ELECTRONIC JOURNAL
OF INTERNATIONAL GROUP
ON RELIABILITY

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

●

ISSN 1932-2321

●

VOL.2 NO.1 (20)
MARCH, 2011

Gnedenko Forum Publications

RELIABILITY:
THEORY&APPLICATIONS

San Diego

ISSN 1932-2321

http://www.gnedenko-forum.org/Journal/index.htm

Journal of International Group on Reliability

RELIABILITY:

THEORY & APPLICATIONS

Vol.2 No.1 (20),
March, 2011

San Diego
2011

Journal Council

Send your paper

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

Theoretical papers have to 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 have to be presented in English in MSWord format. (Times New Roman, 12 pt , 1.5 intervals).
2. The total volume of the paper (with illustrations) can be up to 15 pages.
3. A presented paper has to be spell-checked.
4. For those whose language is not English, we kindly recommend to use professional linguistic proofs before sending a paper to the journal.

* * *

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

the Editor-in-Chief ,
Igor Ushakov
igusha22@gmail.com

or

the Deputy Editor,
Alexander Bochkov
a.bochkov@gmail.com

Table of Contents

Algorithms of methods of an estimation of faultlessness transformers resulted at the set version of distinctive attributes and classification park transformers on groups with statistically various parameters of faultlessness. Algorithms serve one of distinctive features of the automated information control system developed by authors as reliability of transformers a power supply system

For discrete time risk model with dependent financial and insurance risks numerical experiment with recurrent procedure of ruin probability calculation is made. It shows that suggested recurrent procedure is much faster than application of usual Monte-Carlo method.

The IEC standards 61508/61511 require that reliability targets for safety instrumented functions are defined and verified. The reliability targets are given as one out of a possible four safety integrity levels. For each safety integrity level there are many design requirements, including requirements for the probability of failure on demand. Verification of the requirements for the probability of failure on demand is usually based on a quantitative analysis. In this paper we argue that such an approach is better replaced by a semi-quantitative approach. The approach acknowledges that the probability of failure on demand requirement cannot be adequately verified only by reference to an assigned probability number. There is a need for seeing beyond the probability number. The key aspect to include is related to uncertainty.

A number of definitions and interpretations of the risk concept exist. Many of these are probability-based. In this paper we present and discuss a structure for characterising the definitions, which is founded on a clear distinction between (a) risk as a concept based on events, consequences and uncertainties; (b) risk as a modelled, quantitative concept; and (c) risk descriptions. The discussion leads to recommended perspective for conceptualising and assessing risk, which is based on risk defined by (a), and the probability-based definitions of risk can be viewed as related model parameters and/or risk descriptions. Two ways of detailing the framework are outlined: the relative frequency-based approach and the Bayesian approach.

External hazards can provide safety significant contributions to the risk in case of nuclear power plant operation because such hazards have the potential to reduce simultaneously the level of redundancy by damaging redundant systems and lines or their supporting systems. Therefore, risk assessment of all potential external hazards to the plant under consideration is part of the overall safety assessment. In this paper, the procedure for assessing the external hazard aircraft crash is described in more detail. The first step is an appropriate screening procedure in order to determine scope and content of the assessment, taking into account plant- and site-specific conditions. The second step is to determine the methodical approach for those cases where a full scope analysis has to be performed and the inclusion into the used overall risk model. The considerations regarding this hazard do not cover an intended aircraft crash.

Key equipment for the electric power transmission is the transformer. Because of the high failure frequency and the resultant reliability and safety implications in particular of main transformers, an in-depth assessment is necessary. Main transformers are considered as a critical equipment because of the large quantity of oil in contact with high voltage elements. Experience has shown an increasing number of transformer explosions and fires in all types of power plants worldwide. Therefore, these phenomena have been investigated in more detail and are discussed with regard to potential root causes for these events such as potential influence of the age of the transformers. Moreover, possible diagnostic measures to avoid such events and enhance the reliability are shortly described. For investigating the current status of the reliability of transformers different types of databases have been evaluated.

Bayesian statistical decision theory would be questionable when applied directly to non-random uncertainty circumstances. In this paper, we investigate the basic elements of decision analysis oriented to observational data arising from a general uncertainty environment, so that a framework for Bayesian uncertainty decision doctrine is established. Further, we propose a copula-linked uncertainty marginals mechanism for constructing the uncertainty multivariate distributions to represent both observational data and an uncertainty parameter vector. This mechanism paves the way towards the establishment of an uncertainty posterior distribution of the parameter vector given the observational data, based on uncertain measure Axiom 5. Finally, we present an illustrative example of the development of a posterior uncertainty distribution for a parameter given a single observation, step by step. The significance of this paper is to establish for the first time a Bayesian uncertainty data inference and decision framework, which constitutes a critical step towards the establishment of uncertainty statistics and a Bayesian uncertainty decision theory.

The real world phenomena are often facing the co-existence reality of different formality of uncertainty and thus the probabilistic reliability modeling practices are very doubtful. Under complicated uncertainty environments, hybrid variable modeling is important in reliability and risk analysis, which includes Bayesian distributional theory, random fuzzy distributional theory, as well as fuzzy random distributional theory as special distribution families. In this paper, we define a new hybrid lifetime which is specified by a random lifetime distribution with imprecise parameter with an uncertainty distribution. We furthermore define the average chance distribution as a quality index for quantifying the hybrid lifetime and accordingly the average chance reliability is derived.

The general model of the safety of complex technical systems in variable operation conditions linking a semi-Markov modeling of the system operation process with a multi-state approach to system safety analysis and linear programming are applied in maritime transport to safety and risk optimization of a ferry technical system.

The paper demonstrates new approach of rendering the graph point series called gantts. The gantts are placed in the two dimensional graph which contains the information about available production sources in the real manufacturing process. To have the interaction with a user, gantts are accompanied with the description text giving detailed information about each gantt. All gantt descriptions must be displayed without overlapping with each other. To optimize this task, the modified version of the RBF neural network with biases is applied. With respect to the similarity to the RBF structure, the new type of neural network is named RBF 2. We also give the picture of positive and negative attributes of the solution based on the neural network architecture.

Inspection planning is an important activity in process industries, and one of the key tools used for such planning is the risk based inspection (RBI) methodology. The RBI is commonly used in planning of inspections for static mechanical equipment, in particular piping networks. The inspections are prioritized based on risk, expressed as expected values, integrating the likelihood and consequences of failures. In this paper we suggest an extension of the RBI methodology which reflects risk and uncertainties beyond expected values. We argue that such an extension is essential for adequately supporting the inspection planning. A pipeline example from the Norwegian oil and gas industry is presented to illustrate and discuss the suggested approach.

In general, Evolutionary Sets of Cooperating Ship Trajectories combine some of the assumptions of game theory with evolutionary programming and aim to find optimal set of cooperating trajectories of all ships involved in an encounter situation. In a two-ship encounter situation the method enables the operator of an on-board collision-avoidance system to predict the most probable behaviour of a target and to plan the own manoeuvres in advance. In a multi-ship encounter the method may be used to help an operator of a VTS system to coordinate the manoeuvres of all ships. The improvement presented here is a new way of modelling some of the COLREGS rules. Due to this change, the method is now able to find solutions, which are more compliant with COLREGS, more intuitive and consequently – safer from the navigator’s point of view. The paper contains a detailed description of collision-avoidance operators used by the evolutionary method and simulation examples of the method’s results for digital maps.

Drinking water technical system is an essential element of urban infrastructure. The operation of this system is inseparably connected with a risk of failure. The main problem in the risk of failure analysis of water pipe network is the uncertainty of the information collected on the description of failure. In order to consider the uncertainty of information, the theory of fuzzy sets was used. The fuzzification of frequency, severity and the consequences of the incident scenario is basic input for fuzzy risk analysis. The presented model is part of a complex model of risk management of failures in drinking water technical system manly in water pipe network and can be used in practice in system operator’s decision-making process. An adaptation of the fuzzy set theory to analyse risk of failure of water mains is not a standard approach for water works. An effect of the analysis of different sources of risk can be used for the design of a more reliable safety system assurance.

Dear Colleagues!

As you know, our site is named after outstanding statistician, one of the best pupils of Andrei Kolmogorov – Boris Gnedenko. On January the 1st, 2012 there will be the 100th anniversary of the birth of Boris Gnedenko.

We would like to have more members of this informal International Association to this date.

Please invite your colleagues to join our Forum. Send them our invitation.

Welcome to Gnedenko Forum!

Join us!

We join already about 500 specialists

from 44 countries

To join us send your photo
and brief resume to
Professor Igor Ushakov:
igusha22@gmail.com,
or
Dr. Alexander Bochkov
a.bochkov@gmail.com

THE COMPUTER ANALYSIS OF FAULTLESSNESS TRANSFORMERS OF POWER SUPPLY SYSTEMS

Farhadzadeh E.M., Muradaliyev A.Z., Rafiyeva T.K, Ismayilova S.M.
•
Azerbaijan Scientific-Research and Design-Prospecting Institute of Energetic
AZ1012, Ave. H.Zardabi-94
E-mail:fem1939@rambler.ru

ABSTRACT

Algorithms of methods of an estimation of faultlessness transformers resulted at the set version of distinctive attributes and classification park transformers on groups with statistically various parameters of faultlessness. Algorithms serve one of distinctive features of the automated information control system developed by authors as reliability of transformers a power supply system

Perfection of a control system by reliability of the equipment of power supply systems represents important and, simultaneously, a difficult problem. Importance of the decision of this problem mainly caused increasing (in process of ageing the equipment) by cost of technological tests and restoration of deterioration (repair), consequences of damage of the equipment. Difficulty of the decision consists that modern methods of the statistical analysis of retrospective data assume realization in the form of computer technologies and exclude an opportunity the manual account owing to the complexity and bulkiness.

One of the basic directions of overcoming of these difficulties is development and application of the specialized automated information systems (AIS). The universal computerization of the power enterprises of power supply systems allows realize successfully in practice information support of a management and technical officers at the control execution of rules of technical operation, the decision of problems of maintenance service and repair of the power equipment. Experience of practical use such AIS testifies to their high efficiency [1].

Among properties of reliability most studied is property of faultlessness. At the same time, insufficient objectivity of quantitative estimations of parameters of faultlessness (PF) and absence of specifications of faultlessness demands the further researches. So, for example, if to calculate PF under the collected information on refusals of all transformers of a power supply system (further: on a final data set - FDS) we shall receive an estimation which analogue will be «average temperature on hospital». Attempts to concretize this estimation for the set version of attributes (VP) lead to that the number of statistical data about refusals is sharply reduced and, despite of an observable divergence of initial and «specified» values PF, the hypothesis about their casual divergence often does not contradict statistical data. It is possible, certainly, «closed eyes» on these features (that more often occurs), to speak only about an observable divergence of estimations PF, «to explain» the reasons of a divergence and to recommend ways of elimination of these reasons. The methodology of overcoming of the specified difficulties is developed by authors and approved in AIS estimations of a technical condition of power transformers and autotransformers (further: transformers), called as AISTR.

General characteristic AISTR.

To basic features AISTR concern:

1. The system focused on the organization of maintenance service and repair of the transformer (TR) depending on their technical condition;

2. Existing systems, as a rule, are intended for processing statistical given refusals TR, or processing of results of test (for example, data of chromatography analysis of the gases dissolved in transformer oil), or data about the revealed defects at surveys and scheduled repairs TR. Considering interrelation of these data developed AISTR provides processing and sharing of data about non-working conditions TR, results of tests and the surveys, the given restoration of deterioration;

3. Application of the methods developed by authors:
   - protection of a database against casual or deliberate distortion;
   - the account of casual character of estimations PF, calculation PF with optimum accuracy (width a confidential interval) and reliability (with the minimal sum mistakes of the first and second sort);
   - estimations of optimum confidential area of change PF in time or other VP;
   - estimations of parameters of durability according to test TP;
   - planning terms and volumes of major overhauls (MO) TR on an integrated parameter which considers distinction of term of service, operating time after MO, numbers of through currents of short circuit, loading, a technical condition (data of surveys and tests), the importance consequences of refusals;
   - quality assurance of restoration of deterioration;
   - specifications of maximum permissible values of diagnostic parameters;

4. The opportunity the reference to the specifications and technical documentation defining strategy of maintenance service and repair TR, especially, TR, which service life exceeds normative.

Integrated block diagram AISTR resulted on fig.1.

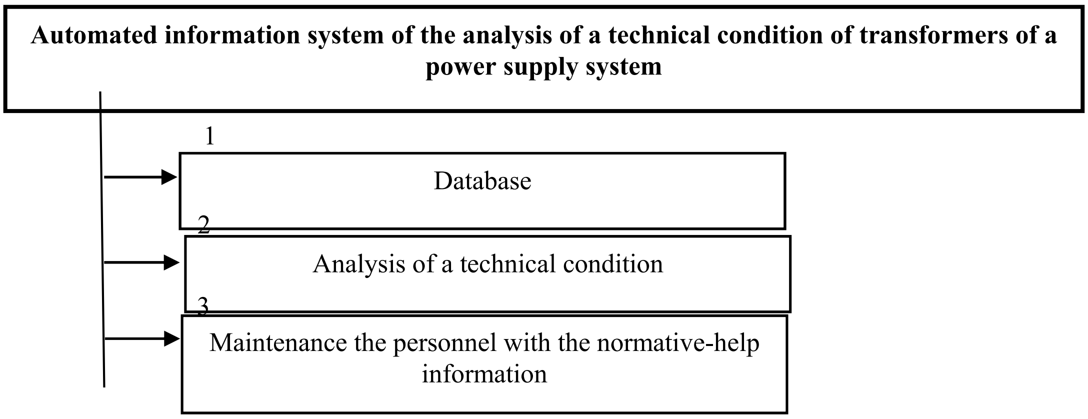


Fig.1 Integrated block diagram of modeling algorithm AISTR

On fig.2, the integrated block diagram of algorithm of a subsystem «Analysis of a technical condition» is resulted. In present clause, we consider some features of algorithm of a subsystem «analysis of faultlessness». The integrated block diagram of algorithm of this subsystem is resulted on fig.3.

Sample of data about non-working conditions on the set interval of time for concrete or same TR (block 2.2.1) is the obligatory document at the analysis of the reasons of damage TR, development of actions decrease in number and duration of not scheduled switching-off TR, the organization of maintenance service and repair. We shall distinguish following switching-off: automatic at short circuits in TR, under emergency or scheduled applications for restoration of deterioration, in a reserve and on an operating mode, false and if necessary restoration of deterioration of the adjacent equipment under the scheme. These switching-off classified on versions of the attributes set in tables of nameplate data and conditions of operation TP.

In the block 2.2.2, the analysis of a kind of switching-off TP is spent. Versions of this attribute are switching-off sudden, under the emergency or scheduled application, and compelled on a mode manually.

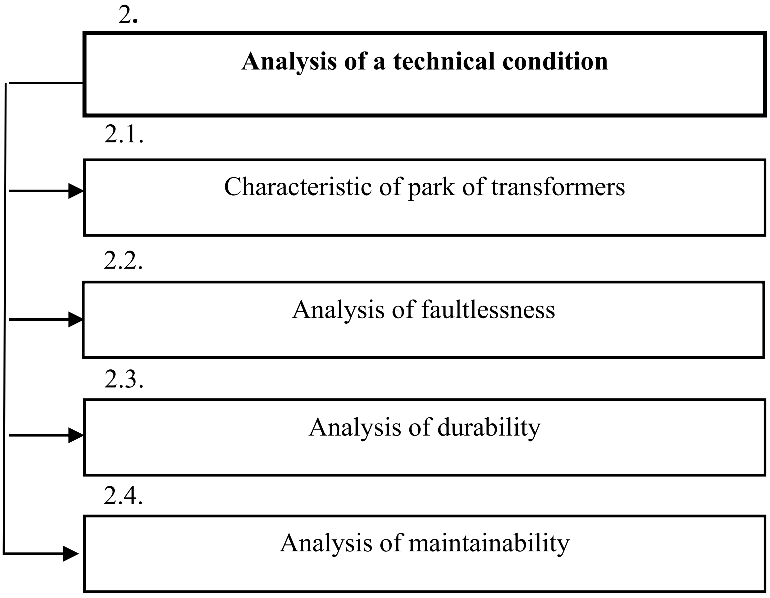


Fig.2 Integrated block diagram of algorithm of a subsystem «Analysis of a technical condition»

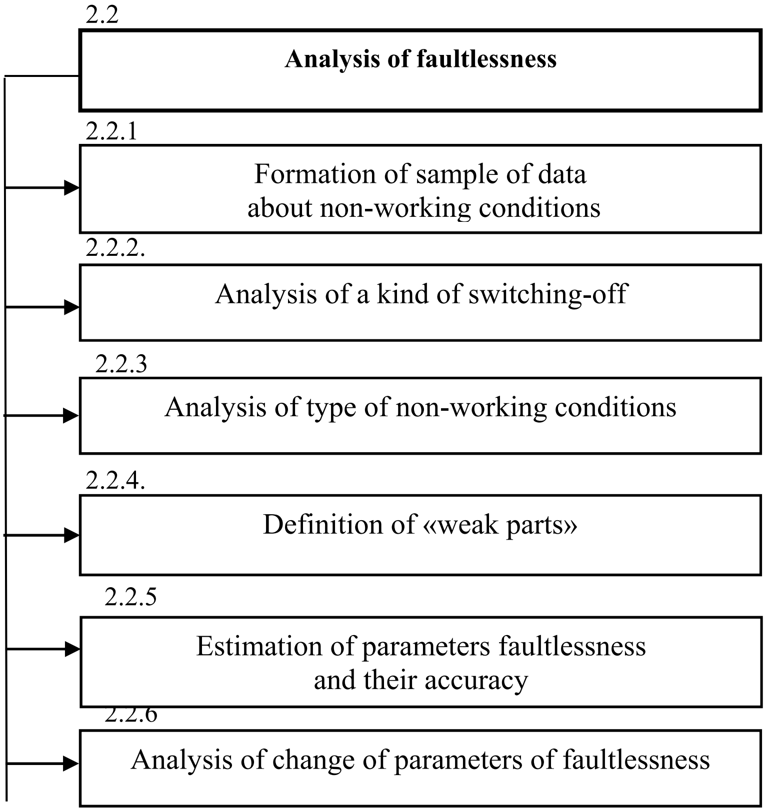


Fig.3 Integrated block diagram of algorithm of a subsystem «Analysis of faultlessness»

In the block 2.2.3 versions of type of conditions (emergency, the emergency or compelled idle time, scheduled repair) are analyzed. We shall specify distinction of these conditions. The emergency condition is the disabled condition TR connected with refusal or defect. At emergency idle, time efficient TR disconnected owing to system failure or refusal of adjacent elements under the scheme. An example of the compelled idle time is switching-off of the efficient block transformer at emergency repair of boiler installation of the power unit. Results of the analysis of observable conditions network TR for number of years of supervision on type's show that from 37% of switching-off only 26% connected with defects TR

In the block 2.2.4 «weak parts» TR defined, i.e. units, which defects most often lead to emergency switching-off TR. At classification of data by emergency switching-off TR on set VP, carried out in blocks 2.2.2-2.2.4, laws of distribution of emergency switching-off can vary. These changes include functional and statistical components. The methodology of comparison of these dependences developed by authors on the basis of methods of imitating modeling and the theory of check of statistical hypotheses allow to calculate optimum area of their change with the minimal probability of the erroneous decision.

The estimation of parameters faultlessness (block 2.2.5) at first sight does not represent any complexity. Formulas of their calculation well known. Thus, it supposed, that the information on refusals has enough for accuracy of calculations comprehensible in practice.

Actually, at attempt estimate parameters of individual faultlessness concrete TR or parameters of faultlessness of the same equipment it appears, that number of refusals and defects a little and it is necessary to solve a problem of accuracy of estimations. The matter is that the estimation parameters of faultlessness TR is not end in it self, and serves for the control non-excess as them of normative values or for comparison parameters of faultlessness of two types TR no excess. And in both cases, despite of significant distinction of parameters observable sometimes, it frequently appears casual, insignificant since is within the limits of accuracy of these estimations.

It is necessary to note, that concept of uniformity, being apprehended literally i.e. as TR one type, yet does not mean uniformity for sample of data about faultlessness these TR as conditions of operation can essentially differ (to differ with loading, service life, an operating time after major overhaul, intensity through currents of short circuit, the importance of consequences damage and so forth). Classification of the information on versions of noted attributes can lead to that, the number of refusals will be, not only it is not enough they will be absent. At the same time, it is clear, that from the point of view of faultlessness the importance of numerous attributes of distinction of nameplate data and conditions of operation is not identical. Hence, «tool» by means of which it would be possible to allocate the most significant attributes for had retrospective information is necessary.

In real conditions, usually it is necessary to solve two types of problems. The first type of problems reduced to estimation PF TR with set VP, and the second type - to ranking TR on non-failure operation.

Method and algorithm of estimation PF TR for set VP.

Includes following sequence of operations:

1. On had data about park TR of a power supply system and about emergency switching-off TR average values PF calculated. We shall consider Main principles of methodology of calculation on most often used PF - specific number of refusals $\left(\lambda_\Sigma^*\right)$ which is calculated as the attitude of number of refusals TR to number TR and duration of supervision;

2. Estimations of the same parameter for each of set VP are calculated. We shall designate these estimations as $\left(\lambda_\Sigma^*\right)$, where i - serial number of the set attributes $i = 1, n$;

3. Ranking $\left(\lambda_{\Sigma}^{*}\right)$ from $i=1,n$ in decreasing order numerical values spent. We shall designate the greatest value through $\lambda_{\max}^{*}=\max\left\{\lambda_{i}^{*}\right\}_{n}$;

4. Methods of imitating modeling and the theory of check of statistical hypotheses in view mistakes of the first and second sort check the assumption (hypothesis H_1) of a casual divergence $\left(\lambda_{\Sigma}^{*}\right)$ and $\lambda_{\max}^{*}$. If hypothesis H_1 proves to be true, it means, that the number of data about refusals is not enough, that any of set VP is not significant also classification FDS inexpedient. If the hypothesis H_1 does not prove to be true, PП for which the estimation $\lambda_{\max}^{*}$ is calculated considered essential.

5. The calculations noted items (2.4), repeat, with that difference, that as $\left(\lambda_{\Sigma}^{*}\right)$ accepted $\lambda_{\max}^{*}$, and the number considered VP becomes on unit less and will be equal (n-1);

6. If on j-that iteration the hypothesis H_1 proves to be true, it means, that from n set VP are significant (j-1), and a required estimation λ is $\lambda_{\max(j-1)}^{*}$. Usually size (j-1) is more, than more than statistical data. It is obvious, that for the others (n-j+1) VP classification of data is inexpedient. That reaches optimum accuracy (width of a confidential interval) and reliability (the minimal value of the sum of mistakes of the first and second sort).

At planning maintenance service and repair TR, the statement of specifications of safety it is important to be able to classify park TR of a power supply system on groups with differing faultlessness and VP. To generate these groups the new method offered.

Method and algorithm of classification TR on faultlessness.

The scheme of algorithm integrated the block is presented on fig.4.

Let's consider some features of calculations. For carrying out of calculations it is necessary to define (block 1): number TR, duration of supervision over their technical condition, number of emergency switching-off, number and a kind of considered attributes, number and type of versions of each of attributes. On these data, is calculated (block 2) the average estimation of specific number of refusals under the known formula:

$$\lambda_{\Sigma}^{*}=\frac{\sum_{i=1}^{K} d_i}{\sum_{i=1}^{K} \Delta T_i}=\frac{D}{M}$$

Where d_i - number of refusals of i-th TR on an interval ΔT_i; to - number TR;

In the block 3, the greatest value among set $\left\{\lambda_{i,j}\right\}_{\substack{i=1,n\\ j=1,r_i}}$ defined. For what, originally pay off $\lambda_{i,j}^{*}$ for all VP, the greatest value $\lambda_{i,j}^{*}$ among versions of each attribute $\left(\lambda_{i,m}^{*}\right)$ defined and, at last, the greatest value among all attributes $\left(\lambda_{m}^{*}\right)$ is defined.

The method of statistical modeling and the theory of check of statistical hypotheses in view of mistakes of the first and second sort checks a hypothesis (H_1) about a casual divergence λ_{Σ}^{*} and $\lambda_{\max}^{*}$. If hypothesis H_1 is true, and significant attributes are absent, calculations come to the end (block 11). If λ_{Σ}^{*} and $\lambda_{\max}^{*}$ differ not casually management is transferred to the block 5 where the number and type of significant attributes and their versions are registered.

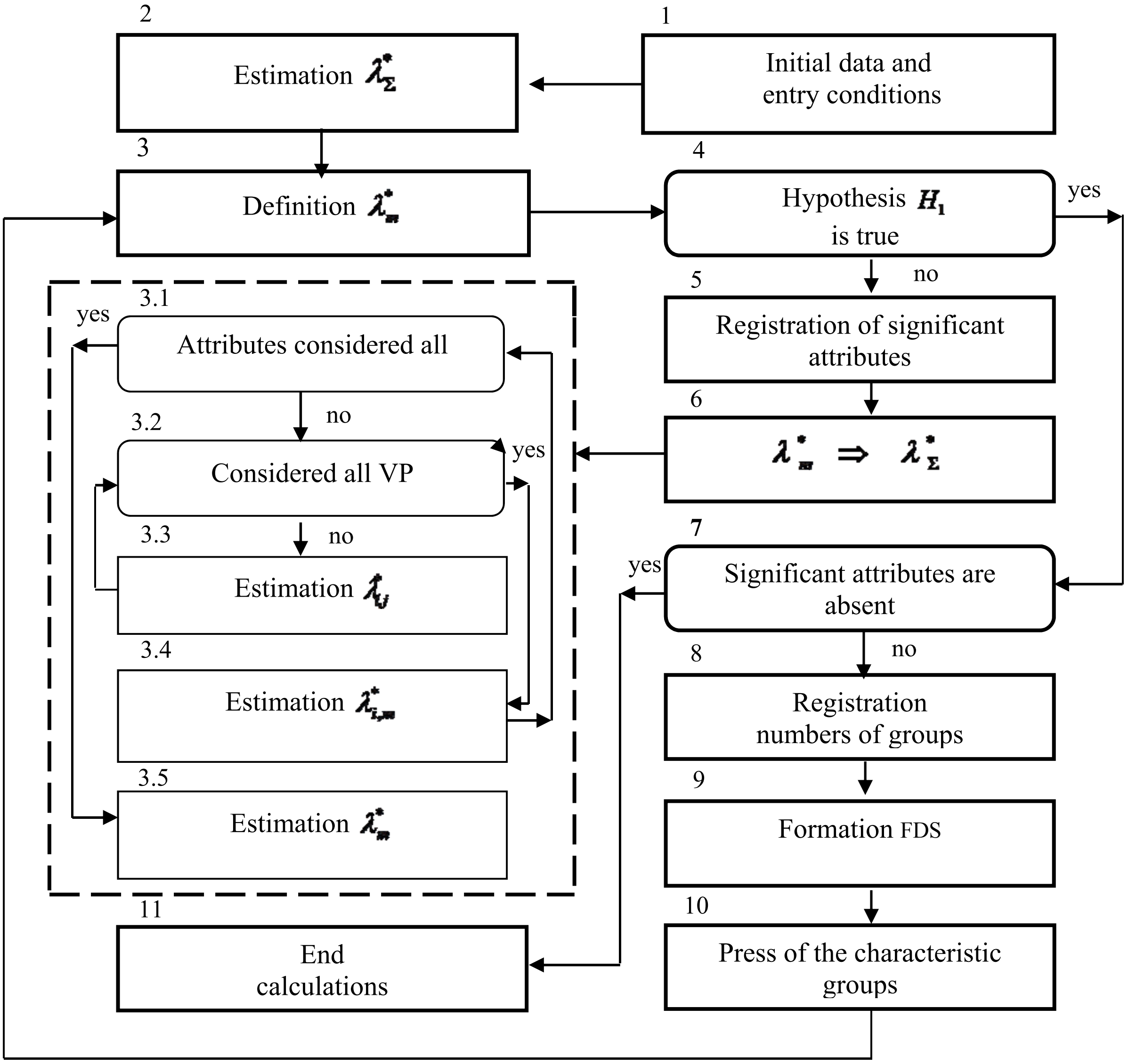


Fig.4. Integrated block diagram of algorithm of classification FDS

In the block 6, as FDS sample of data for which $\lambda^*_{i,j} = \lambda^*_m$ with i=1,n and j=1, r_i is accepted, and management is transferred to the block 3, where again, (but already for sample and (n-1) an attribute) is defined, checked λ^*_m hypothesis H_1, etc. Process of search of significant versions of each attribute comes to the end if the paritynd λ^*_m does not contradict hypothesis H_1. As group TR with distinct from λ^*_Σ faultlessness and known PF is generated, having registered it (block 8), management is transferred to the next stage of formation FDS (block 9). Its essence consists that from general number TR (M) and numbers of refusals (D) are calculated TP (M_m) and refusals (D_m) describing again generated group TR. Further are printed a serial number of this group, significant VP, M_m, D_m, λ^*_m and management is transferred to the block 3. For search of the next group of data. It can be or separate group and remained group TR as a whole.

As possible (at infinite great volume of statistical data) the number of groups equally $L = \prod_{i=1}^{n_n} r_i$, and statistically proved (at real volume of data) - is many times less, the recommended algorithm allows to allocate practically comprehensible number of groups TR with differing VP and to use them at forecasting reliability TR in view of conditions of their operation.

Casual character of distinction observed not only for estimations PF, but also for their change depending on the set combination of attributes. Most often in practice curve changes of specific number of refusals in time (on calendar years, duration of operation, a season and day) are used. However, the curves constructed on retrospective data for all TR, are poorly significant. Therefore, as a rule, classification of the information on set VP is spent. For example, curves for power transformers and autotransformers, TR various classes of a pressure, various capacity and so forth. Here in essential a greater degree are analyzed, than for estimations PF dependence on volume of statistical data is shown. The methodology of the account of this dependence offered by us in and realized in AISTR.

CONCLUSION

1. A necessary condition of an objective estimation of expediency of replacement of transformers of a power supply system, carrying out MO, modernizations, changes of loading and an operating mode, test the opportunity of an operative estimation of their technical condition is. This problem solved in the automated information system of an estimation of a technical condition of transformers developed by authors (AISTR);

2. The subsystem «Analysis of faultlessness» AISTR, alongside with traditional elements of the analysis, represents an opportunity:

- To optimize accuracy of estimation PF TR and laws of their change in time for set combination VP;
- To raise objectivity of comparison of faultlessness various TR (in other words, comparisons of faultlessness TR to various set VP);
- To divide park TR into groups on a condition of identity of their non-failure operation;

3. The analysis of the defects eliminated during switching-off TR under the emergency application, and also the defects which have caused automatic switching-off of transformers owing to of short circuit growing old TR, allows to improve type and periodicity of diagnostic check, to raise quality of scheduled repairs, to specify possible loading and operating mode TR.

LITERATURE

1. Antonenko I.N. Information support of operation power equipment. St.-Petersburg, magazine "ISUP" №1(13), 2007, p.4-7

NUMERICAL EXPERIMENT FOR RUIN PROBABILITY IN RISK MODEL WITH DEPENDENT FINANCIAL AND INSURANCE RISKS

G. Tsitsiashvili, M. Osipova
IAM, FEB RAS, Vladivostok, Russia
●
e-mails: guram@iam.dvo.ru, mao1975@list.ru

ABSTRACT

For discrete time risk model with dependent financial and insurance risks numerical experiment with recurrent procedure of ruin probability calculation is made. It shows that suggested recurrent procedure is much faster than application of usual Monte-Carlo method.

INTRODUCTION

In (Tsitsiashvili, 2010) for discrete time risk model with dependent financial and insurance risks recurrent algorithm of ruin probability calculation is constructed on a base of hyperexponential approximation of insurance losses distributions. Special methods of symbol forms transformation and economical procedure of enumeration of vectors with integer components and fixed sum are developed

In this paper constructed recurrent algorithm is tested in numerical experiment. Comparison analysis showed that suggested algorithm is significantly faster then application of usual Monte-Carlo method.

1. PRELIMINARIES

In (Tsitsiashvili 2010) recurrent discrete time risk model with initial capital $x,\ x \geq 0,$ nonnegative losses Z_n, $n = 1,2,...,$ and inflation factor $Y_n^{-1} > 1$ from $n-1$ to n year:

$$S_0 = x,\ \ S_n = Y_n^{-1} S_{n-1} + 1 - Z_n,\ \ n = 1,2...,$$

is considered. Here $X_n = Z_n - 1$ is insurance risk and Y_n is financial risk. Suppose that the sequence $\{Y_n, n \geq 0\}$ is stationary reversible Markov chain with state set $\left\{r_q^{-1}, q \in Q\right\}$, $Q = \{1,...,m\}$, with transition matrix $\|\pi_{q'q}\|_{q,q' \in Q}$ and with initial distribution

$$P\left(Y_n = r_q^{-1}\right) = p_q > 0,\ \ \sum_{q \in Q} p_q = 1\,.$$

Introduce dependence between financial and insurance risks by a conditional distribution of random variable Z_n

$$\bar{F}_q(t) = P\left(Z_n > t \Big/ Y_n = r_q^{-1} \right),\ \ q \in Q,\ \ t \geq 0.$$

In this model finite horizon ruin probability

$$\psi_n(x) = P\left(\inf\{n = 1, 2, \ldots : S_n \le 0 \mid S_0 = x\} \le n\right).$$

Following recurrent algorithm of its calculation is constructed. Denote (δ_{ij} is Kronecker symbol)

$$1_q = \left(\delta_{1q}, \ldots, \delta_{mq}\right), \quad K = (k_1, \ldots, k_m), \ k_i \in \{0, 1, \ldots\}, \ i = 1, \ldots, m, \quad r^K = \prod_{q \in Q} r_q^{k_q}, \ K_\Sigma = \sum_{q \in Q} k_q.$$

Theorem. *If*

$$r^K \lambda_i \ne \lambda_j, \ i \ge 1, \ j \le l, \ K_\Sigma \ge 1,$$

and

$$\bar{F}_q(t) = \sum_{i=1}^{l} a_{qi} \exp(-\lambda_i t), -\infty < a_{qi} < \infty, \sum_{i=1}^{l} a_{qi} = 1, \ q \in Q, \ t \ge 0,$$

when for $n \ge 1$

$$\psi_n(t) = \begin{cases} \sum\limits_{q \in Q} \sum\limits_{1 \le K_\Sigma \le n} \sum\limits_{i=1}^{l} B_{n,i,q}^K, \ t > 0, \\ 1 + \sum\limits_{q \in Q} \sum\limits_{1 \le K_\Sigma \le n} \sum\limits_{i=1}^{l} B_{n,i,q}^K \left(\exp\left(-r^K \lambda_i t\right) - 1\right), \ t \le 0, \end{cases} \tag{1}$$

and for $q, q' \in Q, \ i = 1, \ldots, l$

$$B_{1,i,q}^{1_q} = p_q a_{qi} \exp(-\lambda_i), \ B_{1,i,q}^{1_{q'}} = 0, \ q \ne q',$$

$$B_{n+1,i,q}^{1_q} = \sum_{q' \in Q} \left[\pi_{q'q} a_{qi} \exp(-\lambda_i) \left(\lambda_i \sum_{1 \le K_\Sigma \le n} \sum_{j=1}^{l} \frac{B_{n,j,q'}^K}{r^K \lambda_j - \lambda_i} + p_{q'} \right) \right], \ q = q', \ B_{n+1,i,q}^{1_{q'}} = 0, \ q \ne q',$$

$$B_{n+1,j,q}^K = -\sum_{q' \in Q} \left[\pi_{q'q} I\left(k_q > 0\right) \exp\left(-r^{K-1_q} \lambda_i\right) \sum_{j=1}^{l} \frac{B_{n,j,q'}^{K-1_q} a_{qj} \lambda_j}{r^{K-1_q} \lambda_i - \lambda_j} \right], \ 1 < K_\Sigma \le n+1.$$

Remark 1. In (Tsitsiashvili 2010) there is some inaccuracy in recurrent formulas for coefficients $B_{n+1,i,q}^{1_q}, B_{n+1,j,q}^K$. Here odd multiplier $p_{q'}$ is cancelled from these formulas.

In (Tsitsiashvili 2010) recurrent algorithm of matrix $\left\|\pi_{q'q}\right\|_{q,q' \in Q}$ .generation is replaced by a random choice of elements of symmetric matrix $\left\|A_{ij}\right\|_{i,j=1}^{m}$ for fixed probabilities $\{p_1, \ldots, p_m\}$ from formulas

$$\max\left[0, \ p_i - \sum_{k=1}^{j-1} A_{ik} - \sum_{s=j+1}^{m} p_s + \sum_{s=1}^{i-1} \sum_{k=j+1}^{m} A_{sk}\right] < A_{ij} < \min\left[p_j - \sum_{k=1}^{i-1} A_{kj}, \ p_i - \sum_{k=1}^{j-1} A_{ik}\right], \ i \le j \le m-1,$$

$$A_{im} = p_i - \sum_{k=1}^{m-1} A_{ik}, \ 1 \le i \le m,$$

then $\pi_{ij} = \dfrac{A_{ij}}{p_i}, \ 1 \le i, j \le m.$

In (Tsitsiashvili 2010) a problem of an enumeration of all vectors of the set

$$\boldsymbol{K} = \{K = (k_1, \ldots, k_m) : \ k_i = 0, 1, \ldots, i = 1, \ldots, m, 1 \le K_\Sigma \le n\}$$

is solved via recurrent calculation of the sets of vectors

$$\boldsymbol{K}_i^j = \{K = (k_1, \ldots, k_j) : \ k_i = 0, 1, \ldots, i = 1, \ldots, j, \ K_\Sigma = i\}, \ 0 \le i \le n, \ 1 \le j \le m,$$

$$\boldsymbol{K_0^1} = \{0\},\ \boldsymbol{K_i^1} = \{i\},\ 1 \le i \le n,\ \boldsymbol{K_i^{j+1}} = \bigcup_{t=0}^{i} \{(K,t) : K \in \boldsymbol{K_{i-t}^j}\},\ 0 \le i \le n, 1 \le j \le m-1.$$

Then

$$\boldsymbol{K} = \bigcup_{i=1}^{n} \boldsymbol{K_i^m}$$

and calculation complexity of this algorithm is not larger than

$$(n+1)^{m+1}.$$

2. NUMERICAL EXPERIMENT

Suppose that $m = 2,\ Q = \{1, 2\},\ p_1 = 0.25,\ p_2 = 0.75,\ r_1 = 1.03,\ r_2 = 1.08,$

$$\pi_{11} = 5/9,\ \pi_{12} = 4/9, \pi_{21} = 4/27, \pi_{22} = 23/27$$

and consider Pareto distributions of insurance losses

$$\bar{F}_1(t) = (1 + 5x)^{-1.2},\ \bar{F}_2(t) = (1 + 0.83x)^{-2.2},\ t > 0.$$

We approximate Pareto distributions by hyperexponential (Anja Feldman & Ward Whitt 1998)

$$\bar{F}_1(t) \approx \sum_{i=1}^{27} a_{1i} \exp(-\lambda_i t),\ \bar{F}_2(t) \approx \sum_{i=1}^{27} a_{2i} \exp(-\lambda_i t),$$

with parameters

i	a_{1i}	a_{2i}	λ_i
1	0.089437	0	23.304
2	0.533823	0	6.516
3	0.307218	0	1.546
4	0.059768	0	0.306
5	0.008462	0	0.057
6	0.001122	0	0.01
7	0.000147	0	0.002
8	0.0000192	0	0.00035
9	2.5×10^{-6}	0	0.000065
10	3.27×10^{-7}	0	0.000012
11	4.27×10^{-9}	0	2.2×10^{-6}
12	5.56×10^{-10}	0	3.9×10^{-7}
13	7.18×10^{-10}	0	6.8×10^{-8}
14	8.37×10^{-11}	0	8.3×10^{-9}
15	0	0.193963	4.491
16	0	0.651199	1.422
17	0	0.147814	0.371
18	0	0.006832	0.076
19	0	0.000188	0.014
20	0	4.61×10^{-6}	0.003
21	0	1.11×10^{-7}	0.0005
22	0	2.65×10^{-9}	0.000088

i	a_{1i}	a_{2i}	λ_i
23	0	6.35×10^{-11}	0.000016
24	0	1.52×10^{-12}	2.9×10^{-6}
25	0	3.36×10^{-14}	5.4×10^{-7}
26	0	8.51×10^{-16}	9.7×10^{-8}
27	0	1.72×10^{-17}	1.58×10^{-8}

Results of numerical experiments for ruin probability $\psi_{10}(x)$ with step h=0.5 by x using Monte-Carlo method with N=10000000 realizations and by the formula (1) are represented on the table

x	Monte-Carlo method	the formula (1)
0.5	0.421138	0.422476
1	0.347786	0.348598
1.5	0.291625	0.292944
2	0.247464	0.249402
2.5	0.212054	0.214347
3	0.183264	0.185541
3.5	0.159391	0.161535
4	0.139627	0.141338
4.5	0.123078	0.124237
5	0.109101	0.109692

Time of calculation by Monte-Carlo method is approximately 2 hours and 15 minutes and time of calculation by the formula (1) is 15 seconds.

Results of numerical experiments for ruin probability $\psi_{50}(x)$ with step h=0.5 by x using Monte-Carlo method with N=10000000 realizations and by the formula (1) are represented on the table

X	Monte-Carlo method	the formula (1)
0.5	0.442933	0.444793
1	0.38219	0.38395
1.5	0.335	0.336951
2	0.296502	0.298993
2.5	0.264418	0.267366
3	0.237423	0.240461
3.5	0.214501	0.217261
4	0.19457	0.197082
4.5	0.177139	0.179427
5	0.162125	0.163917

Time of calculation by Monte-Carlo method is approximately 10 hours and 30 minutes and time of calculation the formula (1) is 26 minutes.

REFERENCES

1. Tsitsiashvili, Gurami 2010. Algorithmic problems in discrete time risk model. *Reliability: Theory and Applications* (3) 1: 29-37.
2. Feldman, Anja & Whitt, Ward 1998. Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. *Performance evaluation* (3) 1: pp. 245-279.

A NEW APPROACH FOR VERIFICATION OF SAFETY INTEGRITY LEVELS

E.B. Abrahamsen
University of Stavanger, Norway
e-mail: eirik.b.abrahamsen@uis.no
•

W. Røed
Proactima AS, Norway
e-mail: wr@proactima.com

ABSTRACT

The IEC standards 61508/61511 require that reliability targets for safety instrumented functions are defined and verified. The reliability targets are given as one out of a possible four safety integrity levels. For each safety integrity level there are many design requirements, including requirements for the probability of failure on demand. Verification of the requirements for the probability of failure on demand is usually based on a quantitative analysis. In this paper we argue that such an approach is better replaced by a semi-quantitative approach. The approach acknowledges that the probability of failure on demand requirement cannot be adequately verified only by reference to an assigned probability number. There is a need for seeing beyond the probability number. The key aspect to include is related to uncertainty.

1 INTRODUCTION

A Safety Instrumented System (SIS) comprises input elements (e.g. pressure transmitters and gas detectors), logic solvers (e.g. relay-based logic and programmable logic controllers) and final elements (e.g. valves, circuits breakers) for the purpose of bringing the plant or an equipment to a safe state if a hazardous event occurs (Lundteigen, 2009). Each SIS has one or more Safety Instrumented Functions (SIF), where every SIF within an SIS has a Safety Integrity Level (SIL). The IEC standards 61508/61511 define four safety integrity levels (SIL 1-SIL 4). The higher the safety integrity level, the more stringent become the requirements. For each safety integrity level there are many design requirements, including requirements for the Probability of Failure on Demand (PFD). The probability of failure on demand for each SIL is given in the IEC standards as shown in Table 1. The levels depend on whether the demand mode of operation is low or high/continuous. Low demand mode embraces systems where the frequency of demands for operation made on safety-related systems is not greater than one per year and not greater than twice the proof-test frequency; otherwise it is classified as a high demand system (IEC, 2003a). An example of a low demand application in subsea production is a down-hole safety valve (DHSV), which remains in open position until a demand occurs. An application in high demand mode can, for example, be the brake system in a car (Rolén, 2007).

Table 1. Safety integrity levels for safety functions.

SIL	Low demand mode	High demand or continuous mode
1	$\geq 10^{-2}$ to $< 10^{-1}$	$\geq 10^{-6}$ to $< 10^{-5}$
2	$\geq 10^{-3}$ to $< 10^{-2}$	$\geq 10^{-7}$ to $< 10^{-6}$
3	$\geq 10^{-4}$ to $< 10^{-3}$	$\geq 10^{-8}$ to $< 10^{-7}$
4	$< 10^{-4}$	$< 10^{-8}$

The IEC standards 61508/61511 require that safety integrity levels for the different safety instrumented functions are verified. Verification of the quantitative part (PFD) of the SIL level for a safety instrumented function is usually done by a calculation of PFD and then by a comparison with the criterion established. If the calculated PFD is higher than the target value, risk reducing measures should be implemented.

This traditional approach for verification of a quantitative SIL seems intuitively appealing. Firstly, a criterion for the probability of failure on demand is given. Then the probability of failure on demand is calculated and compared with the criterion established.

In this paper we do, however, argue that uncertainties should be taken into consideration more extensively than is seen in the traditional approach. The assigned probability for failure on demand is conditioned on a number of assumptions and suppositions. They depend on the background knowledge. Uncertainties are often hidden in the background knowledge, and restricting attention to the assigned probabilities could camouflage factors that could produce surprising outcomes. By jumping directly into probabilities, important uncertainty aspects are easily truncated, meaning that potential surprises could be left unconsidered (Aven, 2008). See also Abrahamsen and Aven (2011) and Abrahamsen et al. (2010). We also find similar ideas underpinning approaches such as the risk governance framework (Renn, 2008) and the risk framework used by the UK Cabinet Office (Cabinet Office, 2002).

In this paper we present and discuss an alternative approach, acknowledging that the calculated probability should not be the only basis for verifying the established quantitative SIL requirements. In the alternative approach the uncertainty aspects are given special attention, and are seen in relation to the assigned probabilities.

The paper is organized as follows. In Section 2 we review and discuss the traditional approach for verification of quantitative SIL requirements. Then, in Section 3, an alternative approach which gives more attention to the uncertainty dimension is presented. Finally, in Section 4, we draw some conclusions.

2 THE TRADITIONAL APPROACH FOR VERIFICATION OF SAFETY INTEGRITY LEVELS

An example from the offshore oil and gas industry is used in this section in order to illustrate the main ideas of the traditional approach for verification of SIL requirements. The example is strongly related to the isolation of subsea well example presented in the OLF-070 Guideline (OLF, 2004).

Isolation of a subsea well is defined as the system needed to isolate one well. For a standard subsea well, the system normally consists of (with reference to Figure 1):

- The emergency shut-down node(s) (ESD), located topside
- Hydraulic bleed down solenoid valves in the hydraulic power unit (HPU), located topside
- Electrical power isolation relays located in the electric power unit (EPU), located topside
- Directional control valves located in the subsea control module (SCM), located topside
- Production wing valve (PWV), production master valve (PMV) and chemical injection valve (CIV) (including actuators) located on the Christmas tree (XT) on the seabed
- Down Hole Safety Valve(s) (DHSV) including actuator(s), located in the well (below seabed)

Isolation of a subsea well can be activated through a hydraulic power unit (HPU) and/or through an electric power unit (EPU); ref. Figure 1.

