 (0.995)

\*The values in the parenthesis is coverage probability.

**Table 7:** Average length and coverage probability of 95% Confidence interval of  $\hat{\gamma}_{M1}$  calculated for different values of the parameters  $\gamma_0 = (10:1:14), \gamma_1 = (0.75:0.25:1.75)$  and  $t=7$

$\gamma_1 \backslash \gamma_0$	10	11	12	13	14
0.75	0.49438 (0.992)	0.83019 (0.992)	1.02628 (0.993)	1.50549 (0.994)	2.09446 (0.994)
1	0.49828 (0.992)	1.02850 (0.993)	1.19715 (0.993)	1.94524 (0.994)	2.45315 (0.994)
1.25	0.59144 (0.992)	1.10866 (0.993)	1.67258 (0.993)	2.09082 (0.994)	2.54534 (0.994)
1.50	0.59262 (0.992)	1.13891 (0.993)	1.70941 (0.993)	2.16158 (0.994)	2.83801 (0.994)
1.75	0.63329 (0.992)	1.16842 (0.993)	1.88625 (0.993)	3.37754 (0.994)	3.46204 (0.994)

\*The values in the parenthesis is coverage probability.

**Table 8:** Average length and coverage probability of 95% Confidence interval of  $\hat{\gamma}_{M1}$  calculated for different values of the parameters  $\gamma_0 = (10:1:14), \gamma_1 = (0.75:0.25:1.75)$  and  $t=8$

$\gamma_1 \backslash \gamma_0$	10	11	12	13	14
0.75	0.47184 (0.992)	0.78585 (0.993)	1.30901 (0.993)	1.44314 (0.993)	1.57704 (0.993)
1	0.51061 (0.992)	0.95328 (0.993)	1.36087 (0.993)	1.74261 (0.993)	2.70185 (0.993)
1.25	0.54229 (0.992)	1.03187 (0.993)	1.39260 (0.993)	2.15898 (0.993)	2.93696 (0.993)
1.50	0.56675 (0.992)	1.07622 (0.993)	1.51162 (0.993)	2.34460 (0.993)	4.63501 (0.993)
1.75	0.56918 (0.992)	1.09768 (0.993)	1.62715 (0.993)	2.96881 (0.993)	5.06561 (0.993)

\*The values in the parenthesis is coverage probability.

## 5. Conclusion

In this research paper, confidence intervals are proposed for the parameters of Poisson type Rayleigh class model. For the proposed model confidence intervals, average length and coverage probability is calculated with the help of simulated data. From the above discussion it is observed that confidence interval for the parameters  $\gamma_0$  and  $\gamma_1$  gives higher coverage probability for particular values of both the parameters for fixed execution time. It is concluded that confidence interval can be preferred for certain values of parameters for fixed execution time.

### Reference:

- [1] D’Cunha, J. G. and. Rao, K. A.(2016). Frequentist Comparison of the Bayesian Credible and Maximum Likelihood Confidence interval for the Median of the Lognormal Distribution for the Censored Data. *International Journal of Scientific and Research Publications*, 6:61-66.
- [2] Dey,S. Dey,T. and Kundu, D. (2014).Two-parameter Rayleigh distribution: different methods of estimation, *American Journal of Mathematical and management Sciences*,33:55-74.
- [3] Fang C.C. and Yeh C.W. (2011). Confidence interval estimation of the software reliability growth models derived from stochastic differential equations, *IEEE International Conference on Industrial Engineering Management Society*, pp. 1843-1847.
- [4] Fang C.C. and Yeh C.W. (2016). Effective confidence interval estimation of fault detection process of software reliability growth model.*International Journal of Systems science*, 47: 2878-2892.
- [5] Hirano K. (1986). Rayleigh distributions, *John Wiley & Sons: New York, NY, USA*.
- [6] Jeske, D.R. Zhang, X., and Pham, L. (2001). Accounting for realities when estimating the field failure rate of software. *In proceedings of the 12<sup>th</sup> International Symposium on Software Reliability Engineering*, 332-339.



- [7] Lalitha, S. and Mishra, A. (1996). Modified maximum likelihood estimation for Rayleigh distribution, *Communications in Statistics-Theory and methods*, 25: 389-401.
- [8] Lee, T.Q., Fang, C.C., and Yeh, C.W., (2013). Confidence Interval Estimation of Software Reliability Growth Models based on Obha's inflection S- shaped model, *Journal of Industrial and Intelligent Information*,1:196-200.
- [9] Merovci, F. and Elbatal, I., (2015). Weibull Rayleigh Distribution: Theory and Applications, *Applied Mathematics and Information Sciences*, 9: 2127-2137.
- [10] Musa J.D. and Okumoto K.(1984). A logarithmic Poisson execution time model for software reliability measurement. *Proc. 7th International Conference on Software Engineering*, Orlando, Florida, 230–238.
- [11] Musa, J., Lannino, A. and Okumoto K. (1987). *Software Reliability: Measurement, Prediction, Application*, McGraw-Hill, New York.
- [12] Rao, B.V.P., Rao, K. G., and Rao, B.S.,(2013). Inverse Rayleigh Software Reliability Growth Model, *International Journal of Computer Applications*, 75:1-5.
- [13] Roy,D. (2004). Discrete Rayleigh Distributions, *IEEE Transactions Reliability*, 52: 255-260.
- [14] Saroj,A. Sonker, P.K., and Kumar, M. (2022). Statistical properties and applications of transformed lifetime distribution: Inverse Muth distribution, *Reliability Theory & Applications*, 17:178-193.
- [15] Seo, J. I. Jeon, J. W. and Kang, S. B. (2016). Exact Interval Inference for the Two parameter Rayleigh Distribution based on the upper record values, *Journal of Probability and Statistics*,1-5.
- [16] Sinha, S.K. and Howlader H.A. (1983). Credible and HPD intervals of the parameter and reliability of Rayleigh distribution, *IEEE Transactions on Reliability*, 32: 217-220
- [17] Sinha, S. K. (1985). *Reliability and life testing*, Wiley, New Delhi.
- [18] Yamada, S., Ohtera, H. and Narihisa, H. (1986). Software Reliability Growth Models with Testing- Effort, *IEEE Transactions on Reliability*, 35: 19 – 23.