In the above-mentioned design, the DHSV(s) is/are located in the well below the seabed, the XT is located on the seabed, and the SCM, HPU, EPU and ESD node systems are located topside. The signals are transferred through an umbilical integrated in the production riser. Activation of the safety function will occur if one of the following valve systems is activated:

- DHSV
- PMV
- PWV and CIV

In order to close the DHSV, the directional control valve for DHSV (DCV_{DHSV}) in the control module must be activated. The DCV_{DHSV} is activated from one of the solenoid valves in the hydraulic power unit. The solenoid values are activated from the ESD Node. To close the PMV or PWV and CIV, the same logic as the one described above follows; see Figure 1.

From the OLF-070 Guideline the requirement for the function "ESD isolation of one subsea well" is SIL category 3, which means that the probability of failure on demand (PFD) should not be higher than 10^{-3}, i.e. SIL 3 can be claimed for the safety function presented if the PFD can be demonstrated to be in the range 10^{-4} to 10^{-3}.

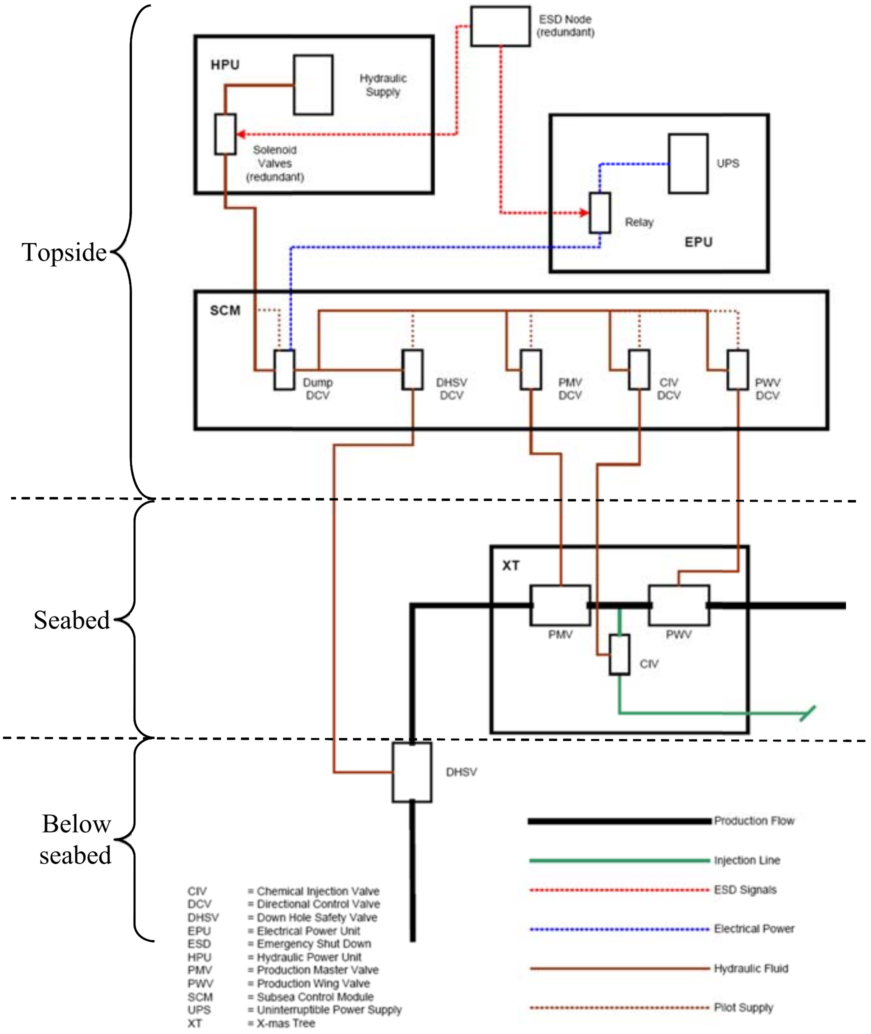


Figure 1. Components included to ensure isolation of one subsea well (typical design based on the OLF-070 Guideline)

The safety function "ESD isolation of one subsea well" can be represented by a reliability block diagram as shown in Figure 2 (OLF, 2004).

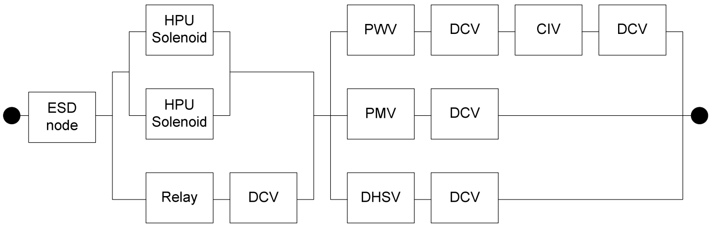


Figure 2. Reliability block diagram for "ESD isolation of subsea well"

Assume that the reliability values for the components included in Figure 2 are as shown in Table 2.

Table 2. Component reliability values used in example calculations (OLF, 2004).

Component	Component redundancy	Calculated PFD
ESD logic	Duplicated	$2.20 \cdot 10^{-4}$
HPU Solenoid	Duplicated	$2.00 \cdot 10^{-4}$
PMV/PWV	Single	$2.20 \cdot 10^{-4}$
CIV	Single	$8.80 \cdot 10^{-4}$
DHSV	Duplicated	$5.50 \cdot 10^{-4}$
DCV	Single	$2.20 \cdot 10^{-4}$
Relay	Single	$1.18 \cdot 10^{-3}$

By using the method shown in the OLF guideline, the calculated system unreliability is $2.2 \cdot 10^{-4}$. Compared to the values presented in Table 1 we conclude that the safety function is within safety integrity level 3, as the calculated PFD is less than 10^{-3} and greater than 10^{-4}.

There are other traditional approaches as well. See for example the approach presented by Hauge et al. (2010). The main idea for verification of the quantitative part of the SIL level is, however, equal; attention is given to the calculated PFD and then compared with a target value.

3 A NEW APPROACH FOR VERIFICATION OF SAFETY INTEGRITY LEVELS

The assigned probability provides a useful insight for decision makers, but there is a need for a broader reflection of uncertainties. The point is that the above calculations express conditional probabilities. In mathematical terms this can be expressed as P(failure on demand|K) where K is the background information and knowledge. The background knowledge covers historical system performance data, system performance characteristics and knowledge about the phenomena in question. Assumptions and presuppositions are an important part of this information and knowledge. The background knowledge can be viewed as frame conditions of the analysis, and the produced probabilities must always be seen in relation to these conditions. Thus, different analysts could come up with different values, depending on the assumptions and presuppositions made. The differences could be very large. Hence, uncertainty needs to be considered, beyond the assigned probability number.

The assigned probability (P) for the safety function should be seen in relation to uncertainties (U). The point is that probability is a tool to express uncertainty. It is, however, not a perfect tool, and we should not restrict verification of SIL only to the probabilistic world. The probabilities are conditional on specific background knowledge (K), and they could produce poor predictions. Surprises relative to the assigned probabilities may occur, and by just addressing probabilities such surprises may be overlooked.

We argue that there are important aspects of uncertainty that should be taken into consideration when a conclusion is made on the SIL level. In particular there are uncertainties on the non-technical aspects that are not taken into consideration in the PFD calculation methods applied by the industry. In the common implementation, there is a close link between the PFD calculation results and the SIL level conclusion. We argue that uncertainties should be taken into consideration before a conclusion is made on the SIL level. In practice, this could be done qualitatively in a workshop subsequent to the quantitative SIL verification analysis, but prior to the SIL level conclusion. This principle is presented in Figure 3 below illustrating both the traditional approach and the approach suggested in this paper. We will come back to an example of how information about the uncertainties could be taken into consideration.

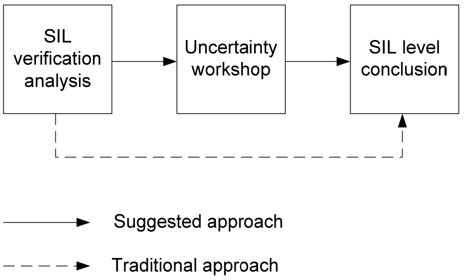


Figure 3. Main principles of the suggested approach

To reflect the uncertainties to the decision makers we recommend that the uncertainties should be classified within one of the three categories: high, medium or low. The categorisation process should be based on some guidelines or criteria to ensure consistency. The following descriptions could serve as a guideline (Flage and Aven, 2009):

Low uncertainty:
All of the following conditions are met:
- The assumptions made in calculations of P are seen as very reasonable
- Much reliable data are available
- There is broad agreement among experts

High uncertainty:
One or more of the following conditions are met:
- The assumptions made in calculations of P represent strong simplifications
- Data are not available, or are unreliable
- There is lack of agreement/consensus among experts

Medium uncertainty:
Conditions between those characterising high and low uncertainty

Note that the degree of uncertainty must be seen in relation to the effect/influence the uncertainty has on the assigned probability. For example, a high degree of uncertainty combined with high effect/influence on the assigned probability number will lead to a conclusion that the uncertainty factor is high. However, if the degree of uncertainty is high but the assigned probability number is relatively insensitive to changes in the uncertain quantities, then the uncertainty classified in the diagram could be low or medium.

As already mentioned, the uncertainty evaluations should be carried out in a workshop. An example of how the results from the workshop could be presented is shown in Table 3.

Based on the discussion in the workshop, documented in Table 3, many aspects with high uncertainty have been identified. The uncertainty factor which is considered most important is 'experience with subcontractors'. The calculated probability number (PFD) is based on the assumption that the subcontractors have a high level of experience from the Norwegian Continental Shelf. This is not necessarily the case. Changes in assumptions related to this factor will have a significant influence on the calculated probability number. The calculated probability may be considered to be less than 10^{-3} even for small changes in the assumptions related to the factor 'experience with subcontractors'.

Table 3. Uncertainty evaluation example

Main categories	**Sub-categories**	**Evaluation**	**Uncertainty categorisation**
Human aspects (M)	Competence and experience	Well-educated personnel. But some operations have never been carried out before by the present crew	High
	Operator training	Operators will be trained in advance to operations being carried out	Medium
Technical aspects (T)	Environmental aspects	Harsh climate at location	Medium
	Internal: Fluid composition	High uncertainties on fluid composition. May result in corrosion and other challenges	High
	New or well-known technology	New equipment: Limited experience with the equipment to be installed subsea	High
	Well characteristics	Challenging well due to high pressures and unknown reservoir characteristics	High
Operational aspects (O)	Experience with subcontractors	New subcontractor (first operation). Limited experience from Norwegian Continental Shelf	High
	Maintenance	No specific challenges identified	Low
	Documentation	No specific challenges identified	Low

With no attention on the uncertainty dimension, we conclude that the SIL requirement is within SIL 3 as the calculated probability number is within the range 10^{-4} to 10^{-3}. Taking the uncertainty dimension into account, the safety integrity level for the safety function considered may be judged not to be within SIL 3, even if the calculated probability is within this category; ref. Figure 4. In this case additional risk reducing measures should be implemented prior to the operation. These could be measures in order to reduce the PFD or means to reduce the uncertainty factors to such an extent that an updated evaluation concludes on SIL3.

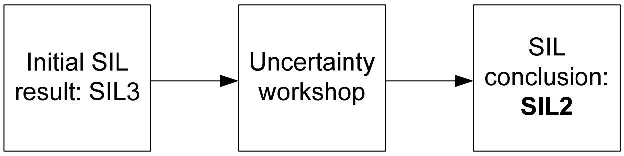


Figure 4. Application example

4 CONCLUSION

The common approach for verification of a safety function's safety integrity level is usually based on probability calculations only. In this paper we argue that such an approach is better replaced by an approach including uncertainty assessment qualitatively in a workshop. This approach acknowledges that the probability requirement for a safety function cannot be adequately verified only by reference to an assigned probability number. There is a need for seeing beyond the probability number. The key aspect to include is related to uncertainty. An example has been included in order to illustrate the ideas.

ACKNOWLEDGEMENT

The work on the paper has been funded by the Research Council of Norway through the RAMONA research programme. The financial support is gratefully acknowledged.

REFERENCES

1. Abrahamsen EB, Aven T. 2011. Safety oriented bubble diagrams in project risk management. International Journal of Performability Engineering 7(1): 91-96.
2. Abrahamsen EB, Aven T, Iversen RS. 2010. Integrated framework for safety management and uncertainty management. Journal of Risk and Reliability 224(2): 97-103.
3. Aven T. 2008. Risk analysis – Assessing uncertainties beyond expected values and probabilities. Wiley: N.J.
4. Cabinet Office. 2002. Risk: improving government's capability to handle risk and uncertainty. Strategy unit report. UK.
5. Flage R, Aven T. 2009. Expressing and communicating uncertainty in relation to quantitative risk analysis. Risk & Reliability – Theory & Application 2(13): 9-18.
6. Hauge S, Lundteigen MA, Hokstad P, Håbrekke S. 2010. Reliability prediction method for safety instrumented systems, PDS method Handbook 2010 edition. SINTEF.
7. IEC – International electrotechnical commission. 2003a. IEC 61508. Functional safety of electric/electronic/programmable electronic safety-related systems. International Electrotechnical Commission; Geneva.
8. IEC – International electrotechnical commission. 2003b. IEC 61511. Functional safety – safety instrumented systems for the process industry. International Electrotechnical Commission; Geneva.

9. Lundteigen MA. 2009. Safety instrumented systems in the oil and gas industry: Concepts and methods for safety and reliability assessments in design and operation. Doctoral thesis at NTNU.
10. OLF – The Norwegian Oil Industry Association. 2004. OLF-070. Application of IEC 61508 and IEC 61511 in the Norwegian petroleum industry. Technical report. The Norwegian Oil Industry Association, Stavanger, Norway.
11. Renn O. 2008. Risk governance: coping with uncertainty in a complex world. London: Earthscan.
12. Rolén H. 2007. Partial and imperfect testing of safety instrumented functions. Master thesis at NTNU.

ON HOW TO CONCEPTUALISE AND DESCRIBE RISK

T. Aven
•
University of Stavanger, Stavanger, Norway
terje.aven@uis.no

ABSTRACT

A number of definitions and interpretations of the risk concept exist. Many of these are probability-based. In this paper we present and discuss a structure for characterising the definitions, which is founded on a clear distinction between (a) risk as a concept based on events, consequences and uncertainties; (b) risk as a modelled, quantitative concept; and (c) risk descriptions. The discussion leads to recommended perspective for conceptualising and assessing risk, which is based on risk defined by (a), and the probability-based definitions of risk can be viewed as related model parameters and/or risk descriptions. Two ways of detailing the framework are outlined: the relative frequency-based approach and the Bayesian approach.

1 INTRODUCTION

Risk is a fundamental concept for most scientific disciplines, but no consensus exists on how to define and interpret risk. Some definitions are based on probabilities, some on expected values, and others on uncertainty. Some consider risk as subjective and epistemic, dependent on the available knowledge, whereas others grant risk an ontological status independent of the assessors. The situation is chaotic and leads to poor communication. We are also afraid that it hampers effective risk management as well as the development of the risk field, as many of these definitions and interpretations lack proper scientific support and justification.

Of course, business needs a different set of risk methods, procedures and models than, for example, medicine and engineering. But there is no reason why these areas should have completely different perspectives on how to think when approaching risk and uncertainty, when the basic challenge is the same---to conceptualise that the future performance of a system or an activity could lead to outcomes different from those desired and planned, or not in line with stated objectives.

Think of an activity in the future, say the operation of an offshore installation for oil and gas processing. We all agree that there is some risk associated with this operation. For example, fire and explosions could occur leading to fatalities, oil spills, economic loss, etc. But it is not straightforward to explain what we mean by this risk if we require a precise definition and would like to use the concept in scientific studies. Risk analysts would introduce a set-up which directly or indirectly defines how risk is understood and assessed. The set-up would typically be probability-based, with probabilities interpreted either as relative frequencies or as subjective probabilities. An example would be the traditional statistical approach which considers risk as a relative frequency-interpreted probability or probability distribution, and the aim of the risk assessment is to accurately estimate this risk using models and hard data. All such set-ups can be challenged, as not being able to reflect risk in a proper way. Important risk aspects could be camouflaged or hidden by the set-up. Discussions of the set-up are therefore important, not only from a theoretical point of view but also from a practical risk management perspective. Many researchers have contributed to this discussion, e.g. Reid (1992) and Stirling (2007). Reid argues that there is a common tendency to underestimate the uncertainties in risk assessments. According to Stirling (2007), using risk assessment when strong knowledge about the probabilities and outcomes does not exist, is irrational, unscientific and potentially misleading. Many other critical comments could have been added, but for the purpose of the present paper it is sufficient to conclude that there is a discussion in the scientific literature about the ability of the set-up of risk assessments to adequately reflect risk.

To be able to make judgments about this issue we need to clarify what risk is and how risk can and should be described. This is the topic of the present paper. A main purpose of the paper is to present a structure for characterising the various definitions of risk in a scientific context. This structure is based on a clear distinction between (a) risk as a concept based on events, consequences and uncertainties; (b) risk as a modelled, quantitative concept; and (c) risk descriptions. Examples of these categories are:

- Uncertainty about the occurrence of future events and their consequences (a)
- Frequentist-interpreted probability P_f of an event (b)
- Estimates of P_f (c)
- A subjective probability P_s (c).

From thus structure we establish a framework that integrates the (a), (b) and (c) definitions to obtain a hierarchy with the a) definitions as the overall risk concept. To further specify the framework we need to distinguish between the relative frequency-based approach and the Bayesian approach. This framework is the main contribution of the present paper.

In the paper we identify several definitions of risk that can be used as an overall, common definition. They all belong to the category (a). Many attempts have been made to establish a unified risk perspective, but none of these have obtained broad acceptance in practice. There could be many reasons for this. Firstly, the scientific work on risk may not have reached a sufficiently mature level for establishing such a definition. The exploring phase is not completed. Secondly, the scientific literature has a focus on the generation of new ideas and suggestions, and on a critique of other contributions. By its nature, it is hard to obtain broad consensus on scientific issues in general and risk definitions in particular. And thirdly, the standardisation organisations have not been able to produce sufficient broad and well-defined definitions which could be accepted by the scientific expertise on risk.

Consider for example the latest proposal from the International Standardisation Organisation (ISO 2009) for defining risk: Risk is the effect of uncertainty on objectives. What does this mean? Risk has to do with uncertainty, but is it the *effect* of uncertainty? And risk is related to objectives, but what if objectives are not defined? Then we have no risk? Asking experts on risk, there is no doubt that this definition would lead to numerous different interpretations. The definition is not sufficiently precise, and one may certainly also question its rationale as indicated.

The present paper is partly based on Aven (2010a). For some reflections on the ontological status of the concept of risk, see Aven et al. (2010).

2 A CLASIFICATION OF RISK DEFINITIONS BASED ON THE PROPOSED STRUCTURE

As stressed above, there exist a number of definitions of risk. Here are some typical examples (list based on Aven and Renn 2009a):

1) Risk equals the expected loss (Verma and Verter 2007, Willis 2007).
2) Risk equals the expected disutility (Campbell 2005).

3) Risk is a measure of the probability and severity of adverse effects (Lowrance 1976).
4) Risk is the combination of probability and extent of consequences (Ale 2002).
5) Risk is equal to the triplet (s_i, p_i, c_i), where s_i is the *i*th scenario, p_i is the probability of that scenario, and c_i is the consequence of the ith scenario, i =1,2, …N (Kaplan and Garrick 1981).
6) Risk refers to uncertainty of outcome, of actions and events (Cabinet Office 2002).
7) Risk is a situation or event where something of human value (including humans themselves) is at stake and where the outcome is uncertain (Rosa 1998, 2003).
8) Risk is an uncertain consequence of an event or an activity with respect to something that humans value (Renn 2005).
9) Risk is the effect of uncertainty on objectives (ISO 2009).
10) Risk is equal to the two-dimensional combination of events/consequences and associated uncertainties (Aven 2007, Aven 2008).
11) Risk is uncertainty about and severity of the consequences (or outcomes) of an activity with respect to something that humans value (Aven and Renn 2009a).

For the measures that are based on probabilities and expected values, we may generate two versions, one where the probabilities are interpreted as relative frequencies (and the expected values as averages), and one where the probabilities are subjective probabilities (and the expected value is interpreted as the centre of gravity of the probability distribution). We write definitions x_f and x_s, respectively, to separate the two categories, x =1, 2, …, 5. Consider as an example category 1, risk defined as the expected loss. According to definition 1_f, risk is understood as the average loss when considering an infinite number of similar situations, whereas 1_s means that risk is the centre of gravity of the subjective probability distribution of the loss. Following the suggested structure for characterising the various risk definitions we have to place these definitions in one of the categories (a), (b) (c), defined in the previous section.

The result is that definition 1_f is in category (b) and 1_s is in category (c), as risk in the former case is based on the model of an infinite number of similar situations and risk in the latter case is a way for the assessor to describe or characterise risk. The expected loss E_s when using subjective probabilities is a risk index based on the background knowledge (K) of the assessor.

This is in line with the rejection of risk as being defined by the expected value, as argued in, for example, Haimes (2004) and Aven (2010b). The expected value does not adequately capture events with low probabilities and high consequences. Take as examples nuclear accidents and terrorism risk, where the possible consequences could be extreme and the probabilities are relatively low. The expected value can be small, say 0.01 fatalities, but extreme events with millions of fatalities may occur, and this needs special attention.

A similar analysis is carried out for the other ten definitions. The result is shown in Table 1.

Table 1. Categorisation of the 11 risk definitions according to the structure (a) – (c). For definition 2 it is assumed that the expectation is taken with respect to a subjective probability distribution.

Risk definition	Category
1_f	b
1_s	c
2	c
3_f	b
3_s	c
4_f	b
4_s	c

Risk definition	Category
5_f	b
5_s	c
6	a
7	a
8	a
9	a
10	a
11	a

Some comments are in place for the various definitions (2-11).

The second definition considers risk as the expected disutility, i.e - Eu(C), where C is the outcomes (consequences) and u(C) the utility function (Campbell 2005). The expectation is based on subjective probabilities. According to this definition, the preferences of the decision-maker are a part of the risk concept. In our view, and this view is shared by many risk experts, the preferences and values should not be a part of the risk concept and the risk assessments (Paté-Cornell 1996). There will be a strong degree of arbitrariness in the choice of the utility function, and some decision-makers would also be reluctant to specify the utility function as it reduces their flexibility to weight different concerns in specific cases. Risk should be possible to describe also in cases where the decision-maker is not able or willing to define his/her utility function.

Definitions 3-5 are all probability-based. The concept of risk comprises events (initiating events, scenarios), consequences (outcomes) and probabilities. Severity is a way of characterising the consequences, and refers to intensity, size, extension, scope and other potential measures of magnitude, and affects something that humans value (lives, the environment, money, etc.). Losses and gains, for example expressed by money or the number of fatalities, are ways of defining the severity of the consequences (Aven and Renn 2009a).

If relative frequency-interpreted probabilities P_f constitute the basis (definitions 3_f, 4_f and 5_f), risk is a modelled, quantitative concept (category (b)) and we may formalise the definitions by writing

$$\text{Risk} = (A,C,P_f),$$

where A represents the events (initiating events, scenarios) and C the consequences of A. Examples of events A are: gas leakage occurring in a process plant, and the occurrence of a terrorist attack. Examples of C are the number of casualties due to leakages, terrorist attacks, etc.

If on the other hand subjective probabilities constitute the basis (definitions 3_s 4_s and 5_s, the definitions must be viewed as risk descriptions as they express the analysts' degree of belief concerning A and C. Also the background knowledge K that the probabilities are based on, should be considered a part of the risk description.

A quick look at definitions 6-11 may give the impression that they are not that different from 3-5. However, there are important principle differences, as will be clear from the coming analysis. Probability is just a tool used to represent or express the uncertainties. The thesis of all the perspectives and definitions 6-11 is that risk should not be limited to (A,C,P). The uncertainties should be highlighted. Consider first definition 10, which we simply refer to as the (A,C,U) definition. Definition 11 may be viewed as a reformulation of this definition, based on the same ideas.

We consider an activity in the future, and something that humans value is at stake (lives, the environment, etc.). Undesirable (and desirable) events and consequences could occur. There are uncertainties about the occurrence of the events, and what will be the consequences (outcome) of these events if they should occur. How many will be killed? What will the value of the stock be?

Risk has two main components: i) the events and their consequences, and ii) uncertainty about these - will the events occur and what will the consequences be? These two components define risk.

According to definition 6, risk refers to uncertainty of outcome, of actions and events (Cabinet Office 2002). Hence, strictly speaking risk is not (A,C,U) but only U. As an example, consider the number of fatalities in traffic next year in a specific country. Then the uncertainty is rather small, as the number of fatalities shows rather small variations from year to year. Thus following this definition of risk, we must conclude that the risk is small, even though the number of fatalities is many thousands each year. Clearly, this definition of risk fails to capture an essential aspect, the consequence dimension. Uncertainty cannot be isolated from the intensity, size, extension etc. of the consequences.

According to definition 7, risk is a situation or event where something of human value (including humans themselves) is at stake and where the outcome is uncertain (Rosa 1998, 2003). Hence, strictly speaking risk is A, and not (A,C,U). However, Rosa expresses risk using the description (A,C,U), and refers to probability as a tool to describe the uncertainties. The Rosa (1998, 2003) definition is thoroughly discussed by Aven and Renn (2009a). The conclusion is that compared to common terminology, the Rosa definition leads to conceptual difficulties that are incompatible with the everyday use of risk in most applications. By considering risk as an event (A), we cannot conclude, for example, about the risk being high or low, or compare different options with respect to risk. The same conclusion is made for definition 8, which says that risk is an uncertain consequence of an event or an activity with respect to something that humans value (Renn 2005). This definition is similar to Rosa's definition but the event A is replaced by the consequence C.

We have already commented on definition 9 that risk is the effect of uncertainty on objectives (ISO 2009). This definition seems to be in line with the two previous definitions. Alternatively, we may interpret the suggested ISO definition as (A,C,U), where the consequences (the effect) are seen in relation to the objectives.

Based on the perspectives and definitions 6-11, various types of risk descriptions (category (c)) can be specified, for example by using subjective probabilities. But we can also introduce modelled, quantitative risk concepts (category (b)). The result is a hierarchy of concepts which together provide a holistic system for conceptualising and describing risk. The next section will present the details of this system.

3 RECOMMENDATIONS FOR HOW TO CONCEPTUALISE AND DESCRIBE RISK

If we search for a widespread agreement on one definition we have to look among the categories (a). The others have to be excluded as they are based on either a model or an assignment of uncertainty using the tool, subjective probability. Risk should also exist as a concept without modelling and subjective probability assignments. We face risk when we drive a car or run a business, also when probabilities are not introduced. For risk assessment we need the probabilities, but not as a general concept of risk. In this way we obtain a sharp distinction between risk as a concept and risk descriptions (assessments).

The discussion in the previous section led to two candidates among the a-definitions; the (A,C,U) definitions (10-11) and the (A,C) definitions (7-8). The latter group means that the common risk terminology has to be revamped and we therefore prefer to use the (A,C,U) definition.

We will use this risk concept as a pillar for a recommended framework for conceptualising and describing risk. The next stage would then be to specify how to describe risk in this framework. To be able to do this we need to distinguish between a relative frequency-based approach and a Bayesian perspective as will be demonstrated by the following analysis. The main elements of the frameworks for these two approaches are shown in Figures 1 and 2. We first look at the relative frequency case.

In this case we introduce relative frequency-interpreted probabilities P_f. These are in general unknown. Risk assessment is introduced to describe the risk, to estimate P_f. The description covers an estimate P_f^* of P_f, as well as assessments of uncertainties about P_f^* and P_f. Thus, if the relative frequency perspective to risk is the starting point, we are led to a risk description:

i) Risk description in the relative frequency case = $(A,C,P_f^*,U(P_f^*),U, K)$,

where $U(P_f^*)$ refers to an uncertainty description of P_f^* relative to the true value P_f, U refers to uncertainty factors not covered by $U(P_f^*)$, and K is the background knowledge that the estimate and uncertainty description is based on. We may refer to $U(P_f^*)$ as a second-order uncertainty description.

One way of reflecting $U(P_f^*)$ is to use confidence intervals. These intervals describe the variation in the data available, but do not reflect other types of uncertainties, in particular uncertainties as a result of more or less relevant data.

If we use subjective probabilities P_s to express our uncertainties about P_f, the risk description takes the form:

i)' Risk description = $(A,C,U, P_f^*,P_s(P_f),K)$,

where K now is the background knowledge that the estimate P_f^* and the probability distribution P_s is based on. Kaplan & Garrick (1981), see also Kaplan (1997), refer to this distribution as the second level definition of risk – it is combined with the first level (A,C,P) definition. When including the second level definition the perspective is referred to as the probability of frequency approach. The risk description i)' can be viewed as an extended probability of frequency approach, as it covers all the elements of the probability of frequency approach and in addition address uncertainties U not reflected by the P_s.

The U covers in general factors not included in $U(P_f^*)$ or $P_s(P_f)$. Examples include the relevancy of the data when using confidence intervals and the fact that the subjective probabilities could produce poor predictions. The background knowledge K could be poor. Probability assignments are conditioned on a number of assumptions and suppositions, and these could turn out to be wrong. One may assign a low probability of health problems occurring as a result of some new chemicals, but these probabilities could produce poor predictions of the actual number of people that experience such problems. Or one may assign a probability of fatalities occurring on an offshore installation based on the assumption that the installation structure will withstand a certain accidental load; in real-life the structure could however fail at a lower load level: the assigned probability did not reflect this uncertainty.

The analysts need to clarify what is uncertain and subject to the uncertainty assessment and what constitutes the background knowledge. From a theoretical point of view, one may think that it is possible (and desirable) to remove all such uncertainties from the background knowledge, but in a practical risk assessment context that is impossible. The assessment of the uncertainty factors would normally be qualitative, see approach indicated below.

Next we consider the Bayesian case. A risk description based on this definition would cover the following components:

ii)' Risk description = (A,C,U,P_s,K),

where P_s is a subjective probability expressing U based on the background knowledge K. This description covers probability distributions of A and C, as well as predictions of A and C, for example a predictor C* given by the expected value of C, unconditionally or conditional on the occurrence of A, i.e. C* = EC or C* = E[C|A].

Using the description ii)' there are no second-order probabilities, as talking about uncertainties of a subjective probability has no meaning. A subjective probability P(A)= P(A|K) is interpreted as a knowledge-based probability with reference to a standard expressing the analysts' uncertainty about the occurrence of the event A given the background knowledge K. Following this interpretation the assessor compares his/her uncertainty (degree of belief) about the occurrence of the event *A* with the standard of drawing at random a favourable ball from an urn that contains *P*(*A*) · 100 % favourable balls (Lindley 2006). The traditional betting interpretation of a subjective probability (Singpurwalla 2006) can also be used, but we prefer the reference to a standard definition as it does not mix uncertainty assessments with our attitude to money (Aven 2003, Aven 2010c). According to the betting interpretation, the probability of an event is the price at which the person assigning the probability is neutral between buying and selling a ticket that is worth one unit of payment if the event occurs, and worthless if not (Singpurwalla 2006).

Also in the Bayesian context we establish relative frequencies, but they are referred to as chances and not probabilities. A chance is the limit of a frequency of similar (formally exchangeable) random events. More generally we introduce probability models with unknown parameters. A chance is an example of such a parameter. By the Bayesian updating machinery, knowledge about the parameters is described first by the prior distribution, then updated to produce the posterior distribution to reflect observations. Finally, this distribution is used to generate the predictive distribution of the events A and consequences C. These predictive distributions then incorporate the variation reflected by the probability model (and the chances) and the epistemic uncertainties about the true value of the parameters.

If probability models and chances are introduced, the Bayesian approach looks similar to the extended probability of frequency approach. However, there is a difference. In the relative frequency case, probabilities P_f always need to be defined. They constitute the foundation of the approach. In the Bayesian case, chances are only defined when exchangeable sequences can be justified. Chances need some sort of model stability (Bergman 2009): populations of similar units need to be constructed (formally an infinite set of exchangeable random variables). We will for example not define a chance p of an attack (Aven & Renn 2009b). It has no meaning. Subjective probabilities can however be used.

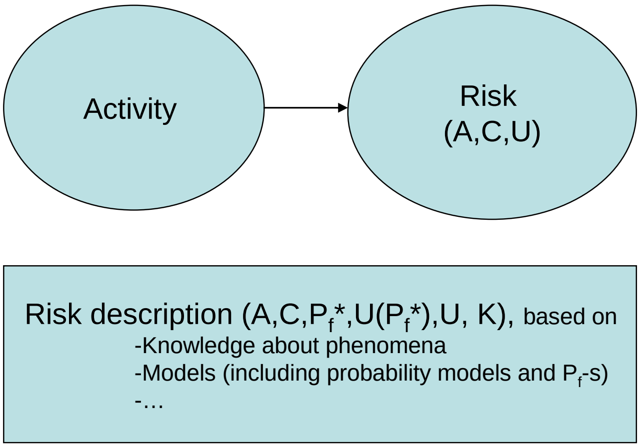


Figure 1. The main elements of the recommended risk framework when it is based on relative frequencies

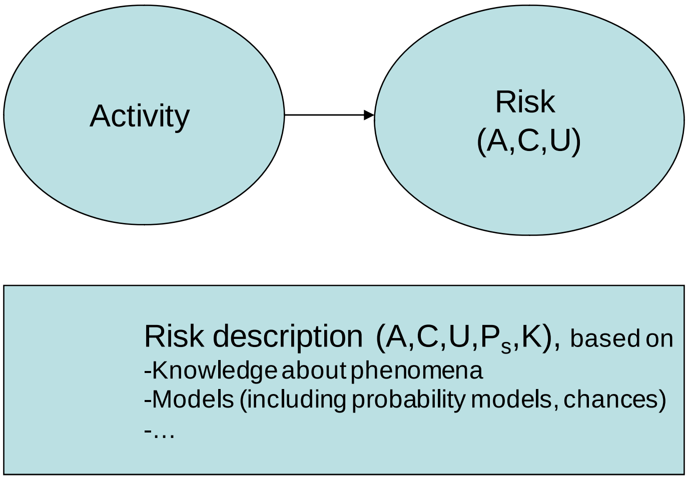


Figure *2.* The main elements of the recommended risk framework when it is based on the Bayesian approach

Models, including probability models, are used in both cases. Instead of estimating P_f or chances we estimate g(q), where g is the model and q is a vector of parameters of the model. An event tree and a fault tree are two simple examples of such models. We may also use models to simplify and/or give rigour to the specification of the subjective probabilities P_s.

The U in the risk description covers an assessment of uncertainties that are not captured by the probabilistic analysis. A simple approach for how to do this assessment is outlined in Flage and Aven (2009). Assume that the events A are described by probabilities P(A|K) and conditional expected impact values given the occurrence of A, E[C|A,K], are determined. For each A, a list of uncertainty factors is identified. These factors are assessed with respect to uncertainty and classified as “high” if one or more of the following conditions are met:

- The assumptions made represent strong simplifications.
- Data are not available, or are unreliable.
- There is lack of agreement/consensus among experts
- The phenomena involved are not well understood

In addition, the expected impact values must be sensitive to changes of the factor. The point is that a factor could be subject to large uncertainties but it is not considered important unless changes in the factor affect the impact.

4 CONCLUSIONS AND FINAL REMARKS

In this paper we have presented a new framework for conceptualising and assessing risk. Compared to earlier analyses of the risk concept (Aven 2009a,b, Aven and Renn 2009a), the framework provides a structure for both the relative frequency approach and the Bayesian approach within the same overall risk concept (A,C,U). In this way two approaches are developed and specified. There could be different opinions on which approach should be preferred, but only the Bayesian approach would work in cases where relative frequency-interpreted probabilities (chances) cannot be meaningfully defined. In this sense the Bayesian approach is more general than the relative frequency approach.

By this framework it is acknowledged that risk is more than probabilities, probability distributions and expected values. The uncertainty dimension of risk extends beyond the probabilities. In this way the framework provides important input for making judgments about the quality of risk assessments. If a risk assessment is restricted to probabilities, important aspects of risk may be overlooked.

It may be a challenge to reveal and describe all the uncertainties. Qualitative approaches can be used, and further research is required to develop methods for proper identification and analysis of the uncertainties. But this is not the issue here. In this paper we address the overall conceptual structure of risk and risk assessment, not the analysis methods. Before such conceptual structures can be established it is difficult to develop suitable methods, as the methods would depend on the aim of the analyses. Risk analysis is a young discipline and has been characterised by many weakly justified risk perspectives and also by lack of consistency in approaches. The aim of the present paper has been to contribute to rectifying these problems by suggesting an overall holistic framework for conceptualising and assessing risk that could provide improved structure and guidance on how to think in a risk analysis context.

Acknowledgments

The work has been funded by The Research Council of Norway through the SAMRISK and PETROMAKS research programmes. The financial support is gratefully acknowledged.

References

1. Ale, B.J.M. 2002. Risk assessment practices in The Netherlands. *Safety Science* 40:105-126.
2. Aven, T. 2003. *Foundations of Risk Analysis*. Chichester: Wiley.
3. Aven, T. 2007. A unified framework for risk and vulnerability analysis and management covering both safety and security. *Reliability Engineering and System Safety* 92: 745-754.
4. Aven, T. 2008. *Risk Analysis*. Chichester: Wiley.
5. Aven, T. 2009a. Perspectives on risk in a decision-making context – Review and discussion. *Safety Science* 47: 798–806.
6. Aven, T. 2009b. Safety is the antonym of risk for some perspectives of risk. *Safety Science*, 7, 925–930.
7. Aven, T. 2010a. A holistic framework for conceptualising and describing risk. In Kolowrocki, K., Soszynska, J. & Zio, E. (eds.) Proceedings SSARS 2010, Journal of Polish Safety and Reliability Association. Vol 1, pp. 7-14.
8. Aven, T. 2010b. *Misconceptions of Risk*. Chichester: Wiley.
9. Aven, T. 2010c. On the need for restricting the probabilistic analysis in risk assessments to variability. *Risk Analysis*, 30(3):354–360. With discussions 381-384.
10. Aven, T. & Renn, O. 2009a. On risk defined as an event where the outcome is uncertain. *Journal of Risk Research*, 12: 1-11.
11. Aven, T. & Renn, O. 2009b. The role of quantitative risk assessments for characterizing risk and uncertainty and delineating appropriate risk management options, with special emphasis on terrorism risk. *Risk Analysis*. 29: 587-600.
12. Aven, T., Renn, O. & Rosa, E. 2010. On the ontological status of the concept of risk. Paper submitted for possible publication. .
13. Aven, T. & Vinnem, J.E. 2007. *Risk Management, with Applications from the Offshore Oil and Gas Industry*. Springer Verlag. N.Y.
14. Bergman, B. 2009. Conceptualistic pragmatism: a framework for Bayesian analysis? *IIE Transactions*, 41: 86-93.

15. Cabinet Office 2002. *Risk: improving government's capability to handle risk and uncertainty*. Strategy unit report. UK.
16. Campbell, S. 2005. Determining overall risk. *Journal of Risk Research*, 8: 569-581.
17. Flage, R & Aven, T. 2009. Expressing and communicating uncertainty in relation to quantitative risk analysis. *Reliability & Risk Analysis: Theory & Application*, 2 (13): 9-18.
18. Haimes, Y.Y. 2004. *Risk Modelling, Assessment, and Management*. 2nd ed. N.Y: Wiley.
19. ISO 2009. Guide 73:2009 Risk management — Vocabulary.
20. Kaplan, S. 1997. Words of risk. *Risk Analysis*, 17: 407-417.
21. Kaplan, S. & Garrick, B.J. 1981. On the quantitative definition of risk. *Risk Analysis,* 1: 11-27.
22. Lindley, D. 2006. *Understanding Uncertainty*. Chichester: Wiley.
23. Lowrance, W. 1976. *Of Acceptable Risk – Science and the Determination of Safety*. Los Altos, CA:
24. William Kaufmann Inc.
25. Paté-Cornell, M.E. 1996. Uncertainties in risk analysis: Six levels of treatment, *Reliability Engineering and System Safety* 54(2-3): 95-111.
26. Reid, S.G. 1992. Acceptable risk. In: D.I. Blockley (ed.): *Engineering Safety*, New York: McGraw-Hill, pp. 138–166.
27. Renn, O. 2005. Risk Governance. White paper no. 1, International Risk Governance Council, Geneva.
28. Rosa, E.A. 1998. Metatheoretical Foundations for Post-Normal Risk. *Journal of Risk Research*, 1: 15-44.
29. Rosa, E.A. 2003. The Logical Structure of the Social Amplification of Risk Framework (SARF); Metatheoretical Foundations and Policy Implications. In: N. Pidgeon, R.E. Kasperson and P. Slovic (eds.): *The Social Amplification of Risk*. Cambridge University Press, Cambridge, UK, pp. 47-79.
30. Singpurwalla, N. 2006. *Reliability and Risk. A Bayesian Perspective*. Chichester: Wiley.
31. Stirling, A. 2007. Science, Precaution and Risk Assessment: towards more measured and constructive policy debate. *European Molecular Biology Organisation Reports* 8: 309-315.
32. Verma, M. & Verter, V. 2007. Railroad transportation of dangerous goods: Population exposure to airborne toxins. *Computers & Operations Research* 34:1287-1303.
33. Willis, H.H. 2007. Guiding resource allocations based on terrorism risk. *Risk Analysis* 27(3): 597-606.

RISK ASSESSMENT OF AIRCRAFT CRASH ONTO A NUCLEAR POWER PLANT

H.-P. Berg
•
Bundesamt für Strahlenschutz, Salzgitter, Germany
e-mail: hberg@bfs.de

ABSTRACT

External hazards can provide safety significant contributions to the risk in case of nuclear power plant operation because such hazards have the potential to reduce simultaneously the level of redundancy by damaging redundant systems and lines or their supporting systems. Therefore, risk assessment of all potential external hazards to the plant under consideration is part of the overall safety assessment. In this paper, the procedure for assessing the external hazard aircraft crash is described in more detail. The first step is an appropriate screening procedure in order to determine scope and content of the assessment, taking into account plant- and site-specific conditions. The second step is to determine the methodical approach for those cases where a full scope analysis has to be performed and the inclusion into the used overall risk model. The considerations regarding this hazard do not cover an intended aircraft crash.

1 INTRODUCTION

International experience has shown that internal hazards such as fire and flooding as well as external hazards such as earthquakes, flooding and air craft crash can be safety significant contributors to the risk in case of plants with the potential high hazardous risk such as process plants or nuclear power plants. This risk results from the fact that such hazards potentially can reduce simultaneously the level of redundancy – implemeted for increasing the overal reliability and safety of the plant – by damaging redundant components, systems and lines or their supporting systems (energy, water etc.).

Therefore, arrangements should be implemented by the operator of the plant for assessing the vulnerability of plant and structures, determining how the safe operation of a plant is affected, and introducing measures to prevent the hazard at all, to prevent that it develops and to mitigate against its effects in case it nevertheless develops. These arrangements and their effectiveness and efficiency has to be justified to the regulatory body and approved.

Methods to analyse operating plants with a higher risk potential systematically with respect to the adequacy of their existing safety protection equipment against hazards can be deterministic as well as probabilistic.

In particular in case of probabilistic analyses, the assessment can be very detailed and time consuming. Therefore, it is necessary to develop procedures to screen out, e.g., rooms or buildings of a plant where no further analysis is required or to have a graded procedure for the respective hazard taking into account plant- and site-specific conditions.

Since October 2005, a revised guideline as well as revised and extended technical documents are issued in Germany which describe the methods and data to be used in performing probabilistic safety assessment in the frame of comprehensive safety reviews for nuclear power plants (Berg 2005) which have to be performed every ten years to achieve a current overall snapshot of the safety level of the plant.

In these documents, probabilistic considerations of aircraft crash, external flooding, earthquake and explosion pressure waves are required and described in (FAK PSA 2005). Also on international level, new recommendations are issued such as the Safety Guide on Level 1 PSA (IAEA 2010) and the updated US Standard (ASME 2008) taking into account external hazards to a larger extent.

In this paper, the procedure for probabilistic assessment of the external hazard aircraft crash is described in more detail. The first step is to define an appropriate screening procedure in order to determine scope and content of the assessment to be performed. The second step is to determine the methodical approach for those cases where a full scope analysis has to be performed. The third step is to include this hazard analysis into the used overall risk model of the plant under consideration.

The consideration regarding the external hazard aircraft crash usually does not cover an intended aircraft crash for already operating plants.

However, the aspect of an intended aircraft crash onto a plant with a potential high hazardous risk is also discussed since many years although a probabilistic evaluation of the behaviour of terrorists is very difficult to address.

As a result of the difficulties to adequately model an intended aircraft crash, new nuclear power plant projects take into account such an event in their design. For example, in the design of the European Pressurized Reactor as currently constructed in Finland (see Figure 1), the reactor building as well as the surrounding auxiliary buildings which house safety related equipment are structurally strengthened to the extent needed for surviving the impact of a large commercial aircraft (Kuznetsov 2007). Moreover, redundant safety equipment is spatially separated in a strictly manner.

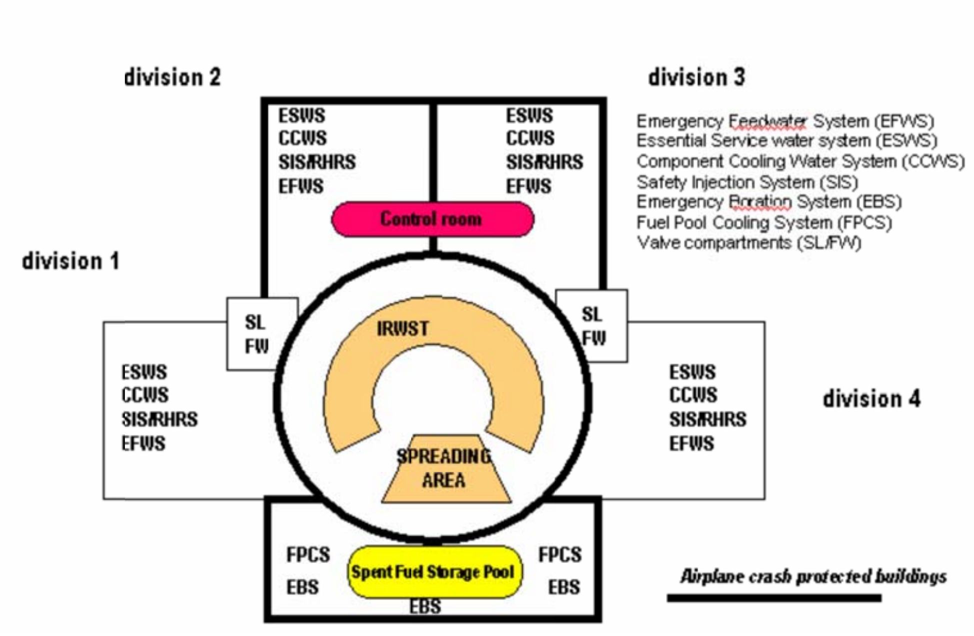


Figure. 1. Layout of the European Pressurized Reactor showing features for protection against aircraft impact

Both, crashes of military aircrafts and of commercial aircrafts contribute to the plant risk. The location of the nuclear power plant under consideration is important, both with respect to the distance from a nearby airport and to a close-by air lane. Moreover, it has to be taken into account whether the plant is situated in an area of landing and take-off traffic.

2 FLIGHT SITUATION IN GERMANY

Because of the central position of Germany within Europe, there is a close-meshed net of civil air lanes with a high density of flights. Although the military flight activities have changed after 1990 due to the new political situation in Eastern Europe, German and, in particular, US Air Force units are stationed in Germany and, in addition, a lot of air traffic resulting from military units stationed outside Germany but crossing the German airspace has to be considered. Thus, there might be a non-negligible hazard due to aircraft crashing onto nuclear sites.

German nuclear power plants can be divided in three generations with respect to air craft crash depending on the load assumptions which had been the basis for the structural design of the building structures to protect against hints of aircrafts or wracked aircraft parts.

The consequences of hints in case of buildings which are not protected depend on the plant specific layout of buildings and systems, in particular the missing strict spatial separation of redundant safety equipment.

These design differences are reflected and evaluated within the safety assessment performed in the frame of comprehensive (periodic) safety reviews.

3 SCREENING

In the following guidance is given to perform a probabilistic safety analysis of nuclear power plants for the initiating event aircraft crash. A conservative approach in form of a rough analysis is described which allows the estimation of an upper limit for the frequency of plant hazard states caused by an aircraft crash.

Further methods are described which are appropriate to replace the conservative considerations of the rough analysis by more detailed validation procedures. Application of these methods with a larger analysis effort lead to a more realistic validation compared to the rough analysis.

Requirements with respect to aircraft crash are laid down in a document of the German Reactor Safety Commission (RSK 1984). A load function for buildings to be protected (reactor building etc.) has been defined mainly based on theoretical calculations assuming an impact of the military aircraft “Phantom F4” (see Figure 2) which was the mostly used aircraft in the military fleet at that time.

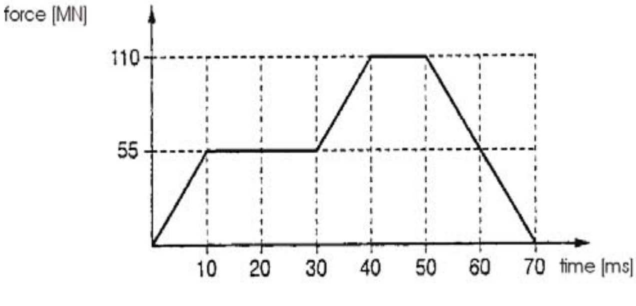


Figure 2. Load time diagram

Table 1 provides on overview of the graded screening process with respect to aircraft crashes applied in probabilistic safety assessments in the frame of periodic safety reviews for German nuclear power plants (Berg 2010a).

Table 1. The graded process of evidence regarding aircraft crash impact.

Criterion	Extent of analysis
Structures are designed according to the diagram in Fig. 1 and not located in a military zone for fly maneuver drills	Analysis is not necessary
Contribution is negligible compared to the other contributions in the PSA	A conservative rough-analysis regarding the consequences of impact on important areas A, B, C where A: e.g. primary circuit B: e.g. turbine building C: separated emergency building
Not negligible	Detailed probabilistic analysis of all plant areas, e.g. by using Monte-Carlo-methods

4 DETERMINATION OF THE FREQUENCY OF AN AIRCRAFT CRASH

The plant-specific determination of the frequency for the occurrence of an aircraft crash is performed on the basis of flight accident statistics valid for the respective location, taking into account the types of aircrafts and the weight classes which can be set.

The following input information is needed:

- the air traffic lanes in the near field of the plant,
- data concerning civil and military small and middle airports (in the range of about 50 km) and large airports (in the range up to 150 km) such as distance and adjustment of the starting and take-off runways.

The crash frequencies are determined separately in three different traffic categories:

- The landing and take-off phase,
- the air lane traffic and waiting loop traffic,
- the free air traffic.

The aircrafts can be grouped into different weight classes. One example is shown in Table 2. Furthermore, the weight classes can be correlated to accidents.

Table 2. Aircraft crash rates outside the airports onto the ground according to (Hoffmann et al. 1997).

Weight class (Mg)		Aircraft crashes per km flown
1	> 20	$2.08 \cdot 10^{-10}$
	5.7 – 20	$3.21 \cdot 10^{-09}$
2	2 – 5.7	$5.44 \cdot 10^{-08}$
3	< 2	$1.11 \cdot 10^{-07}$

The annual frequency of impact for each weight class and flight phase is calculated, based on global crash rates valid for Germany as reported in (BFU 2010) and further processed as shown in see Table 2, the local flight density and the impact area of the plant or the silhouette areas of the buildings.

Conservatively, crash angles are assumed as 30° to the horizon. The silhouette areas result from the projection of the crash angle over the up-stations of the building onto the power plant surface (over the four directions). However, the definition of the silhouette area is not unique and allows different calculations as shown in equations (1) and (2), illustrated in Figure 3.

$$F_{Za} = 6\sqrt{3} + 2 \approx 12 \qquad (1)$$

$$F_{Zb} = 6\sqrt{3} + 2 + 3\pi \approx 22 \qquad (2)$$

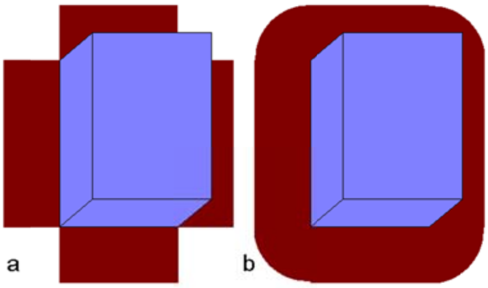


Figure 3. Determination of the silhouette area

As already explained, further hits can result from wracked aircraft parts, even if the aircraft crashes outside the defined silhouette areas.

An area of 1000 m outside the silhouette areas is assumed as a possible hazard range. Tthe probability of a hit by a wracked part is estimated as 20%. More than one hit per crash is not assumed.

Relevant for the generation of such wracked parts are mainly aircrafts of the weight classes 1 and 2 as well as fast flying military aircrafts. The wracked part is treated like an aircraft of the weight class 3.

5 CRASH FREQUENCY IN COMMERCIAL AIR TRAFFIC

Most of the air crashes occur during take-off and landing phases of flight. Hence existence of an airport in the site vicinity increases the potential of aircraft crash hazard on the plant under consideration (e. g.. a nuclear power plant or a chemical plant). The location of site with respect to the distance to major and minor air fields, including military airports is identified. In case of establishing a new installation with a higher hazardous risk, some countries like the Netherlands and United Kingdom define prescribed quantitative criteria (Berg 2010b). If the distances of the planned plant are greater than the prescribed value for respective types of airports, the location of site is considered as acceptable.

If the site falls within this prescribed value for different types of airfields, a comprehensive probabilistic study of aircraft crashes onto the installation considering flight frequencies are carried out. If this probability is not acceptably low the site is considered unsuitable for establishing the planned installation.

With the advances made in assessing the hazards from aircraft impact including the effects of mechanical impact, fire and explosion, one can now estimate the consequences of impact. However, this needs to be investigated in detail in a case-by-case basis.

Up to 10 km from the end of the runway, the crash rate decreases exponentially with distance R. For each weight class and different angular segment (or sector) of flight directions, this decrease has been calculated based on data of worldwide crash vectors (see Figure 4).

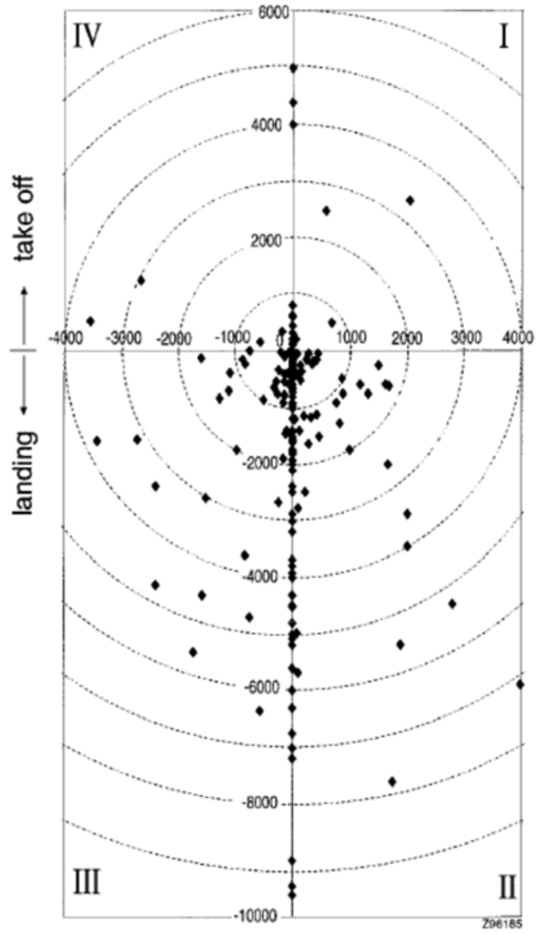


Figure 4. Worldwide crash vectors of aircraft crashes during landing and take-off phase (weight class 2).

The landing range before the touch-down and the take-off range after the start are subdivided into three different sectors (see Figure 5):

Sector 1: ± 15° to the landing and take-off axis (30° sector angle),

Sector 2: outside of segment 1 to ± 45° to the landing and take-off axis (90° sector angle),

Sector 3: outside of sector 2 to ± 90° to the landing and take-off axis (180° sector angle).

The number of crashes $h_{i,j}$ within definite angular sectors are summed up for rings of ΔR= 1000 m and approximately expressed by the following formula (Fürste & Glupe 1974):

$$h_{i,j} = \frac{a_{i,j}}{c_i} \cdot \exp\left(-b_{i,j} \cdot R\right) \tag{3}$$

with

i = Weight class,

j = Annular segment with corresponding C_j as portion of the total sector as provided in Table 3,
$a_{i,j}$; $b_{i,j}$ = Constants to approximate the observed number of crashes in weight class i and sector j,
c_i = number of landing and take-offs in the weight class I and within the observation time.

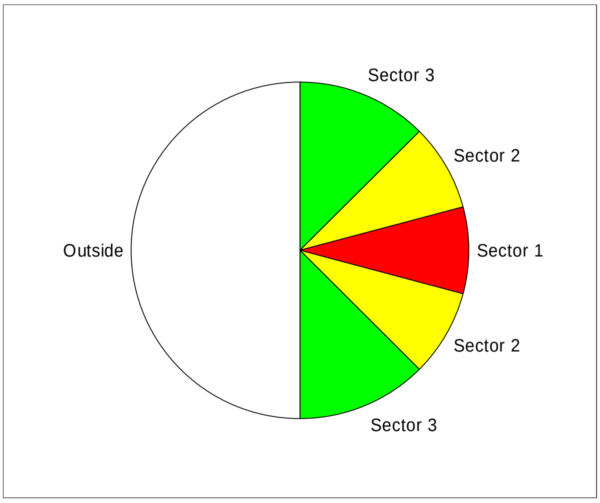
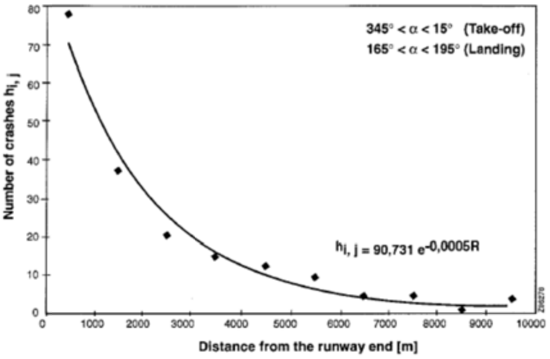


Figure 5. Subdivision of sectors for the landing and take-off range

Based on the number of yearly flying operations (take-offs and landings) of the airport to be considered, the number of the yearly crashes $H_{i,j}$ can be calculated for an impact area of the plant in a given annulus:

$$H_{i,j} = h_{i,j} \cdot \frac{d_i}{d_{global,i} \cdot \Delta t} \cdot \frac{F_{NPP}}{F_\alpha} \tag{4}$$



$$H_{i,j} = h_{i,j} \cdot \frac{d_i}{d_{global,i} \cdot \Delta t} \cdot \frac{F_{NPP}}{F_{\alpha i}}$$

$h_{i,j}$ = Number of crashes within definite angular segment j at distance R (ΔR = 1000 m) and weight class i (Δt = 19 years)
d_i = Number of flying operations per year at the airport considered
$d_{global,i}$ = Number of global flying operations per year
F_{NPP}, $F_{\alpha i}$ = Area of NPP, annular segment
i, j = Weight class, annular segment
Δt = Time span analysed (19 years)
$H_{i,j}$ = Number of theoretical yearly crashes within the NPP area

Figure 6. Exponential decrease of the number of worldwide crash vectors in the annular segment with the distance from the runway

One result for an observation time of 19 years and the number of global flying operations per year according to (ICAO 1995) is shown in Figure 6.

For the calculation of the crash frequency at lateral locations to air lanes, it has been assumed that the frequency distribution will be governed by a Gaussian distribution at right angles to the air lane with a standard deviation of σ.

The fraction of the expected number of crashes onto the area of the NPP is explained in Figure 7.

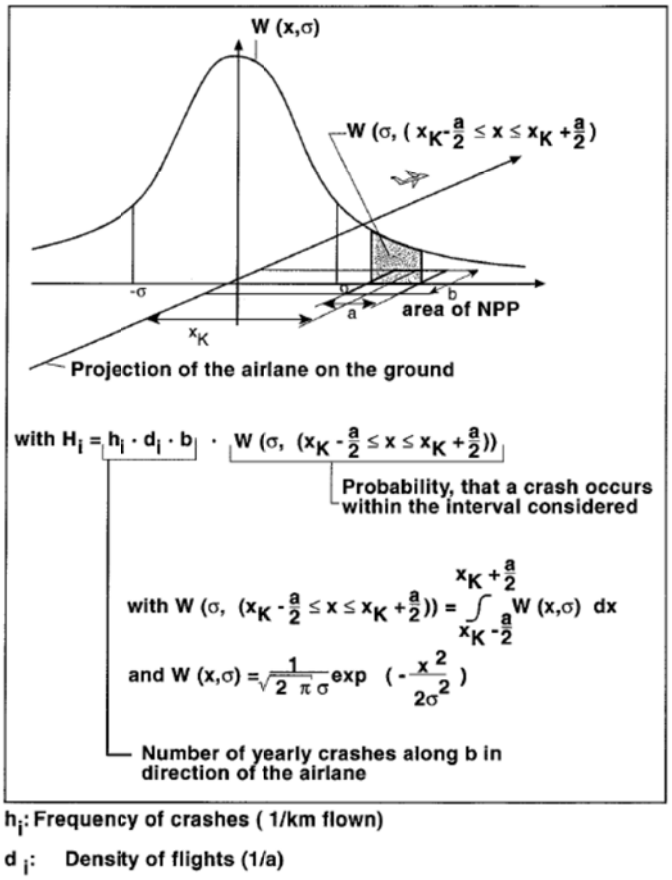


Figure 7. Determination of the crash frequency of aircrafts flying along air lanes

The aircraft crash rates outside the airports in correlation to the weight class has been listed in Table 3.

Passenger aircrafts are only part of the aircrafts in weight class 1. Resulting from data of the German Federal Bureau of Aircraft Accident Investigation (BFU) a lower crash frequency in the range of 10^{-12} to 10^{-11} per flight km can be assumed.

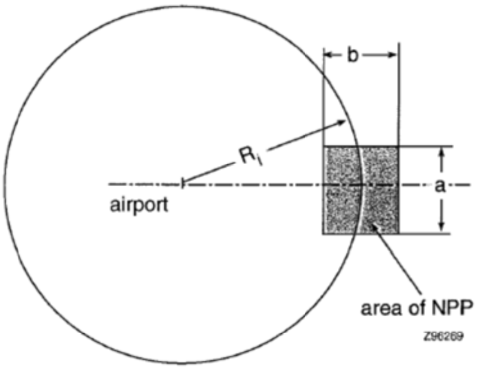
Table 3. Correlation between sectors and aircraft weight classes.

Sector	**Weight class 1a > 20 Mg**			**Weight class 1b 14 up to 20 Mg**			**Weight class 1c 5,7 up to 14 Mg**		
	$a_{1a,j}$	$b_{1a,j}$	c_{1a}	$a_{1b,j}$	$b_{1b,j}$	c_{1b}	$a_{1c,j}$	$b_{1c,j}$	c_{1c}
1	90.7	0.0005	$2.7 \cdot 10^8$	10.4	0.0004	$7.1 \cdot 10^7$	13.3	0.0006	$6.3 \cdot 10^7$
2	17.4	0.0005		3.4	0.0005		9.2	0.0006	
3	10.6	0.0006							

Outside a distance of 10 km from the airport each flight direction becomes equally probable within the free air traffic. The number of crashes H_i will be calculated by multiplying the global crash rate with the local flight density (the number of take-offs and landings of the considered airport):

$$H_{i,j} = \frac{a}{2\pi R_i} \cdot \frac{b}{v_j} \cdot d_{i,j} \cdot h_j \tag{5}$$

With equation (5) the contributions to the overall crash frequency are calculated as shown in Figure 8.



$$H_{i,j} = \frac{a}{2\pi R_i} \cdot \frac{b}{v_j} \cdot h_j \cdot d_{i,j}$$

$H_{i,j}$ = Number of crashes per year within the NPP area of airplanes of airport i and in weight class j
a, b = Dimension of the NPP area
R_i = Distance of NPP from airport i
v_j = Average flying speed in weight class j
h_j = Number of crashes of airplanes in the free air traffic (weight class j)
$d_{i,j}$ = Number of flying operations at the airport i considered (take-offs and landings; weight class j)

Figure 8. Determination of the crash frequency of aircrafts in the free air traffic

The average flying speed in the respective weight class as described in Figure 8 (Hoffmann et a. 1997) is given in Table 4.

Table 4. Average flying speed correlated to weight classes.

Weight class (Mg)		**Average flying speed km/h**
1	> 20	833.4
	5.7 – 20	463.0
2	– 5.7	333,4
3	< 2	203.7

6 CRASH FREQUENCY IN MILITARY AIR TRAFFIC

The following considerations are not valid for large military aircrafts, which are used to transport military equipment, goods or soldiers. They use the air lanes of the commercial air traffic and have to be treated for the flying operation outside the landing and take-off phase as described earlier.

The plant-specific crash frequency of military aircrafts has to be calculated according to the procedures applied for the free air traffic taking into account the crash frequency per flying hour and the number of take-offs and landings on the neighbouring military airports.

In addition, the statistics of the local crash history which took place in the square of 30 km x 30 km around the power plant area is to be evaluated.

Figure 9 shows the statistics of crashes of fast flying military aircrafts from Germany and abroad with more than 7.5 Mg in Germany for the time frame of 1984 to 2000 and of 2000 up to 2008.

As one can see, changes in the military flying operation resulted in a significant reduction of the crash frequency. This is the result of twofold changes: due to the political situation less military flights have taken place, but also the aircraft type, mainly used in the eighties, was replaced by a modern and more reliable type.

Therefore, for current calculations it has been recommended to take into account only events since 1991.

For the hazard analysis, the buildings are divided into classes to reflect the degree of protection against aircraft crash impact: one which is designed against air plane crash and another which is not specifically designed against it. It will be distinguished between a direct hit frequency and a penetration frequency.

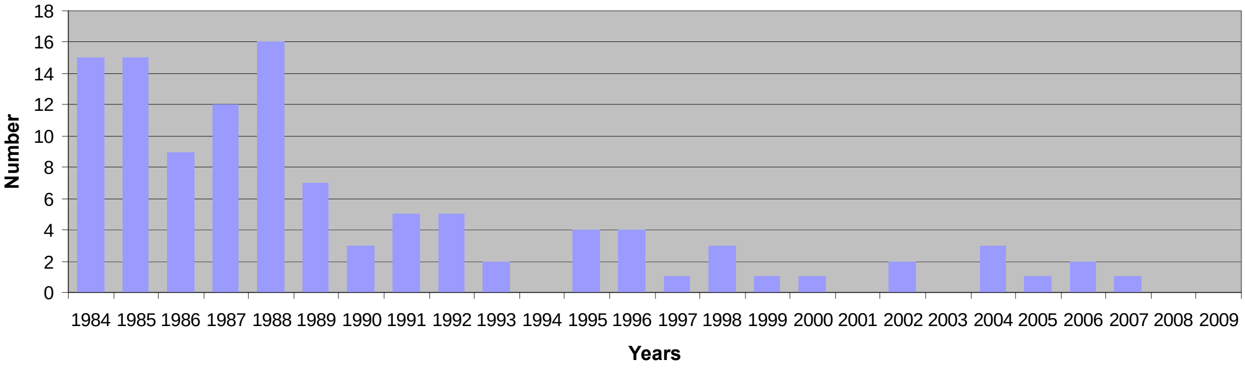


Figure 9. Number of military aircraft crashes since 1984 up to 2009

In case the kinetic energy of the projectile is greater than the penetration energy of the outer shell a total damage of the building with all equipment in it is postulated.

In the detailed assessment, a plant-specific probability for the penetration can be determined, using, e.g., Monte Carlo procedures which allow calculations with a large number of possible impact points and impact angles to determine the position where design loads of the buildings are exceeded.

A transient or an initiating event leading to a core damage situation will be caused only, if systems in other buildings, necessary to mitigate that event, fail stochastically. For that case, the core melt frequency will be calculated in an event tree analysis.

For penetrations leading directly to core melt accidents, the initiating frequency is assumed to be equal to the core melt frequency. For the buildings, not specifically designed against air plane crashes, additional hits by parts of the wracked aircraft are taken into account.

7 CONCLUDING REMARKS

Aircraft crash onto a nuclear power plant or a chemical plant is an external hazard which has to be taken into account in a comprehensive safety assessment.

Methods to analyse plants systematically regarding the adequacy of their existing protection equipment against hazards can be deterministic as well as probabilistic.

In case of probabilistic safety assessments experience has shown that is reasonable to have procedures to screen out, e.g., rooms or buildings of the plant under consideration where no further analysis is required or to establish a safety graded procedure taking into account plant- and site-specific conditions such as design of buildings against aircraft crash impact as well as distance to smaller and larger airports including the current travel situation for commercial and military aircrafts.

This information has to be site-specific and has to be collected from the respective national organizations (e.g., for commercial flights from the National Department of Civil Aviation).

Some guidance to perform frequency calculations is given in (USDoE 2006), describing the determination of the number of operations, aircraft crash rates and crash location probability including some (unfortunately older) data.

More recently, a guide to assess aircraft accidents and incidents with the focus on fires and explosions has been published (NFPA 2010).

For the free air traffic, also international data bases can be taken into account (see for example references (ICAO 2007) and (NTSB 2009a and b). However, such databases have to be used very carefully because they cover all aircraft crash statistics including those from countries which are well known for a high risk of aircraft crashes due to the age of the aircrafts used, the reduced maintenance activities and an unsufficient flight control system.

In the past the flights along fixed air lanes (i.e. airspace monitored by air traffic controller and marked by radio navigation equipment) as discussed in section 3 were the regular case; a deviation was only allowed in exceptional cases (e. .g. in case of a storm). Today only 20 % of the flights are follwing the prescribed air lanes as stated by the German Aeronautical Information Service. The air traffic controller try to allocate an optimal short lane the pilots independently from the usually used air lanes (Felbermeier 2010).

8 OUTLOOK

Looking ahead, the km flown are increasing from year to year, however, up to now there is no systemetic increase of the crash rates per year.

As described aircraft crash rates are determined according to weight classes because of the different impact and resulting consequences for the plant under consideration. New and bigger aircrafts are on the market and partially already in operation. These aircrafts are constructed with new material such as fibre reinforced composite materials which leads to a reduction of the weight of the structures.

On the other hand, during the take-off phase the amount of fuel is bigger compared to older aircrafts.

As explained earlier, external hazards like aircraft crash are analysed in the frame of periodic safety reviews in case of operating nuclear power plants. This review on the one hand investigates the current plant safety status based on the operational experiences in the last ten years, but also should look forward for the next ten years period.

Because elements of the of periodic safety review might also be used in the frame of assessing the safety level with respect to an extended life time, the confidence in the results of the probabilistic safety assessment has to be justified to a larger extent, in case of the external hazard aircraft crash by taking into account the prognosis of aircraft movements in the next ten to twenty years which may enhance the aircraft crash frequency to be assumed.

One earlier forecast of the expected traffic increase is provided in Figure 10 reflecting the political situation after the changes in Eastern Europe.

Figure 10 illustrates that this strong increase of the expected number of flights in 2020 compared to 1997 and the additional airways needed for these flights particularly affect the German airspace.

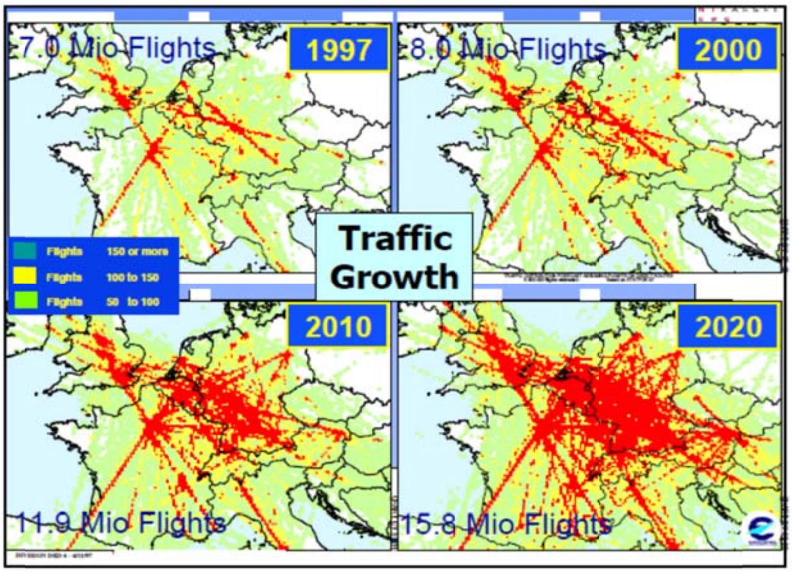


Figure 10. Traffic growth and expected increase of air traffic in 2020

A more recent study describes the forecast of aircraft traffic by 2030 (EUROCONTROL 2008). With the available capacity, the new forecast is that by 2030 there will be 1.7 to 2.2 times the number of flights in Europe seen in 2007.

Figure 11 shows the areas where extra flights per day are expected in 2020 compared to 2007. Also in this new forecast, Germany seems to be the most impacted area in Europe.

The terrorist attacks of September 11th 2001 prompted new activities and investigations of aircrafts impacting nuclear power plants. New ideas – beside the existing building concepts against aircraft crash – were born to prevent the impact at the plant, for example to obscure the plant by fog or to erect additional structures in front of the plant as shown, for example in (Eibl 2003). These investigations are still ongoing.

New technologies originating from automotive crash analysis have been developed and are used by few researchers to investigate the crash of a plane into a building. The Airbus 320 represents the type of city liners which are most frequently flown. This aircraft was chosen for a crash simulation.

The outer surface corresponds to the real structure. The stiffness of the different surface components as fuselage or wing is modelled by shell elements with equivalent thickness. Because no detailed information about the structural parts is available from the aircraft producers, most of the stiffening components are represented in a simplified manner, whereas important stiffening parts, e.g. wing box, are modelled realistically. The turbines are defined as rigid body with point mass.

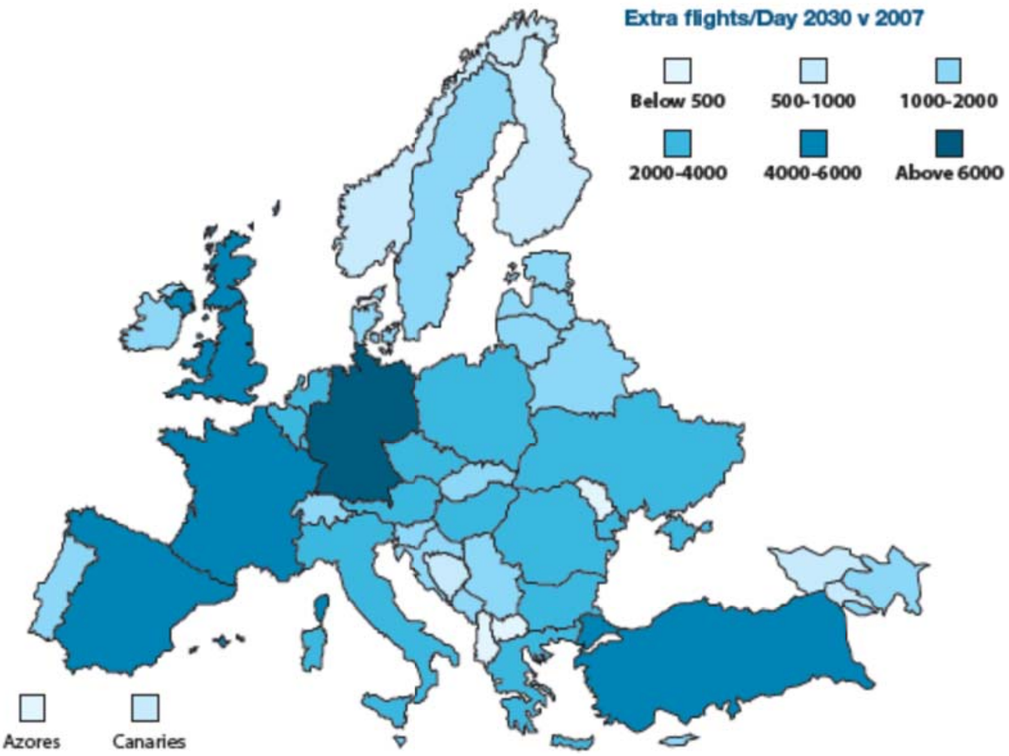


Figure 11. Extra flights through the airspace in 2030 compared to 2007

Figure 12 shows the end state of the simulation where almost two-thirds of the fuselage is collapsed. The right wing is broken into two halves. While one part is still connected to the fuselage, the other shows a tailspin behaviour. This asymmetric behaviour for a symmetric model and load is caused by small numerical effects, which have a great influence onto the buckling behaviour. This leads finally to the effect that the right turbine has contact before the left one what causes the cut of the right wing.

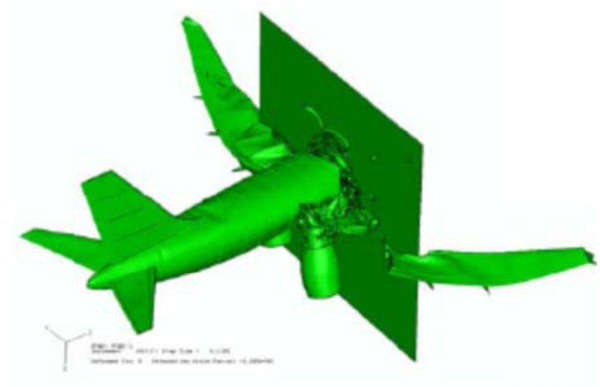


Figure 12. Final state of aircraft crash simulation

More details can be found in (Henkel & Klein 2007): However, further investigations require a better data base of the structure of the airplane. The investigations then can be extended for coupled analysis of airplane and structure.

9 REFERENCES

American Society of Mechanical Engineers (ASME). (2008) *Standard for Level 1/Large Early Release Frequency Probabilistic Risk Assessment for Nuclear Power Plants,* ASME/ANS RA-S-2008.

Berg, H.P. (2005). Screening procedures for the probabilistic analyses of internal and external hazards. *Proceedings of the 9th International Conference on Structural Safety and Reliability (ICOSSAR 2005),* Rome, 2005, 3663 – 3670.

Berg, H.P. (2010a). Aircraft crash onto a nuclear power plant - screening procedure and approach for a probabilistic analysis. *Journal of Polish Safety and Reliability Association, Summer Safety & Reliability Seminars*, Volume 1, Gdynia, Poland, June 20 – 26, 2010, 27 – 33.

Berg, H.P. (2010b). Quantitative safety goals and criteria as a basis for decision making, *Reliability & Risk Analysis: Theory and Applications,* Vol. 1, No. 2, June 2010, 62 – 78.

Bundesstelle für Flugunfalluntersuchung (BFU). (2010). *Jahresbericht 2008, Unfälle und Störungen beim Betrieb ziviler Luftfahrzeuge*. Bericht V 173, Braunschweig, Juli 2010 (published in German).

Facharbeitskreis Probabilistische Sicherheitsanalyse für Kernkraftwerke (FAK PSA). (2005). *Methoden zur probabilistischen Sicherheitsanalyse für Kernkraftwerke, Stand: August 2005.* BfS-SCHR-37/05, Salzgitter, October 2005 (published in German).

European Organisation for the Safety of Air Navigation (EUROCONTROL). (2008). *Challenges of growth 2008*, Summary report, November 2008.

Eibl, J. (2003). Airplane impact on nuclear power plants, *Transactions of the 17th International Conference on Structural Mechanics in Reactor Technology (SMiRT 17),* Prague, Czech Republic, Paper J03/6.

Felbermeier, B. (2010). Probabilistische Behandlung der Gefährdung durch unfallbedingten Flugzeugabsturz, Tagungsunterlagen des Behördenseminars zu übergreifenden Einwirkungen von außen, GRS Köln, September 2010 (published in German).

Fürste, W. & Glupe, G. (1974). Absturzhäufigkeit von Flugzeugen auf Kernkraftwerke. *TÜ (Technische Überwachung) 15*, Nr. 4, April, 115 – 119 (published in German).

Henkel, F.-O. & Klein, D (2007). Variants of analysis of the load case airplane crash, *Transactions of the 19th International Conference on Structural Mechanics in Reactor Technology (SMiRT 19*), Toronto, Canada, Paper J03/2.

Hoffmann, H. H., Skoff, G. & Berg, H.P. (1997). Probabilistic safety analysis of external events in German nuclear power plants. *Transactions of the 14th International Conference on Structural Mechanics in Reactor Technology (SMiRT 14*), Lyon, France, 1997, Vol. 10, 87 – 94.

International Civil Aviation Organization (ICAO). (2007). *Annual report of the Council for 2006,* Montreal, Canada.

International Atomic Energy Agency (IAEA). (2010). *Development and Application of Level 1 Probabilistic Safety Assessment for Nuclear Power Plants*, Special Safety Guide SSG-3, IAEA, Vienna, April 2010.

International Civil Aviation Organization ICAO). (1995). *Civil Aviation Statistics of the World 1994*. Doc. 9180/20, 20th Annual Edition, Montreal, Quebec, Canada.

Kuznetsov, V. (2007). Advanced nuclear plant design options to cope with external events. *Transactions of the 19th International Conference on Structural Mechanics in Reactor Technology (SMiRT 19),* Toronto, Canada, Paper S01/2.

National Fire Protection Association (NFPA). (2010). *Guide for Aircraft Accident / Incident Response Assessment.* NFPA 422, 2010 Edition, NFPA, 1 Batterymarch Park, Quincy, MA.

National Transport Safety Board (NTSB). (2009a). *Accidents and Accident Rates 1983 – 2008*. US Aviation, www.ntsb.gov/aviation/stats.htm.

National Transport Safety Board (NTSB). (2009b). *Annual Review of Aviation Aircraft Data 2005.* NTSB/ARF-09/01, National Transport Safety Board, Washington.

Reactor Safety Commission (RSK). (1984). *Guidelines for Pressurized Water Reactors.* 3rd Edition, October 18, 1981, amended 1982 and 1984.

US Department of Energy(USDoE). (2006). *Accident Analysis for Aircraft Crash into Hazardous Facilities.* DOE-STD-3014-2006.

RELIABILITY OF MAIN TRANSFORMERS

H.-P. Berg, N. Fritze
•
Bundesamt für Strahlenschutz, Salzgitter, Germany
e-mail: hberg@bfs.de

ABSTRACT

Key equipment for the electric power transmission is the transformer. Because of the high failure frequency and the resultant reliability and safety implications in particular of main transformers, an in-depth assessment is necessary. Main transformers are considered as a critical equipment because of the large quantity of oil in contact with high voltage elements. Experience has shown an increasing number of transformer explosions and fires in all types of power plants worldwide. Therefore, these phenomena have been investigated in more detail and are discussed with regard to potential root causes for these events such as potential influence of the age of the transformers. Moreover, possible diagnostic measures to avoid such events and enhance the reliability are shortly described. For investigating the current status of the reliability of transformers different types of databases have been evaluated.

1 INTRODUCTION

A broad spectrum of events such as design defects, voltage surges, lightning strikes, structural damage, rapid unexpected deterioration of insulation, sabotage, and even maintenance errors can lead to transformer fires and explosions. Experience has shown that the consequences of such events can be severe.

In particular, a fire of an oil-cooled transformer that contains several thousand litres of combustible insulating oil can result in severe damage to nearby power plant structural components such as concrete walls and damage or destroy electrical components such as nearby transformers, bus work, and circuit breakers (US Department of the Interior 2005). A one-year research project led to the discovery of 730 transformer explosions in the USA only. Many experts anticipate that the number of failures per year will increase significantly in the near future to 2%. In addition, the shorter lifetime of new transformers will sharply increase above this rate after 2010. Because about 115 000 large transformers are in operation in the US and about 400 000 worldwide, the number of impacted transformers is high, even when only in some cases fire and explosion lead to a total damage (Berg & Fritze 2010).

Power transformers with an upper voltage of more than 100 kV are necessary for the undisturbed operations of a developed society. In electricity generation plants, power transformers transform the voltage of the generator to a higher level for the transmission of electricity in the main grid. The voltage of the main grid must again be transformed to a lower voltage, so that the electrical energy can be utilized in numerous purposes (Valta 2007).

Electric power is normally generated in a power station at 11 to 25kV. In order to enable the transmission lines to carry the electricity efficiently over long distances, the low generator voltage has to be increased to a higher transmission voltage by a step-up transformer, i.e. 750 kV, 400kV, 220kV or 110kV as necessary. Supported by tall metal towers, the lines transporting these voltages can run into hundreds of kilometres. The grid voltage has then to be reduced to a sub-transmission voltage, typically 26kV, 33kV or 69kV, in terminal stations (also known as power substations).

Sub-transmission lines supply power from terminal stations to large industrial customers and other lower voltage terminal stations, where the voltage is stepped down to 11kV for load points through a distribution network lines. Finally, the transmission voltage is reduced to the level adapted for household use, i.e. 415V (3-phase) or 240V (1-phase) at distribution substations adjacent to the residential, commercial and small to medium industrial customers. Figure 1 shows a typical electrical network system, in which power is transformed to the voltages most suitable for the different parts of the system.

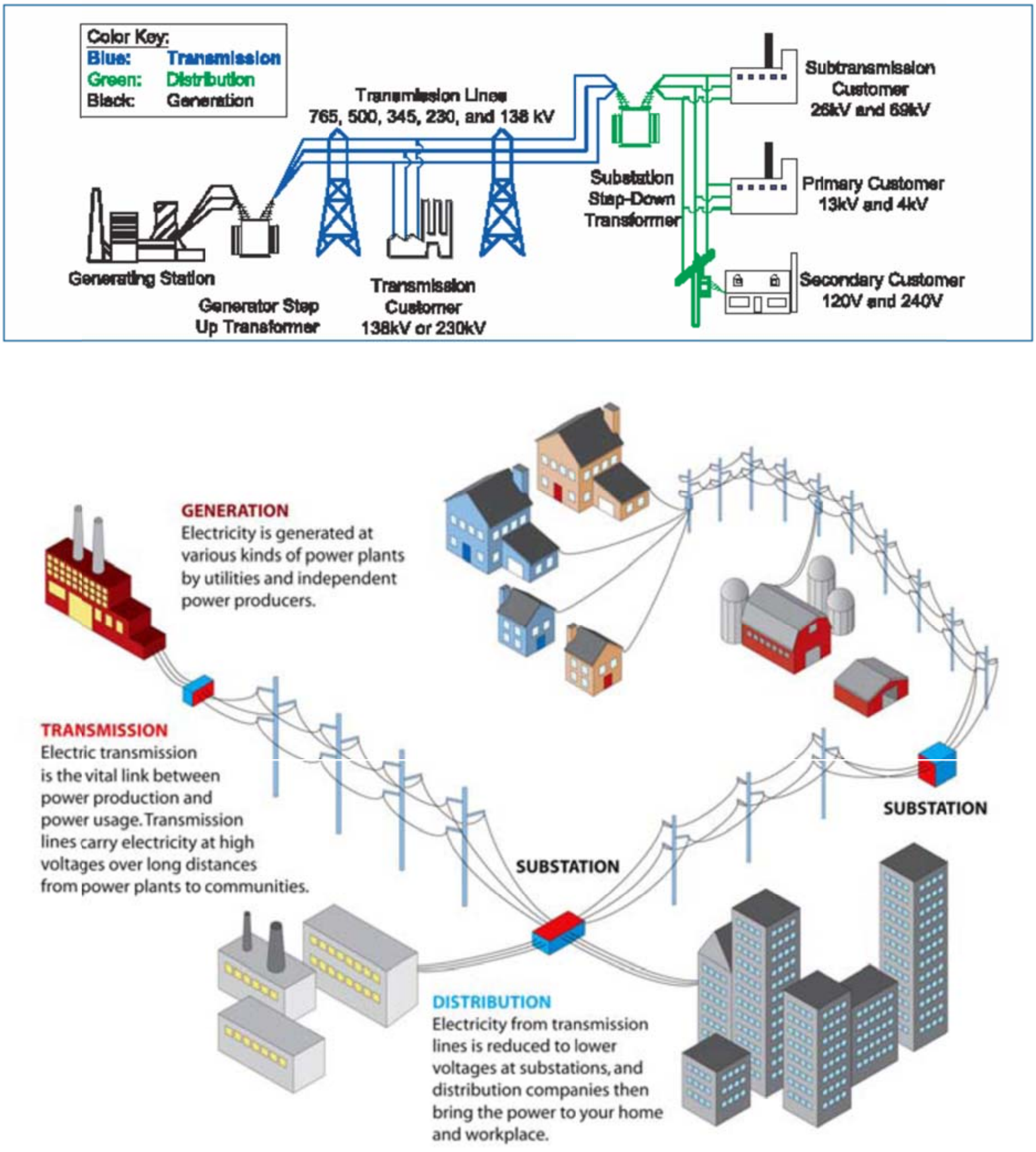


Figure 1. Typical electrical power network

At every point where there is a change in voltage, a transformer is needed that steps the voltage either up or down. There are essentially five levels of voltages (United States Department of Energy 2006) used for transmitting and distributing AC power (Table 1): Ultra-High Voltage

(UHV, 1100 kV), Extra-High Voltage (EHV, 345 to 765 kV), High Voltage (HV, 115 to 230 kV), medium (or sub-transmission) voltage (MV, 34.5 to 115 kV), and distribution voltage (2.5 to 35 kV). The UHV, EHV, HV, and MV equipment is mainly located at power plants or at electric power substations in the electric grid, while distribution-level transformers are located in the distribution network on poles, in buildings, in service vaults, or on outdoor pads.

Table 1. AC voltage classes

Transmission Voltages		**Distribution Voltages**	
Class	**kV**	**Class**	**kV**
Medium voltage (MV)	34,5	2.5	2.4
	46	5	4.16
	69	8.66	7.2
	115	15	12.47
High voltage (HV)	115	25	22.9
	138	35	32.5
	161		
	230		
Extra-High voltage (EHV)	345		
	500		
	765		
Ultra-High voltage (UHV)	1100		

For the different activities of changing voltage, the following two types of transformers are commonly used:

- dry type transformers and
- liquid insulated transformers.

Dry type transformers are transformers containing solid or gas insulation material. The fire hazard of dry type transformers is generally considered to be lower compared to liquid type transformers because the amount of combustible materials present in the transformers is limited.

Liquid insulated type transformers are usually subdivided from a fire hazard point of view into less flammable liquid transformers and flammable liquid transformers.

Less flammable liquid (e.g. silicone oil, ester) is expected to have a high fire point (above 300°C) and, hence, such a transformer is more difficult to ignite. Transformers which are insulated with flammable liquid such as mineral oil are considered to have the highest fire hazard because of the combustible liquid oil and its relatively lower fire point (100°C to 170°C).

Today, liquid-filled main transformers are widely used in power distribution systems. Most are outdoors, where the risk of property damage associated with a flammable liquid dielectric is lower.

When flammability must be reduced, alternative liquids are needed which lead to other problems (toxicity) resulting from PCB. Therefore, the majority of main transformer is still oil-insulated.

2 MAIN COMPONENTS OF A TRANSFORMER

The major components of a transformer are the coils (windings), the core, the tank or casing, the radiator, and the bushings as shown in Figure 2. Generally, transformer coils are made of copper because it has a lower resistance and is more efficient compared to other metals. Each winding is wrapped with an insulating material such as paper. The primary winding is usually wound around the transformer core and the secondary winding is then wound on top of the primary winding. Between each layer of the windings, another layer of insulating material is wrapped to provide extra insulation between the windings. There are ten major transformer components (Ng 2007).

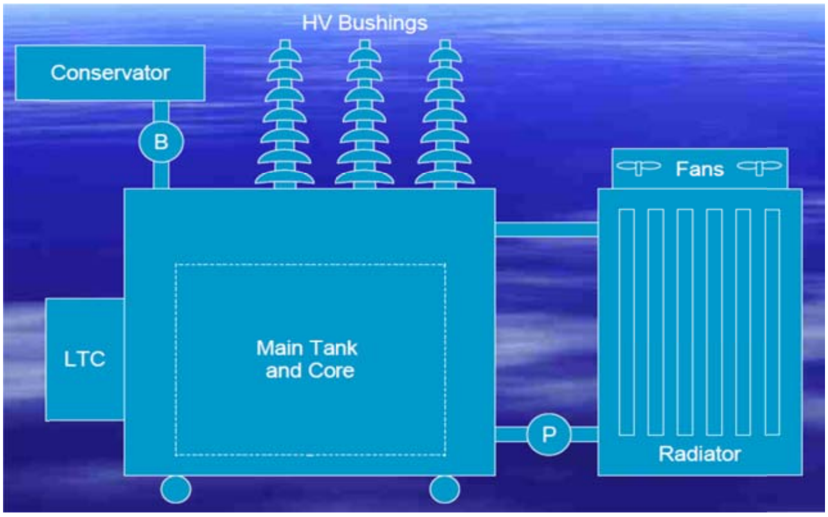


Figure 2. Main transformer components

These components can briefly be described as follows:

1) Core is a ferromagnetic material (commonly soft iron or laminated steel) that provides a path of high magnetic permeability from the primary circuit to the secondary circuit.
2) Windings allow a secondary voltage to be induced in the secondary circuit from the alternating current (AC) voltage in the primary circuit. The change in magnetic field in the transformer core caused by applying primary AC voltage causes an induced magnetic field and, hence, voltage on the secondary winding.
3) Tank or casing, which is usually a reinforced rectangular structure in these transformers, contains the dielectric material, the core and the windings.
4) Dielectric material is a substance that is a poor conductor of electricity but an efficient supporter of electrostatic fields. It can be fluid oils, dry solids or gases.
5) The expansion tank or conservator containing dry air or dry inert gas is maintained above the fluid level.
6) Bushing is an insulating structure that provides a conducting path though its centre, its primary function is to insulate the entrance for an energised conductor into the tank.
7) Pressboard barriers, between the coils and between the coils and core, are installed to increase the dielectric integrity of the transformer.
8) The tap changer is a connection point along a transformer winding that allows the number of turns to be selected, or so-called voltage regulating device.
9) The radiator provides a heat transfer path to dissipate the internal heat generated in the transformer.
10) The pressure relief device is used to protect the tank against excessive pressure release inside a transformer tank.

An oil-insulated transformer is made up of a steel tank, which includes windings and the transformer's iron core. During the manufacturing phase, the windings are covered with insulation paper and electrical insulating board. The steel tank is full of transformer oil and it impregnates the insulation paper, during which time the combination of paper and oil and the electrical insulating board form a necessary electrical insulation.

Basic core and winding configurations differ little between dry and oil-insulated transformers. However, air is a much poorer insulator than dielectric fluid; hence, clearances between conducting surfaces can be much smaller in a liquid-filled transformer, allowing operating voltages to be much higher than with dry-type design.

To ensure that the transformer can operate without failure for at least 30 years and that the life expectancy of the transformer can be correctly estimated the properties of the transformer oil and insulating paper must be kept at a specific level.

3 RESULTS FROM INTERNATIONAL DATABASES

3.1 OECD FIRE Database

One application of the OECD FIRE Database has been an analysis of events associated with explosions (Berg et al. 2009 and 2010a) base on the database issued March 2009. A query in the Database on the potential combinations of fire and explosion events has indicated a significant number of explosion induced fires. Most of such event combinations occurred at transformers on-site, but outside of the NPP buildings or in compartments with electrical equipment. Approximately 50 % of the fires were extinguished in the early (incipient) fire phase before the fire had fully developed. As a consequence of these indications, improvements concerning the fire protection of transformers are intended in Germany. As there is no specific coded field in the database to indicate explosions, the main source of information is provided by the event description field.

Basis for the results provided in this section is the version March 2010 of the OECD FIRE Database. The 24 reported explosions amount to 6.5 % of all events reported up to date (see Figure 3).

Concerning the process of explosion distinction should be made between an explosion as a process of rapid combustion (chemical explosion) and an explosion as a physical process resulting from a sudden gas pressure rise by a high energy electric (arcing) fault (HEAF).

A chemical explosion was found for only three events (solvent vapor, diesel fuel, hydrogen). In the other 20 cases, HEAF events obviously took place at the same time indicating a physical explosion. In some of these cases the electric fault might have caused a fuel pyrolysis or fuel spread and acted as an ignition source for a chemical explosion, thus a HEAF event and a chemical explosion may have taken place simultaneously.

In one event, a fire led to the explosion of diesel fuel vapor while in another event a fire and an explosion occurred independently from each other in parallel. In all other cases explosions induced the fire).

Concerning the buildings/locations where the events took place it was found that 13 (54 %) events took place outside buildings, five inside electrical buildings.

A majority of 54 % of the reported explosions started at transformers. The other 11 events took place at electrical cabinets, other electrical equipment, or process equipment (three each representing 13 %). External fire brigades were needed in four of 24 cases (17 %). The 24 events were also evaluated concerning the fire duration with the following results shown in Table 2.

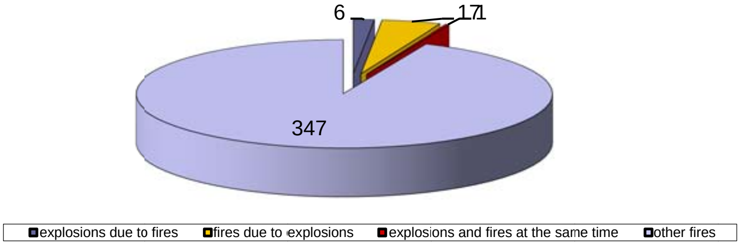


Figure 3. Results from the OECD FIRE Database

Table 2. Fire duration

Fire duration	**Number of events**
0 - < 5 min	11
15 - < 30 min	4
30 - < 60 min	4
> 60 min	3

For the remaining two events no information on the fire duration is provided. This is in good agreement with the fire durations recorded for all events, where for approx. 59 % of the events (154 out of 261 events with fire duration provided) a fire duration of less than 15 min could be found.

As one can see from Figure 4, fires of high voltage transformers contribute to about 8,9 % of all fires contained in the OECD FIRE database.

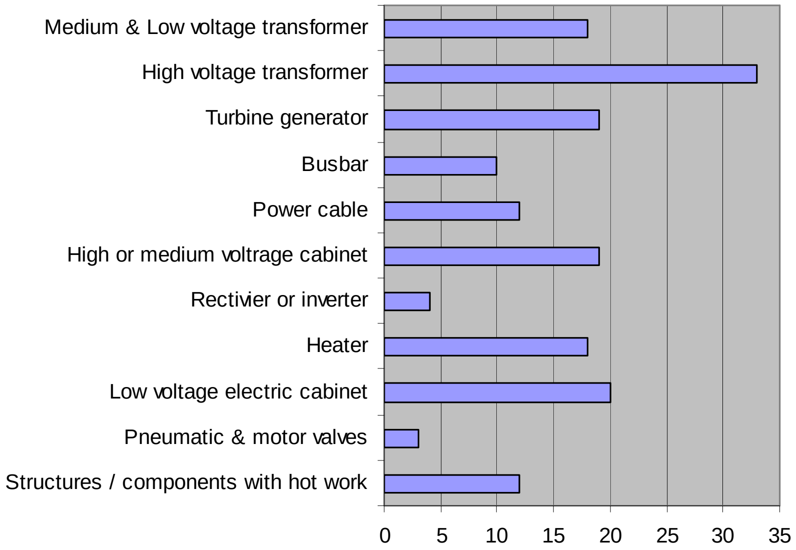


Figure 4. Components where the fire started

3.2 Statistics from EPRI Database

Within the document on fire PRA methodology (EPRI and USNRC 2005) some generic fire frequencies are provided based on the operational experience of US nuclear power plants:

- Catastrophic fire at transformer yard (includes events with rupture of transformer tank, oil spill and burning oil splattered a distance from the transformer): 6.0E-03 / reactor year.
- Non Catastrophic fire at transformer yard (includes events without oil spill outside transformer tank): 1.2E-02 / reactor year.
- Other fires at transformer yard (includes events associated with the transformers but not the transformers themselves): 2.2E-03 / reactor year.
- The above given mean values are based on 1674 reactor years and about 35 fire events in total.

The transformer yard fire frequencies estimated in the above mentioned report are comparable to the operating experience shown by OECD FIRE database. The number of transformer yard fires collected in the OECD FIRE project is also adequate for qualitative purposes. Quantitative analysis of the OECD FIRE data would still require additional information about number of transformer under consideration in each NPP to avoid using reactor years, which causes some uncertainty. Also additional information on the amount of burned transformer oil would be welcome to realize the necessary performance of fixed extinguishing systems and operative fire fighting measures (Lehto et al. 2010).

3.3 Statistics from non-nuclear industry

Transformers are used for stepping up or down the voltage.. High voltage equipment is mainly located at power plants and at substations representing high voltage electric systems facilities used to switch generators, equipment and circuits or lines of the system on and out. Substations can be large with several transformers and dozens of switches.

In 2003, the International Association of Engineering Insurers (IMIA) presented a research, which contained an analysis of transformer failures, which have occurred in IMIA member countries (see Bartley 2003 and Bartley 2005). During the period 1997 – 2001 a total of 94 failures occurred.

These 94 failures have been divided in Table 3 below according to age.

Table 3. Division of failure according to age of transformer 1997 – 2001

Age	Number of failures
0 – 5 years	9
6 – 10 years	6
11 – 15 years	9
16 – 20 years	9
21 – 25 years	10
Over 25 years	16
Age unknown	35

Insulation failures were the leading cause of failure in this study. The average age of the transformers that failed due to insulation was set to 18 years, in some cases leading to transformer fire and explosion. However, the high number of failures where the age is not known may indicate a tendency to higher transformer failures due to ageing.

During the normal use of a transformer, oil and insulation paper becomes old and at some phase they are no longer able to fulfil their tasks concerning electrical and mechanical strength.

The damage databases provide clear observations that transformer damages often arise due to defects in insulation that originate in the interior of the transformer. It is, therefore, necessary to monitor the ageing phenomena so that reliable information concerning potential faults can be obtained during the earliest phase possible.

More recent industry data show that in case of substation transformer 20 % of failures result in a fire. In Los Angeles, 97 transformer fires occurred in the first three month of 2006, averaging more than one per day.

According to (Lord & Hodge 2008) the contribution of the different main components (as shown in Figure 2) to major failures are winding and tap changes with about 25 % each, whereas high voltage (HV) bushings are the cause in about 20 % to 40 % of failures depending on the underlying statistical basis. However, HV bushings provide the highest contribution to all transformer fires with about 70 % (see Figure 5).

These results are furthermore supported by the experience provided in Table 4 (Foata & Nguyen 2010) where the failure statistics for 735 kV transformers over 25 years is collected.

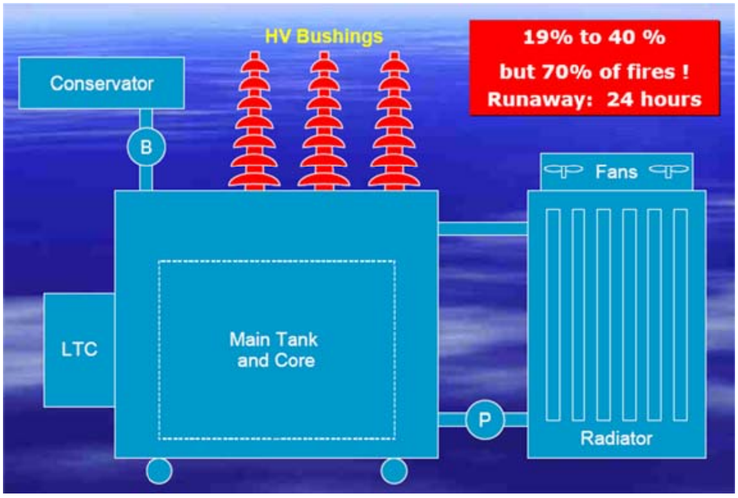


Figure 5. Contribution of Major Failures

This database contains 175 transformer failures that resulted in 110 high energy arcs causing in total 44 tank ruptures and 18 fires. In 13 of the 18 fires, the component HV bushings contribute to the transformer fires.

Table 4. Failure statistics for 735 kV transformers over 25 years

Component	Faults	Ruptures	Fires
HV bushing	41	19	13
Windings	57	21	3
Core	3	2	1
OLTC	2	1	0
Others	7	1	1

The large contribution of bushing failures to the transformer fire risk is also a result presented in the last transformer session of the International Council of Large Electric Systems (Foaka 2010) as shown in Figure 6.

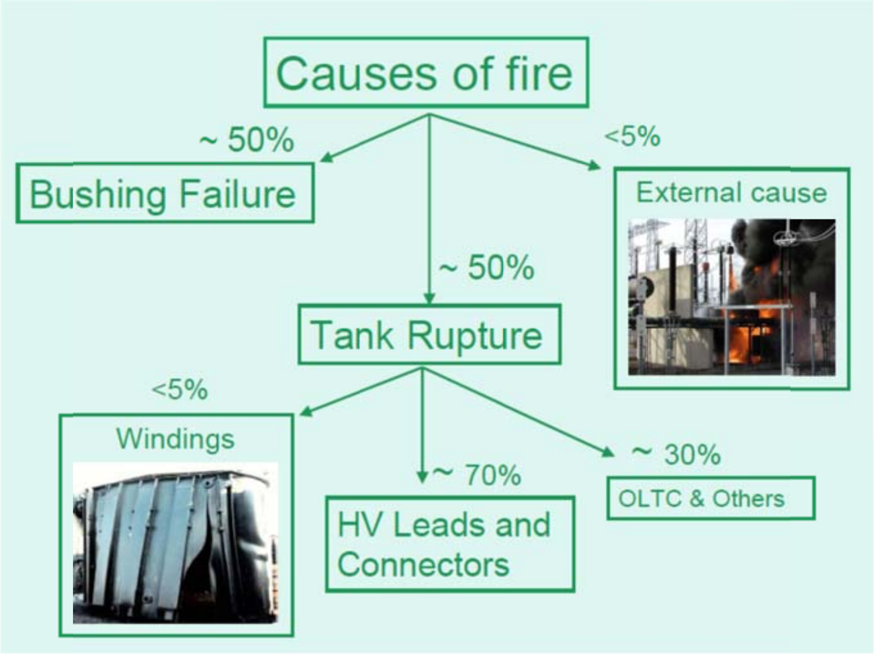


Figure 6. Causes of transformer fires (Foata 2010)

As in the case of the statistics provided in (Foata & Nguyen 2010) also other power utilities traditionally collect data and information on failure causes (Minhas et al. 1999). This information shows that in smaller transformers ageing related failures are dominant. In the medium power rating class, tap changer failures constitute the highest failure rate.

In the large transformers insulation coordination failures are the most common cause in the early service life of transformer (Mirzai et al. 2006), the influence of ageing could not be justified at present due to a lack of data.

However, one result provided in (Foata 2010) shows a correlation between the fire rate per year and the voltage which depicts an increasing probability for a transformer failure resulting in a fire for larger power transformers.

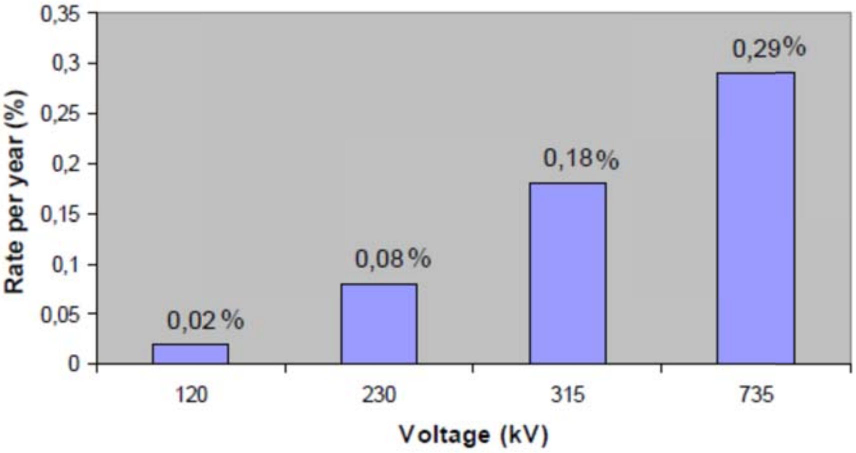


Figure 7. Fire probability vs. voltage

Figure 8 shows an example of a destroyed transformer, the root cause was an electrical arc, fortunately not resulting in a fire or explosion.

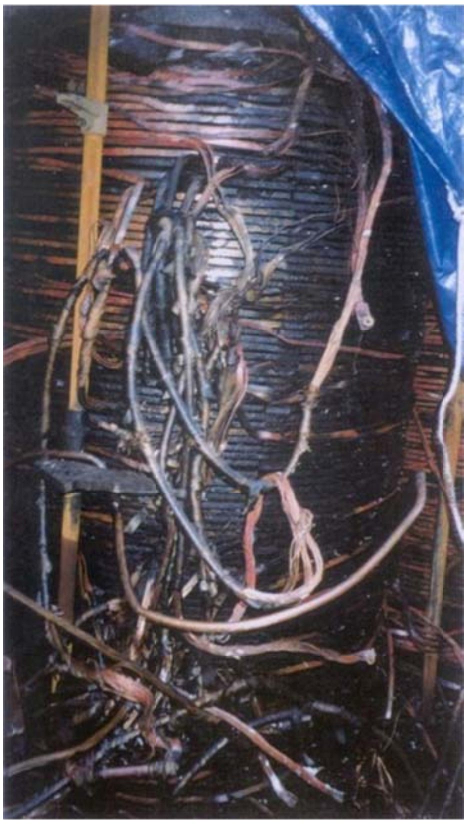


Figure 8. Example of a destroyed transformer

10 EXAMPLE OF A TRANSFORMER FIRE IN A NUCLEAR POWER PLANT

A fire started in the transformer building leading to a short circuit to occur in the transformer. The resulting electric arc – a spark – set fire to the oil in the transformer.

A simplified diagram showing the main components of station service supply and the grid connection of the nuclear power plant is provided in Figure 9.

The short circuit was recognised by the differential protection of the generator transformer, and the circuit-breaker between the 380-kV grid connection and the affected generator transformer (AC01) as well as the generator circuit-breaker upstream of the unaffected transformer (AQ02) were opened. At the same time, de-excitation of the generators was actuated. The short circuit was thereby isolated. In addition, two of the four station service supply bus bars (3BC and 4BD) were switched to the 110-kV standby grid (VE in Figure 9).

Within another 500 ms, the generator protection system caused the circuit-breaker between the second 380-kV grid connection and the intact generator transformer (AC02) to open. Subsequently the two other station service supply bus bars (2BB and 1BA) were also switched to the standby grid. After approx. 1.7 s, station service supply was re-established by the standby grid. Due to the two short under voltage on station service supply bus bars (signal "Voltage of unit bus bars BB and BC<70%") the reactor protection system triggered reactor trip.

Due to the damage caused by the fire in the transformer, the plant was shutdown. The fire of the transformer showed the normal behaviour of a big oil-filled transformer housing, the fire lacks combustion air and produces a large amount of smoke (see Figure 10).

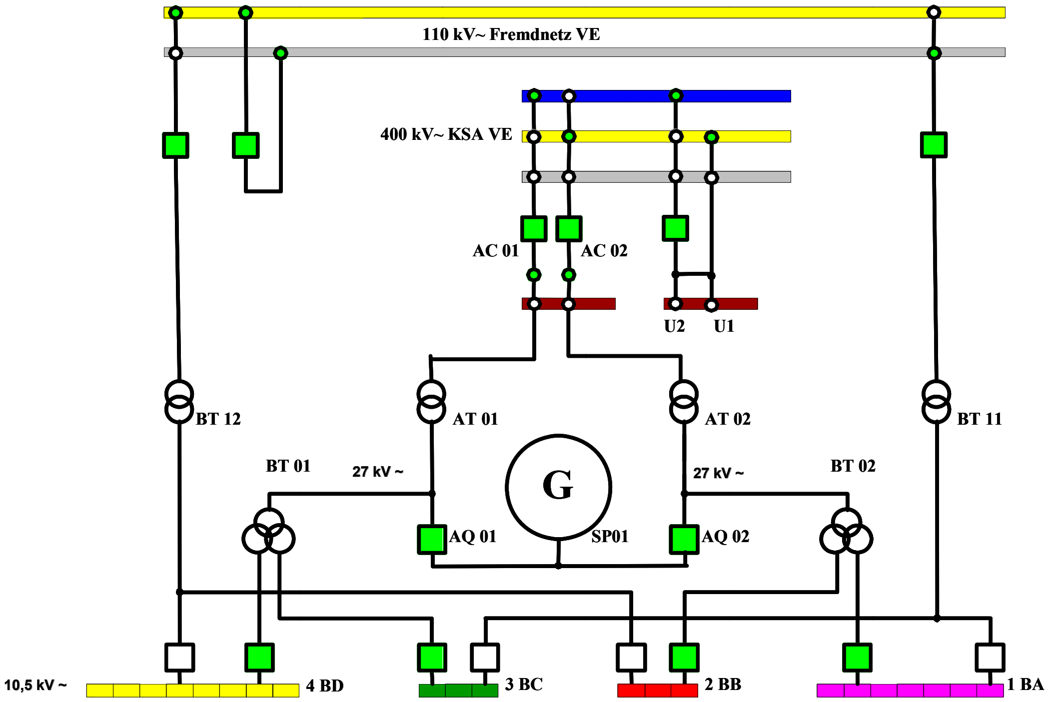


Figure 9. Simplified diagram of the station service supply and the grid connection of the nuclear power plant

The fire extinguishing activities start with an automatic fire extinguishing system, followed by activities of the on-site fire brigade, later supported by external local fire fighters (see Figure 11).

Figure 10. Flames and smoke occurring at the 2007 generator transformer fire at a German NPP

After the end of the fire fighting operations, a foam attack and later a flooding of the transformer vessel has been started to cool down the spools.

The time, until the fire in the transformer housing was extinguished, lasted nearly seven hours, approx. 70.000 kg transformer oil were ignited.

The long duration of the extinguishing phase is due to the large amount of fire loads involved and the exceptional heat capacity of the transformer core and windings (approx. 350.000 kg of iron and copper).

Figure 11. Fire extinguishing activities supported by the external local fire fighters

All fire fighting equipment worked as designed. Because of the non chloride oil the influence on the environment is low.

The root cause analysis of this event was not successful due to the total damage of the transformer. Therefore, one alternative is to perform a simulation of the event which has to couple electromagnetic, thermal and hydrodynamic phenomena. For that purpose, one has to:

- determine the magnetic field created by the inductance and/or arc in the surrounding field versus the injected current per phase;
- calculate the induced currents and the Joule and Eddy current local dissipated power;
- calculate the temperature by using the resulting above values as heat sources.
- Such a calculation can be done using four sub models as described in (Scheurer et al. 2007a and 2007b).

The analysis has been based on the information that an electric arc was the starting point of the event sequence. The electrical arc is a high temperature plasma. Therefore, at this heat level, the oil cracking process generates sufficient gas to create an overpressure. The vessel maximum tolerated pressure was determined to be above atmospheric pressure, but the pressure relief valves are inefficient for such pressures.

11 APPLICATION OF DIFFERENT METHODOLOGIES

In order to determine the risk of potential transformer failures leading to damage, a typical HAZOP analysis should be performed and possible risk categories can be defined.

Figure 12 shows exemplary input data needed for evaluating possible risks and risk categories to be taken into account (Stiegemeier 2007).

The risk of a short circuit failure is based on the assessment of the short circuit strengths of the windings and clamping structure. The thermal condition of the winding is based on the condition of the paper insulation; brittle insulation is more likely to fail under the mechanical and electrical stress conditions. The risk of dielectric failure is based on the assessment of the dielectric withstand capability of the transformer insulation system (oil, paper, etc.) and the electrical stress imposed by the power system and naturally occurring events. Accessory failures are failures of a bushing, pump or tap changer which may cause a failure of the transformer. Miscellaneous risk covers other failures including random ones.

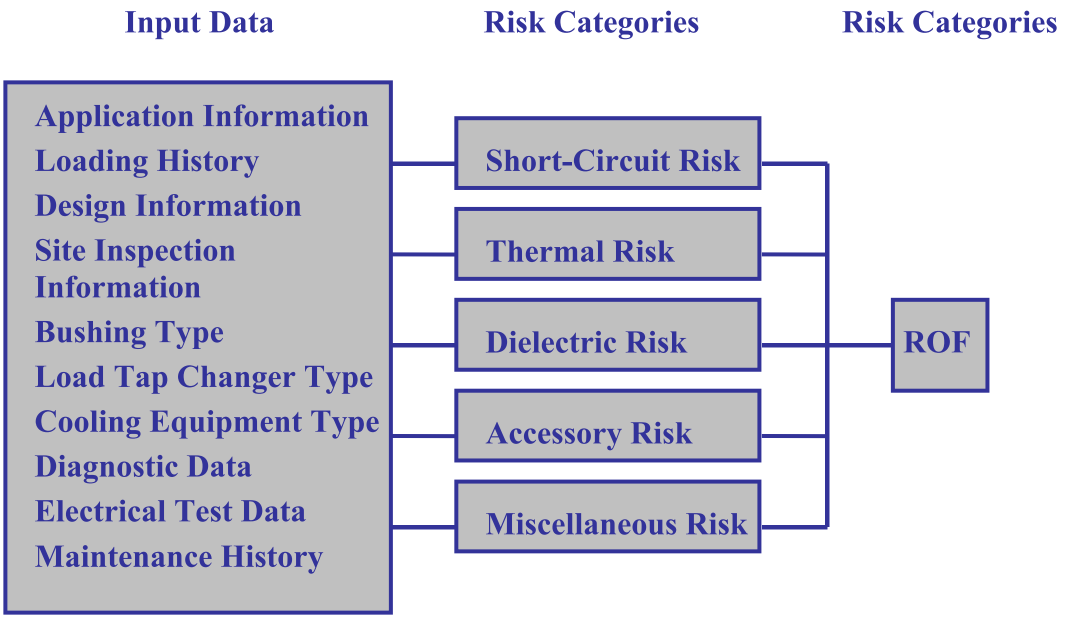


Figure 12. Risk of failure resulting from different risk categories

Such risk considerations can be performed for a single unit but also for a fleet of transformers if the boundary conditions are comparable.

In order to assess the risk of power transformer failures caused by external faults such as short circuit, a fuzzy risk index has been developed and applied (Flores et al. 2009). The risk index is obtained by comparing the condition of the insulation paper with the probability that the transformer withstands the short circuit current. This probability and the value of the degree of polymerization of the insulating paper are used as inputs of a fuzzy logic system in order to assess the failure risk.

Recently, the failure mode and effect analysis methodology has been applied to transformers, as a first step for a comprehensive project on lifetime modelling and management (Franzen & Karlsson 2007). The fault trees developed for the transformer result form discussions with experts on transformers and from a literature study. In order to analyze the transformer and to develop a fault tree the transformer has been subdivided into different sub-components such as windings, bushings, insulation, cooling and tap changer.

The objective of the above mentioned project is to develop a quantitative probabilistic model, based on both failure statistics and measurements, for the lifetime of transformer components. First models for lifetime estimation of transformers and measurement techniques will be studied. Then an improved model will be developed. Finally, the model developed will be implemented into a maintenance planning problem. Work has also been carried out on developing a statistical method for lifetime estimation based on results from Dissolved Gas Analyses. The project started in June 2009 and will be completed in September 2014.

12 CONCLUDING REMARKS

6.1 Further investigation

It has been found that main transformer failures require an in-depth assessment because of the high failure frequency and the resultant reliability and safety implications (USNRC 2010).

A lot of events in all types of power plants and substations has shown that ageing transformers are a matter of concern. Thus, transformer age might be an important factor to consider when identifying candidates for replacement or rehabilitation. Age is one important indicator of remaining life and upgrade potential to current state-of-the art materials. During transformer life, structural strength and insulating properties of materials used for support and electrical insulation (especially paper) deteriorate (US Department of the Interior 2003). Ageing reduces both mechanical and dielectric strength. All transformers are subject to faults with high radial and compressive forces. Clamping and isolation can then not longer withstand short circuit forces which can result in explosions and fires.

Although actual service life varies widely depending on the manufacturer, design, quality of assembly, materials used, maintenance, and operating conditions, the designed life of a transformer is about 40 years, but in practice industry has noted that they last 20 to 30 years.

However, in some cases the transformer are younger as in the case of the transformer fire at the Diablo Canyon plant in 2008 where the transformer was only nine years old.

The most mostly applied method for obtaining this information is to take oil samples from the transformer oil and carry out a so called Dissolved Gas Analysis. Certain gases are formed in transformer oil as a result of the transformer's age but they are also formed as a result of different over-loading situations, partial discharges and electric arc phenomena, etc. This method will now implemented in several nuclear power plants to avoid recurrence of a fire event.

However, the effectiveness of the current practice of oil sampling to predict the failure of power transformers has been checked within a research project. It was found that the current method of oil sampling using dissolved gas analysis alone is not as effective as usually perceived. An average of only 1,7% of transformer failures were actually predicted by this method. Thus, alternative mitigating strategies have to be developed to manage the risk of transformer failures (Visser & Brihmohan 2008).

One approach might be a combined use of gas and optical sensing technologies for the testing of transformer oil. The performance of such a method was evaluated to-date only on a small database of transformer oil samples and has to be further validated (Amrulloh, Abeyratne & Ekanayake 2010).

A further aspect which needs to be taken into account is the fact that the detection method Dissolved Gas Analysis is not able to measure the amount of the gases that are inside the solid insulation.

However, temperature variations can cause the generated gases to migrate into the solid insulation or more gases come out from the solid insulation into the liquid. This could generate error in Dissolved Gas Analysis measurements or trigger a false alarm. A mathematical model can be used to convert the Dissolved Gas Analysis results to the real amount of gas present in the system based on the current gas concentration in the oil and the system temperature. A possible approach is provided in (Shahsiah, Degeneff & Nelson 2006).

6.2 Countermeasures

The four main types of transformer failures are well known:

- arcing or high current break down;
- low energy sparking or partial discharges;

- localized overheating or hot spots;
- general overheating due to inadequate cooling or sustained overloading.

Therefore, protecting transformers against explosion and fire has becomes a high priority taking into account

- worldwide privatization programs of electricity production and distribution companies have resulted in a reduction of investments,
- today's competitive markets demand longer life, greater production, which results in ageing equipment and overloaded transformers.

However, transformer failures and transformer fires are not only important for operational reasons but could lead to significant safety-relevant consequences.

Therefore, a working group of the International Council of Large Electric Systems was initiated in 2007 which deals with transformer fire safety practices. Results of this working group are expected in 2011.

Detection techniques serve as a warning system to developing abnormalities in a transformer or one of its components. Detection techniques are comprised of parametric measurements and visual inspection.

The parametric measurements most often used are the current, the voltage, the internal pressure of the tank, the oil level, the oil and winding temperature, gas in oil analysis, and winding power factor, to name a few. The least frequently used measurements include the load tap changer acoustic vibration, acoustic surveillance of partial discharge, etc.

Visual inspection of the transformer exterior reveals important condition information. For example, valves positioned incorrectly, plugged radiators, stuck temperature indicators and level gauges, and noisy oil pumps or fans. Oil leaks can often be seen which may indicate a potential for oil contamination, loss of insulation, or even environmental problems. Physical inspection requires staff onsite experienced in these techniques.

Existing diagnosis concepts for power transformers are traditionally categorized by the underlying measurement technique (online vs. offline). The subdivision into physical subsystems (e.g. mechanic subsystem, dielectric subsystem, thermal subsystem) is a first step for a model-based approach. Interpretation methods for measurement results and the integration of the subsystems into a common diagnosis scheme are missing links on the way to a model-based diagnosis concept (Hribernik et al. 2008).

A further approach in addition to protective measures is the implementation of a structured description for different scenarios which can occur and their consequences in a plant, in particular in the case of power increase.

An example is a coal-fired plant in Australia with four 660 MW generating units which has planned a capacity increase to 750 MW for each unit. In the framework of a comprehensive risk analysis of this project, a specific fire safety study has been performed for four main scenarios: steam generator oil fire, generator hydrogen fire, boiler explosion and fire as well as generator transformer explosion and fire (Fire Risk Solution 2009).

Figure 13 summarises in a simplified manner the steps of the assumed scenario "transformer explosion resulting in fire" including the normal fire control processes in place and the worst case for a transformer fire if the foreseen measures fail.

Such a flow diagram should be complemented by a more detailed list of risk reduction strategies in place (technical and procedural measures).

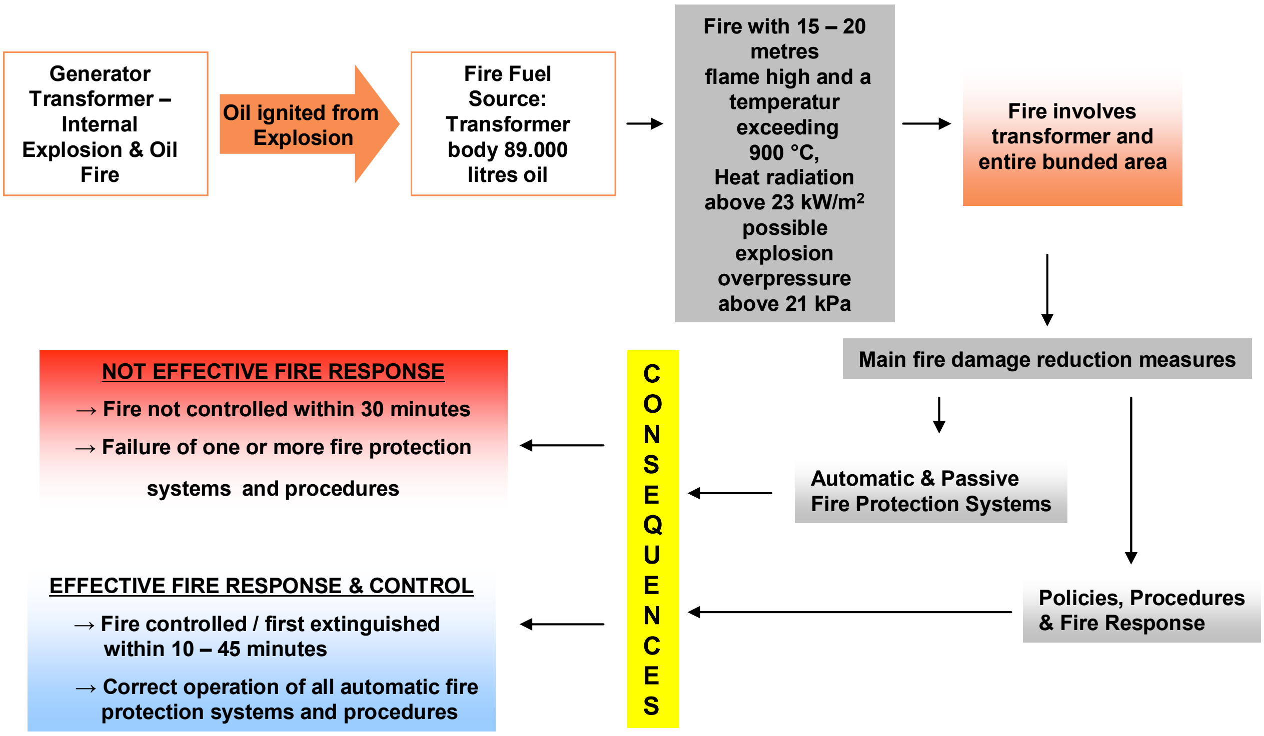


Figure 13. Flow diagram for an assumed generator transformer internal explosion and resulting fire.

13 OUTLOOK

Transformer are considered as vulnerable equipment because of the large quantity of oil in contact with high voltage elements since they can result in dangerous spillages, expensive damages and possible environmental pollution.

In particular, worldwide experience has shown an increasing number of transformer explosions and fires in all types of power plants. Therefore, these phenomena have been investigated and discussed in more detail in this paper with regard to causes for these events, potential influence of the age of the transformers and possible diagnostic measures in order to avoid such events. For that purpose, different types of databases have been evaluated.

However, the different databases have to be used very carefully, since the underlying criteria are not known in all cases or they are different, which requires a careful interpretation. Even the OECD FIRE Database providing data in a well structured manner is not homogeneous enough due to different reporting criteria in the member countries ranging from reporting every type of fire to reporting only fires with safety significant consequences according to their national regulations. In addition, databases of insurance companies or industries provided only a selected picture, e.g. collecting data in IMIA member countries (see Bartley 2003 and 2005) or investigating manufacturer specific transformer types (Petersson et al. 2008). Moreover, the population of transformer investigated is sometimes different.

Both offline and online diagnostics of transformers can be extremely successful to avoid significant events.

Besides monitoring the condition of the transformer, it might be possible to limit the consequences of a transformer explosion, e.g. by protective walls surrounding the transformers to limit the propagation of the explosions while sprinklers extinguish the induced fire. Nevertheless, despite of these equipments, transformers still explode like in the case of the example shown in section 4.

A further important task for receiving the required risk informed insights is to compare the distinguishing parameters of transformers such as their insulation type, number of phases, adjustability, core/coil configuration and winding configurations, oil content, design against overpressure, maintenance and monitoring features. In addition, the fire extinguishing systems installed in the locations of transformers have to be considered which may also affect the fire duration. Such investigations are intended in the future. For that purpose, exemplary experiences are also helpful.

As a result of the German transformer fire outlined in this paper the regulatory authority is mulling the inclusion of transformers outside reactor buildings into the routine supervisory activities although these transformers are operational components with no direct safety significance.

14 REFERENCES

Amrulloh, Y., Abeyratne, U. & Ekanayake, C. (2010). Gas and optical sensing technology for the field assessment of transformer oil. *International Journal of Emerging Electric Power Systems*, Vol. 11, Issue 4, August 2010.

Bartley, W.H. (2003). Analysis of transformer failures. *Proceedings 36th Annual Conference of the International Association of Engineering Insurers*, Stockholm 2003.

Bartley, W.H. (2005). Analysis of transformer failures …the insurance company perspective. *Proceedings AVO New Zealand International Conference*, Methven, New Zealand, April 19 – 21, 2005.

Berg, H.P., Fritze, N. (2010). Power plant transformer explosion and fire. *Journal of Polish Safety and Reliability Association, Summer Safety & Reliability Seminars*, Volume 1, Gdynia, Poland, June 20 – 26, 2010, 35 - 42

Berg, H.P., Forell, B., Fritze, N. & Röwekamp, M. (2009). First national applications of the OECD FIRE Database. *Proceedings of SMIRT 20, 11th International Seminar of Fire Safety in Nuclear Power Plants and Installations*, August 17 – 19, 2009, GRS-A-3496, Paper 3.19.

Berg, H.P., Forell, B., Fritze, N. & Röwekamp, M. (2010a). Exemplary applications of the OECD FIRE Database. *Compact of the Annual Meeting on Nuclear Technology*, 4 – 6 May 2010, Berlin, paper 306.

Berg, H.P., Fritze, N., Forell, B. & Röwekamp, M. (2010b). Risk oriented insights in transformer at nuclear installations. *Proceedings of ESREL 2010, September 5 -11, 2010*, 351 – 361.

EPRI and USNRC. (2005). *EPRI/NRC-RES Fire PRS Methodology for Nuclear Power Facilities; Volume 2: Detailed Methodology*. EPRI TR-101/1089 and NUREG/CR-6850, Electric Power Research Institute and U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research.

Flores, W.C., Mombello, E., Jardini, J. A. & Ratta, G. (2009). Fuzzy risk index for power transformer failures due to external short-circuits. *Electrical Power Systems Research*, 79 (4): 539 – 549.

Foata, M. (2010). Transformer fire risk and mitigation. *CIGRE A2 Transformers Session*, 2010.

Foata, M. & Nguyen, V. N. (2010). Transformer tank rupture and mitigation – Hydro-Quebec perspective, Presentation, September 29, 2010.

Franzen, A. & Karlsson, S. (2007). *Failure modes and effects analysis of transformers*. Royal Institute of Technology, KTH, Stockholm: TRITA-EE 2007:040.

Fire Risk Solution. (2009). *Earing Energy – Power Station Upgrade Fire Safety Study,* March 2009.

Hribernik, W., Kubicek, B., Pascoli, G., & Fröhlich, K. (2008). Verification of a model-based diagnosis system for on-line detection of the moisture content of power transformer insulations using finite element calculations. *International Conference on Condition Monitoring and Diagnosis, Bejing, China, April 21-24, 2008.*

Lehto, M., Marttila, J. & Virolainen, R. (2010). Risk informed insights from transformer fires at NPPs. *Proceeding of PSAM 10 - 10th International Probabilistic Safety Assessment & Management Conference, June 7-1, 2010.*

Lord, T. & Hodge, G. (2008). On-line monitoring of transformers as a strategic tool.

Minhas, M., Reynders, J. & de Clerk, P. (1999). Failure in power system transformers and appropriate monitoring techniques. *High Voltage Engineering Symposium, 22 – 27 August 1999*, Conference publication no. 467, IEEE.

Mirzai, M., Gholami, A. & Aminifar, F. (2006). Failures analysis and reliability calculation for power transformers. *J. Electrical Systems*, 2-1, 1 – 12.

Ng, K.-L. A. (2007). *Risk assessment of transformer fire protection in a typical New Zealand high-rise building*. Thesis, University of Canterbury, Christchurch, New Zealand.

Petterson, L., Melzer, L., Bengtsson, C., Fantana, N. (2008). Transformer challenges. *ABB Review* 1: 29 – 33.

Scheurer, D. et al. (2007a). Study and design of power plant transformer explosion and fire prevention Part I, *Electricity Today*, October 2007.

Scheurer, D. et al. (2007b). Study and design of power plant transformer explosion and fire prevention Part II, *Electricity Today*, November 2007.

Shahsiah, A., Degeneff, R.C. & Nelson, J.K. (2007). Modeling dynamic propagation of characteristic gases in power transformers oil-paper insulation, *IEEE Transactions on Dielectrics and Electrical Insulation*., Vol. 14, Issue 3, 710 – 717.

Stiegemeier, C.L. (2007). Power transformer analytical assessment - The micro approach. *IEEE/IPES Transformer Committee Tutoria*l, March 13, 2007, Dallas.

United States Department of Energy (2006). *Benefits of Using Mobile Transformers and Mobile Substations for Rapidly Restoring Electrical Service,* Report to the United States Congress pursuant to section 1816 of the Energy Policy Act of 2005, August 2006.

United States Department of the Interior (2003). *Transformer Diagnostics*. Facility Instructions, Standards and Techniques, Volume 3 – 31, June 2003.

United States Department of the Interior (2005). *Transformer Fire Protection*. Facility Instructions, Standards and Techniques, Volume 3 – 32, January 2005.

U.S. Nuclear Regulatory Commission (2010). *Resolution of Generic safety issues: issue 107: main transformer failures (Rev. 3)*. NUREG-933, Main report with supplements 1 – 32, January 15, 2010.

Valta, V. (2007). Oil-insulated power transformers. *If's Risk Management Journal,* 2/2007, 13 – 15.

Visser, K. & Brihmohan, Y. (2008). The effectiveness of oil analysis for failure prediction of power transformers. *Electricity Distribution Maintenance Summit*, June 10, 2008.

BAYESIAN UNCERTAINTY DECISION ANALYSIS

Y. H. Cui, R. Guo, T. Dunne
•
University of Cape Town, Rondebosch, 7701, Cape Town, South Africa
e-mail: rguo@worldonline.co.za

D. Guo
•
South African National Biodiversity Institute, Cape Town, South Africa
e-mail: D.Guo@sanbi.org.za

ABSTRACT

Bayesian statistical decision theory would be questionable when applied directly to non-random uncertainty circumstances. In this paper, we investigate the basic elements of decision analysis oriented to observational data arising from a general uncertainty environment, so that a framework for Bayesian uncertainty decision doctrine is established. Further, we propose a copula-linked uncertainty marginals mechanism for constructing the uncertainty multivariate distributions to represent both observational data and an uncertainty parameter vector. This mechanism paves the way towards the establishment of an uncertainty posterior distribution of the parameter vector given the observational data, based on uncertain measure Axiom 5. Finally, we present an illustrative example of the development of a posterior uncertainty distribution for a parameter given a single observation, step by step. The significance of this paper is to establish for the first time a Bayesian uncertainty data inference and decision framework, which constitutes a critical step towards the establishment of uncertainty statistics and a Bayesian uncertainty decision theory.

1 INTRODUCTION

Any applied mathematical model is proposed to reflect a particular aspect of the real natural world. Decision making moving from analysis of the collected data (information) to reach a final decision is actually a process to resolve the uncertainty being faced. In the real world, there are many forms of *uncertainty* surrounding us, but thus far we may only deal successfully with uncertainty as randomness or fuzziness within information. How should we solve the problems with other kind of uncertainty in real business life? For example, recently, a "Made in Japan" crisis was triggered by a Toyota Prius brake fault event and quickly spread widely over other industries widely. At the first glance, it may seem a trivial event has been exaggerated by journalists. It is a well-known fact that Japanese manufacturing arms itself to the teeth with statistical quality control. There is no reason to ascribe the fault event to an absence of total quality management. Nevertheless, we cannot deny what happened, and infer that the event indicates that some unaddressed problem exists there. The only possible answer is the methodology used to manage the quality imperative does not match the real quality problem faced. In other words, while the existing quality control and decision making doctrine, which is based on probability theoretical foundation, addressing random uncertainty problems, is powerful, nevertheless for other forms of uncertainty problems, the existing theory and methodologies may be inadequate. The law of the real world tells us that each specific form of uncertainty must be addressed by the corresponding specific uncertainty doctrine and methodology. There is no universal law for addressing all the forms of uncertainties.

In this paper, we first review the basic elements of Liu's (2007, 2009, 2010) uncertain measure theory in Section 2, and further investigate a copula-linked uncertainty marginals approach, to construct multivariate uncertainty distributions. The purpose is to represent both observational data and an uncertainty parameter vector. In Section 3, we note the basic elements of an observational-data oriented decision analysis under general uncertainty environments, in contrast to probabilistic Bayesian decision theory (Lee (1989), Cheng (1981), Bernardo & Smith (1994)), in order to establish a framework for a Bayesian uncertainty decision theoretic foundation. In Section 4, we propose a method to construct a posterior uncertainty distribution of parameter vector given the observational data, in terms of uncertain measure Axiom 5, see Liu (2010). In Section 5 we present an illustrative example, namely the development of a posterior uncertainty distribution for a parameter given a single observation, step by step. Section 6 concludes this paper.

2 UNCERTAIN MEASURE FOUNDATION

Uncertain measure (Liu (2007, 2009, 2010)) is an axiomatically defined set function mapping from a σ-algebra of a given space (set) to the unit interval [0,1], which provides a measuring grade system of an uncertain event (a reflection of an uncertainty phenomenon) and enables the formal definition of an uncertain variable and its uncertainty distribution.

Let Ξ be a nonempty set (space), and $\mathcal{A}(\Xi)$ the σ-algebra on Ξ. Each element, let us say, $A \subset \Xi, A \in \mathcal{A}(\Xi)$ is called an uncertain event. A number denoted as $\mathcal{M}\{A\}$, $0 \le \mathcal{M}\{A\} \le 1$, is assigned to the event $A \in \mathcal{A}(\Xi)$, which indicates the uncertain measuring grade with which event $A \in \mathcal{A}(\Xi)$ occurs. The normed set function $\mathcal{M}\{A\}$ satisfies following axioms given by Liu (2007, 2009, 2010):

Axiom 1: (Normality) $\mathcal{M}\{\Xi\} = 1$.

Axiom 2: (Monotonicity) $\mathcal{M}\{\cdot\}$ is non-decreasing, i.e., whenever $A \subset B$, $\mathcal{M}\{A\} \le \mathcal{M}\{B\}$.

Axiom 3: (Self-Duality) $\mathcal{M}\{\cdot\}$ is self-dual, i.e., for any $A \in \mathcal{A}(\Xi)$, $\mathcal{M}\{A\} + \mathcal{M}\{A^c\} = 1$.

Axiom 4: (σ- Subadditivity) $\mathcal{M}\left\{\bigcup_{i=1}^{\infty} A_i\right\} \le \sum_{i=1}^{\infty} \mathcal{M}\{A_i\}$ for any countable event sequence $\{A_i\}$.

Axiom 5: (Product Measure) Let $(\mathrm{X}_k, \mathcal{A}_{\mathrm{X}_k}, \mathcal{M}_k)$ be the k^{th} uncertain space, $k = 1,2,\ldots,n$. Then product uncertain measure $\mathcal{M}$ on the product measurable space $(\mathrm{X}, \mathcal{A}_{\mathrm{X}})$ is defined by

$$\mathcal{M} = \mathcal{M}_1 \wedge \mathcal{M}_2 \wedge \cdots \wedge \mathcal{M}_n = \min_{1 \le k \le n}\{\mathcal{M}_k\} \tag{1}$$

where

$$\Xi = \Xi_1 \times \Xi_2 \times \cdots \times \Xi_n = \prod_{k=1}^{n} \Xi_k \tag{2}$$

and

$$\mathcal{A}_{\Xi} = \mathcal{A}_{\Xi_1} \times \mathcal{A}_{\Xi_2} \times \cdots \times \mathcal{A}_{\Xi_n} = \prod_{k=1}^{n} \mathcal{A}_{\Xi_k} \tag{3}$$

That is, for each product uncertain event $\Lambda \subset \mathcal{A}_{\mathrm{X}}$ (i.e, $\Lambda = \Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \mathcal{A}_{\mathrm{X}_1} \times \mathcal{A}_{\mathrm{X}_2} \times \cdots \times \mathcal{A}_{\mathrm{X}_n} = \mathcal{A}_{\mathrm{X}}$), the uncertain measure of the event Λ is

$$\mathcal{M}\{\Lambda\} = \begin{cases} \sup\limits_{A_1 \times \cdots \times A_n \subset \Lambda} \min\limits_{1 \le k \le n} \mathcal{M}\{\Lambda_k\} & \text{if } \sup\limits_{A_1 \times \cdots \times A_n \subset \Lambda} \min\limits_{1 \le k \le n} \mathcal{M}\{\Lambda_k\} > 0.5 \\ 1 - \sup\limits_{A_1 \times \cdots \times A_n \subset \Lambda^c} \min\limits_{1 \le k \le n} \mathcal{M}\{\Lambda_k\} & \text{if } \sup\limits_{A_1 \times \cdots \times A_n \subset \Lambda^c} \min\limits_{1 \le k \le n} \mathcal{M}\{\Lambda_k\} > 0.5 \\ 0.5 & \text{otherwise} \end{cases} \tag{4}$$

Definition 2.1: (Liu (2007, 2009, 2010)) Any set function $\mathrm{D}:\mathrm{A}(\mathrm{X})\rightarrow[0,1]$ which satisfies Axioms 1-4 is called an uncertain measure. The triple $(\mathrm{X},\mathrm{A}(\mathrm{X}),\mathrm{D})$ is called the uncertain measure space.

Definition 2.2: An uncertain variable ξ is a measurable mapping, i.e., $\xi:(\Xi,\mathrm{A}(\Xi))\rightarrow(\mathrm{R},\mathrm{B}(\mathrm{R}))$, where $\mathrm{B}(\mathrm{R})$ denotes the Borel σ-algebra on $\mathrm{R}\equiv(-\infty,+\infty)$.

Remark 2.3: The fundamental difference between a random variable and an uncertain variable is the σ-additivity: the probability measure obeys σ-additivity (Kolmogorov (1950), Primas (1999)) and the uncertain measure (Liu (2007, 2009, 2010), Liu (2008)) obeys σ-subadditivity. The way of specifying measure inevitably has impacts on the behaviour of the measurable function over the triple, and hence on the mathematical characterization of the theories. For example, in contrast to probability theory, no "uncertainty density function" can be defined and then be entered into an integral of density to characterise an uncertainty distribution. Because an uncertain measure is permitted to be σ-subadditive, any set of uncertainty distributions derived from integration, being necessarily σ-additive, will necessarily be incomplete.

Definition 2.4: (Liu (2007, 2009, 2010)) The uncertain distribution $\Psi:\mathrm{R}\rightarrow[0,1]$ of an uncertain variable ξ on $(\mathrm{X},\mathrm{A}(\mathrm{X}),\mathrm{D})$ is

$$\Psi_{\xi}(x)=\lambda\{\tau\in\Xi\,|\,\xi(\tau)\leq x\} \tag{5}$$

Theorem 2.5: (Peng and Iwamura (2010)) The necessary and sufficient conditions for a function $\Psi:\mathbb{R}\rightarrow[0,1]$ be an uncertainty distribution function is that Ψ is non-decreasing function and

$$0\leq\Psi(x)\leq 1,\ \forall x\in\mathbb{R} \tag{6}$$

The function Ψ is referred to an uncertainty distribution function.

Remark 2.6: A probability distribution $F_X(x)$ requires right-continuity and $F_X(-\infty)=0, F_X(+\infty)=1$ in addition to those requirements of the uncertainty distribution function, while an uncertainty distribution is not limited by any continuity and $\Psi_{\xi}(-\infty)=0,\Psi_{\xi}(+\infty)=1$ requirements. This relaxation enables an uncertainty distribution to model even the most complicated pattern in real world data. The following definition reveals an essential characteristic of the uncertainty distribution.

Definition 2.7: Let ξ be an uncertainty variable, which takes values from a subset, denoted as E, of the real line $\mathbb{R}$, with n discontinuity points collected in an ascending order as set $\mathrm{D}\equiv\{c_1,\cdots,c_n\}$. The uncertainty distribution , Ψ ,of the variable ξ is specified as follows:

1. On the set $\mathrm{D}\equiv\{c_0,c_1,\cdots,c_n\}$,

$$\begin{gathered}\Psi(c_i-)=\psi_{i-},\ \Psi(c_i)=\psi_i,\ \Psi(c_i+)=\psi_{i+}\\ i=1,2,\cdots,n\end{gathered} \tag{7}$$

where $\psi_{i-}<\psi_i<\psi_{i+},\ \psi_{1-}\geq 0,\ \psi_{n+}\leq 1,\ i=1,2,\cdots,n$;

2. At the inner points of the sub-intervals (c_{i-1},c_i), $i=1,2,\cdots,n$, the uncertainty distribution Ψ is continuous

$$\Psi(z)=\begin{cases}\psi_{i-1+} & \text{if } z\downarrow c_{i-1}\\ \Lambda_i(z) & \text{if } z\in(c_{i-1},c_i)\\ \psi_{i-} & \text{if } z\uparrow c_i\end{cases} \tag{8}$$

where the function Λ_i is positive, non-decreasing, and bounded by ψ_{i-1+} and ψ_{i-}, i.e., $\psi_{i-1+} \le \Lambda_i \le \psi_{i-}$, $i = 1,2,\cdots,n$. Then Ψ is an uncertainty distribution of the essential form and ξ is called an essential uncertain variable.

Remark 2.8: The aim of this paper is to develop an observational-data oriented decision making doctrine. Whenever an observation is obtained, this specific observation should not be regarded as an isolated real number (or a real-valued vector), rather, it should be regarded as a representative from a population typically specified by a hypothesized uncertainty distribution. This approach matches the standard viewpoint in the statistical community. It is also a convention that the term "population" (Cheng 1981)) is equivalent to the term distribution, or to the term random variable. In the new uncertainty theory, this statistical convention continues. We formally state this convention as a definition on observational data.

Definition 2.9: An observation is a real number, (or more broadly, a symbol, or an interval, or a real-valued vector, a statement, etc), which is a representative of a population or equivalently of an uncertainty distribution under a given scheme comprising set and σ-algebra.

Remark 2.10: The uncertainty distribution is unknown but exists objectively. A workable solution is to hypothesize a family of uncertainty distributions of a specified functional form with unknown parameter θ, where the family is denoted by $\{\Psi_\xi^\theta, \theta \in \Theta\}$.

Definition 2.11: (Liu (2007, 2009, 2010)) Let multivariate uncertainty variable $(\xi_1,\xi_2,\cdots,\xi_d)$ be defined on an uncertain measure space $(\mathrm{X},\mathrm{A}(\mathrm{X}),\mathrm{D})$, then the multivariate function $\Psi_{\xi_1,\xi_2,\cdots,\xi_d} : D \to [0,1]$ is called an multivariate uncertainty distribution if

$$\Psi_{\xi_1,\xi_2,\cdots,\xi_d}(x_1,x_2,\cdots,x_d) = \mathrm{D}\{\xi_1 \le x_1, \xi_2 \le x_2, \cdots, \xi_d \le x_d\} \tag{9}$$

To present a concrete form of a multivariate uncertainty distribution, Guo et al. (2010) propose a copula-linked uncertainty marginals approach.

Definition 2.12: Let $(\xi_1,\xi_2,\cdots,\xi_d)$ be a multivariate uncertainty variable with joint uncertainty distribution $\Psi_{\xi_1,\xi_2,\cdots,\xi_d}(x_1,x_2,\cdots,x_d)$, in which all the marginal uncertainty distributions $\Psi_{\xi_1}(\cdot),\Psi_{\xi_2}(\cdot),\cdots,\Psi_{\xi_d}(\cdot)$ exist and are regular (i.e., $\Psi_{\xi_i}^{-1}(\cdot)$ exists, $i = 1,2,\cdots,d$). Then the uncertainty copula is defined by

$$C\left(\Psi_{\xi_1}(x_1),\Psi_{\xi_2}(x_2),\cdots,\Psi_{\xi_d}(x_d)\right) = \Psi_{\xi_1,\xi_2,\cdots,\xi_d}(x_1,x_2,\cdots,x_d) \tag{10}$$

We use a bivariate uncertainty distribution as an illustrative multivariate example.

Example 2.13: Let bivariate uncertainty variable (ξ_1,ξ_2) have marginal uncertainty distributions $\Psi_{\xi_1}(\cdot)$ and $\Psi_{\xi_2}(\cdot)$ respectively. The Farlie-Gumbel-Morgenstern (FGM) copula is defined by

$$C(u_1,u_2) = u_1u_2\left(1 + v\,(1-u_1)(1-u_2)\right),\ v \in [-1,1] \tag{11}$$

Further, let the bivariate uncertainty variable (ξ_1,ξ_2) have marginal uncertainty distributions $\Psi_{\xi_1}(\cdot)$ and $\Psi_{\xi_2}(\cdot)$ respectively, where

$$\Psi_{\xi_i}(x_i) = \frac{1}{1 + \exp\left(-\dfrac{\pi}{\sqrt{3}\sigma_i}(x_i - \theta_i)\right)},\ i = 1,2 \tag{12}$$

Then the bivariate FGM-Normal joint uncertainty distribution is

$$\Psi_{x_1,x_2}(x_1,x_2)=\sum_{i=1}^{2}\frac{1}{1+\exp\left(-\frac{\pi}{\sqrt{3}\sigma_i}(x_i-\theta_i)\right)}\left[1+\nu\sum_{i=1}^{2}\frac{\exp\left(-\frac{\pi}{\sqrt{3}\sigma_i}(x_i-\theta_i)\right)}{1+\exp\left(-\frac{\pi}{\sqrt{3}\sigma_i}(x_i-\theta_i)\right)}\right] \tag{13}$$

Finally, it is necessary to prepare the uncertainty expectation and the variance of an uncertainty variable to support the development of an uncertainty decision doctrine.

Definition 2.14: (Liu (2007, 2009, 2010)) Let ξ be an uncertainty variable defined on the uncertain space $(\Xi, \mathrm{A}(\Xi), \text{ƛ})$, then the expectation of ξ is

$$\mathrm{E}[\xi]=\int_{0}^{+\infty}\text{ƛ}\{\xi\ge r\}dr-\int_{-\infty}^{0}\text{ƛ}\{\xi\le r\}dr \tag{14}$$

provided at least one of the two integrals is finite.

Definition 2.15: (Liu (2007, 2009, 2010)) Let ξ be an uncertainty variable with finite expectation $\mathrm{E}[x]$, then the variance of ξ is

$$\mathrm{V}[\xi]=\mathrm{E}\left[(\xi-\mathrm{E}[\xi])^2\right] \tag{15}$$

Theorem 2.16: Let ξ be an uncertainty variable on uncertain measure space $(\mathrm{X},\mathrm{A}(\mathrm{X}),\mathrm{D})$ and h be a monotonic non-decreasing function $h:\Re\to\Re^{+}$, then the expectation of $h(x)$ is

$$\mathrm{E}[h(\xi)]=\int_{0}^{+\infty}h(r)\text{ƛ}\{\xi\ge r\}dr-\int_{-\infty}^{0}h(r)\text{ƛ}\{\xi\le r\}dr \tag{16}$$

3 ELEMENTS OF BAYESIAN DECISION THEORY

A decision theory is built upon a mathematical foundation, which provides a framework (or guidelines) for decision making according to a specified criterion, based on the observational data with a distribution of the assumed uncertainty type, e.g.,

1. The statistical decision is based on probability (measure) theory, which addresses the random uncertainty;
2. The fuzzy decision theory deals with fuzziness;
3. The uncertainty decision theory deals with a general uncertainty different from randomness or fuzziness.

Recall that the statistical decision theory is established on the axiomatic foundation of probability measure.

The basic elements of statistical decision are: (1) Sample space and distributional family; (2) Decision space; (3) Loss function and decision function.

It is necessary to point out the basic elements, namely *state, action,* and *loss* in statistical decision theory (Lee (1989), Cheng (1981), Bernardo & Smith (1994)), are still the essential elements in the Bayesian uncertainty decision theory.

Firstly, in statistical decision theory, the state, termed "state of nature" is regarded as objectively in existence, at least in some consensus sense. In contrast, in any general uncertainty environments, the state may include subjective, judgmental or even phenomenological events or factors. Note here the conceptual interpretations that *state* acquires across the decision environments, i.e., "reality" in front of the decision makers, along with possible virtual actions, and virtual loss. The differentiation between the *state of nature* in the statistical decision theory and the *state* in the uncertainty decision theory is critical. The former reflects more or less reflecting the "truth" for the frequentist school, while the uncertainty decision theory is a mixture of subjective and objective reflections.

Secondly, the connotations of *action* in the uncertainty decision theory is virtual, in that some elements are of a precautionary nature and do not correspond to any specific state element. The nature of the mapping is from multiple states to multiple actions.. However, the inclusion of virtual action elements is extremely important, because the top decision maker does not need to deal with routine decisions of day-to-day operations but with the extreme event(s) or the most important event decision(s).

Thirdly, the *loss* mechanism in both decision theories is the same. An uncertainty decision is a selection, which minimizes the loss function $l(\theta,a)$ of an action a from action space A for given state θ in the state space Θ. However, the social loss and environmental loss extract more and more attention from the public, NGO's and the governmental agencies. In the new uncertainty decision theory, the safety factor state, the health factor state, and the environmental factor state should be automatically assigned uncertain measure grades because of their intrinsic features. In the uncertainty decision theory, an action is made in terms of observational data, denoted by x, which is described by an uncertainty distribution $\Psi(x\,|\,\theta)$. Based on observational data x (i.e., representative of population $\Psi(x\,|\,\theta)$), a decision is actually a mapping from data space D into action space A. In other words,

$$a:\mathrm{D}\to\mathrm{A} \tag{17}$$

which can be expressed by

$$a=d(x) \tag{18}$$

The loss $l(\theta,d(x))$ is measurable on the joint uncertainty space.

Definition 3.1: The expected value of the loss with respect to the uncertainty distribution of observational data x

$$R(q,d)=\mathrm{E}_q\left[l(q,d(x))\right] \tag{19}$$

is called a risk function.

The uncertainty distribution of observational data x depends on state θ, because the dependence of $R(\theta,d)$ on θ enters explicitly from $l(\theta,a)$ and also through the state θ in the distribution function $\Psi(x\,|\,\theta)$ for x. Therefore the uncertainty distribution of the data determines the fundamental characteristics of observational-data oriented uncertainty decision theory, which deserves further exposure.

Let us consider the uncertainty decision problem for a given uncertainty distribution. Assume a state space $\Theta=\mathrm{R}$, and a continuous action space $\mathrm{A}\equiv\mathrm{R}$, and the loss function defined by

$$l(\theta,a)=w(\theta)(\theta-a)^2 \tag{20}$$

i.e., a quadratic loss function is assumed

Definition 3.2: (Uncertainty Bayes loss) Given a continuous state space Θ, the uncertainty variable θ is defined on uncertain space $(\Theta,\mathrm{B}(\Theta),\lambda_\theta)$, where $\lambda_\theta(\cdot)$ is an uncertain measure. The uncertain distribution $\Psi_\theta(\cdot)$ is defined on $(\Theta,\mathrm{B}(\Theta)))$. Then the average of loss with respect to state space for a given action $a\in\mathrm{A}$, is the quantity

$$B(a)=\mathrm{E}\left[l(\theta,a)\right]=\int_\Theta l(y,a)\,d\Psi_\theta(y) \tag{21}$$

and is called the uncertainty Bayes loss for a given action a.

Definition 3.3: (Uncertainty Bayes risk) The uncertainty Bayes risk is defined by

$$B(d) = \mathrm{E}\left[l\left(\theta, d(x)\right)\right] = \int_{\Theta} l\left(\theta, d(x)\right) d\Psi(\theta) \tag{22}$$

Definition 3.4: (Uncertainty Bayes rule) A Bayes decision rule, denoted as d^B, is a rule such that the Bayes risk is minimized, i.e.,

$$B\left(d^B\right) = \min_{d \in \mathrm{D}}\left\{B(d)\right\} \tag{23}$$

Example 3.5: Given a continuous state space $\Theta = \mathrm{R}$, the uncertainty variable θ is defined on an uncertain space $\left(\mathrm{R}, \mathrm{B}(\mathrm{R}), \lambda_{\partial}\right)$, where $\lambda_{\theta}(\cdot)$ is properly defined. The uncertainty distribution is assumed to be

$$G_{\theta}(y) = \frac{y-a}{2(b-a)} \vartheta_{[a,b]}(y) + \frac{y+c-2b}{2(c-b)} \vartheta_{[b,c]}(y) \tag{24}$$

Then we derive the average loss with respect to the state space for a given action $a \in \mathrm{A}$, as the uncertainty Bayes loss:

$$B(a) = \mathrm{E}\left[l(\theta, a)\right] = \int_{\Theta} w(y)(y-a)^2 \, d\Psi_{\theta}(y) \tag{25}$$

Set $w(\theta) = w_0$, a constant, then the uncertainty Bayes loss is

$$\begin{aligned}
B(a) &= \frac{w_0}{2(\beta-\alpha)} \int_{\alpha}^{\beta} (y-a)^2 dy + \frac{w_0}{2(\gamma-\beta)} \int_{\beta}^{\gamma} (y-a)^2 dy \\
&= \frac{w_0}{2(\beta-\alpha)}\left[(\beta-a)^3 - (\alpha-a)^3\right] + \frac{w_0}{2(\gamma-\beta)}\left[(\gamma-a)^3 - (\beta-a)^3\right] \\
&= \frac{w_0}{2}\left[(\beta-a)^2 + (\alpha-a)^2 + (\beta-a)(\alpha-a)\right] + \frac{w_0}{2}\left[(\gamma-a)^2 + (\beta-a)^2 + (\gamma-a)(\beta-a)\right] \\
&= \frac{w_0}{2}\left(3a^2 - 3(\alpha+\beta)a + \alpha^2 + \beta^2\right) + \frac{w_0}{2}\left(3a^2 - 3(\gamma+\beta)a + \gamma^2 + \beta^2\right) \\
&= w_0\left(3a^2 - \frac{3}{2}(\alpha + 2\beta + \gamma)a + \frac{1}{2}\left(\alpha^2 + 2\beta^2 + \gamma^2\right)\right)
\end{aligned} \tag{26}$$

With an appropriate specification of decision function in term of data, the uncertain Bayesian decision analysis can be formulated.

4 A POSTERIOR UNCERTAINTY DISTRIBUTION

When (Q, B_Q, l_q) is an uncertain (prior) space and (X, B_X, P_q) is a probability space, we actually use random sample information to make inferences on the uncertain parameter q. The critical step in the probabilistic Bayesian inference is to develop the posterior distribution for parameter q. We strongly believe that the Bayesian uncertainty inference requires parallel manipulations.

Let (Q, B_Q) be a parameter measurable space, (X, B_X) be a sample measurable space.

Definition 4.1: An uncertain measure defined on (Q, B_Q) is called an uncertain prior measure, denoted as l_q. The space (Q, B_Q, l_q) is called an uncertain prior space, the uncertain distribution $G(y) = l_q\{q \text{ J } y\}$ is called an uncertain prior distribution.

An uncertainty variable, denoted by x, is defined on a measurable space (X, B_X) with uncertainty distributional family $\{Y_q, q \text{ O } Q\}$ where Q is a parameter space. Formally,

Definition 4.2: The uncertainty observations are representatives of an uncertainty variable x, which is called an uncertainty population, or alternatively, called as an uncertainty distribution $\Psi_{\xi}\left(x|\theta\right),\theta\in\Theta$. The uncertainty variable x is defined on $\left(\mathcal{X},\mathcal{A}_{\mathcal{X}},\lambda_{\mathcal{X}}\right)$. The uncertainty distribution is

$$\Psi_{\xi}\left(x|\theta\right)=\lambda_{\xi}\left\{\xi\le x|\theta\right\} \tag{27}$$

Remark 4.3: The uncertainty observations are presented by observers or experts, while the prior distribution (prior uncertain measure) is offered by knowledgeable experts on the observers' behaviors. In probabilistic Bayesian statistics it is typically assumed that the prior and the likelihood are independent of each other. In Bayesian uncertainty doctrine we continue to follow this convention without any theoretical justification, although the independence between prior and likelihood is debatable.

Remark 4.4: The joint cumulative distribution of the observational data $\left(x_1,x_2,\cdots,x_n\right)$ may be specified by a hypothesized copula functional according to the features of the data, as

$$\begin{aligned}&\Psi_{\xi_1,\xi_2,\cdots,\xi_n}\left(x_1,x_2,\cdots,x_n|\theta\right)\\&=\lambda_{\xi_1,\xi_2,\cdots,\xi_n}\left\{\xi_1\le x_1,\xi_2\le x_2,\cdots,\xi_n\le x_n\mid\theta\right\}\\&=C_{\underline{\varpi}}\left(\Psi_{\xi_1}\left(x_1\right),\Psi_{\xi_2}\left(x_2\right),\cdots,\Psi_{\xi_n}\left(x_n\right)\right)\end{aligned} \tag{28}$$

where $\Psi_{\xi_1}\left(\cdot\right),\Psi_{\xi_2}\left(\cdot\right),\cdots,\Psi_{\xi_n}\left(\cdot\right)$ are given marginal uncertainty distributions and $\underline{v}$ is unknown parameter vector. In contrast, within probability theory, the multivariate joint distribution function is $F_{X_1,X_2,\cdots,X_n}\left(x_1,x_2,\cdots,x_n|\theta\right)$. Given a population $F\left(x|\theta\right)$, the *i.i.d.* random sampling observations have a joint distribution function

$$F_{X_1,X_2,\cdots,X_n}\left(x_1,x_2,\cdots,x_n|\theta\right)=\prod_{k=1}^{n}F_{X_k}\left(x_k|\theta\right) \tag{29}$$

Also, the joint density (i.e., the likelihood function) is

$$f_{X_1,X_2,\cdots,X_n}\left(x_1,x_2,\cdots,x_n|\theta\right)=\frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2}\cdots\frac{\partial}{\partial x_n}F_{X_1,X_2,\cdots,X_n}\left(x_1,x_2,\cdots,x_n|\theta\right)=\prod_{k=1}^{n}f_{X_k}\left(x_k|\theta\right) \tag{30}$$

Now, let us continue our arguments on the posterior uncertainty distribution of θ. For the convenience, let us assume that a pair of observations (x_1,x_2) is obtained from the bivariate uncertainty variable, denoted by (x_1,x_2), which is defined by a hypothesized bivariate FGM-normal uncertainty distribution

$$\Psi_{x_1,x_2}(x_1,x_2)=\prod_{i=1}^{2}\frac{1}{1+\exp\left(-\frac{\pi}{\sqrt{3}\sigma_i}(x_i-\theta_i)\right)}\left[1+\nu\prod_{i=1}^{2}\frac{\exp\left(-\frac{\pi}{\sqrt{3}\sigma_i}(x_i-\theta_i)\right)}{1+\exp\left(-\frac{\pi}{\sqrt{3}\sigma_i}(x_i-\theta_i)\right)}\right] \tag{31}$$

with marginals

$$\Psi_{x_i}(x_i)=\frac{1}{1+\exp\left(-\frac{\pi}{\sqrt{3}\sigma_i}(x_i-\theta_i)\right)},\ i=1,2 \tag{32}$$

Then the bivariate uncertainty distribution has parameter vector $\theta=(\theta_1,\sigma_1,\theta_2,\sigma_2,\nu)$. For simplification only, we set $\theta=\theta_1=\theta_2$, and assume that both $\sigma_0=\sigma_1=\sigma_2$, and ν are known. Then what we aim to derive the posterior distribution of parameter θ, i.e., $\Psi_{\theta}(y\mid x_1,x_2,\sigma_0,\nu_0)$.

For the parameters, there is again a specification issue of the joint multivariate uncertainty prior distribution. For example, the FGM-normal bivariate uncertainty distribution $Y_{x_1,x_2}(x_1,x_2)$ has five parameters, i.e., $q = (q_1, s_1, q_2, s_2, v)$. The full specification of prior vector needs a five-dimensional copula, $C_j(v_1,v_2,v_3,v_4,v_5)$ with marginals $p_i(v_i), i = 1,2,L,5$, and the prior takes a form

$$Y_{v_1,v_2,L,v_5}(y_1,y_2,L,y_5) = C_j\left(p_1(v_1), p_2(v_2), L, p_5(v_5)\right) \tag{33}$$

Definition 4.5: Let $\Psi(x_1,x_2,\cdots,x_n,\underline{\theta})$ be the joint uncertainty distribution of the uncertainty observations $(x_1,x_2,\cdots,x_n)$ together with the parameter vector $\underline{q}$. Note the event

$$\Lambda = \left\{\underline{X} \le \underline{x}, \underline{\theta} \le \underline{y}\right\} \tag{34}$$

Then the joint distribution of $\underline{X}$ and $\underline{\theta}$ defined on (X, A_X, D_X) and (Q, A_Q, D_q) respectively, according to Axiom 5, Liu (2010), is the joint uncertainty measure defined by

$$\begin{aligned}
&\Psi_{\underline{X},\underline{\theta}}(\underline{x},\underline{y}) = \lambda_{\underline{X},\underline{\theta}}\left\{\underline{X} \le \underline{x}, \underline{\theta} \le \underline{y}\right\} \\
&= \begin{cases}
\sup\limits_{A_1 \times A_2 \subset \Lambda} \min\left(\lambda_{\underline{\theta}}\{A_1\}, \lambda_{\underline{X}}\{A_1\}\right) & \text{if if } \sup\limits_{A_1 \times A_2 \subset \Lambda} \min\left(\lambda_{\underline{\theta}}\{A_1\}, \lambda_{\underline{X}}\{A_1\}\right) > 0.5 \\
1 - \sup\limits_{A_1 \times A_2 \subset \Lambda^C} \min\left(\lambda_{\underline{\theta}}\{A_1\}, \lambda_{\underline{X}}\{A_1\}\right) & \text{if if } \sup\limits_{A_1 \times A_2 \subset \Lambda^C} \min\left(\lambda_{\underline{\theta}}\{A_1\}, \lambda_{\underline{X}}\{A_1\}\right) > 0.5 \\
0.5 & \text{otherwise}
\end{cases}
\end{aligned} \tag{35}$$

Definition 4.6: We denote $\Psi_{\underline{X}}(\underline{x}) = \Psi_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n)$ as the absolute joint uncertainty distribution.

$$\Psi_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n) = \sup_{\underline{y}\in\Theta}\left(\Psi_{X_1,X_2,\cdots,X_n}\left(x_1,x_2,\cdots,x_n,\underline{y}\right|\right)\right) \tag{36}$$

For example, if a pair of bivariate FGM-normal uncertainty observation (x_1,x_2) is obtained, then

$$\Psi_{X_1,X_2}(x_1,x_2) = u_1u_2\left(1+\varpi(1-u_1)(1-u_2)\right) \tag{37}$$

where

$$u_i = Y_{x_i}(x_i) = \frac{1}{1+\exp\left(-\frac{p}{\sqrt{3}s_i}(x_i - q_i)\right)}, \quad i = 1,2 \tag{38}$$

Finally, we define a Bayesian uncertainty posterior for $\underline{q}$.

Definition 4.7: We denote $\Psi_{\underline{\theta}}\left(\underline{y}\middle|x_1,x_2,\cdots,x_n\right)$ as the posterior uncertainty distribution under the Maximum Uncertainty Principle, The MUP posterior uncertainty distribution is thus

$$\Psi_{\underline{\theta}}\left(\underline{y}\middle|x_1,x_2,\cdots,x_n\right) = \frac{\Psi_{\xi_1,\xi_2,\cdots,\xi_n,\underline{\theta}}\left(x_1,x_2,\cdots,x_n,\underline{y}\right)}{\Psi_{\xi_1,\xi_2,\cdots,\xi_n}\left(x_1,x_2,\cdots,x_n\right)} \tag{39}$$

5 A BAYESIAN POSTERIOR UNCERTAINTY EXAMPLE

In this section, we take the uncertainty zigzag distribution as the uncertainty prior, and Liu's (2007, 2009, 2010) normal distribution as uncertainty observation distribution, and in a step by step manner, illustrate the construction of an posterior uncertainty distribution.

The observational data is assumed to be a representative value of the population observed, which can be specified by a hypothesized uncertainty distribution. For our example, the hypothesized uncertainty distribution is Liu's (2007, 2009, 2010) uncertainty normal distribution:

$$\Psi\left(x\middle|\theta,\sigma_0\right)=\frac{1}{1+\exp\left(-\frac{\pi}{\sqrt{3}\sigma_0}\left(x-\theta\right)\right)} \tag{40}$$

As an illustration, it is assumed that the standard deviation is given, denoted by s_0, the only unknown is parameter q, the mean or expectation of the uncertainty distribution. Because in Bayesian treatments the parameter in the distribution function is no longer an unknown real number, the parameter is treated as an uncertainty variable, Its distribution is supposed to be uniquely specified by a given uncertainty measure D_q defined on an uncertain measurable space (Q,A(Q)). In practice, the unknown parameter is usually specified by an uncertainty distribution. Although the uncertainty distribution can induce an uncertain measure on Borel measurable space (Y,B(Y)), nevertheless, it is unique in the sense of an equivalence class. The uncertainty prior distribution is assumed to be

$$G_\theta\left(y\right)=\frac{y-a}{2(b-a)}\vartheta_{[a,b]}\left(y\right)+\frac{y+c-2b}{2\left(c-b\right)}\vartheta_{[b,c]}\left(y\right) \tag{41}$$

where

$$\vartheta_{[a,b]}\left(y\right)=\begin{cases}1 & \text{if} \quad a<y\le b\\ 0 & \quad \text{otherwise}\end{cases} \tag{42}$$

We further assume that a single observation $x=2.3$ is taken from hypothesized Liu's uncertainty normal distribution $\Psi_X\left(x\middle|\theta,2.00\right)=1\Big/\left(1+\exp\left(-\pi\left(x-\theta\right)\big/2\sqrt{3}\right)\right)$ and the uncertainty prior parameter $(a,b,c)=(0,2,3)$, i.e., the uncertainty prior distribution is

$$G_\theta\left(y\right)=\frac{1}{4}y\vartheta_{[0,2]}\left(y\right)+\frac{1}{2}\left(y-1\right)\vartheta_{[b,c]}\left(y\right) \tag{43}$$

The Axiom 5 based posterior uncertainty distribution of θ given uncertainty observation $x=2.3$ is

$$\Psi_\theta\left(y\middle|x=2.3,\sigma_0=2\right)=\frac{\Psi_{\xi,\theta}\left(2.3,y\right)}{\sup\limits_{y\in\Theta}\left(\Psi_{\xi,\theta}\left(2.3,y\right)\right)} \tag{44}$$

where

$$\Psi_{\xi,\theta}\left(2.3,y\right)=\begin{cases}\sup\limits_{A_1\times A_2\subset\Lambda}\min\left(G_\theta\left(y\right),\Psi_\xi\left(2.3\middle|y,2.0\right)\right) & \text{if} \quad \sup\limits_{A_1\times A_2\subset\Lambda}\min\left(G_\theta\left(y\right),\Psi_\xi\left(2.3\middle|y,2.0\right)\right)>0.5\\ 1-\sup\limits_{A_1\times A_2\subset\Lambda^C}\min\left(G_\theta\left(y\right),\Psi_\xi\left(2.3\middle|y,2.0\right)\right) & \text{if} \quad \sup\limits_{A_1\times A_2\subset\Lambda^C}\min\left(G_\theta\left(y\right),\Psi_\xi\left(2.3\middle|y,\sigma_0\right)\right)>0.5\\ 0.5 & \text{otherwise}\end{cases} \tag{45}$$

and

$$\sup_{y\in\Theta}\left(\Psi_{\xi,\theta}\left(2.3,y\right)\right)=0.60392 \tag{46}$$

The plot of the posterior uncertainty distribution $\Psi_\theta\left(y\mid x=2.3,\sigma_0=2\right)$ is shown in Figure 1.

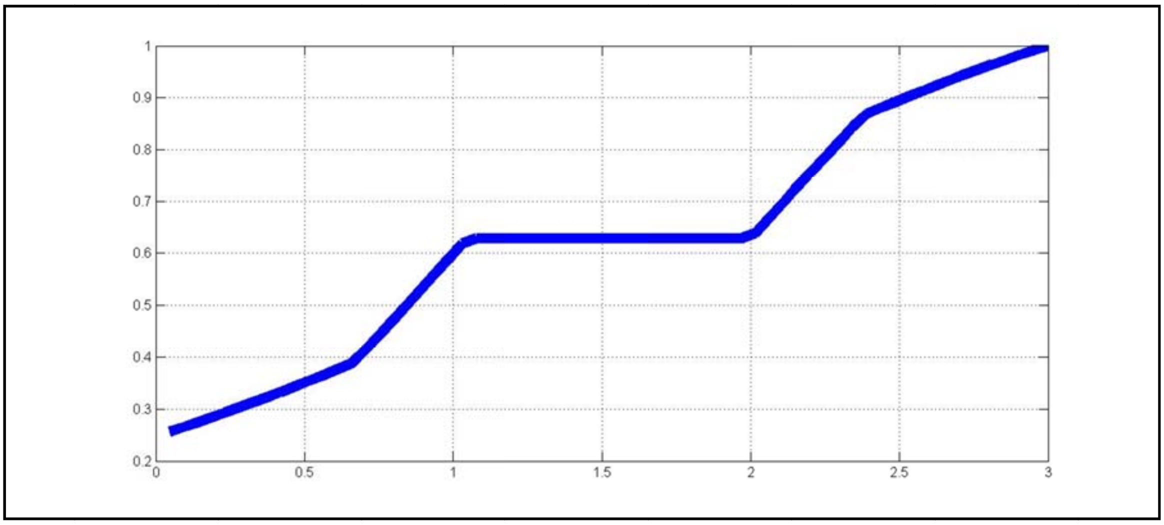


Figure 1. Posterior uncertainty distribution $\Psi_\theta\left(y\middle|x=2.3,\sigma_0=2\right)$

Once we find the posterior uncertainty distribution $\Psi_\theta\left(y\middle|x=2.3,\sigma_0=2\right)$, it is very natural to calculate the uncertainty posterior mean $\mathrm{E}[q\,|\,x=2.3]$and variance$V[q\,|\,x=2.3]$, and carry on further Bayesian inferential analysis.

6 CONCLUSIONS

The text should fit exactly into the type area of 210 × 297 mm (8.27" × 11.69"). For correct settings of margins in the Page Setup dialog box (File menu) see Table 1.

In this paper, a Bayesian uncertainty decision theoretical framework is proposed under the uncertain measure foundation, which paves the way toward data-oriented inferential uncertainty statistics.

The contributions of this paper are listed as follows: (1) for the first time a concrete uncertainty multivariate uncertainty distribution, in terms of an uncertainty copula with uncertainty marginals, is presented in the uncertainty theory literature; (2) for the first time an uncertainty product measure data-oriented posterior uncertainty distribution is developed from axiom 5; and (3) a detailed illustrative example is given in stepwise manner.

Definitely, the treatments in this paper are debatable. Particularly, the independence between prior measure and observational data uncertainty distribution measure may be contested. Also, we have not explored the necessary and sufficient conditions for an uncertainty copula constructed multivariate uncertainty distribution to be uncertainty measure. There are many unaddressed questions ahead of us for the new Bayesian uncertainty decision theory to become fully applicable in industries and business.

7 ACKNOWLEDGEMENTS

This research is supported by VC's Incentive for NRF Rated Researchers, University of Cape Town, and South African National Research Foundation Grant IFR2010042200062, and IFR2009090800013 as well.

The research is also partially supported by Tsinghua University during Professor Guo and Mr YH Cui's visit to the Uncertainty Theory Laboratory, Department of Mathematical Science, Tsinghua University from July 18 to October 14, 2010.

The authors would like to thank Professor B.D. Liu and his PhD students: X.W. Chen, W. Dai, K. Yao, and Z. X. Peng for their invaluable debates and inputs.

8 REFERENCES

1. Lee, P. M. (1989) Bayesian Statistics: An Introduction. Oxford University Press, New York, USA.
2. Cheng, X. R. (1981) Introductory Mathematical Statistical Theory. The Scientific Publishing House, Beijing, China (in Chinese).
3. Bernardo, J. M. & Smith A. F. M. (1994) Bayesian Theory. John Wiley & Sons Ltd, England.
4. Guo, R., Cui, Y. H. & Guo, D. (2010) Uncertainty Linear Regression models. (Draft paper submitted to Journal of uncertain ststems)
5. Kolmogorov, A. N. (1950) Foundations of the Theory of Probability, translated by Nathan Morrison, Publisher: Chelsea, New York.
6. Liu, B. D. (2007) Uncertainty Theory: An Introduction to its Axiomatic Foundations, 2nd Edition, Springer-Verlag Heidelberg, Berlin.
7. Liu, B. D. (2009) Some research problems in uncertainty theory, Journal of Uncertain Systems, Vol.3, No.1, 3-10.
8. Liu, B. D. (2010) Uncertainty Theory: A Branch of Mathematics of Modelling Human Uncertainty, Springer-Verlag, Berlin.
9. Liu, Y. H. (2008) How to generate uncertain measures, Proceedings of Tenth National Youth Conference on Information and Management Sciences, August 3-7, 2008, Luoyang, pp.23-26.
10. Peng, Z. X. & Iwamura K. (2010) A sufficient and necessary condition of uncertainty distribution. Journal of Interdisciplinary Mathematics, to be published.
11. Primas, H. (1999) Basic Elements and Problems of Probability Theory, *Journal of Scientific Exploration*, Vol. 13, No. 4, pp. 579–613.
12. Zadeh, L. A. (1965) Fuzzy Sets, *Information and Control*, Vol. 8, 338-353.

RT&A # 01 (20)
(Vol.2) 2011, March

HYBRID RELIABILITY MODELLING WITH IMPRECISE PARAMETER

R. Guo, Y. H. Cui, C. Thiart
●
University of Cape Town, Rondebosch, 7701,
Cape Town, South Africa
e-mail: rguo@worldonline.co.za

D. Guo
●
South African National Biodiversity Institute,
Cape Town, South Africa
e-mail: D.Guo@sanbi.org.za

ABSTRACT

The real world phenomena are often facing the co-existence reality of different formality of uncertainty and thus the probabilistic reliability modeling practices are very doubtful. Under complicated uncertainty environments, hybrid variable modeling is important in reliability and risk analysis, which includes Bayesian distributional theory, random fuzzy distributional theory, as well as fuzzy random distributional theory as special distribution families. In this paper, we define a new hybrid lifetime which is specified by a random lifetime distribution with imprecise parameter with an uncertainty distribution. We furthermore define the average chance distribution as a quality index for quantifying the hybrid lifetime and accordingly the average chance reliability is derived.

1 INTRODUCTION

System reliability, as a quality index, is the capability to complete the specified functions accurately in mutually harmonious manner under the specified conditions within specified period. The quality and reliability engineering facilitates the specification of the system reliability function on the ground of probability and statistics theory. The Toyota crisis does not only tear off the brand image of quality but also shake the belief of existing quality and reliability engineering practices and the underlying probability and statistics theory, which treat the random uncertainty. Uncertainty in real world is intrinsic and diversified in formality. For example, the vagueness is another form of uncertainty, which is more and more aware of in today's industrial environments, just as Carvalho & Machado (2006) commented, "In a global market, companies must deal with a high rate of changes in business environment. … The parameters, variables and restrictions of the production system are inherently vagueness." Therefore quality and reliability engineering is no longer a blind exercise of applying the traditional techniques from existing probabilistic reliability engineering literature.

The coexistence of randomness and other forms of uncertainty in reliability concept is intrinsic and inherent and therefore modern reliability analysis inevitably engages hybrid lifetime modeling.

Accordingly, the methodology to solve the reliability of hybrid lifetime should be developed in terms of the basic concept of general uncertain measure theory.

The remaining structure of the paper is stated as follows: Section Two serves reviewing Liu's axiomatic uncertain measure and defines the concept of impreciseness in terms of uncertainty distribution; Section Three is utilized to establish the hybrid variable theory. Particularly, the hybrid variable is constituted by a random lifetime with an imprecise parameter governed by an uncertainty distribution; Section Four defines the average chance measure for hybrid variable; Section Five is used to investigate the construction of hybrid variable; while in Section Six the commonly used lifetime models for construction of hybrid lifetime models are discussed; Section Seven uses exponential lifetime with imprecise uncertainty parameter for develop the average chance reliability as an illustrative examples; and Section Eight concludes the paper.

2 UNCERTAIN MEASURE AND IMPRECISENESS

Uncertain measure (Liu (2010)) is an axiomatically defined set function mapping from a σ-algebra of a given space (set) to the unit interval [0,1], which provides a measuring grade system of an uncertain phenomenon and facilitates the formal definition of an uncertain variable.

Let Ξ be a nonempty set (space), and $\mathrm{A}(\Xi)$ the σ-algebra on Ξ. Each element, let us say, $A\subset\Xi, A\in\mathrm{A}(\Xi)$ is called an uncertain event. A number denoted as $ƛ\{A\}$, $0\le ƛ\{A\}\le 1$, is assigned to event $A\in\mathrm{A}(\Xi)$, which indicates the uncertain measuring grade with which event $A\in\mathrm{A}(\Xi)$ occurs. The normal set function $ƛ\{A\}$ satisfies following axioms given by Liu (2007, 2009, 2010):

Axiom 1: (Normality) $ƛ\{\Xi\}=1$.

Axiom 2: (Monotonicity) $ƛ\{\cdot\}$ is non-decreasing, i.e., whenever $A\subset B$, $ƛ\{A\}\le ƛ\{B\}$.

Axiom 3: (Self-Duality) $ƛ\{\cdot\}$ is self-dual, i.e., for any $A\in\mathrm{A}(\Xi)$, $ƛ\{A\}+ƛ\{A^c\}=1$.

Axiom 4: (σ-Subadditivity) $ƛ\left\{\bigcup_{i=1}^{\infty}A_i\right\}\le\sum_{i=1}^{\infty}ƛ\{A_i\}$ for any countable event sequence $\{A_i\}$.

Axiom 5: (Product Measure) Let $\left(\mathrm{X}_k,\mathrm{A}_{\mathrm{X}_k},\mathrm{D}_k\right)$be the k^{th} uncertain space, $k=1,2,\mathrm{L},n$. Then product uncertain measure Don the product measurable space$(\mathrm{X},\mathrm{A}_{\mathrm{X}})$is defined by

$$ƛ=ƛ_1\wedge ƛ_2\wedge\cdots\wedge ƛ_n=\min_{1\le k\le n}\{ƛ_k\} \tag{1}$$

where

$$\Xi=\Xi_1\times\Xi_2\times\cdots\times\Xi_n=\prod_{k=1}^{n}\Xi_k \tag{2}$$

and

$$\mathrm{A}_{\Xi}=\mathrm{A}_{\Xi_1}\times\mathrm{A}_{\Xi_2}\times\cdots\times\mathrm{A}_{\Xi_n}=\prod_{k=1}^{n}\mathrm{A}_{\Xi_k} \tag{3}$$

That is, for each product uncertain event $\mathrm{L}\,\mathrm{O}\,\mathrm{A}_{\mathrm{X}}$ (i.e, $\mathrm{L}=\mathrm{L}_1\,\mathrm{r}\,\mathrm{L}_2\,\mathrm{r}\,\mathrm{L}\,\mathrm{r}\,\mathrm{L}_n\,\mathrm{O}\,\mathrm{A}_{\mathrm{X}_1}\,\mathrm{r}\,\mathrm{A}_{\mathrm{X}_2}\,\mathrm{r}\,\mathrm{L}\,\mathrm{r}\,\mathrm{A}_{\mathrm{X}_n}=\mathrm{A}_{\mathrm{X}}$), the uncertain measure of the event L is

$$ƛ\{\Lambda\}=\begin{cases}\sup\limits_{A_1\times\cdots\times A_n\subset\Lambda}\ \min\limits_{1\le k\le n}ƛ\{\Lambda_k\} & \text{if}\ \sup\limits_{A_1\times\cdots\times A_n\subset\Lambda}\ \min\limits_{1\le k\le n}ƛ\{\Lambda_k\}>0.5\\ 1-\sup\limits_{A_1\times\cdots\times A_n\subset\Lambda^c}\ \min\limits_{1\le k\le n}ƛ\{\Lambda_k\} & \text{if}\ \sup\limits_{A_1\times\cdots\times A_n\subset\Lambda^c}\ \min\limits_{1\le k\le n}ƛ\{\Lambda_k\}>0.5\\ 0.5 & \text{otherwise}\end{cases} \tag{4}$$

Definition 2.1: (Liu (2007, 2009, 2010)) Any set function $\mathrm{D}:\mathrm{A}(\mathrm{X})\text{®}\,[0,1]$ which satisfies Axioms 1-4 is called an uncertain measure. The triple $(\mathrm{X},\mathrm{A}(\mathrm{X}),\mathrm{D})$ is called the uncertain measure space.

Definition 2.2: (Liu (2007, 2009, 2010)) An uncertain variable ξ is a measurable mapping, i.e., $\xi:(\Xi,\mathrm{A}(\Xi))\to(\mathrm{R},\mathrm{B}(\mathrm{R}))$, where $\mathrm{B}(\mathrm{R})$ denotes the Borel σ-algebra on $\mathrm{R}\equiv(-\infty,+\infty)$.

Remark 2.3: The fundamental difference between a random variable and an uncertain variable is the σ-additivity: the probability measure obeys σ-additivity (Kolmogorov (1950), Primas (1999)) and the uncertain measure (Kaufmann (1975), Liu (2007, 2009, 2010)) obeys σ-subadditivity. The way of specifying measure inevitably has impacts on the behaviour of the measurable function over the triple, and hence on the mathematical characterization of the theories. For example, in contrast to probability theory, no “uncertainty density function” can be defined and then be entered into an integral of density to characterise an uncertainty distribution. Because an uncertain measure is permitted to be σ-subadditive, any set of uncertainty distributions derived from integration, being necessarily σ-additive, will necessarily be incomplete.

Definition 2.4: (Liu (2007, 2009, 2010)) The uncertain distribution $\Psi : \mathrm{R} \to [0,1]$ of an uncertain variable ξ on $(\mathrm{X}, \mathrm{A}(\mathrm{X}), \mathrm{D})$ is

$$\Psi_{\xi}(x) = \lambdabar\left\{\tau \in \Xi \middle| \xi(\tau) \le x\right\} \tag{5}$$

Theorem 2.5: (Peng and Iwamura (2010)) The necessary and sufficient conditions for a function $\Psi : \mathbb{R} \to [0,1]$ be an uncertainty distribution function is that Ψ is non-decreasing function and

$$0 \le \Psi(x) \le 1,\ \forall x \in \mathbb{R} \tag{6}$$

The function Ψ is referred to an uncertainty distribution function.

Remark 2.6: A probability distribution $F_X(x)$ requires right-continuity and $F_X(-\infty) = 0, F_X(+\infty) = 1$ in addition to those requirements of the uncertainty distribution function, while an uncertainty distribution is not limited by any continuity and $\Psi_{\xi}(-\infty) = 0, \Psi_{\xi}(+\infty) = 1$ requirements. This relaxation enables an uncertainty distribution to model even the most complicated pattern in real world data. The following definition reveals an essential characteristic of the uncertainty distribution.

Definition 2.7: Let ξ be an uncertainty variable, which takes values from a subset, denoted as E, of the real line $\mathbb{R}$, with n discontinuity points collected in an ascending order as set $\mathrm{D} \equiv \{c_1, \cdots, c_n\}$. The uncertainty distribution , Ψ ,of the variable ξ is specified as follows:

1. On the set $\mathrm{D} \equiv \{c_0, c_1, \cdots, c_n\}$,

$$\begin{aligned} &\Psi(c_i -) = \psi_{i-},\ \Psi(c_i) = \psi_i,\ \Psi(c_i +) = \psi_{i+} \\ &i = 1, 2, \cdots, n \end{aligned} \tag{7}$$

where $\psi_{i-} < \psi_i < \psi_{i+},\ \psi_{1-} \ge 0,\ \psi_{n+} \le 1,\ i = 1, 2, \cdots, n$;

2. At the inner points of the sub-intervals (c_{i-1}, c_i), $i = 1, 2, \cdots, n$, the uncertainty distribution Ψ is continuous

$$\Psi(z) = \begin{cases} \psi_{i-1+} & \text{if } z \downarrow c_{i-1} \\ \Lambda_i(z) & \text{if } z \in (c_{i-1}, c_i) \\ \psi_{i-} & \text{if } z \uparrow c_i \end{cases} \tag{8}$$

where the function Λ_i is positive, non-decreasing, and bounded by ψ_{i-1+} and ψ_{i-}, i.e., $\psi_{i-1+} \le \Lambda_i \le \psi_{i-}$, $i = 1, 2, \cdots, n$. Then Ψ is an uncertainty distribution of the essential form and ξ is called an essential uncertain variable.

Remark 2.8: Whenever an "observation" is obtained, this specific observation should not be regarded as an isolated real number (or a real-valued vector), rather, it should be regarded as a representative from a "population" typically specified by a hypothesized uncertainty distribution. This approach matches the standard viewpoint in the statistical community, see wikipedia (2010). It is also a convention that the term “population” (Bernardo & Smith (1994), Lee (1989)) is equivalent to the term distribution, or to the term random variable. In the new uncertainty theory, this statistical convention should be retained. We formally state this convention as a definition on observational data.

Definition 2.9: An observation is a real number, (or more broadly, a symbol, or an interval, or a real-valued vector, a statement, etc), which is a representative of a population or equivalently of an uncertainty distribution under a given scheme comprising set and σ-algebra.

Remark 2.10: The uncertainty distribution is unknown but exists objectively. A workable solution is to hypothesize a family of uncertainty distributions of a specified functional form with unknown parameter θ, where the family is denoted by $\{\Psi_{\xi}^{\theta}, \theta \in \Theta\}$.

Definition 2.11: (Liu (2007, 2009, 2010)) Let multivariate uncertainty variable $(\xi_1, \xi_2, \cdots, \xi_d)$ be defined on an uncertain measure space $(\Xi, \mathfrak{A}(\Xi), \mathcal{M})$, then the multivariate function $\Psi_{\xi_1, \xi_2, \cdots, \xi_d} : D \to [0,1]$ is called an multivariate uncertainty distribution if

$$\Psi_{\xi_1, \xi_2, \cdots, \xi_d}(x_1, x_2, \cdots, x_d) = \mathcal{M}\{\xi_1 \le x_1, \xi_2 \le x_2, \cdots, \xi_d \le x_d\} \tag{9}$$

To present a concrete form of a multivariate uncertainty distribution, Guo et al. (2010) propose a copula-linked uncertainty marginals approach.

Definition 2.12: Let $(\xi_1, \xi_2, \cdots, \xi_d)$ be a multivariate uncertainty variable with joint uncertainty distribution $\Psi_{\xi_1, \xi_2, \cdots, \xi_d}(x_1, x_2, \cdots, x_d)$, in which all the marginal uncertainty distributions $\Psi_{\xi_1}(\cdot), \Psi_{\xi_2}(\cdot), \cdots, \Psi_{\xi_d}(\cdot)$ exist and are regular (i.e., $\Psi_{\xi_i}^{-1}(\cdot)$ exists, $i = 1, 2, \cdots, d$). Then the uncertainty copula is defined by

$$C\left(\Psi_{\xi_1}(x_1), \Psi_{\xi_2}(x_2), \cdots, \Psi_{\xi_d}(x_d)\right) = \Psi_{\xi_1, \xi_2, \cdots, \xi_d}(x_1, x_2, \cdots, x_d) \tag{10}$$

We use a bivariate uncertainty distribution as an illustrative multivariate example.

Example 2.13: Let bivariate uncertainty variable (ξ_1, ξ_2) have marginal uncertainty distributions $\Psi_{\xi_1}(\cdot)$ and $\Psi_{\xi_2}(\cdot)$ respectively. The Farlie-Gumbel-Morgenstern (FGM) copula is defined by

$$C(u_1, u_2) = u_1 u_2 \left(1 + \nu (1 - u_1)(1 - u_2)\right), \ \nu \in [-1, 1] \tag{11}$$

Further, let the bivariate uncertainty variable (ξ_1, ξ_2) have marginal uncertainty distributions $\Psi_{\xi_1}(\cdot)$ and $\Psi_{\xi_2}(\cdot)$ respectively, where

$$\Psi_{\xi_i}(x_i) = \frac{1}{1 + \exp\left(-\frac{\pi}{\sqrt{3}\sigma_i}(x_i - \theta_i)\right)}, \ i = 1, 2 \tag{12}$$

Then the bivariate FGM-Normal joint uncertainty distribution is

$$\Psi_{\xi_1, \xi_2}(x_1, x_2) = \prod_{i=1}^{2} \frac{1}{1 + \exp\left(-\frac{\pi}{\sqrt{3}\sigma_i}(x_i - \theta_i)\right)} \left(1 + \nu \prod_{i=1}^{2} \frac{\exp\left(-\frac{\pi}{\sqrt{3}\sigma_i}(x_i - \theta_i)\right)}{1 + \exp\left(-\frac{\pi}{\sqrt{3}\sigma_i}(x_i - \theta_i)\right)}\right) \tag{13}$$

Finally, it is critical to define impreciseness with mathematical rigor. To achieve this goal, we review the discussions on randomness concept in statistics first for comparison purpose. Randomness in classical (i.e., probabilistic) statistics is referred to a term with an intrinsic property "governed by or involving equal chances for each of the actual or hypothetical members of a population; (also) produced or obtained by such a process, and therefore unpredictable in detail".

Randomness is "closely connected, therefore, with the concepts of chance, probability, and information entropy, randomness implies a lack of predictability. More formally, in statistics, a random process is a repeating process whose outcomes follow no describable deterministic pattern, but follow a probability distribution, such that the relative probability of the occurrence of each outcome can be approximated or calculated", see wikipedia (2010). In other words, randomness is an intrinsic property of a variable or an observation being characterized by a probability measure. Just as Kolmogrov (1950) emphasized probability measure specification is the prerequisite to randomness.

Remark 2.14: Parallel to revelation of the connotation of randomness. Impreciseness in uncertainty statistics is referred to a term with an intrinsic property governed by an uncertain measure or an uncertainty distribution for each of the actual or hypothetical members of an uncertainty population; (also) produced or obtained by such a process, and therefore unpredictable in detail. An uncertainty process is a repeating process whose outcomes follow no describable deterministic pattern, but follow an uncertainty distribution, such that the uncertain measure of the occurrence of each outcome can be only approximated or calculated.

Definition 2.15: Impreciseness is an intrinsic property of a variable or an expert's knowledge being specified by an uncertainty measure.

Remark 2.16: Impreciseness exists in engineering, business and research practices. Just as Utikin & Gurov (2000) as well as Walley (1991) argued strongly that “it very often happens that probabilities cannot be determined exactly, either due to measurement imperfections, or due to more fundamental reasons, such as insufficient available information, ... , or "is of a linguistic nature, i.e. the information is conveyed by statements in natural language”, …, a part of “the reliability assessments may be supplied by experts” or reliability “assessments may be made by the user of the system during the experimental service”. Thus it is an unarguable fact that impreciseness exists intrinsically in expert’s knowledge on the real world.

Definition 2.17: Let ξ be a uncertainty quantity of impreciseness on an uncertainty measure space $\left(\Xi, \mathcal{A}(\Xi), \lambda\right)$. The uncertainty distribution of ξ is $\Psi_{\xi}(x) = \lambda\left\{\tau \in \Xi \mid \xi(\tau) \le x\right\}$.

Remark 2.18: An imprecise variable ξ is an uncertainty variable and thus is a measurable mapping, i.e., $\xi : D \to \mathbb{R}$, $D \subseteq \mathbb{R}$. An observation of an imprecise variable is a real number, (or more broadly, a symbol, or an interval, or a real-valued vector, a statement, etc), which is a representative of the population or equivalently of an uncertainty distribution $\Psi_{\xi}(\cdot)$ under a given scheme comprising set and σ-algebra. The single value of a variable with impreciseness should not be understood as an isolated real number rather an interval or a set.

3 HYBRID VARIABLE THEORY

Since Zadeh (1965, 1978) proposed fuzzy set theory, fuzzy random fuzzy set, a special case of hybrid variable, soon proposed by Kaufmann (1975). Liu (2007) defined that a random fuzzy variable, another special case of hybrid variable, is a mapping from the credibility space $\left(Q, 2^{Q}, Cr\right)$ to a set of random variables. Let us start with a general hybrid variable definition.

Definition 3.1: (Liu (2007)) A hybrid variable is a real-valued measurable mapping, i.e., $\eta : \left(\Xi, \mathcal{A}\right) \to \left(\mathrm{R}, \mathcal{B}\right)$.

Remark 3.2: It is obvious that the order of the formation of a hybrid variable does matter. For example, Random fuzzy variable (Liu (2007)) and fuzzy random variable (Kaufmann (1975)) are two types of hybrid variable, even with the same component uncertain variables. Therefore, it is necessary to define them separately when specifying the hybrid variable with different uncertain variables.

Definition 3.3: A random-uncertain hybrid variable is a measurable mapping η from product space $(X, A(X), D)r\ (W, F(W), \Pr)$ into $(R, B(R), n)$, which is called as hybrid variable of Type I; An uncertain-random hybrid variable is a measurable mapping η from product space $(W, F(W), \Pr)r\ (X, A(X), D)$ into $(R, B(R), n)$, which is called hybrid variable of Type II.

In the remaining of the paper, we only deal with hybrid variable of Type I, i.e., random-uncertain hybrid variable. Therefore, for convenience we simply use the term hybrid variable. For reliability engineers and managers armed with introductory probability and statistics, this definition will be difficult to understand. For a more intuitive understanding, we would like to present a definition similar to that of stochastic process in probability theory and expect readers who are familiar with the basic concept of stochastic processes can understand our comparative definition.

Definition 3.4: A hybrid variable (of Type I), denoted by $\xi = \left\{ X_{\beta(\tau)}, \tau \in \Xi \right\}$, is a collection of random variables X_β defined on the common probability space $(\Omega, F, \Pr)$ and indexed by an uncertain variable $\beta(\tau)$ defined on the uncertainty space $(X, A(X), D)$.

Similar to the interpretation of a stochastic process $X = \left\{ X_t, t \in \mathbf{R}^+ \right\}$, a hybrid variable is also a bivariate mapping from $(\Omega \times \Xi, F \times A)$ to the space (R, B). As to the index set, in stochastic process theory, index set used is referred to as time typically, which is a positive (scalar variable), while in the random fuzzy variable theory, the "index" is an uncertain variable β. Using uncertain parameter as index is not starting in hybrid variable definition. In stochastic process theory we already know that the stochastic process $X = \left\{ X_{\tau(\omega)}, \omega \in \Omega \right\}$ uses stopping time $t(w)$, $w\,O\,W$, which is an random variable as its index.

4 AVERAGE MEASURE FOR A HYBRID VARIABLE

Hybrid variable can be quantified in terms of chance measure concept, see Liu (2007).

Definition 4.1: Let ξ be a random-uncertain hybrid variable and B a Borel set of real numbers. Then the chance measure of random fuzzy event $\{\xi \in B\}$ is a function mapping from $(0,1]$ to $[0,1]$,

$$\mathrm{Ch}\,\{x\,O\,B\}(a) = \sup_{D\{A\}i\ a}\ \inf_{q O A} \Pr\{q : x(q)\,O\,B\} \tag{14}$$

However, we notice the potential mathematical complexity associated with the chance measure formulation, see Liu (2008). Therefore, it is necessary to explore a convenient way to deal with the chance measure specification. Recall that in probability theory, the distribution of a random variable ξ on probability space $(W, A, \Pr)$, $F_\xi(\cdot)$ links to the probability measure of event $"\ \{w : x(w) J\ x\} O A$

$$F_x(x) = \Pr\{w : x(w) J\ x\}. \tag{15}$$

In random-uncertain hybrid variable theory, we may say that that average chance measure plays an equivalent role similar to probability measure, denoted as $\Pr$, in probability theory.

Definition 4.2: Let x be a random-uncertain hybrid variable, then the *average* chance measure, denoted as $\mathrm{ch}\ \{\}$, of a random-uncertain event $\{t\ O X : x(t) J\ x\}$, is

$$\mathrm{ch}\{\xi \le x\} = \int_0^1 \Delta\{\tau \in \Xi \mid \Pr\{\xi(\tau) \le x\} \ge \alpha\} \mathrm{d}\alpha \tag{16}$$

Then function $\Psi(\cdot)$ is called as average chance distribution if and only if

$$\Psi(x) = \mathrm{ch}\{\xi \le x\} \tag{17}$$

Now, we are required to establish a theoretical framework in terms of average chance measure concepts. Once the average chance measure for the basic event form $\{\xi \le x\}$ is given, then the average chance measure for any event A should be established in terms of the basic event $\{\xi \le x\}$. In this way, we may define average chance measure for an arbitrary event A. The triple space $(\Omega \times \Xi, \mathsf{F} \times \mathsf{A}, \mathrm{ch})$ is called the average chance space.

Proposition 4.3: Let $\mathrm{ch}(\cdot)$ be an average chance measure on a product measure space $(\Omega \times \Xi, \mathsf{F}(\Omega) \times \mathsf{A}(\Xi))$. Then

(i) $\mathrm{ch}\{\emptyset\} = 0$ and $\mathrm{ch}\{\Omega \times \Xi\} = 1$;

(ii) (Normality) $\forall A \in \mathsf{F} \times \mathsf{A}$, $0 \le \mathrm{ch}\{A\} \le 1$;

(iii) (Self-Duality) For $\forall A \in \mathsf{F} \times \mathsf{A}$, then $\mathrm{ch}\{A^c\} = 1 - \mathrm{ch}\{A\}$

(iv) (Weak monotone increasing) For $\forall A \subset B$, $A, B \in \mathsf{F} \times \mathsf{A}$, $\mathrm{ch}\{A\} \le \mathrm{ch}\{B\}$;

(v) (Semi-Continuity) For $\forall A_n, A \in \mathsf{F} \times \mathsf{A}$, $n = 1, 2, \cdots$, if $A_n \to A$, then

$$\lim_{A_n \to A} \mathrm{ch}\{A_n\} = \mathrm{ch}\{A\} \tag{18}$$

if and only if one of the following conditions holds:

(a) $\Delta\{A_n\} \le 0.5 \,\&\, A_n \uparrow A$,

(b) $\lim_{n \to \infty} \Delta\{A_n\} < 0.5 \,\&\, A_n \uparrow A$,

(c) $\Delta\{A_n\} \ge 0.5 \,\&\, A_n \downarrow A$, and

(d) $\lim_{n \to \infty} \Delta\{A_n\} > 0.5 \,\&\, A_n \downarrow A$.

(vi) (Sub-Additivity) For $\forall A \subset B$, $A, B \in \mathsf{F} \times \mathsf{A}$,

$$\mathrm{ch}\{A \cup B\} \le \mathrm{ch}\{A\} + \mathrm{ch}\{B\} \tag{19}$$

Proposition 4.4: Let $\Psi_\xi(\cdot)$ be average chance distribution of (random-uncertain) hybrid variable ξ on the chance measure space $(\Omega \times \Xi, \mathsf{F} \times \mathsf{A}, \mathrm{ch})$. Then

(i) $\Psi_\xi(-\infty) = 0$ and $\Psi_\xi(+\infty) = 1$;

(ii) For $\forall x \in \Re = (-\infty, +\infty)$, $0 \le \Phi_\xi(x) \le 1$;

(iii)Nonnegative real-valued function $\psi_\xi(\cdot)$ is called average chance density for a (random-uncertain) hybrid variable ξ if for $\psi_\xi(x) \ge 0$, $x \in \Re$ and

$$\Psi_\xi(x) = \int_{-\infty}^{x} \psi_\xi(u) \mathrm{d}u \tag{20}$$

5 CONSTRUCTION OF HYBRID VARIABLE

Liu (2007) mentioned an exponentially distributed random fuzzy variable ξ has a density function

$$\phi(x)=\begin{cases}\dfrac{1}{\beta}\exp\left(-\dfrac{x}{\beta}\right) & \text{if } x\geq 0\\ 0 & \text{otherwise}\end{cases} \tag{21}$$

if the value of β is assumed to be a fuzzy variable, then ξ is a random fuzzy variable. Similarly, let parameter β be an uncertain variable following a distribution function $\Lambda_\beta(\cdot)$, and the probability density is defined by Equation (11), then the random-uncertain hybrid variable ξ is said to be exponentially distributed. This example hints a constructive definition for specifying hybrid variable, i.e., random-uncertain variable or equivalently, the average chance distribution.

Definition 5.1: Let $\{F(x;b(t)), t \text{ OX}\}$ be a family of probability distributions on the probability space $(\text{W}, \text{A}, \Pr)$ with a common uncertain parameter β on the uncertain measure space $(\text{X}, \text{A}(\text{X}), \text{D})$, then the average distribution derived from $(F(x, \text{b}), \text{D})$ defines a (random-uncertain) hybrid variable x.

Theorem 5.2: Let ξ be a random-uncertain hybrid variable. If the expectation $E_P\left[\xi(\tau_0)\right]$ exists for any given $\tau_0\in\Xi$, then $E_P\left[\xi(\cdot)\right]$ is an uncertain variable.

6 RANDOM LIFETIME WITH IMPRECISE PARAMETER

Analyzing hybrid lifetimes, or survival times, or failure times, is the focus of lifetime modeling and analysis under randomness and general uncertainty co-existence environments. Different from the statistical lifetime modeling and analysis, where the random lifetimes are concerned, also different from the uncertainty lifetime modeling and analysis, where the uncertainty lifetimes are concerned, hybrid lifetime modeling analysis provides a general guideline with a rigorous theoretical foundation.

A (random-uncertain) hybrid lifetime, denoted by x, which is a special case of hybrid (of Type I), takes only a positive real values. In other words, hybrid lifetime is a bivariate mapping from $(\Omega\times\Xi, \text{F}\times\Xi)$ to the space $\left(\text{R}^+, \text{B}\left(\text{R}^+\right)\right)$.

6.1 Basic construction of continuous hybrid lifetimes

It is well-known fact the probability distribution contains the full information on system lifetime and there are many related concepts, particularly, hazard function reveals an aspect of lifetime distribution, which links to the physical structure of a system.

Theorem 6.1: Let x be a continuous hybrid lifetime having probability distribution function $F(t;\text{b}(\text{t}))$, where the imprecise parameter b is defined on the uncertain measure space $(\text{X}, \text{A}(\text{X}), \text{D})$. Then function $\text{P}(t;\text{b}) = \text{L}(F(t;\text{b}))$ can uniquely define the hybrid lifetime x if the operator or function L is invertible.

Table 1 lists four commonly used operators or functions.

Table 1. Examples of operators or functions

Name	Form of $\Pi(t;\beta)$	$\Lambda(\cdot)$
Survival function	$\bar{F}(t;\beta)=1-F(t;\beta)$	$F(t;\beta)=1-\bar{F}(t;\beta)$
Density function	$f(t;\beta)=\mathrm{d}F(t;\beta)/\mathrm{d}t$	$F(t;\beta)=\int_0^t f(u;\beta)\mathrm{d}u$
Hazard function	$h(t;\beta)=f(t)/(1-F(t;\beta))$	$F(t;\beta)=1-\exp\left(-\int_0^t h(u;\beta)\mathrm{d}u\right)$
Moment generating function	$m(\theta;\beta)=\int_0^{+\infty} e^{\theta t}\mathrm{d}F(t;\beta)$	$F(t;\beta)=\int_0^t\left(\frac{1}{2\pi i}\int_{s-i\infty}^{s+i\infty} m(s;\beta)e^{su}\mathrm{d}s\right)\mathrm{d}u$

6.2 Continuous hybrid lifetime models

In statistical lifetime modeling and analysis, the elementary lifetime models are exponential, Weibull, Log-normal, gamma, Cox-Lewis, bathtub, and etc. These are essential for the construction of hybrid lifetimes. Table 2 lists these models.

In Table 2, $I(\beta,\lambda t)$ denotes the incomplete gamma function of the first-type and $\Phi(\cdot)$ represents the cumulative distribution of a standard normal variable.

Table 2. Commonly used distributional lifetime models

Name	Probability density & hazard function	
Exponential	density	$\beta\exp(-\beta t)$
	hazard	β
Weibull	density	$(\beta/\eta)(t/\eta)^{\beta-1}\exp\left(-(t/\eta)^{\beta}\right)$
	hazard	$(\beta/\eta)(t/\eta)^{\beta-1}$
Extreme - value	density	$(1/u)\exp((t-b)/u)\exp(-\exp((t-b)/u))$
	hazard	$(1/u)\exp((t-b)/u)$
Log-Normal	density	$\left(1/(\sqrt{2\pi}\sigma t)\right)\exp\left(-(\ln t-\mu)^2/2\sigma^2\right)$
	hazard	$\left(\left(1/(\sqrt{2\pi}\sigma t)\right)\exp\left(-(\ln t-\mu)^2/2\sigma^2\right)\right)/(1-\Phi((\ln t-\mu)/\sigma$
Gamma	density	$\left(\lambda(\lambda t)^{\beta-1}/\Gamma(\beta)\right)e^{-\lambda t}$
	hazard	$\left(\lambda(\lambda t)^{\beta-1}/\Gamma(\beta)\right)e^{-\lambda t}/(1-I(\beta,\lambda t))$
Bathtub	density	$(\beta/\eta)(t/\eta)^{\beta-1}\exp\left((t/\eta)^{\beta-1}\right)\exp\left(-\exp\left((t/\eta)^{\beta-1}\right)\right)$
	hazard	$(\beta/\eta)(t/\eta)^{\beta-1}\exp\left((t/\eta)^{\beta-1}\right)$

6.3 Proportional hazard models

Covariate models play very important roles in lifetime analysis. Cox (1972) initiated proportional hazards (abbreviated as PH) model as following:

$$h(t;b,g)=h_0(t;b)V(g^T y) \tag{22}$$

where $h_0(t;b)$ is called the baseline hazard function having a fuzzy parameter b defined on the credibility measure space $(\mathrm{X},\mathrm{A},\mathrm{D})$, while $V:\mathbf{R}\to\mathbf{R}^{+}$ with

$$\gamma^T y = \gamma_0 + \gamma_1 y_1 + \cdots + \gamma_p y_p \tag{23}$$

where $y = (1, y_1, \cdots, y_p)^T$ is covariate vector and $\gamma = (\gamma_0, \gamma_1, \cdots, \gamma_p)^T$ is covariate effect parameter vector. A typically function of $V: \mathbf{R} \to \mathbf{R}^+$ used is the exponential function $V(x) = e^x$. It is easy to show that the accumulated hazard if covariate y is not time-dependent is

$$H(t; \beta, \gamma) = H_0(t; \beta) e^{\gamma^T y}. \tag{24}$$

And therefore the average chance distribution with covariate y is

$$\Phi(t, y) = \int_0^1 \mathcal{M}\left\{(t_1, t_2): H_0(t; \beta(t_1)) e^{\gamma^T y(t_2)} \geq -\ln(1-\alpha)\right\} d\alpha \tag{25}$$

where covariate y is assumed to be uncertain distributed but parameter γ is assumed to be determined. Other options are also possible to be formulated.

7 EXPONENTIAL RANDOM VARIABLE WITH IMPRECISE PARAMETER

The purpose to have this section is double-folded: (a) exponential hybrid lifetime is an important member for system lifetime analysis; (b) the arguments for deriving the average chance distribution are demonstration in line with hybrid variable reliability analysis. Bearing this agenda in mind, the following step-by-step developments will be very beneficial.

Let us use exponentially distributed hybrid lifetime which has probability density

$$f(t; \beta) = \begin{cases} 0 & t \leq 0 \\ \beta e^{-\beta t} & t > 0 \end{cases} \tag{26}$$

where the imprecise parameter β has a five-piece-wise linear uncertainty distribution function (Liu (2007))

$$\Lambda(x) = \mathcal{M}\{\xi \leq x\} = \begin{cases} 0 \quad \text{if } x \leq a \\ \dfrac{x-a}{2(b-a)} \quad \text{if } a < x \leq b \\ 0.5 \quad \text{if } b < x \leq c \\ \dfrac{x+d-2c}{2(d-c)} \quad \text{if c} < x \leq d \\ 1 \quad \text{if } x > d \end{cases}. \tag{27}$$

Note that

$$\Pr\{\xi(\theta) \leq t\} = 1 - e^{\beta(\theta) t} \tag{28}$$

Therefore the event $\{\theta : \Pr\{\xi(\theta) \leq t\} \geq \alpha\}$ is an uncertain event and is equivalent to the uncertain event $\{\theta : \beta(\theta) \geq -\ln(1-\alpha)/t\}$. As a critical step toward the derivation of the average chance distribution, it is necessary to calculate the uncertain measure for the uncertain event $\{\theta : \beta(\theta) \geq -\ln(1-\alpha)/t\}$, i.e., obtain the expression for

$$\mathcal{M}\{\theta : \beta(\theta) \geq -\ln(1-\alpha)/t\} \tag{29}$$

Accordingly the range for integration with respect to α can be determined as shown in Table 3. Recall that the expression of $x = -\ln(1-\alpha)/t$ appears in Equation (29), which facilitates the link between intermediate variable α and average chance measure.

Table 3. Range analysis for α

Range for x	α and credibility measure expression	
$-\infty < x \le a$	Range for α	$0 \le \alpha \le 1-e^{-at}$
	$\Delta\{\theta:\beta(\theta) \ge -\ln(1-\alpha)/t\}$	1
$a < x \le b$	Range for α	$1-e^{-at} < \alpha \le 1-e^{-bt}$
	$\Delta\{\theta:\beta(\theta) \ge -\ln(1-\alpha)/t\}$	$1-(x-a)/(2(b-a))$
$b < x \le c$	Range for α	$1-e^{-bt} < \alpha \le 1-e^{-ct}$
	$\Delta\{\theta:\beta(\theta) \ge -\ln(1-\alpha)/t\}$	0.5
$c < x \le d$	Range for α	$1-e^{-ct} < \alpha \le 1-e^{-dt}$
	$\Delta\{\theta:\beta(\theta) \ge -\ln(1-\alpha)/t\}$	$(d-x)/(2(d-c))$
$d < x < +\infty$	Range for α	$1-e^{-dt} < \alpha \le 1$
	$\Delta\{\theta:\beta(\theta) \ge -\ln(1-\alpha)/t\}$	0

The average chance distribution for the exponentially distributed hybrid lifetime is then derived by splitting the integration into five terms according to the range of α and the corresponding mathematical expression for the uncertain measure $\Delta\{\theta:\beta(\theta) \ge -\ln(1-\alpha)/t\}$, which is detailed in Table 3. Then the exponential random fuzzy lifetime has an average chance distribution function:

$$\begin{aligned}\Psi_\xi(t) &= \int_0^1 \Delta\{\theta:\beta(\theta) \ge -\ln(1-\alpha)/t\}d\alpha \\ &= 1+\frac{e^{-bt}-e^{-at}}{2(b-a)t}+\frac{e^{-dt}-e^{-ct}}{2(d-c)t}\end{aligned} \tag{30}$$

and the average chance density is

$$\begin{aligned}\psi_\xi(t) &= \frac{e^{-at}-e^{-bt}}{2(b-a)t^2}+\frac{be^{-bt}-ae^{-at}}{2(b-a)t} \\ &+\frac{e^{-ct}-e^{-dt}}{2(d-c)t^2}+\frac{ce^{-ct}-de^{-dt}}{2(d-c)t}\end{aligned} \tag{31}$$

Similar to the probabilistic reliability theory, we define a reliability function or survival function for a random fuzzy lifetime and accordingly name it as the average chance reliability function, which is defined accordingly as

$$R_\xi(t) = 1-\Psi_\xi(t) \tag{32}$$

Then, for exponential random fuzzy lifetime, its average chance reliability function is

$$R_\xi(t) = \frac{e^{-at}-e^{-bt}}{2(b-a)t}+\frac{e^{-ct}-e^{-dt}}{2(d-c)t} \tag{33}$$

In standard statistical lifetime modelling and analysis reliability function reveals the system functioning behaviour. The average chance reliability function should play similar roles in hybrid lifetime modelling and analysis. In order to gain an intuitive perceptions on the average chance reliability function, let us assume that the trapezoidal identification function defined by (0.1, 0.15, 0.25, 0.30), i.e., the parameters for specifying the identification function are $a = 0.1$, $b = 0.15$,

$c = 0.25, d = 0.30$. For comparison purpose, we define an exponentially distributed random lifetime with fixed valued parameter, 0.20, which is obtained by

$$m_b = E(b) = 0.20 \tag{34}$$

Then the reliability function for the exponentially distributed random lifetime with parameter $m_b = 0.20$ is

$$R(t;0.20) = \exp(-0.2t) \tag{35}$$

The corresponding average chance reliability function, $R_\xi(t;\beta)$:

$$R_x(t;\mathrm{b}) = \frac{10\left(e^{-at} - e^{-bt}\right)}{t} + \frac{10\left(e^{-ct} - e^{-dt}\right)}{t} \tag{36}$$

Figure 1 gives a graphic comparison between $R_x(t;b)$ and $R(t;0.20)$.

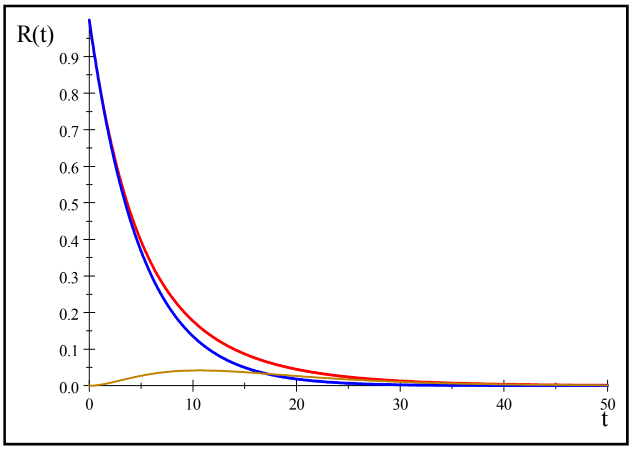


Figure 1. Exponential hybrid lifetime average chance reliability $R_x(t;b)$ (Red), corresponding exponential lifetime reliability $R(t;0.20)$ (Blue), and the difference function $d\left(R_x(t;b), R(t;0.20)\right)$ (Sienna)

Intuitively, we can see that given two systems: the first one is an exponentially distributed hybrid system with trapezoidal uncertain distributed parameter $b = (0.10, 0.15, 0.25, 0.30)$ and the second one is an exponentially distributed random system with parameter $m_b = 0.20$, the first one enjoys a higher reliability than that of the second one. Definitely, a rigorous mathematical proof should be pursued before stating this impression as a general statement.

However, the purpose for us to develop hybrid lifetime analysis theory is a serious effort to facilitate a foundation for analyzing reliability data collected from system performance.

8 CONCLUDING REMARKS

In this paper, we develop a framework for modeling hybrid lifetimes (of Type I) and the average chance distribution as well as the average chance reliability. The models are constructive. We use exponentially distributed hybrid lifetime with an imprecise parameter having a five-piece-like uncertainty function as an example to illustrate the model developments on hybrid lifetimes. It should mention that for two-parameter with impreciseness, the bivariate copula-linked uncertainty marginals approach can facilitate a bivariate uncertainty distribution for imprecise parameters and further the derivation of the average chance distribution. Guo et al. (2007) demonstrated hybrid variable theory in repairable modeling, although in random fuzzy context. However, many research work need to be done ahead, for example, the parameter estimation, the asymptotic distribution for the estimated parameters, the small sample theory, etc.

9 ACKNOWLEDGEMENTS

This research is supported by VC's Incentive for NRF Rated Researchers, University of Cape Town, and South African National Research Foundation Grant IFR2010042200062, and IFR2009090800013 as well.

The research is also partially supported by Tsinghua University during Professor Guo and Mr YH Cui’s visit to the Uncertainty Theory Laboratory, Department of Mathematical Science, Tsinghua University from July 18 to October 14, 2010.

The authors would like to thank Professor B.D. Liu and his PhD students: X.W. Chen, W. Dai, K. Yao, and Z. X. Peng for their invaluable debates and inputs.

10 REFERENCES

13. Bernardo, J. M. & Smith A. F. M. (1994) Bayesian Theory. John Wiley & Sons Ltd, England.
14. Carvalho, H. & Machado, V. C. (2006) Fuzzy set theory to establish resilient production systems. Proc. of IIE Annual Conference and Exhibition.
15. Cox, D. R. (1972) Regression models and life tables. Journal of Royal Statistical Society, B26, 187-220.
16. Guo, R., Zhao, R. Q., Guo, D. & Dunne, T. (2007) Random Fuzzy Variable Modelling on Repairable System. Journal of Uncertain Systems 1, 3, 222-234, August, 2007.
17. Guo, R., Cui, Y. H. & Guo, D. (2010) Uncertainty Linear Regression Models. (Draft paper submitted to Journal of Uncertain Systems).
18. Kaufmann, A. (1975). Introduction to the theory of fuzzy subsets. Vol.1, Academic Press, New York.
19. Kolmogorov, A. N. (1950) Foundations of the Theory of Probability, translated by Nathan Morrison, Publisher: Chelsea, New York.
20. Lee, P. M. (1989) Bayesian Statistics: An Introduction. Oxford University Press, New York, USA.
21. Liu, B. D. (2007) Uncertainty Theory: An Introduction to its Axiomatic Foundations, 2nd Edition, Springer-Verlag Heidelberg, Berlin.
22. Liu, B. D. (2009) Some research problems in uncertainty theory, Journal of Uncertain Systems, Vol.3, No.1, 3-10.
23. Liu, B. D. (2010) Uncertainty Theory: A Branch of Mathematics of Modelling Human Uncertainty, Springer-Verlag, Berlin.
24. Liu, Y. H. (2008) How to generate uncertain measures, Proceedings of Tenth National Youth Conference on Information and Management Sciences, August 3-7, 2008, Luoyang, pp.23-26.
25. Peng, Z. X. & Iwamura, K. (2010) A sufficient and necessary condition of uncertainty distribution. Journal of Interdisciplinary Mathematics, to be published.
26. Primas, H. (1999) Basic Elements and Problems of Probability Theory, Journal of Scientific Exploration, Vol. 13, No. 4, pp. 579-613.
27. Utkin, L. V. & Gurov, S. V. (2000) New Reliability Models Based on Imprecise Probabilities, Advanced Signal Processing Technology by Soft Computing. Fuzzy Logic Systems Institute

(FLSI) Soft Computing Series - Vol. 1, pp. 110-139, edited by Charles Hsu, eISBN: 9789812792105. Publisher, World Scientific. November 2000.
28. Walley, P. (1991) Statistical Reasoning with Imprecise Probabilities. London: Chapman and Hall. ISBN: 0412286602.
29. wikipedia, (2010) http://en.wikipedia.org/wiki/Imprecise_probability; http://en.wikipedia.org/wiki/Randomness; http://en.wikipedia.org/wiki/Statistics; http://en.wikipedia.org/wiki/Uncertainty_Theory
30. Zadeh, L. A. (1965) Fuzzy sets. Information and Control 8, 338-353.
31. Zadeh, L. A. (1978) Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1, 3-28.

SAFETY OPTIMIZATION OF A FERRY TECHNICAL SYSTEM IN VARIABLE OPERATION CONDITIONS

Krzysztof Kolowrocki, Joanna Soszynska-Budny
•
Gdynia Maritime University, Gdynia, Poland
e-mail: katmatkk@am.gdynia.pl, joannas@am.gdynia.pl

ABSTRACT

The general model of the safety of complex technical systems in variable operation conditions linking a semi-Markov modeling of the system operation process with a multi-state approach to system safety analysis and linear programming are applied in maritime transport to safety and risk optimization of a ferry technical system.

1 INTRODUCTION

Most real technical systems are very complex and it is difficult to analyze and optimize their safety. Large numbers of components and subsystems and their operating complexity cause that the evaluation and optimization of their safety is complicated. The complexity of the systems' operation processes and their influence on changing in time the systems' structures and their components' safety characteristics is often very difficult to fix and to analyze. A convenient tool for solving this problem is a semi-markov (Grabski 2002) modeling of the system operation processes linked with a multi-state approach to the system safety analysis (Kolowrocki, Soszynska 2008, Kolowrocki, Soszynska 2009) and a linear programming for the system safety optimization (Kolowrocki, Soszynska 2010). This approach to system safety investigation is based on the multi-state system reliability analysis considered for instance in (Aven 1985, Kolowrocki 2004) and on semi-markov operation processes modeling discussed for instance in (Soszynska 2006, Soszynska 2007). An application of the proposed approach to safety analysis and optimization of maritime ferry technical system is presented in this paper.

2 THE FERRY TECHNICAL SYSTEM SAFETY AND RISK

The considered maritime ferry is a passenger Ro-Ro ship operating in Baltic Sea between Gdynia and Karlskrona ports on regular everyday line. We assume that the ferry is composed of a number of main subsystems having an essential influence on its safety. These subsystems are illustrated in Figure 1.

On the scheme of the ferry presented in Figure 1, there are distinguished its following subsystems:

S_1 - a navigational subsystem,
S_2 - a propulsion and controlling subsystem,
S_3 - a loading and unloading subsystem,
S_4 - a hull subsystem,
S_5 - an anchoring and mooring subsystem,
S_6 - a protection and rescue subsystem,
S_7 - a social subsystem.

In our further analysis of the ferry safety we omit the protection and rescue subsystem S_6 and the social subsystem S_7 and we consider its strictly technical subsystems S_1, S_2, S_3, S_4 and S_5 only, further called the ferry technical system (Kolowrocki, Soszynska 2009).

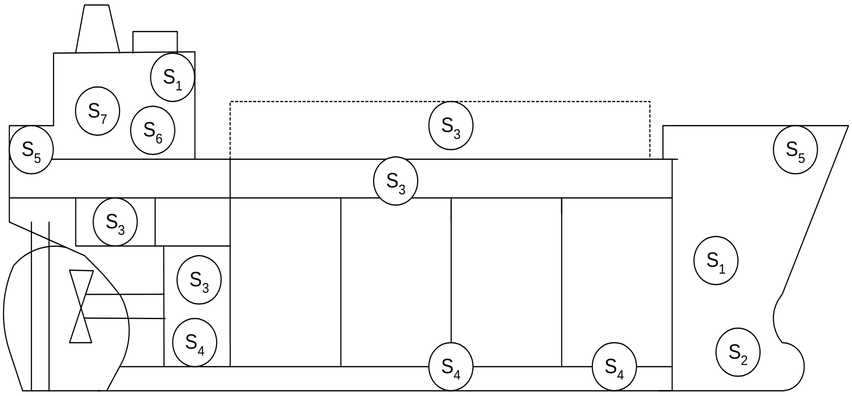


Figure 1. Subsystems having an essential influence on the ferry technical system safety

We assume that the ferry technical system safety structure and the subsystems and components safety depend on its changing in time operation states (Kolowrocki, Soszynska 2010).

Taking into account the experts' opinion on the operation process of the considered ferry, we distinguish the following as its eighteen operation states:

- an operation state z_1 – loading at Gdynia Port,
- an operation state z_2 – unmooring operations at Gdynia Port,
- an operation state z_3 – leaving Gdynia Port and navigation to "GD" buoy,
- an operation state z_4 – navigation at restricted waters from "GD" buoy to the end of Traffic Separation Scheme,
- an operation state z_5 – navigation at open waters from the end of Traffic Separation Scheme to "Angoring" buoy,
- an operation state z_6 – navigation at restricted waters from "Angoring" buoy to "Verko" Berth at Karlskrona,
- an operation state z_7 – mooring operations at Karlskrona Port,
- an operation state z_8 – unloading at Karlskrona Port,
- an operation state z_9 – loading at Karlskrona Port,
- an operation state z_{10} – unmooring operations at Karlskrona Port,
- an operation state z_{11} – ferry turning at Karlskrona Port,
- an operation state z_{12} – leaving Karlskrona Port and navigation at restricted waters to "Angoring" buoy,
- an operation state z_{13} – navigation at open waters from "Angoring" buoy to the entering Traffic Separation Scheme,
- an operation state z_{14} – navigation at restricted waters from the entering Traffic Separation Scheme to "GD" buoy,
- an operation state z_{15} – navigation from "GD" buoy to turning area,
- an operation state z_{16} – ferry turning at Gdynia Port,
- an operation state z_{17} – mooring operations at Gdynia Port,
- an operation state z_{18} – unloading at Gdynia Port.

Additionally, as in (Kolowrocki, Soszynska 2009, 2010), we assume that subsystems S_υ, $\upsilon = 1,2,...,5,$ of the ferry technical system are composed of five-state, components, and their safety states are 0,1,2,3 and 4. Consequently the components conditional multi-state safety function is the vector (Kolowrocki, Soszynska 2009)

$$[s_{ij}^{(\upsilon)}(t,\cdot)]^{(b)} = [1,\ [s_{ij}^{(\upsilon)}(t,1)]^{(b)}, [s_{ij}^{(\upsilon)}(t,2)]^{(b)}, [s_{ij}^{(\upsilon)}(t,3)]^{(b)}, [s_{ij}^{(\upsilon)}(t,4)]^{(b)}],\ b = 1,2,...,18,$$

with the exponential co-ordinates

$$[s_{ij}^{(\upsilon)}(t,1)]^{(b)} = \exp[-[\lambda_{ij}^{(\upsilon)}(1)]^{(b)}t],\ [s_{ij}^{(\upsilon)}(t,2)]^{(b)} = \exp[-[\lambda_{ij}^{(\upsilon)}(2)]^{(b)}t],$$

$$[s_{ij}^{(\upsilon)}(t,3)]^{(b)} = \exp[-[\lambda_{ij}^{(\upsilon)}(3)]^{(b)}t],\ [s_{ij}^{(\upsilon)}(t,4)]^{(b)} = \exp[-[\lambda_{ij}^{(\vartheta)}(4)]^{(b)}t],\ b = 1,2,...,18,$$

Further, assuming that the ferry technical system is in the safety state subset $\{u, u+1,...,4\}$ $u = 0,1,2,3,4,$ if all its subsystems are in this subset o safety states, we conclude that the ferry is five-state series system (Kolowrocki, Soszynska 2009) of subsystems S_1, S_2, S_3, S_4, S_5 and S_6.

The ferry operation process is very regular in the sense that the operation state changes are from the particular state z_b, $b = 1,2,...,17,$ to the neighboring state z_{b+1}, $b = 1,2,...,17,$ and from z_{18} to z_1 only. Therefore, the probabilities of transitions between the operation states are given by (Kolowrocki, Soszynska 2009)

$$[p_{bl}] = \begin{bmatrix} 0\,1\,0 \ldots 0\,0 \\ 0\,0\,1 \ldots 0\,0 \\ \ldots \\ 0\,0\,0 \ldots 0\,1 \\ 1\,0\,0 \ldots 0\,0 \end{bmatrix}.$$

On the basis of statistical data coming from experts the matrix of the density functions of the ferry technical system operation process conditional sojourn times θ_{bl} $b, l = 1,2,...,18,$ defined in (Kolowrocki, Soszynska 2009), can be evaluated.

Next, the mean values $M_{bl} = E[\theta_{bl}],$ $b, l = 1,2,...,18,$ $b \neq l,$ of the system operation process conditional sojourn times θ_{bl} in particular operation states can be determined and they are:

$$M_{12} = 54.67,\ M_{23} = 2.57,\ M_{34} = 37.33,\ M_{45} = 52.27,\ M_{56} = 526.43,\ M_{67} = 37.16,$$

$$M_{78} = 7.02,\ M_{89} = 23.26,\ M_{910} = 53.69,\ M_{1011} = 2.86,\ M_{1112} = 4.38,\ M_{1213} = 24.12,$$

$$M_{1314} = 508.60,\ M_{1415} = 50.14,\ M_{1516} = 34.43,\ M_{1617} = 4.59,\ M_{1718} = 7.92,\ M_{181} = 18.74.$$

Hence, according to (2) (Kolowrocki, Soszynska 2010), the mean values of the unconditional sojourn times in the operation states are:

$$M_1 = 54.67,\ M_2 = 2.57,\ M_3 = 37.33,\ M_4 = 52.27,\ M_5 = 526.43,\ M_6 = 37.16,\ M_7 = 7.02,$$

$$M_8 = 23.26,\ M_9 = 53.69,\ M_{10} = 2.86,\ M_{11} = 4.38,\ M_{12} = 24.12,\ M_{13} = 508.60,\ M_{14} = 50.14,$$

$$M_{15} = 34.43,\ M_{16} = 4.59,\ M_{17} = 7.92,\ M_{18} = 18.74.$$

Since from the system of equations given in (Kolowrocki, Soszynska 2009, 2010) taking here the form

$$\begin{cases} [\pi_b]_{1x18} = [\pi_b]_{1x18}[p_{bl}]_{18x18} \\ \sum\limits_{b=1}^{v} \pi_b = 1, \end{cases}$$

we get

$\pi_b \cong 0.056$ for $b = 1,2,...,18.$

Thus, according to the results contained in (Kolowrocki, Soszynska 2009, 2010), the long term proportion of transients p_b at the operational states z_b, can be approximated by

$$p_1 = 0.038,\ p_2 = 0.002,\ p_3 = 0.026,\ p_4 = 0.036,\ p_5 = 0.363,\ p_6 = 0.026,\ p_7 = 0.005,$$

$$p_8 = 0.016,\ p_9 = 0.037,\ p_{10} = 0.002,\ p_{11} = 0.003,\ p_{12} = 0.016,\ p_{13} = 0.351,\ p_{14} = 0.034,$$

$$p_{15} = 0.024,\ p_{16} = 0.003,\ p_{17} = 0.005,\ p_{18} = 0.013. \qquad (1)$$

Under the assumption that the changes of the ferry operation states have an influence on the subsystems S_υ, $\upsilon = 1,2,...,5,$ components safety and on the ferry technical system safety structures as well, on the basis of expert opinions and statistical data given in (Soszynska et al. 2009), the ferry technical system safety structures and their components safety functions and the ferry technical system conditional safety functions at different operation states can be determined (Kolowrocki, Soszynska 2010). Namely, in the case when the system operation time is large enough, the unconditional fife-state safety function of the ferry technical system is given by the vector

$$\boldsymbol{s}(t,\ \cdot) = [1,\ \boldsymbol{s}(t,1),\ \boldsymbol{s}(t,2),\ \boldsymbol{s}(t,3),\ \boldsymbol{s}(t,4)],\ t \geq 0, \qquad (2)$$

where, after considering the values of p_b, $b = 1,2,...,18,$ given by (1), its co-ordinates are

$$\boldsymbol{s}(t,u) = 0.038 \cdot [\boldsymbol{s}(t,u)]^{(1)} + 0.002 \cdot [\boldsymbol{s}(t,u)]^{(2)} + 0.026 \cdot [\boldsymbol{s}(t,u)]^{(3)} + 0.036 \cdot [\boldsymbol{s}(t,u)]^{(4)}$$

$$+ 0.363 \cdot [\boldsymbol{s}(t,u)]^{(5)} + 0.026 \cdot [\boldsymbol{s}(t,u)]^{(6)} + 0.005 \cdot [\boldsymbol{s}(t,u)]^{(7)} + 0.016 \cdot [\boldsymbol{s}(t,u)]^{(8)}$$

$$+ 0.037 \cdot [\boldsymbol{s}(t,u)]^{(9)} + 0.002 \cdot [\boldsymbol{s}(t,u)]^{(10)} + 0.003 \cdot [\boldsymbol{s}(t,u)]^{(11)} + 0.016 \cdot [\boldsymbol{s}(t,u)]^{(12)}$$

$$+ 0.351 \cdot [\boldsymbol{s}(t,u)]^{(13)} + 0.034 \cdot [\boldsymbol{s}(t,u)]^{(14)} + 0.024 \cdot [\boldsymbol{s}(t,u)]^{(15)} + 0.003 \cdot [\boldsymbol{s}(t,u)]^{(16)}$$

$$+0.005\cdot[\bar{s}(t,u)]^{(17)}+0.013\cdot[s(t,u)]^{(18)},$$

(3)

for $t \ge 0$, $u = 1,2,3,4$, where $[s(t,u)]^{(b)}$, $b = 1,2,...,18$, are the system conditional safety functions at particular operation states z_b, $b = 1,2,...,18$, determined in (Kolowrocki, Soszynska 2010).

The safety function of the ferry technical system is presented in Figure 2

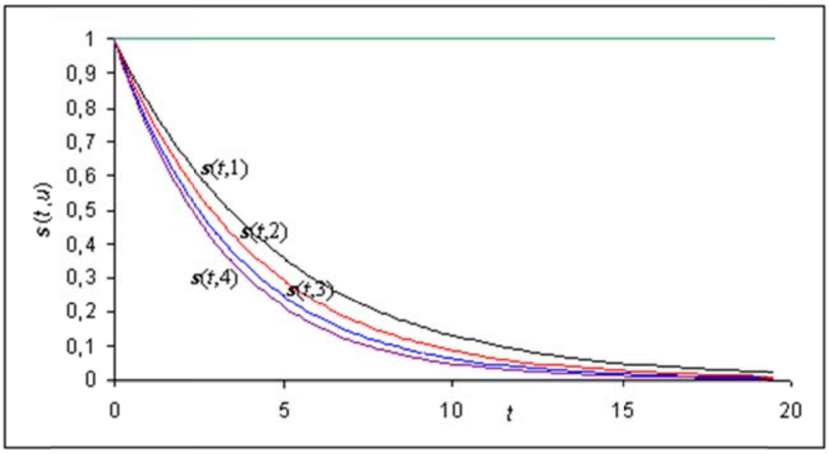


Figure 2. Graph of the safety function $[s(t,\cdot)]$ coordinates

From (3), the mean values of the ferry technical system unconditional lifetimes in the safety state subsets {1,2,3,4}, {2,3,4}, {3,4}, {4} respectively are:

$\mu(1) \cong 4.70$ years,

$$\begin{aligned}\mu(2) \cong\ & 0.038\cdot 6.45+0.002\cdot 2.43+0.026\cdot 3.9+0.036\cdot 3.80+0.363\cdot 3.80+0.026\cdot 3.24\\ &+0.005\cdot 2.43+0.016\cdot 7.69+0.037\cdot 7.69+0.002\cdot 2.43+0.003\cdot 3.37+0.016\cdot 3.80\\ &+0.351\cdot 3.80+0.034\cdot 3.80+0.024\cdot 3.90+0.003\cdot 3.37+0.005\cdot 2.43+0.013\cdot 6.45\\ &\cong 4.11 \text{ years,}\end{aligned} \tag{4}$$

$\mu(3) \cong 3.66$ years, $\mu(4) \cong 3.29$ years.

From the above (Kolowrocki, Soszynska 2009, 2010), the mean values of the system lifetimes in the particular safety states are:

$$\bar{\mu}(1)=\mu(1)-\mu(2)=0.59,\ \bar{\mu}(2)=\mu(2)-\mu(3)=0.77 \text{ years,}$$

$$\bar{\mu}(3)=\mu(3)-\mu(4)=0.45,\ \bar{\mu}(4)=\mu(4)=2.29 \text{ years.} \tag{5}$$

If we assume that the critical safety state is $r = 2$, then the system risk function (Kolowrocki, Soszynska 2009, 2010), is given by

$$\boldsymbol{r}(t) = 1-s(t,2) \tag{6}$$

where $s(t,2)$ is given by (3) for $u = 2$.

The moment when the system risk function exceeds a permitted level, for instance $\delta = 0.05$, is

$$\tau = \boldsymbol{r}^{-1}(\delta) \cong 0.21 \text{ year}. \tag{7}$$

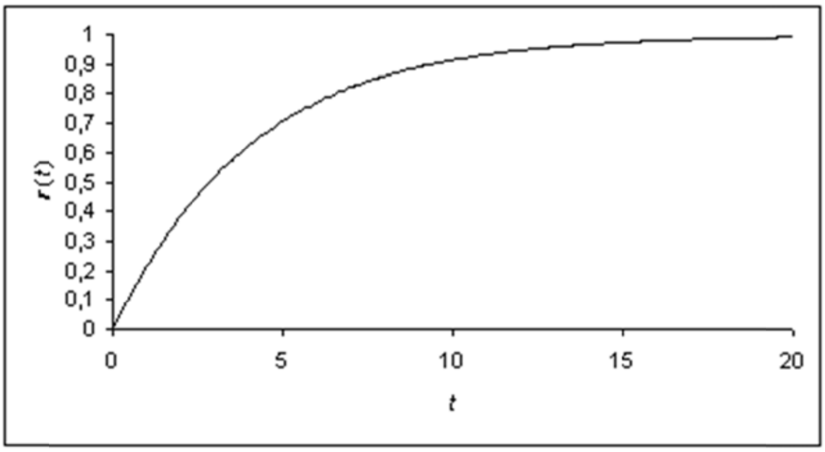


Figure 3. Graph of the ferry technical system risk function $r(t)$

3 OPTIMIZATION OF THE FERRY TECHNICAL SYSTEM OPERATION PROCESS

Considering the equation (3), it is natural to assume that the system operation process has a significant influence on the system safety. This influence is also clearly expressed in the equation (4), for the mean values of the system unconditional lifetimes in the safety state subsets.

The objective function defined in (Kolowrocki, Soszynska 2010), in this case as the system critical state is $r = 2$, takes the form

$$\mu(2) = p_1 \cdot 6.45 + p_2 \cdot 2.43 + p_3 \cdot 3.90 + p_4 \cdot 3.80 + p_5 \cdot 3.80 + p_6 \cdot 3.24$$

$$+ p_7 \cdot 2.43 + p_8 \cdot 7.69 + p_9 \cdot 7.69 + p_{10} \cdot 2.43 + p_{11} \cdot 3.37 + p_{12} \cdot 3.80$$

$$+ p_{13} \cdot 3.80 + p_{14} \cdot 3.80 + p_{15} \cdot 3.90 + p_{16} \cdot 3.37 + p_{17} \cdot 2.43 + p_{18} \cdot 6.45. \tag{8}$$

Since the lower $\breve{p}_b$ and upper $\hat{p}_b$ bounds of the unknown transient probabilities p_b, $b = 1,2,...,18$, coming from experts are respectively:

$$\breve{p}_1 = 0.0006,\ \breve{p}_2 = 0.001,\ \breve{p}_3 = 0.018,\ \breve{p}_4 = 0.027,\ \breve{p}_5 = 0.286,\ \breve{p}_6 = 0.018,$$

$$\breve{p}_7 = 0.002,\ \breve{p}_8 = 0.001,\ \breve{p}_9 = 0.001,\ \breve{p}_{10} = 0.001,\ \breve{p}_{11} = 0.002,\ \breve{p}_{12} = 0.013,$$

$$\breve{p}_{13} = 0.286,\ \breve{p}_{14} = 0.025,\ \breve{p}_{15} = 0.018,\ \breve{p}_{16} = 0.002,\ \breve{p}_{17} = 0.002,\ \breve{p}_{18} = 0.001,$$

$$\hat{p}_1 = 0.056,\ \hat{p}_2 = 0.002,\ \hat{p}_3 = 0.027,\ \hat{p}_4 = 0.056,\ \hat{p}_5 = 0.780,\ \hat{p}_6 = 0.024,$$

$$\hat{p}_7 = 0.018,\ \hat{p}_8 = 0.018,\ \hat{p}_9 = 0.056,\ \hat{p}_{10} = 0.003,\ \hat{p}_{11} = 0.004,\ \hat{p}_{12} = 0.024,$$

$$\hat{p}_{13} = 0.780,\ \hat{p}_{14} = 0.043,\ \hat{p}_{15} = 0.024,\ \hat{p}_{16} = 0.004,\ \hat{p}_{17} = 0.007,\ \hat{p}_{18} = 0.018,$$

then we assume, for the objective function defined by (8), the following bounds constraints

$$0.0006 \le p_1 \le 0.056,\ 0.001 \le p_2 \le 0.002,\ 0.018 \le p_3 \le 0.027,$$

$$0.027 \le p_4 \le 0.056,\ 0.286 \le p_5 \le 0.780,\ 0.018 \le p_6 \le 0.024,$$

$$0.002 \le p_7 \le 0.018,\ 0.001 \le p_8 \le 0.018,\ 0.001 \le p_9 \le 0.056,$$

$$0.001 \le p_{10} \le 0.003,\ 0.002 \le p_{11} \le 0.004,\ 0.013 \le p_{12} \le 0.024,$$

$$0.286 \le p_{13} \le 0.780,\ 0.025 \le p_{14} \le 0.043,\ 0.018 \le p_{15} \le 0.024,$$

$$0.002 \le p_{16} \le 0.004,\ 0.002 \le p_{17} \le 0.007,\ 0.001 \le p_{18} \le 0.018,$$

(9)

$$\sum_{b=1}^{18} p_b = 1,$$

Now, in order to find the optimal values $\dot{p}_b$ of the transient probabilities p_b, $b = 1,2,...,18,$ that maximize the objective function (8), w arrange the system conditional lifetimes mean values $\mu_b(2)$, $b = 1,2,...,18,$ in non-increasing order

$$\mu_8(2) \ge \mu_9(2) \ge \mu_1(2) \ge \mu_{18}(2) \ge \mu_3(2) \ge \mu_{15}(2) \ge \mu_4(2) \ge \mu_5(2) \ge \mu_{12}(2)$$

$$\ge \mu_{13}(2) \ge \mu_{14}(2) \ge \mu_{11}(2) \ge \mu_{16}(2) \ge \mu_6(2) \ge \mu_2(2) \ge \mu_7(2) \ge \mu_{10}(2) \ge \mu_{17}(2),$$

and we substitute

$$x_1 = p_8 = 0.016,\ x_2 = p_9 = 0.037,\ x_3 = p_1 = 0.038,\ x_4 = p_{18} = 0.013,$$

$$x_5 = p_3 = 0.026,\ x_6 = p_{15} = 0.024,\ x_7 = p_4 = 0.036,\ x_8 = p_5 = 0.363,$$

$$x_9 = p_{12} = 0.016,\ x_{10} = p_{13} = 0.351,\ x_{11} = p_{14} = 0.034,\ x_{12} = p_{11} = 0.003,$$

$$x_{13} = p_{16} = 0.003,\ x_{14} = p_6 = 0.026,\ x_{15} = p_2 = 0.002,\ x_{16} = p_7 = 0.005,$$

$$x_{17} = p_{10} = 0.002,\ x_{18} = p_{17} = 0.005. \quad (10)$$

Afterwards, we maximize with respect to x_i, $i = 1,2,...,18,$ the linear form (8) that after considering the substitution (10) takes the form

$$\mu(2) = x_1 \cdot 7.69 + x_2 \cdot 7.69 + x_3 \cdot 6.45 + x_4 \cdot 6.45 + x_5 \cdot 3.90 + x_6 \cdot 3.90$$

$$+ x_7 \cdot 3.80 + x_8 \cdot 3.80 + x_9 \cdot 3.80 + x_{10} \cdot 3.80 + x_{11} \cdot 3.80 + x_{12} \cdot 3.37 + x_{13} \cdot 3.37$$

$$+ x_{14} \cdot 3.24 + x_{15} \cdot 2.43 + x_{16} \cdot 2.43 + x_{17} \cdot 2.43 + x_{18} \cdot 2.43, \qquad (11)$$

with the following bound constraints

$$0.001 \le x_1 \le 0.018,\ 0.001 \le x_2 \le 0.056,\ 0.0006 \le x_3 \le 0.056,$$

$$0.001 \le x_4 \le 0.018,\ 0.018 \le x_5 \le 0.027,\ 0.018 \le x_6 \le 0.024,$$

$$0.027 \le x_7 \le 0.056,\ 0.286 \le x_8 \le 0.780,\ 0.013 \le x_9 \le 0.024,$$

$$0.286 \le x_{10} \le 0.780,\ 0.025 \le x_{11} \le 0.043,\ 0.002 \le x_{12} \le 0.004,$$

$$0.002 \le x_{13} \le 0.004,\ 0.018 \le x_{14} \le 0.024,\ 0.001 \le x_{15} \le 0.002,$$

$$0.002 \le x_{16} \le 0.018,\ 0.001 \le x_{17} \le 0.003,\ 0.002 \le x_{18} \le 0.007,$$

$$\sum_{i=1}^{18} x_i = 1.$$

Further, according to the procedure given in (Kolowrocki, Soszynska 2010), we calculate

$$\breve{x} = \sum_{i=1}^{18} \breve{x}_i = 0.7046,\ \hat{y} = 1 - \breve{x} = 1 -\ 0.7046 = 0.2954 \qquad (12)$$

and we find

$$\breve{x}^0 = 0,\ \hat{x}^0 = 0,\ \hat{x}^0 - \breve{x}^0 = 0,$$

$$\breve{x}^1 = 0.001\ \ \hat{x}^1 = 0.018,\ \hat{x}^1 - \breve{x}^1 = 0.017$$

$$\breve{x}^2 = 0.002,\ \hat{x}^2 = 0.074,\ \hat{x}^2 - \breve{x}^2 = 0.072,$$

$$\breve{x}^3 = 0.0026,\ \hat{x}^3 = 0.13,\ \hat{x}^3 - \breve{x}^3 = 0.1274,$$

$$\breve{x}^4 = 0.0036,\ \hat{x}^4 = 0.148,\ \hat{x}^4 - \breve{x}^4 = 0.1444,$$

$$\breve{x}^5 = 0.0216,\ \hat{x}^5 = 0.175,\ \hat{x}^5 - \breve{x}^5 = 0.1534,$$

$$\breve{x}^6 = 0.0396,\ \hat{x}^6 = 0.199,\ \hat{x}^6 - \breve{x}^6 = 0.1594,$$

$$\breve{x}^7 = 0.0666,\ \hat{x}^7 = 0.255,\ \hat{x}^7 - \breve{x}^7 = 0.1884,$$

$$\breve{x}^{8} = 0.3526,\ \hat{x}^{8} = 1.035,\ \hat{x}^{8} - \breve{x}^{8} = 0.6824. \qquad (13)$$

From the above, as according to (13), after considering the inequality

$$\hat{x}^{I} - \breve{x}^{I} < 0.295\ , \qquad (14)$$

it follows that the largest value $I \in \{0,1,...,18\}$ such that this inequality holds is $I = 7$.

Therefore, we fix the optimal solution that maximize linear function (11) according to the rule given in (Kolowrocki, Soszynska 2010). Namely, we get

$$\dot{x}_1 = \hat{x}_1 = 0.018\ ,\ \dot{x}_2 = \hat{x}_2 = 0.056\ ,\ \dot{x}_3 = \hat{x}_3 = 0.056\ ,\ \dot{x}_4 = \hat{x}_4 = 0.018\ ,$$

$$\dot{x}_5 = \hat{x}_5 = 0.027\ ,\ \dot{x}_6 = \hat{x}_6 = 0.024\ ,\ \dot{x}_7 = \hat{x}_7 = 0.056\ ,$$

$$\dot{x}_8 = \hat{y} - \hat{x}^{7} + \breve{x}^{7} + \breve{x}_8 = 0.2954 - 0.255 + 0.0666 + 0.286 = 0.393,$$

$$\dot{x}_9 = \breve{x}_9 = 0.013,\ \dot{x}_{10} = \breve{x}_{10} = 0.286,\ \dot{x}_{11} = \breve{x}_{11} = 0.025,\ \dot{x}_{12} = \breve{x}_{12} = 0.002,$$

$$\dot{x}_{13} = \breve{x}_{13} = 0.002,\ \dot{x}_{14} = \breve{x}_{14} = 0.018,\ \dot{x}_{15} = \breve{x}_{15} = 0.001,\ \dot{x}_{16} = \breve{x}_{16} = 0.002,$$

$$\dot{x}_{17} = \breve{x}_{17} = 0.001,\ \dot{x}_{18} = \breve{x}_{18} = 0.002.$$

Finally, after making the substitution inverse to (10), we get the optimal transient probabilities

$$\dot{p}_8 = \dot{x}_1 = 0.018\ ,\ \dot{p}_9 = \dot{x}_2 = 0.056\ ,\ \dot{p}_1 = \dot{x}_3 = 0.056\ ,\ \dot{p}_{18} = \dot{x}_4 = 0.018,$$

$$\dot{p}_3 = \dot{x}_5 = 0.027\ ,\ \dot{p}_{15} = \dot{x}_6 = 0.024,\ \dot{p}_4 = \dot{x}_7 = 0.056,\ \dot{p}_5 = \dot{x}_8 = 0.393,$$

$$\dot{p}_{12} = \dot{x}_9 = 0.013,\ \dot{p}_{13} = \dot{x}_{10} = 0.286,\ \dot{p}_{14} = \dot{x}_{11} = 0.025,\ \dot{p}_{11} = \dot{x}_{12} = 0.002,$$

$$\dot{p}_{16} = \dot{x}_{13} = 0.002,\ \dot{p}_6 = \dot{x}_{14} = 0.018,\ \dot{p}_{12} = \dot{x}_{15} = 0.001,\ \dot{p}_7 = \dot{x}_{16} = 0.002,$$

$$\dot{p}_{10} = \dot{x}_{17} = 0.001,\ \dot{p}_{17} = \dot{x}_{18} = 0.002, \qquad (15)$$

that maximize the system mean lifetime in the safety state subset $\{2,3,4\}$ expressed by the linear form (8) giving its optimal value

$$\dot{\mu}(2) \cong 0.056 \cdot 6.45 + 0.001 \cdot 2.43 + 0.027 \cdot 3.90 + 0.056 \cdot 3.80 + 0.393 \cdot 3.80$$

$$+ 0.018 \cdot 3.24 + 0.002 \cdot 2.43 + 0.018 \cdot 7.69 + 0.056 \cdot 7.69 + 0.001 \cdot 2.43$$

$$+ 0.002 \cdot 3.37 + 0.013 \cdot 3.80 + 0.286 \cdot 3.80 + 0.025 \cdot 3.80 + 0.024 \cdot 3.90$$

$$+ 0.002 \cdot 3.37 + 0.002 \cdot 2.43 + 0.018 \cdot 6.45 = 4.27. \qquad (16)$$

4 THE FERRY TECHNICAL SYSTEM OPTIMAL SAFETY CHARACTERISTICS

Further, using the optimal transient probabilities (15), we obtain the optimal solution for the mean value of the system unconditional lifetime in the safety state subset $\{1,2,3,\}$, $\{3,4\}$ and $\{4\}$

$$\dot{\mu}(1) \cong 4.92,\ \dot{\mu}(3 \cong) 3.79,\ \dot{\mu}(4) \cong 3.42, \tag{17}$$

and the optimal solutions for the mean values of the system unconditional lifetimes in the particular safety states 1,2,3 and 4 are as follows

$$\dot{\bar{\mu}}(1) = 0.65,\ \dot{\bar{\mu}}(2) = 0.48,\ \dot{\bar{\mu}}(3) = 0.37,\quad \dot{\bar{\mu}}(4) = 3.42. \tag{18}$$

Moreover, according to (2), the corresponding optimal unconditional multistate safety function of the system is of the form of the vector

$$\dot{\boldsymbol{s}}(t,\cdot) = [1,\ \dot{\boldsymbol{s}}(t,1), \dot{\boldsymbol{s}}(t,2), \dot{\boldsymbol{s}}(t,3), \dot{\boldsymbol{s}}(t,4)],\ t \geq 0, \tag{19}$$

with the coordinates given by

$$\dot{\boldsymbol{s}}(t,1) = 0.056 \cdot [\boldsymbol{s}(t,u)]^{(1)} + 0.001 \cdot [\boldsymbol{s}(t,u)]^{(2)} + 0.027 \cdot [\boldsymbol{s}(t,u)]^{(3)} + 0.056 \cdot [\boldsymbol{s}(t,u)]^{(4)}$$

$$+ 0.393 \cdot [\boldsymbol{s}(t,u)]^{(5)} + 0.018 \cdot [\boldsymbol{s}(t,u)]^{(6)} + 0.002 \cdot [\boldsymbol{s}(t,u)]^{(7)} + 0.018 \cdot [\boldsymbol{s}(t,u)]^{(8)}$$

$$+ 0.056 \cdot [\boldsymbol{s}(t,u)]^{(9)} + 0.001 \cdot [\boldsymbol{s}(t,u)]^{(10)} + 0.002 \cdot [\boldsymbol{s}(t,u)]^{(11)} + 0.013 \cdot [\boldsymbol{s}(t,u)]^{(12)}$$

$$+ 0.286 \cdot [\boldsymbol{s}(t,u)]^{(13)} + 0.025 \cdot [\boldsymbol{s}(t,u)]^{(14)} + 0.024 \cdot [\boldsymbol{s}(t,u)]^{(15)} + 0.002 \cdot [\boldsymbol{s}(t,u)]^{(16)}$$

$$+ 0.002 \cdot [\bar{\boldsymbol{s}}(t,u)]^{(17)} + 0.018 \cdot [\boldsymbol{s}(t,u)]^{(18)}, \tag{20}$$

for $t \geq 0$, $u = 1,2,3,4$, where $[\boldsymbol{s}(t,u)]^{(b)}$, $b = 1,2,...,18$, are the system conditional safety functions at particular operation states z_b, $b = 1,2,...,18$, given in (Kolowrocki, Soszynska 2010).

The safety function of the ferry technical system is presented in Figure 4.

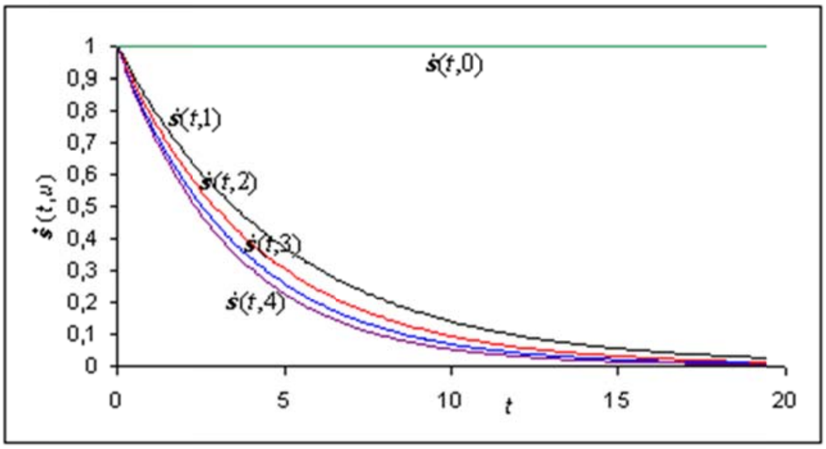


Figure 4. Graph of the optimal safety function $[\dot{s}(t,\cdot)]$ coordinates

If the critical safety state is $r = 2$, then the system risk function (Kolowrocki, Soszynska 2010) is given by

$$\dot{\boldsymbol{r}}(t) = 1 - \dot{\boldsymbol{s}}(t,2) \text{ for } t \geq 0, \tag{21}$$

where $\dot{\boldsymbol{s}}(t,2)$ is given by (20) for $u = 2$.

Hence, the moment when the optimal system risk function exceeds a permitted level, for instance $\delta = 0.05$, is

$$\dot{\tau} = \dot{\boldsymbol{r}}^{-1}(\delta) \cong 0.22 \text{ year}. \tag{22}$$

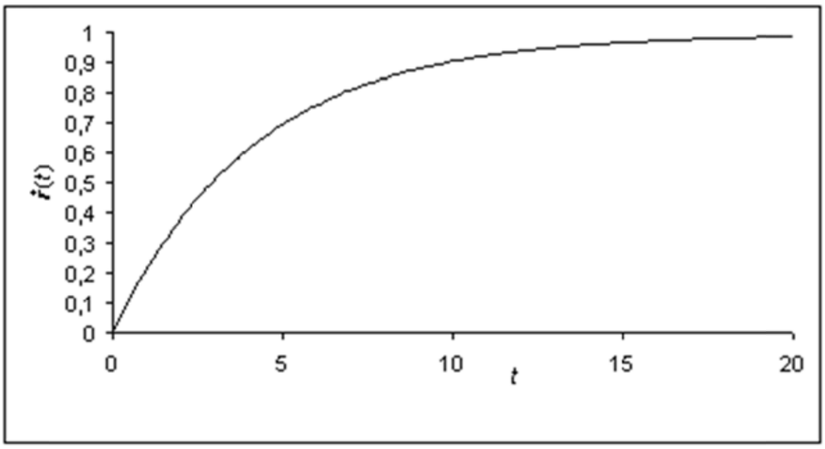


Figure 5. The graph of the ferry technical system optimal risk function $\dot{\boldsymbol{r}}(t)$

The comparison of the ferry technical system safety characteristics after its operation process optimization given by (16)-(22) with the corresponding characteristics before this optimization determined by (2)-(7) justifies the sensibility of this action.

5 CONCLUSION

The joint model of the safety of complex technical systems in variable operation conditions linking a semi-Markov modeling of the system operation processes with a multi-state approach to system safety analysis was applied to the maritime ferry technical system safety characteristics evaluation. Next, the final results obtained from this joint model and a linear programming were used to perform this complex technical system safety optimization. These tools practical application to safety and risk evaluation and optimization of a technical system of a ferry operating in variable operation conditions at the Baltic Sea waters and the results achieved are interesting for safety practitioners from maritime transport industry and from other industrial sectors as well.

Acknowledgements

The paper describes part of the work in the Poland-Singapore Joint Research Project titled "Safety and Reliability of Complex Industrial Systems and Processes" supported by grants from the Poland's Ministry of Science and Higher Education (MSHE grant No. 63/N-Singapore/2007/0) and the Agency for Science, Technology and Research of Singapore (A*STAR SERC grant No. 072 1340050).

6 REFERENCES

Aven, T. 1985. Reliability evaluation of multi-state systems with multi-state components. *IEEE Transactions on reliability* 34, pp 473-479.

Grabski, F. 2002. *Semi-Markov Models of Systems Reliability and Operations*. Warsaw: Systems Research Institute, Polish Academy of Science.

Kolowrocki, K. 2004. *Reliability of Large Systems*. Elsevier, Amsterdam - Boston - Heidelberg - London - New York - Oxford - Paris - San Diego - San Francisco - Singapore - Sydney - Tokyo.

Kolowrocki, K., Soszyńska, J. 2006. Reliability and availability of complex systems. *Quality and Reliability Engineering International* Vol. 22, Issue 1, J. Wiley & Sons Ltd. 79-99.

Kolowrocki, K., Soszynska, J. 2008. A general model of ship technical systems safety. *Proc. International Conference on Industrial Engineering and Engineering Management, IEEM 2008*, Singapore, 1346-1350.

Kolowrocki, K., Soszynska, J. 2009. Safety and risk evaluation of Stena Baltica ferry in Variable operation conditions. *Electronic Journal of Reliability & Risk Analysis: Theory & Applications* Vol.2, No.4 168-180.

Kolowrocki, K., Soszynska, J. 2010. Safety and risk optimization of a ferry technical system. Summer Safety and Reliability Seminars – SSARS 2010, *Journal of Polish Safety and Reliability Association,* Vol. 1, 159-172.

Lisnianski, A., Levitin, G. 2003. *Multi-state System Reliability. Assessment, Optimisation and Applications.* World Scientific Publishing Co., New Jersey, London, Singapore , Hong Kong.

Soszynska, J. 2006. Reliability evaluation of a port oil transportation system in variable operation conditions. *International Journal of Pressure Vessels and Piping* Vol. 83, Issue 4, 304-310.

Soszynska, J. 2007. Systems reliability analysis in variable operation conditions. *International Journal of Reliability, Quality and Safety Engineering. System Reliability and Safety* Vol. 14, No 6, 1-19.

Soszyńska, J., Kołowrocki, K., Kamiński, P., Judziński, M. & Milczek, B. 2009. Data mining for identification and prediction of safety and reliability characteristics of complex industrial systems and processes WP6. Safety and risk analysis and evaluation of a Stena Baltica ferry in variable operation conditions. WP6 - Sub-Task 6.2.5 – English – 30.11.2009. Poland-Singapore Joint Project. MSHE Decision No. 63/N-Singapore/2007/0.

RBF NETWORKS FOR IMAGING POINTS

Jakub Nedbalek
•
VSB-Technical University of Ostrava,
Ostrava-Poruba, Czech Republic
e-mail: jakubnedbalek@centrum.cz

ABSTRACT

The paper demonstrates new approach of rendering the graph point series called gantts. The gantts are placed in the two dimensional graph which contains the information about available production sources in the real manufacturing process. To have the interaction with a user, gantts are accompanied with the description text giving detailed information about each gantt. All gantt descriptions must be displayed without overlapping with each other. To optimize this task, the modified version of the RBF neural network with biases is applied. With respect to the similarity to the RBF structure, the new type of neural network is named RBF 2. We also give the picture of positive and negative attributes of the solution based on the neural network architecture.

1 INTRODUCTION

The human operator can see the software application form containing the tree structure and the graph. The tree structure content includes the items of the production process. These items differ in its kind. The superior element of the production items is operation, which belongs to the specific document called job card.

The two dimensional graph contains the inferior elements of the production item called the *source*. These items have subordinate items, that are called the *calendars,* describe the capacitive availability of the *source* item in the specific time interval. Generally, the *x*-axis depicts the constant time interval selected in advance and the *y*-axis shows the sources names, which depends on the topically chosen tree item. If we choose the superior operation or the job card type item in the tree, the graph must show all sources subordinated to the specifically chosen operation or job card.

Each calendar is depicted as a *gantt*. Gantt is the special chart series type; it’s a dash with predefined color. For gantt series type, we ignore its *y* size and take into account only length in *x*. Gantts cannot overlap with each other; every single gantt has its unique position (Gantt, H.L. 1974).

2 DESCRIPTION

In the usual practice, user almost always wants to have some type of gantt description to be pictured. The reason to do this is that the gantts accompanied with the text give us better lay out of the planned calendars and their assignment to the specified job card. Text labeled gantts stop being “anonymous” and enable user to have convenient feedback.

If the number of gantts in a graph is small enough, maximally up to 100, there is no problem to develop the algorithm of displaying the gantt descriptions without overlapping each other. If the number of gantts is small enough, we do not have to take into account the elapsed computation time, because the computation takes negligible amount of time.

Problems with inadequate long computation time occur in the real world usually. If there is around 10^3 gantts in the graph, the computation can take up to one minute (for the pc, where test took place) in the case of badly optimized algorithm. Too long time makes the screen “stack” without any logical response to the operator. It is necessary to know, that the procedure of displaying gantt descriptions is called not only if the user choose the tree item, but also in the arbitrary operations in the graph itself. User operations in the graph comprise zooming, scrolling and other similar actions that urge to refresh the position of gantt descriptions because these actions also change the position of gantts themselves.

There is also another problem with a large number of gantts in the graph. If we have too many gantts in the graph, there is a high chance of descriptions overlapping. That provides an uneasy survey for the user.

The target is to create the algorithm, which correctly assigns the description to the gantt without any overlapping of descriptions themselves. In the algorithm, there should be also taken into account the time necessary to compute the result. This means that the algorithm should be written with regard to the operator who needs to obtain results as quickly as possible. It is also suitable to depict as many non-overlapping gantt descriptions in the graph as possible to give the operator completed information about available items.

The paper is corrects the potential mistakes in the previous solution (Nedbalek 2009) and shows new prospects of the RBF 2 structure.

Our solution is applied in the *Production* module of the K2® ERP system. With respect to the computer language in which the software itself is written, the neural network is implemented in Delphi code.

3 THE SOLUTION PROPOSAL

The main task of our neural network is to classify, which gantt belongs to the input point according to the criteria of least distance between the point and gantt. The point equals to the last point of the gantt description to which we need to find either a) the nearest non – colliding gantt with higher *x* coordinate (forward direction) or b) the nearest gantt description on a lower *y* coordinate (backward direction). Result of the forward check is the gantt (with higher *x* coordinate), which description we want to display (the gantt is not covered with previously displayed description). The result of the backward check is the gantt (with lower *y* coordinate), which description could collide with the already displayed description that is obtained by the forward check. Whether the description will be shown in the graph depends on the result of the geometry comparison of border points.

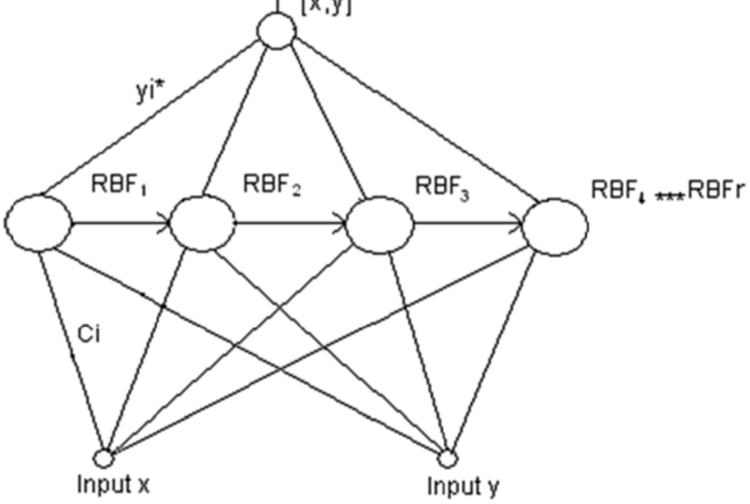


Figure 1. The RBF 2 neural network

To tackle the problem of descriptions overlapping, it is possible to apply the RBF 2 neural network. Its architecture was evolved from the well known RBF (Chen et al. 1991, Yee & Haykin 2001). See the Figure 1.

This network consists of three layers – distributive, hidden and output one. We aim our attention on the hidden one.

The hidden layer contains the neurons with the activation function. In the most cases, the Gauss function is used for the RBF

$$y^* = \exp^{\left(-\frac{\varphi^2}{\sigma^2}\right)} \tag{1}$$

where:

$$\varphi = \sqrt{\sum_{i=1}^{n} (x_i - c_i)^2} \tag{2}$$

is the distance between the $\bar{c}$ center of the neuron and the arbitrary $\bar{x}$ input [4]. The $\bar{c}$ center denotes to the vector, to which the neuron is trained and describes the pattern that is compared with an input. Practically, the $\bar{c}$ center refers to the gantt position.

The size of σ defines the dispersion in a Gauss function.

The second possible activation function is directly

$$y^* = \varphi \quad \text{(Snorek 2004).} \tag{3}$$

Both relations (1) and (3) were checked out. Finally, with respect to the time spare, the modified relation (3) was applied. The modification consists in the fact, that the square can be neglected in the case, if we know when the $\bar{x}$ input is less than $\bar{c}$ center (we know that because the gantts have ascending order according to their position). The sum across *i* is also ignored as we take into account only one coordinate (distance is counted separately for each coordinate – the collision can occur for *x* and *y* independently).

Our modified network activation function for the *x* axis equals

$$\varphi = x - c \ \text{ for } x \ge c, \tag{4}$$

$$\varphi = Grafx_{\max} - Grafx_{\min} + 1 \ \text{ for } \quad c > x, \tag{5}$$

for the backward and

$$\varphi = c - x \ \text{ for } c \ge x, \tag{6}$$

$$\varphi = Grafx_{\max} - Grafx_{\min} + 1 \ \text{ for } \quad x > c, \tag{7}$$

for the forward check. $Grafx_{max}$ is the maximum and $Grafx_{min}$ is the minimum of *x*-axis. The maximal φ is greater than the *x* size of the graph that is always greater than the absolute distance of *c-x*.

To check the *y* coordinate, we will use relations:

$$\varphi = 0 \ \text{ for } h > c - y \tag{8}$$

and

$$\varphi = 1 \quad \text{for } h < c - y, \tag{9}$$

where h is the constant height of gantt description. Activation function for y coordinate is consequently included only in the backward check. It is because the forward check is used to find the first gantt, which does not collide with the last displayed description on the same y line. If there is no such gantt, the first gantt on the nearest subsequent y line will be used.

The idea of relation (3) was applied in (Nedbalek 2009). If we want to take into account the neuron biases, we can simply add to our activation function

$$y^{*} = \varphi + \Theta \tag{10}$$

where Θ stands for the bias. The amendment in (10) merely means that the result of activation function will be ignored (neuron will stay in a passive state) in case that the bias is not set to a predefined value.

The next difference of the RBF 2 is the presence of unidirectional bind between two neighbouring neurons. It means, that the neuron memorizes the $\bar{c}$ center [x,y] of the previous neuron (not vice versa). This improvement is used in the training of the neural output layer.

Neurons of the hidden layer are implemented by a memory table (similar to database one, but stores all data in the memory, not physically on the hard drive) containing the index (sequence number) of neurons, their $\bar{c}$ [x,y] position, the bias flag, the position of a previous neuron and its bias. Each neuron defines directly the position of a gantt that is also the first point of description.

The training of the RBF 2 can be categorized in two parts – training of the hidden and the output layer.

Training of the hidden layer can be realized the way a pattern (gantt) is chosen from the set and is defined directly as a prototype. Its position describes exactly the $\bar{c}$ center of the neuron. This simple method provides the fastest way of training. Another advantage of this method consists in reflecting the data position within the input space. The hidden layer is trained backwardly according to the formulas (4), (5), (8) and (9) and forwardly according to the (6) – (9). The training process is executed during the neural network creation when the input data are disposed. The RBF 2 is forced to be trained again whenever the mutual position of gantts in the graph is actualized (i.e. operator chose another item in the tree, etc.)

When we train the output layer, we need to set the weight according to the hidden layer. If the hidden layer neuron has the least distance between its $\bar{c}$ center and an arbitrary $\bar{x}$ input, then the weight from this input will equal one. In all other cases, the weight will be set to zero. This process of seeking the minimal distance is activated whenever there is any need to call for the forward or backward check.

Finding the minimal distance and the appropriate gantt, defined by the $\bar{c}$ center position, is optimized for time. This is enabled by the mentioned unidirectional connection between neighboring neurons. Due to that, the output neuron in each cycle across the memory table is able to "see" two neurons of hidden layer in one moment.

In the previous variant (Nedbalek 2009) of RBF 2, we neglect the network biases. In the present solution, the neuron bias is taken into account. This practically means that in the output layer training phase (looking for the minimal distance between the $\bar{c}$ centre an arbitrary $\bar{x}$ input) and only for the backward check, we skip all neurons which do not have their description presence flag set – its value equals zero. The flag is additionally activated (set to nonzero value) for the neurons that successfully passed their collision (forward and backward) test and were defined as appropriate to hold and show the gantt description. Acquired information about presence of the description is later used in the backward check for another gantt in the sequence, when we need to decide, whether it is possible to display the description. The reason, why to use biases only in backward check is that in the forward check (from our gantt forth), there are no descriptions yet.

Adding the description presence flag as a neuron bias makes our output layer training phase faster and gives the user better feedback.

4 RESULTS

We obtained all results from the actions that user typically performs – selecting tree nodes of *source* and *operation* type (some of them repeatedly). This is followed by automatic repainting of the graph surface.

Table 1. contains the results for each of three methods. The first one, that is the original method, implements the forward and backward testing without a time optimisation, which is provided by the RBF 2. On contrary to the RBF 2 (a), the (b) version uses the neuron biases.

Table 1. The results for all methods

	Origin. method [s]	RBF 2 (a) [s]	RBF 2 (b) [s]
1. Sources aver.	30,8095	18,5220	9,5889
std.dev.	8,2622	0,1459	2,2745
2. Operations aver.	27,5607	19,4707	9,1866
std.dev.	5,1690	2,9958	1,9079
3. Operations 2x aver.	57,4603	42,6339	18,8508
std.dev.	11,5381	8,2758	3,3455
4. Operations 3x aver.	100,8735	68,2888	28,3871
std.dev.	21,3945	12,2965	7,0934
5. Main node aver.	31,1623	29,6636	5,3722
std.dev.	7,6225	5,1430	1,2236

The abbreviations in the first column from left denote: aver.- average, std.dev. – standard deviation; sources – user clicked on all nodes of the *source* type in the tree component; operations – user clicked on all nodes of the *operation* type; 2x (3x) – user clicked on the specific node type two times (three times).

Numbers in the Table 1. mean the time necessary to compute all possible collisions among the gantt descriptions summed with the time necessary to depict the gantts with their descriptions on the graph surface. Simply expressed, it is the time which the graph needs to respond on the user action.

Making all these comparison tests makes sense only in case the set of displayed gantt descriptions is the same for all methods. This assumption was verified and successfully fulfilled.

*Figure 2.*shows the benefit of both versions of the RBF 2 which are compared to the original method without time optimization.

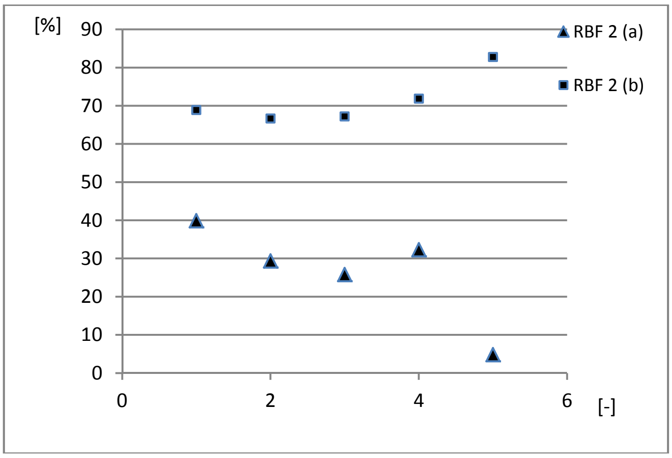


Figure 2. Time saving for the modifications of RBF 2
(a -no biases, b-with biases) with respect to the original method (see Table 1.)

Figure 2. demonstrates the time saving of computing time in percentual ratio (y axis) of the RBF 2 method referred to the former (original) method. The RBF 2 is depicted for both versions (a – no biases, b – with biases). The x axis includes the number of the user action (see Table 1., first column from left). As we can see, the RBF 2 with the neural biases gives us better results. For instance, if we apply the RBF 2 ver. (b), we can achieve the result in computation of all possible gantt description collisions with approximately 70% of time saving with respect to the original method. The RBF 2 ver. (a) for the same computation enables typically 30% of time saving.

5 CONCLUSION

The paper demonstrates the new enhancement of the RBF 2 neural network – based method of displaying the data in a manufacturing graph. The enhancement consists in applying the neural biases, which were omitted in the previous version of the RBF 2 (Nedbalek 2009). This method improves the contemporary state of the RBF 2 architecture, which requires more computation time to display all gantt descriptions. The RBF 2 structure itself was derived from the well known RBF, to optimize our task.

The results were measured on the database, which was obtained from the real manufacturing process. This activity indirectly ensured that the graph should contain the high number of gantts. This is required because the RBF 2 must pass the stress tests in order to be applied in the practice. In other words, working with the real data provides us a good picture about the real conditions. Although we dispose of convictive results of the RBF 2 behavior, it still undergoes the stress tests presently.

The only obvious disadvantage of the RBF 2 neural network is a lack of generality of the output layer training function. That is caused by the fact that the RBF 2 neural network was developed for time optimization of gantt descriptions overlapping. To make the RBF 2 more universal, the output layer training should be altered. However, we can presume that the price for universality will be a partial loss of the time optimization ability.

ACKNOWLEDGEMENT

The author of this paper would like to thank for the financial support on behalf of the research and development project num. 1M06047 (CQR), which is subsidized by the Ministry of Education of the CR.

REFERENCES

[1] Gantt, H.L.:Work, Wages and Profit. The Engineering magazine, New York, 1910.http://en.wikipedia.org/wiki/Gantt
[2] Nedbalek, J.: The quality of graphical data display. ESREL 2009, Reliability, Risk & Safety: Theory and Applications, ISBN 978-0-415-55509-8, 2010 Taylor & Francis Group, pg. 2041-2044.
[3] Chen, S., Cowan, C.F.N. & Grant, P. M. March 1991. Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks. IEEE Transactions on Neural Networks, Vol. 2(2): 302-309.
[4] Yee, Paul V. and Haykin, S.: Regularized Radial Basis Function Networks: Theory and Applications, John Wiley, 2001.
[5] Šnorek, M.: Neuronové sítě a neuropočítače. ČVUT, Praha, 2004.
[6] Nedbálek, J.: New type of neural networks for rendering graph points. Summer Safety and Reliability Seminars, SSARS, Gdansk-Sopot, Poland, 2010.

AN EXTENDED METHODOLOGY FOR RISK BASED INSPECTION PLANNING

J. T. Selvik
•
University of Stavanger* and IRIS (International Research Institute of Stavanger)
IRIS Energy - Risk Management, Stavanger, Norway
e-mail: jts@iris.no

P. Scarf
•
University of Salford
Business Operations, Greater Manchester, Salford, UK
e-mail: P.A.Scarf@salford.ac.uk

T. Aven
•
University of Stavanger
*Department of Industrial Economics, Risk Management
and Planning, Stavanger, Norway
e-mail: terje.aven@uis.no

ABSTRACT

Inspection planning is an important activity in process industries, and one of the key tools used for such planning is the risk based inspection (RBI) methodology. The RBI is commonly used in planning of inspections for static mechanical equipment, in particular piping networks. The inspections are prioritized based on risk, expressed as expected values, integrating the likelihood and consequences of failures. In this paper we suggest an extension of the RBI methodology which reflects risk and uncertainties beyond expected values. We argue that such an extension is essential for adequately supporting the inspection planning. A pipeline example from the Norwegian oil and gas industry is presented to illustrate and discuss the suggested approach.

1 INTRODUCTION

Inspections are widely used in the process industries to reduce risks related to failures on static mechanical equipment, for example on pipelines which accounts for the greatest proportion of equipment damage in petrochemical plants (Tien et al. 2007). Use of inspections is essential if availability and high performance is to be achieved, but preventive maintenance (PM), such as inspections, is expensive and contributes to a relatively large share of the total operational costs. The inspections imply direct costs and also risks for maintenance introduced failures. Maintenance planning is about balancing these concerns.

To aid the decision-makers in their inspection planning, different types of tools are available. One of these tools is addressed in this paper, the risk based inspection (RBI) methodology; a methodology commonly used within the chemical, petrochemical, the oil & gas and the refinery industries. Successful implementation of RBI is demonstrated by many authors (Bragatto et al. 2008); see e.g. Poulassichidis (2009), Herzog & Jackson (2009), Chang et al. (2005), Patel (2005), Landet et al. (2000), Nilsson (2003), Wintle et al (2001), Ablitt & Speck (2005) and Hagemeijer & Kerkveld (1999).

The risk based inspection methodology, as indicated by its name, assesses risk to support the inspection planning. Risk is computed for the relevant pressurized equipment and the failure mode loss of containment, caused by either material deterioration or external influence (such as dropped objects). Risk of failure (RoF) is assessed quantitatively following a two-dimensional risk perspective, comprising the probability of failure (PoF) and the consequence of failure (CoF), and is typically expressed as the product of the two; see for example Chang et al. (2005).

Based on the risk values calculated, the risk based inspection methodology provides recommendation on what, when, where and how to inspect, and also what should be documented. There exist different versions of the RBI methodology, reflecting variations in preferred approach for modelling of the material degradation and the probabilistic treatment. Some assessors promote an expert-based (subjective) approach, for example Kallen & Noortwijk (2005) who adopts a Bayesian approach for handling errors in equipment wall thickness measurements. Others prefer a more traditional frequency-based probability assessment approach. However, although variations exist, the fundamental pillars are shared; they are defined by technical standards such as API (2002, 2008) and DNV (2009); see also Jovanovic (2003), Kallen (2002) and Khan et al. (2004).

The following mix of qualitative, semi-quantitative and quantitative elements summarise the fundamental pillars of the RBI methodology:

- An inductive analysis of potential failures, for example a failure mode and effects analysis (FMEA), is typically used to screen and assess the consequences of the system, see e.g. Hagemeijer and Kerkveld (1999).
- Calculation of RoF, which is traditionally a part of a quantitative risk assessment (QRA), and includes modelling of the degradation process; see e.g. Chien et al. (2009), Chang et al. (2005) and Santosh et al. (2006).
- Application of a qualitative or semi-quantitative risk matrix to express the risk level and relationship between PoF and CoF; see e.g. Patel (2005) and Truchon et al. (2007).
- Use of the ALARP principle for planning of intervals based on the assessed risk; see e.g. Simpson (2007) and Khan et al. (2006).

Our prime concern related to the use of the above elements, is how uncertainties are addressed. The traditional RBI assesses risk as a combination of probabilities and failure events and consequences (or losses), but such a risk perspective fail to bring into account all the relevant uncertainties. The risk assessments are based on background knowledge, and this knowledge may include assumptions that could conceal uncertainties not addressed by the probabilistic assessments. For example, for the assessments of pipeline degradation there are assumptions made on the presence of erosive sources, such as the size and concentration of sand particles in the fluid stream. The probabilities produced to assess the risks are conditioned on these assumptions.

To take such uncertainties as indicated above into account, a broader risk perspective is needed. One way to do this is to apply a risk perspective presented in Aven (2008a), where probability is replaced with uncertainty in the definition of risk. In this perspective probability is a tool used to describe the uncertainties, and is conditional on the background knowledge. By using such a risk perspective we are able to shift the methodological focus from probabilities and expected values to uncertainties. To highlight this shift, we name this adjusted methodology "extended risk based inspection" or as ERBI for short.

The purpose of the present paper is to motivate the use of this extended risk based inspection methodology, and to describe its main features. A pipeline example from the oil and gas industry is used to illustrate the applicability of the suggested methodology. The aim is to determine the inspection interval for a 15 inch carbon pipeline located in the Norwegian sector of the North Sea. The 9 km pipeline has welding points each 12 meters, and transports a corrosive multiphase well stream from multiple subsea production facilities. The pipe is covered with a protective layer, an inner coating, to avoid damage on the carbon steel from inside corrosion, as illustrated by Figure 1. A similar case is discussed by Castanier & Rausand (2006), where a classical PF interval model is used for the pipeline maintenance optimization. See also Tien & Tsai (2007) and Bjørnoy et al. (2001).

The structure of the remaining part of the paper is as follows. The next section presents a brief description of the traditional RBI methodology, demonstrated on the example presented above. Section 3 explains the new extended (ERBI) methodology. The methodology is then discussed in Section 4, where it is compared to the standard RBI. The example presented is used as a basis for the comparison. The last section, Section 5, provides some conclusions.

2 DESCRIPTION OF THE RISK BASED INSPECTION METHODOLOGY

The risk based inspection (RBI) methodology comprises the following four phases:

1. Equipment screening
2. Detailed risk assessment
3. Inspection interval assessment
4. Implementation, evaluation and updating

In this section we will give a brief presentation of these phases (see Sections 2.1- 2.4) using the described pipeline example as an illustration. The methodological description of the RBI is based on the available standards API (2000, 2008) and DNV (2009).

Figure 1. Illustration of pipeline section (3x12 meter) with inner coating

Before starting on the RBI assessments, a project team is designed, to ensure that the adequate capacities are included and relevant information is available for the assessments. For the collection and use of data in the oil & gas industry, we refer to the ISO 14224 standard (ISO 2006).

2.1 Equipment screening

A screening is performed at the initial phase, for example by use of FMEA or risk matrices, to be avoid unnecessary assessments of equipment of low risk. Equipment assigned low consequences and low probability of failure is excluded from the detailed risk assessments in the next phase.

In order to perform the screening, the equipment is grouped into hierarchal levels. DNV (2009) recommends the use of five levels, defined in accordance to ISO (2006), but fewer levels may be appropriate, if the number of items is low and no special concerns are present.

The screening is performed for three categories of failure consequences; operational (production availability), environmental and safety consequences. Redundancy, hidden failures and non-operational consequences are not assessed. However, main focus is normally on personnel safety.

2.2 Detailed risk assessment

The detailed risk assessment is performed in two steps:

1. Separate probability of failure and consequence of failure assessments
2. Assessment of the risk of failure based on the results from the first step

To describe the steps we refer to the case presented in Section 1. The crude assessments indicate that the pipeline has potential failure consequences that require a more detailed assessment before inspection intervals can be specified. Available historical data show that few similar failure events have occurred, but that those that have occurred have been critical to the production. The detailed assessments will provide a more precise risk picture for the determination of the inspection intervals.

First we assess the probability of failure (PoF), or more specifically, the probability of the occurrence of the failure mode loss of containment. Several databases are available for this assessment, including integrity and reliability databases. This part of the assessment is challenging. An understanding of the failure and degradation mechanisms of the equipment is needed to find a model that produces the expected failure rates. Much of the variation in available literature on RBI is related to alternative ways of improving the modelling of the equipment degradation. DNV (2009) for example, has suggested three different models for this purpose; an insignificant rate model, a rate model and a susceptibility model:

- Insignificant rate model: To be used if no degradation is expected. A fixed probability of failure equal to 10^{-5} per year is used. It is assumed that time of the assessment is irrelevant for the risk of failure.
- Rate model: To be used if wall thickness is decreasing with time (the most common scenario). The rate modelling includes factors such as wall thickness as a function of time, the material and fluid properties and the operating conditions.
- Susceptibility model: To be used if external events may lead to a suddenly increased probability of failure. Such events could be a dropped object causing pipeline rupture. It is a difficult task to model such events, and knowledge on environmental and operating conditions, and also monitoring capacities and routines are of relevance to the modelled probability of failure.

For schematic illustration of the models described above, see DNV (2009). If none of the above models are applicable further investigations would be required.

For the pipeline example we find the rate model to be applicable as the high sand concentration in the fluid stream cause significant erosion to the pipe walls. The input parameters in the model are determined by a combined use of historical data and engineering judgements, and by summarizing the probabilities for all potential failure events, the annual failure probability for the pipeline is placed in the range 10^{-4}- 10^{-3}. For technical details on how to model the degradation (for example fatigue assessments) and determine the PoF, we refer to DNV (2009).

Next we assess the consequence of failure (CoF) for the pipeline case, by combining the three categories referred to in Section 2.1. Regardless of the equipment addressed, the failure consequences are to a large extent determined by the operating conditions and system design. For the pipeline, the consequence is dependant on the leakage volume or rate (dispersion), fluid properties, and the ignition potential. For calculation of the expected consequences for the operational, safety and environmental impacts, an event tree is useful to summarize and weight outcomes. Alternatively, API (2000, 2008) refer to use of consequence relevant factors for the calculations, where CoF is a function of factors for production loss, pressure, explosion damage potential, toxicity, production effect, location, recovery time, non-production effect and safety system effect. In many cases, as another alternative, a qualitative expert judgement is used to assess the consequences, (Santosh et al. 2006). Often qualitative categories are used, as in our example where five categories were defined: insignificant, minor effect, local effect, major effect and massive effect. In the analysis we assign pipeline failures to have major effect, as a leakage would shut down the entire production.

Based on the assessed RoF and CoF a risk decision matrix may be produced, as shown for the pipeline example in Table 1. It is seen that a two year inspection interval for the pipeline example is recommended.

2.3 Inspection interval assessment

The risk decision matrix specifies the inspection intervals as a function of probability and consequences. Equipment assessed to have a low RoF are prescribed corrective maintenance (CM). Equipment assessed to have a high RoF are prescribed to have rather frequent inspections, for example once every year.

In cases where significant variation exists in the failure consequences between the operational, safety and environmental categories, separate matrices are often used. And the minimum inspection interval across the separate matrices is then chosen.

Table 1. Example of RBI decision risk matrix (DNV 2009). Recommended time between inspections (in years).

	CoF ranking				
PoF ranking	Insignificant	Minor effect	Local effect	Major effect	Massive effect
$>10^{-2}$	0	4	2	1	1
10^{-3}- 10^{-2}	0	4	2	1	1
10^{-4}- 10^{-3}	0	0	4	2	2
10^{-5}- 10^{-4}	0	0	8	4	4
$<10^{-5}$	0	0	8	8	8

For the use of the results, two different principles are reflected by the referred standards. While DNV (2009) points to use of company risk acceptance criteria, the API (2000, 2008) on the other hand, points to use of the ALARP principle. This principle states that the risk should be reduced to a level that is as low as reasonably practicable, meaning that risk-reducing measures should be implemented or chosen unless it can be demonstrated that there is a gross disproportion between costs and benefits. The common tool to verify ALARP is cost-benefit analysis. Indirectly such link is also provided by DNV (2009), which refers to the NORSOK standard: Z-008 (NORSOK 2001), for planning of maintenance activities in the oil and gas industry. The standard recommends use of cost-benefit assessments to ensure a proper balance between frequency of maintenance and the risks of equipment failures.

2.4 Implementation, evaluation and updating

Decision-making and integrating the results into an inspection plan requires additional considerations to be taken into account. These considerations are strongly dependant on the available inspection resources and existing PM programmes. The implementation and evaluation process is typically a part of the company maintenance management systems (Truchon et al., 2007 and Kallen, 2002), where experience from the inspections will later provide relevant information for updating and evaluating of the inspection programmes. See also Bertolini et al. (2009) and Chien et al. (2009).

2.5 Potential for methodological improvements: Extended uncertainty assessments

Several studies have showed that uncertainties in assumptions made in the RBI assessments are to limited extent reflected by the final results; see for example Geary (2002), Herzog and Jackson (2009) and Simpson (2007). A main source of these uncertainties is related to the choice of models. This in its turn has motivated several adjusted RBI methods to cope with this problem. A main category of such methods are based on fuzzy logic; see for example Khan et al. (2004) and Khan & Haddara (2003). It is argued that risk is difficult to assess due to the complexities involved in modelling of the degradation process (e.g. corrosion rate) and failure consequences, and also due to the model input data (Singh & Markeset 2009). A fuzzy approach is believed to express the relevant uncertainties and produce a more precise method by adjusting modelled material degradation with assessed “trust” values (Singh & Markeset 2009). However, we find the values generated by this approach to be “arbitrary” and not justified, and they are not able to properly address uncertainties in the assumptions made.

The adequate tool for quantifying uncertainties is in our view subjective (knowledge-based) probabilities. If the assessor assigns a probability of an event A, given the background knowledge K, equal to 0.1, i.e. P(A|K) = 0.1, it means that the assessor regards his/her assessment of uncertainty (likelihood, degree of belief) as comparable to randomly drawing one particular ball out of an urn comprising 10 balls. However, we acknowledge the need for qualitatively assessing uncertainties beyond the probabilities as the K could “hide” uncertainties as was noted in Section 1. We need to capture also the risk contributions from potential “surprises” (”black swans” Taleb, 2007).

A proper framework for risk assessment according to this perspective is presented by Aven (2008). In this perspective uncertainty and not probability is the main component of risk. Risk is understood as the two-dimensional combination of:

Events (A) and the consequences of these events (C); A: leakage due to loss of containment, for example pipe rupture; C: the leakage and maintenance consequence

Uncertainties U about A and C (will A occur and what will the consequences C be?)

Such a risk perspective is referred to as the (A, C, U) perspective Aven, 2008a). The key to this risk perspective is the broader risk descriptions highlighting uncertainties “hidden” in the assumptions. These uncertainties are referred to as “uncertainty factors”.

In the following section we present an extension to the risk based inspection based on this risk perspective. It is referred to as the extended risk based inspection (ERBI) methodology. RBI will still be the methodological platform, but the approach to risk and uncertainties will be more comprehensive. Our approach is based on similar ideas as supporting the subjective probability approach by Apeland and Scarf (2003), but the risk perspective is broader by the incorporation of the uncertainty factors.

3 DESCRIPTION OF THE EXTENDED METHODOLOGY

In this section we present the extended risk based inspection (ERBI) methodology as indicated in the previous sections. It is described by eight successive boxes that are placed into a decision framework for determination of the inspection programme, as illustrated by Figure 2 below:

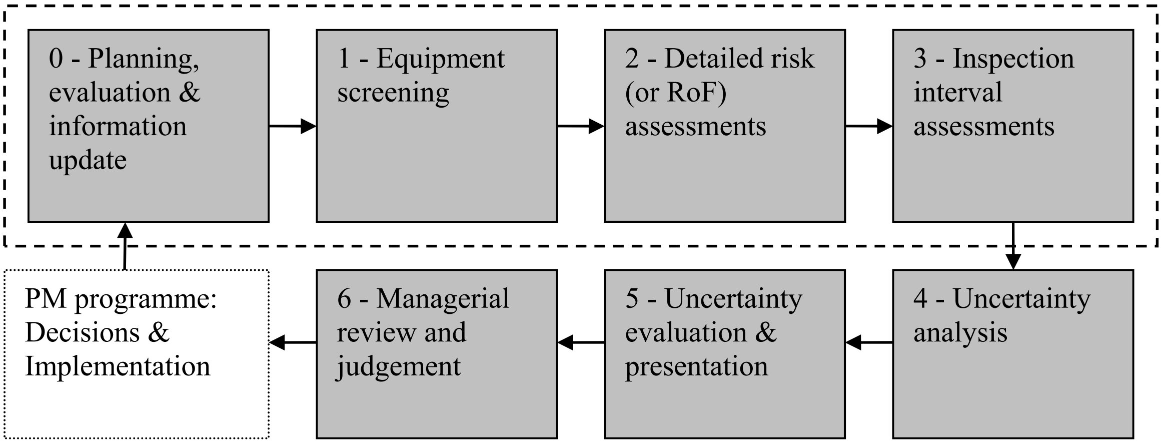


Figure 2. Framework for the extended methodology
(the RBI methodology is indicated by the dashed line)

The first four boxes, 0-3, are described by the phases of the RBI methodology as presented in Sections 2.1-2.4, and provide the already existing methodology to the framework. We then introduce some new assessments in boxes 4 and 5. These are separate uncertainty assessments included in the ERBI methodology, and are additional to those performed as integrated parts of assessments in the RBI phases.

In the fourth box we focus on the uncertainty factors mentioned in the previous section. Many of these factors are derived from the assumptions made in the detailed risk assessments. In line with Aven (2008a), the uncertainty analyses cover the following main tasks:

- Identification of uncertainty factors
- Assessment and categorization of the uncertainty factors with respect to degree of uncertainty
- Assessment and categorization of the uncertainty factors with respect to degree of sensitivity
- Summarization of the uncertainty factors' importance

Scores, high (H), medium (M) or Low (L), are assigned for the tasks 2-4. Below a score system is presented for ranking of the uncertainty, inspired by Flage and Aven (2009), where the interpretation for the scores are as follows:

Low (L) uncertainty:
One or more of the following conditions are met:

- The assumptions made are seen as very reasonable
- Much reliable data are available
- There is broad agreement/consensus among experts.
- The phenomena involved are well understood; the degradation models used are known to give predictions with the required accuracy

High (H) uncertainty:
One or more of the following conditions are met:

- The assumptions made represent strong simplifications
- Data are not available, or are unreliable
- There is lack of agreement/consensus among experts
- The phenomena involved are not well understood; degradation models are non-existent or known/believed to give poor predictions

Medium (M) uncertainty:
Conditions between those characterizing low and high uncertainty.

Scores high (H), medium (M) or Low (L) are also used for the assessment of the sensitivity for the uncertainty factors. The judgement of the sensitivity score is linked to the extent that the factor is able to change the inspection interval. A medium score is assigned if a relatively large change in the base case values is needed to bring about altered conclusions, a low score is assigned if unrealistically changes are needed, and a high score if relatively small changes are needed.

Then after the uncertainties and sensitivities of the uncertainty factors are qualitatively assessed, a summarization of these factors' importance is performed. The importance score is interpreted as the average of the score for the tasks 2-3.

The steps 1-4 provide the input to the uncertainty evaluation of the system studied (see box no. 5). Such evaluation is recommended by for example Khan and Haddara (2003), as part of the communication of results to the management function.

A managerial review and judgement feature is also included, as shown in the sixth box in Figure 2, in line with the decision framework presented in Aven (2008a). The inputs to management from the various assessments are placed into a broader context, where the boundaries and limitations of the various assessments are taken into account, and also additional aspects and inputs are taken into consideration, e.g. manufacturer recommendations and existing PM programmes. The managerial review and judgement may also request revisions or analytic changes should results appear unreasonable.

The next Section presents the results from the uncertainty assessment for the pipeline example.

3.1 Uncertainty assessments in the example

Our focus is on the uncertainty factors that have the potential to change the probabilities (of events and consequences) to such an extent that it may have an effect on the specified inspection intervals. For the pipeline example presented, several critical assumptions made in the detailed RoF calculations were identified. Below we present and list some of the derived uncertainty factors based on these assumptions:

1. The pipeline is properly tested and inspected before and during installation
2. All other items in the assessments are functioning (not only the system considered)
3. Data selection criteria are based on pipe description and fluid type
4. Data are able to describe the pipe material degradation
5. Use of “smart pig” provides accurate sensor readings inside the pipeline
6. External failure events may be ignored
7. Inspection results are representative for the whole pipeline length

These uncertainty factors are briefly described in the following, in the order above.

The first uncertainty factor to be addressed is the assumption that the pipeline, including the welding between the pipes, are adequately tested and inspected prior to production start up. Due to the pipe being produced with a corrosion resistant alloy layer, an inspection challenging type of welding was required to connect the pipes, and thus requiring a new and alternative inspection method instead of using traditional ultra-sonic inspections. It is assumed that this new method ensures detection of weaknesses in the pipe and welding, but although the methodology was verified during the pipeline qualification process, sparse experience exists on this inspection method and its limitations. As for the assessed risk, this would be considerably higher should the inspection method prove inadequate.

The probability of failure assessments were carried out assuming that only one failure event occurs at the time. It is assumed that all the other items are working perfectly. None of the other items are then in a failure state, are waiting for maintenance or have hidden failures. However, real life may very well be different, and may also have relevance for the assessed consequences.

The risk assessments are, to a large extent, based on data found in company internal databases. However, the selection criterion used may lead to failure probabilities that do not reflect the inner diameter of the pipe and number of welding points, and also the erosive properties of the fluid. It is uncertain to what extent the criterion adopted has included pipelines that are subject to similar conditions.

There are limited amounts of relevant data available to predict the performance of the equipment. The data represent newer pipeline systems, and for these few events have occurred. The relevant items’ sizes and material property combinations are not found in older data. Thus, one may question to what extent the data are dominated by the items’ “childhood events”.

The pipeline is regularly pigged by use of a device called a “smart pig”. The purpose of such a device is to clean the inside of the pipe during production as the smart pig is mechanically sent through the pipeline, while at the same time being able to monitor pipeline inside parameters, for example the inner diameter and temperature. The reliability and accuracy of this smart pig is not evaluated by the risk assessments. As this smart pig ensures the integrity of the inside protective layer (the inner coating) of the pipeline, the assumption that use of smart pig provides accurate sensor readings, may lead us to ignore potential damage inside the pipe.

It is assumed that the pipeline is located in an area with limited traffic and exposure to dropped objects; however this is an assumption based on the fact that most of the production and maintenance activities are performed close to the production vessel and riser base. But there may be other vessel operating in the vicinity that may cause damage to the pipe. Such events are very difficult to model, and are assumed to have a negligible risk affect, even though such events exist in the company internal reliability data.

It is assumed that inspection results are representative for whole pipeline length, although only parts of the pipeline will be subject to thorough inspections. In the example presented a relatively short pipeline length is used (only 9 km) to simplify the assessments. In actual pipeline networks the length is often found to exceed 100 km. For example, the Gassco operated pipeline "Zeepipe I" exporting gas from Sleipner on the Norwegian Continental Shelf to Zeebrugge in Belgium has a total length of 813 km. For an overview of various pipeline networks, see Gassco (2010).

Now, having identified a list of uncertainty factors, we next assess and categorize these with respect to degrees of uncertainty and sensitivity, which combined provides a basis for making a judgement of importance. The results are shown in Table 2, based on the score system presented in Table 1.

Table 2. Pipeline example uncertainty assessment

Uncertainty factor	Degree of uncertainty	Degree of sensitivity	Degree of importance
No. 1	M	H	M - H
No. 2	M	L	M - L
No. 3	M	M	M
No. 4	H	H	H
No. 5	L	H	M
No. 6	L	M	L – M
No. 7	H	M	H - M

Table 2 shows that both the uncertainty factors 4 and 7 are classified with high uncertainty. Of these two uncertainty factors, only factor 4 is classified with high importance as the mobilization times alone do not have a high enough potential to change the assessed interval.

The importance classification points to factors that should, if time and resources allow, be considered and prioritized for further assessments and follow-up.

The uncertainty analysis is qualitatively combined with the results of the prior assessments, including the subjective probability assessments for failure events and consequences in the detailed RoF assessments which are evaluated as a basis for communication to management. The evaluations highlight which uncertainties to give weight to in the presentation of the results.

3.2 Managerial review and judgement

The results presented in Table 2 provide additional decision support to that which is included in the standard RBI process. When these results are communicated to management, weight is also given to the limitations in the traditional RBI assessments. The managerial review and judgement reflects that decisions under uncertainty and risk need to be made, and it is a management task to weight these uncertainties and risk, and balance different concerns such as for example risks and costs.

Although management performs review and judgement late in the ERBI process, it does not exclude their involvements in earlier phases, which is often considered a key success factor in project management, so also for the implementation of the ERBI process.

4. DISCUSSION – COMPARISON OF THE TWO METHODOLOGIES

Applying the extended methodology will not necessarily result in different decisions compared to the risk based inspection (RBI) methodology. For the pipeline example, the RBI risk assessments and the first parts of the extended methodology (ERBI) led to a two year inspection plan for the steel pipeline. The importance of the factors identified may; however, change this, the conclusion being that the pipeline should be subject to a different inspection frequency - to prevent failure events.

Consider, for example, a specific segment of the pipeline in the example presented: a 12 meter middle section say, located about 4,500 meters from the riser base. Imagine that during testing this section was somehow neglected and a crack is present in the welding. Consider then the detailed risk assessment performed and assumption 4: ‘Data are able to describe the pipe material degradation’. As the type of pipe and welding in uncommon and limited data exist, there is high uncertainty related to this assumption. The sensitivity is also high, meaning that motivating a cautious policy - a more frequent inspection programme could be justified. By giving the assumptions attention, the decision-makers might see the need for further assessments based on alternative or revised assumptions.

In a project development, a number of assumptions are made and these need to be followed up in coming project phases. The ERBI provides a methodology for assessing the importance of the various assumptions and support the decision-making.

The managerial review and judgement allows for quality assurance and second opinions in the ERBI process. Also other risk methodologies, for example reliability centered maintenance (RCM), which is frequently used for assessment of preventive maintenance tasks and intervals for various equipment in the Norwegian oil & gas industry, could be included in the overall considerations in the managerial review and judgement.

Within the ERBI framework, as stated in Section 2.5, we apply subjective probabilities as a quantitative measure of uncertainty. These probabilities are used both in the equipment screening and the detailed risk assessments. In the RBI a relative frequency-based perspective for the probabilities are often used. The way probability and risk are understood strongly influences the presentation and communication of the results.

The additional assessments produce some increase in the time needed to perform the process, as well as the resources required. However, the extra time and costs due to the uncertainty factor assessments should not be very large compared to the overall costs used for RBI. By creating awareness to the relevance of the uncertainty factors, the process of identifying these could be efficiently integrated into the risk assessments. Nevertheless, if uncertainty and risk are to be adequately incorporated in the assessments, some extra resources are required.

5. CONCLUSIONS

RBI is a systematic analysis method for planning inspection intervals of static mechanical equipment used in the process industries. Risk of failures of the relevant system items is assessed in order to identify and determine suitable intervals. Over the years, RBI has gained reputation for being a successful method, but also for having some shortcomings. One of these is traced to the limited assessment of uncertainties.

In this paper we present and discuss the ERBI methodology: a methodology based on the existing RBI, which improves the risk and uncertainty assessments by adding some additional features to the existing RBI. A separate uncertainty assessment is added, to address uncertainties “hidden” in assumptions of the risk assessments. In the ERBI methodology the uncertainties are then communicated to management through an extended uncertainty evaluation, which integrates the results from the detailed risk analyses (and the cost-benefit analyses if such are performed) and the separate uncertainty analysis. An essential feature of the presented methodology and decision framework is the managerial review and judgement, which places the decision process into a broader management context. In this step consideration is given to the boundaries and limitations of the tools used.

An example from the oil & gas industry is presented to demonstrate the applicability of ERBI. The approach is, however, general and could also be used for other types of applications. We believe that by applying the methodology, an improved basis can be established for informing decision makers compared to the traditional RBI method, as the importance of risk and uncertainties is more adequately taken into account.

REFERENCES

1. Ablitt C & Speck J 2005. Experiences in implementing risk based inspection. Presented at the third (MENDT) Middle East Nondestructive Testing Conference & Exhibition, Manama, Bahrain.
2. Apeland S & Scarf PA 2003. A fully subjective approach to modelling inspection maintenance. *European Journal of Operational Research*, 148, 410-425.
3. API 2002. RP 580 – Risk Based Inspection. 1st ed. American Petroleum Institute.
4. API 2008. RP 581 – Risk Based Inspection Technology. 2nd ed. Am. Petroleum Institute.
5. Aven T & Renn O 2009. On risk defined as an event where the outcome is uncertain. *Journal of Risk Research*, 12, 1-11.
6. Aven T 2007. A unified framework for risk and vulnerability analysis and management covering both safety and security. *Reliability Engineering and System Safety*, 92, 745-754.
7. Aven T 2008. A semi-quantitative approach to risk analysis, an alterative to QRAs. *Reliability Engineering and System Safety*, 93(6), 790-797.
8. Aven T 2008. *Risk Analysis: Assessing Uncertainties Beyond Expected Values and Probabilities*, Chichester, UK: John Wiley & Sons, Ltd.
9. Aven T 2010. On how to define, understand and describe risk. *Reliability Engineering and System Safety*, 95, 623-631.
10. Bertolini M, Bevilacqua M, Ciarapica FE & Giacchetta G 2009. Development of risk-based inspection and maintenance procedures for an oil refinery. *Journal of Loss Prevention in the Process Industries*, 22, 244-253.
11. Bjørnoy OH, Marley MJ & Jephson, N 2001. Risk-based inspection principles and their application to corroded pipelines, *Pipes & Pipelines International*, 46 (5), 19-29.
12. Braggato PA, Pittiglio P & Ansaldi S 2008. The management of mechanical integrity inspections at small-sized “Seveso” facilities. *Reliability Engineering and System Safety*, 94, 412-417.
13. Castanier B & Rausand M 2006. Maintenance optimization for subsea pipelines. *International Journal of Pressure Vessels and Piping*, 83, 236-243.
14. Chang M-K, Chang R-R, Shu C-M & Lin K-N 2005. Application of risk based inspection in refinery & process piping. *Journal of Loss Prevention in the Process Industries*, 18, 397-402.
15. Chien C-H, Chen C-H & Chao YJ 2009. A strategy for the risk-based inspection of pressure safety valves, *Reliability Engineering and System Safety*, 94, 810-818.
16. DNV 2009. DNV-RP-G101: Risk Based Inspection of Offshore Topsides Static Mechanical Equipment. Det Norske Veritas (DNV).
17. Flage, R & Aven, T 2009. Expressing and communicating uncertainty in relation to quantitative risk analysis. *Reliability & Risk Analysis: Theory & Application*, 2 (13), 9-18.
18. Gassco 2010. Pipelines and platforms. < www.gassco.no/wps/wcm/connect/gassco-en/gassco/home/var-virksomhet/ror-og-plattformer>
19. Geary W 2002. Risk based inspection: A case study evaluation of onshore process plant, <www.hse.gov.uk/research/hsl_pdf/2002/hsl02-20.pdfs>

20. Hagemeijer PM & Kerkveld G 1998. A methodology for risk-based inspection of pressurized systems. *Journal of Process Mechanical Engineering*, 212 (1), 37-47.
21. Herzog BO & Jackson P 2009. The ladder to success: Risk-based inspection reduces costs, *ABB Review* [the corporate technical journal of the ABB group], 1, 30-33.
22. ISO 2006. 14224: Petroleum, petrochemical and natural gas industries - Collection and exchange of reliability and maintenance data for equipment. 2nd edition. International Organization for Standardization (ISO).
23. Jovanovic A 2003. Risk-based inspection and maintenance in power and process plants in Europe. *Nuclear Engineering and Design*, 226, 165-182.
24. Kallen MJ & van Noortwijk JM 2005. Optimal maintenance decisions under imperfect inspection. *Reliability Engineering and System Safety*, 90, 177-185.
25. Kallen M-J 2002. *Risk based inspection in the process and refining industry.* Thesis for the Master of Science degree at the Technical University of Delft, Delft, the Netherlands.
26. Khan FI & Haddara MM 2003. Risk-based maintenance (RBM): a quantitative approach for maintenance/ inspection scheduling and planning. *Journal of Loss Prevention in the Process Industries*, 16, 561-573.
27. Khan FI, Haddara MM & Bhattacharya SK 2006. Risk-based integrity and inspection modeling (RBIIM) of process components/ system. *Risk Analysis*, 26 (1), 203-221.
28. Khan FI, Sadiq R & Haddara MM 2004. Risk-based inspection and Maintenance (RBIM) multi-attribute decision-making with aggregative risk analysis. *Process Safety and Environmental Protection*, 82 (B6), 398-411.
29. Landet E, Lotsberg I & Sigurdsson G 2000. Risk-based inspection of an FPSO. OTC paper 12146, presented at the 2000 Offshore Technology Conference, Houston, Texas, USA.
30. Nilsson F 2003. Risk based approach to plant life management. *Nuclear Engineering and Design*, 221, 293-300.
31. NORSOK .2001. Z-008: Criticality analysis for maintenance purposes. Rev. 2. Norwegian Technology Centre.
32. Patel RJ 2005. Risk based inspection. Presented at the third MENDT Middle East Nondestructive Testing Conference & Exhibition, Manama, Bahrain.
33. Poulassichidis T 2009. Application of risk based inspection to offshore facilities. SPE paper 124539, presented at the 2009 Annual Technical Conference and Exhibition, New Orleans, Lousiana, USA.
34. Santosh, Gopika Vinod, Shrivastava OP, Saraf RK, Ghosh AK & Kushwaha HS 2006. Reliability analysis of pipelines carrying H2S for risk based inspection of heavy water plants. *Reliability Engineering and System Safety*, 91, 163-170.
35. Simpson J 2007. The application of risk based inspection to pressure vessels and aboveground storage tanks in petroleum fuel refineries. Presented at the fifth Australasian Congress on Applied Mechanics (ACAM 2007), Brisbane, Australia.
36. Singh M & Markeset T 2009. A methodology for risk-based inspection planning of oil and gas pipes based on fuzzy logic framework. *Engineering Failure Analysis*, 16, 2098-2113.
37. Taleb NT 2007. *Black Swan: The Impact of the Highly Improbable*. 1st. ed. New York: Random House.
38. Tien S-W, Hwang W-T & Tsai C-H 2007. Study of a risk-based piping inspection guideline system. *ISA Transactions*, 46, 119-126.
39. Truchon M, Rouhan A & Goyet J 2007. Risk based inspection approach for topside structural components, OTC paper 18912, presented at the 2007 Offshore Technology Conference, Houston, Texas, USA.
40. Wintle JB, Kenzie BW, Amphlett GJ & Smalley S 2001. Best practice for risk based inspection as part of plant integrity management. <www.hse.gov.uk>

COLREGS COMPLIANCE IN EVOLUTIONARY SETS OF COOPERATING SHIP TRAJECTORIES

R. Szlapczynski
University of Technology, Gdańsk, Poland
email: rafal@pg.gda.pl
●
J. Szlapczynska
Maritime University, Gdynia, Poland
email: asiasz@am.gdynia.pl

ABSTRACT

In general, Evolutionary Sets of Cooperating Ship Trajectories combine some of the assumptions of game theory with evolutionary programming and aim to find optimal set of cooperating trajectories of all ships involved in an encounter situation. In a two-ship encounter situation the method enables the operator of an on-board collision-avoidance system to predict the most probable behaviour of a target and to plan the own manoeuvres in advance. In a multi-ship encounter the method may be used to help an operator of a VTS system to coordinate the manoeuvres of all ships. The improvement presented here is a new way of modelling some of the COLREGS rules. Due to this change, the method is now able to find solutions, which are more compliant with COLREGS, more intuitive and consequently – safer from the navigator's point of view. The paper contains a detailed description of collision-avoidance operators used by the evolutionary method and simulation examples of the method's results for digital maps.

1. INTRODUCTION

The main approaches to the problem of planning optimal ship trajectories in encounter situations are based on either differential games or on evolutionary programming. The former method has been introduced by Lisowski (2005) and it assumes that the process of steering a ship in multi-ship encounter situations can be modelled as a differential game played by all ships involved, each having their strategies. Unfortunately, high computational complexity is its serious drawback. The latter approach is the evolutionary method of finding the trajectory of the own ship, proposed by Smierzchalski & Michalewicz (2000). Especially the second approach is recently very popular among researchers – it may be applied for finding an optimal path (Zeng, 2003) as well as an optimal collision avoidance manoeuvre (Tsou et al., 2010). In short, the evolutionary method uses genetic algorithms, which, for a given set of pre-determined input trajectories find a solution that is optimal according to a given fitness function. However, the method's limitation is that it assumes targets motion parameters not to change and if they do change, the own trajectory has to be recomputed.

Therefore, the authors have decided to try a new approach, which combines some of the advantages of both methods: the low computational time, supporting all domain models and handling stationary obstacles (all typical for evolutionary method), with taking into account the changes of motion parameters (changing strategies of the players involved in a game). Instead of finding the optimal own trajectory for the unchanged courses and speeds of targets, an optimal set of safe trajectories of all ships involved is searched for. The method is called evolutionary sets of safe trajectories and one of its earlier versions has been presented in (Szlapczynski, 2009).

One of the important issues of the method is applying to the International Regulations for Preventing Collisions at Sea (Cockroft & Lameijer 1993). The COLREGS rules, which are discussed here are:

- Rule 13 – overtaking: an overtaking vessel must keep well clear of the vessel being overtaken.
- Rule 14 - head-on situations: when two power-driven vessels are meeting head-on both must alter course to starboard so that they pass on the port side of the other.
- Rule 15 - crossing situations: when two power-driven vessels are crossing, the vessel, which has the other on the starboard side must give way.
- Rule 16 - the give-way vessel: the give-way vessel must take early and substantial action to keep well clear.
- Rule 17 - the stand-on vessel: the stand-on vessel may take action to avoid collision if it becomes clear that the give-way vessel is not taking appropriate action.

The main idea of the improvement, presented here is that COLREGS are modelled directly in the fitness function, instead of reflecting them indirectly on many other levels of the method. The rest of the paper is organized as follows. Section 2 describes the foundations of the collision avoidance method based on evolutionary sets of cooperating trajectories. Section 3 focuses on the details of the new approach to COLREGS, followed by same example results, which are shown in section 4. Finally, summary and conclusions are given in section 5.

2. EVOLUTIONARY SETS OF COOPERATING SHIP TRAJECTORIES

Evolutionary Sets of Cooperating Ship Trajectories (Szlapczynski 2009) is a name of a method solving multi-ship encounters. Foundations of the method are presented in the following subsections. The description includes definition of the optimization problem and some aspects of evolutionary engineering applied to the problem.

2.1. Optimisation problem

It is assumed that we are given the following data:

- stationary constraints (obstacles and other constraints modelled as polygons),
- positions, courses and speeds of all ships involved,
- ship domains,
- times necessary for accepting and executing the proposed manoeuvres.

Ship positions and ship motion parameters are provided by ARPA (Automatic Radar Plotting Aid) systems. A ship domain can be determined, based on the ship's length, its motion parameters and the type of water region. Since the shape of a domain is dependant on the type of water region, the author has decided to use a ship domain model by Davis (Davis et al. 1982) for open waters and to use a ship domain model by Coldwell (1982) for restricted waters. As for the last parameter – the necessary time, it is computed on the basis of navigational decision time and the ship's manoeuvring abilities. By default a 6-minute value is used here.

Knowing all the abovementioned parameters, the goal is to find a set of trajectories, which minimizes the average way loss spent on manoeuvring, while fulfilling the following conditions:

- none of the stationary constraints are violated,
- none of the ship domains are violated,
- the minimal acceptable course alteration is not lesser than 15 degrees,
- the maximal acceptable course alteration is not be larger than 60 degrees,

- speed alteration are not to be applied unless necessary (collision cannot be avoided by course alteration up to 60 degrees),
- a ship only manoeuvres, when it is obliged to,
- manoeuvres to starboard are favoured over manoeuvres to port board.

2.2. Evolutionary issues

The general idea of evolutionary programming is shown in Figure 1.

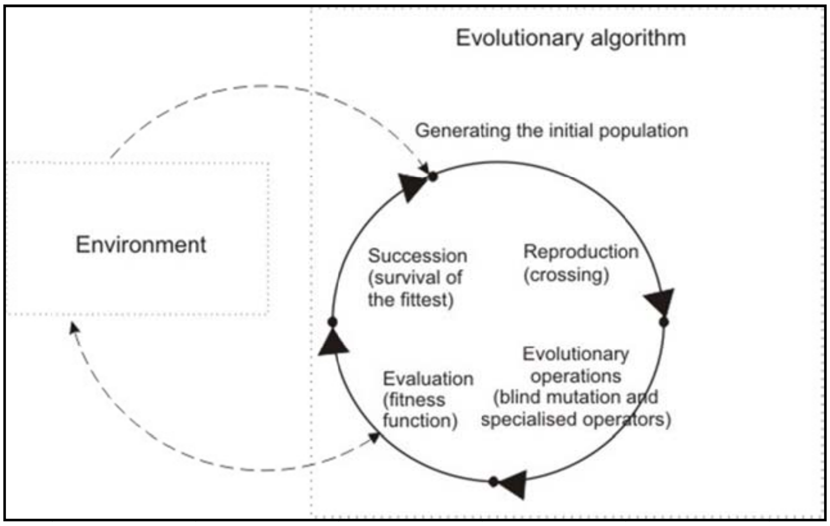


Figure 1. Evolutionary algorithm

First, the initial population of individuals (each being a potential solution to the problem) is generated either randomly or by other methods. Usually none of these individuals is optimal or even close to that. Sometimes none of the individuals is acceptable. This initial population is a subject to subsequent iterations of evolutionary algorithm. Each of these iterations consists of the following steps:

1. Reproduction: sets of parents (usually pairs) are selected from all of the individuals and they are crossed to produce offspring. The offspring inherits some features from each parent.
2. Evolutionary operations: the offspring is modified by means of random mutation operators as well as specialized operators dedicated to the problem.
3. Evaluation: each of the individuals (including parents and the offspring) is assigned a value of a fitness function, which reflects the quality of the solution represented by this individual.
4. Succession: the next generation of individuals is selected. The selection is based on the results of the evaluation. Usually the individuals are chosen randomly, with the probability strictly depending on the fitness function value.

The evolutionary algorithm ends when one of the following happens:

- maximum acceptable time or number of iterations is reached,
- the satisfactorily high value of fitness function has been reached by one of the individuals,
- further evolution brings no improvement.

3. COLREGS COMPLIANCE

3.1. Basic fitness function

The following basic fitness function is used to assess the quality of a solution:

$$fitness = \sum_{i=1}^{n} \left[tr\_fit_i \right], \tag{1}$$

where:

$$tr\_fit_i = \left(\frac{tr\_length_i - way\_loss_i}{tr\_length_i} \right) * sf_i * of_i , \tag{2}$$

sf_i - ship collision factor [/] of the *i*-th ship computed over all prioritised targets:

$$sf_i = \prod_{j=1, j \neq i}^{n} \left(min(fmin_{i,j} , 1) \right) \tag{3}$$

of_i - obstacle collision factor [/] of the *i*-th ship computed over all stationary constraints:

$$of_i = \left(\frac{trajectory\_length_i - trajectory\_cross\_length_i}{trajectory\_length_i} \right)^2 \tag{4}$$

n – the number of ships [/],
m – the number of stationary constraints [/],
i – the index of the current ship [/],
j – the index of a target ship [/],
k – the index of a stationary constraint [/],
$fmin_{i,j}$ – the approach factor value for an encounter of ships *i* and *j* [/],
$trajectory\_length_i$ – the total length of the i-th ship’s trajectory [nautical miles]
$trajectory\_cross\_length_i$ – the total length of the parts of the i-th ship’s trajectory, which violate stationary constraints [nautical miles]

This fitness function focuses on way loss and safe distances between ships, with COLREGS only being applied via ship domain models used to compute the approach factor value (Szlapczynski 2006b). The impact of ship domain model on COLREGS compliance is as follows. Domain shape affects the size of necessary course alteration manoeuvres to starboard and port board, thus affecting way loss and indirectly – fitness function values assigned to different trajectories. Therefore applying asymmetrical ship domain, whose port board area is larger than starboard area, favours manoeuvres to starboard over manoeuvres to port board. Also, larger bow area makes it less likely to cross ahead of stand-on targets. Apart from ship domains, two other means of reaching compliance with COLREGS have been applied:

1. Only collisions with prioritised ships were taken into account so as not to encourage unnecessary or unlawful manoeuvres from so-called “stand-on” vessels.
2. Manoeuvres to starboard were encouraged by a larger probability of course alteration to starboard than port board in mutation and specialised operators:
   - node shift,
   - node insert,

- segment shift
- segment insert in and mutation.

3.2 Penalties for breaking COLREGS

Once the basic fitness function has been computed according to the formulas from Section 3.1, the penalties are applied according to the following rules:

1. On open waters:
    a. if a ship is not obliged to give way, any manoeuvre it performs is penalized,
    b. if a ship is obliged to give way, and does not perform a manoeuvre it is penalized,
    c. all manoeuvres to port board are penalized.

2. On restricted waters: every trajectory node, which is a part of a manoeuvre, contains special information on the reason why this particular node has been inserted or shifted: land or other stationary obstacle avoidance, target avoidance or accidental manoeuvre generated by evolutionary mechanisms. Based on this penalties are applied as follows:
    a. if a ship does not initially have to give way to any target and its first manoeuvre has reason other than stationary obstacle avoidance, it is penalized,
    b. any manoeuvre to port board of reason other than stationary obstacle avoidance is penalized.

For normalized initial fitness function values, the penalties resulting from the unlawful manoeuvres have been set to 0.05. The penalties are additive that is a manoeuvre might be penalized twice. For example a manoeuvre to port board form a stand-on ship would be first penalized for performing any manoeuvre at all (rule 1a) and then, additionally for altering its course to port board (rule 1c).

4. RESULTS OF THE NEW APPROACH: SCENARIOS AND EXAMPLES

This section presents simulation results returned by a software application designed by the authors. The application implements evolutionary sets of cooperating ship trajectories including the abovementioned COLREGS compliance mechanisms. Following subsections present encounter examples on open and restricted waters for various ships configurations.

4.1 Open water basic scenario #*1*

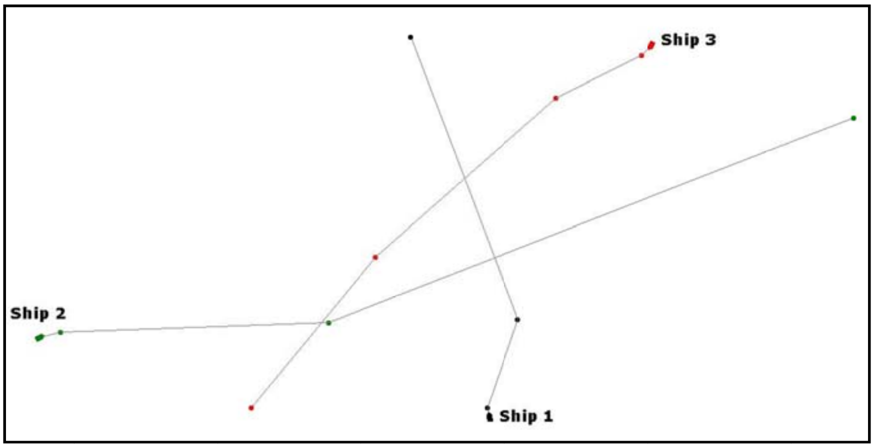


Figure 2. Open water basic scenario #1 – simulation starting screenshot

Table 1. Open water basic scenario #1 – ship positions & resulting fitness values

	Origin position	Destination position	V [kn]	Resulting trajectory fitness value [/]	Resulting general fitness value [/]
Ship 1	20° 34’ 10” E 58° 24’ 29” N	20° 31’ 37” E 58° 36’ 46” E	11.46	0.9595	0.9796
Ship 2	20° 19’ 34” E 58° 27’ 02” N	20° 45’ 58” E 58° 34’ 09” N	14.42	0.9842	
Ship 3	20° 39’ 26” E 58° 36’ 29” N	20° 26’ 27” E 58° 24’ 47” N	12.55	0.9726	

In the scenario presented in Figure 2 all three ships have similar situation of having one ship starboard and one port-board. Thus all these ships have to manoeuvre as follows: ship 1 gives way to ship 3, ship 3 gives way to ship 2 and ship 2 gives way to ship 1. The resulting trajectories assure that all the ships manoeuvre safely and there are no ahead crossings. Due to the specific positions and speeds (Table 1) ship 1 has the largest (the smallest fitness value) and ship 2 the smallest way loss (the largest fitness value).

4.2 Open water basic scenario *#2*

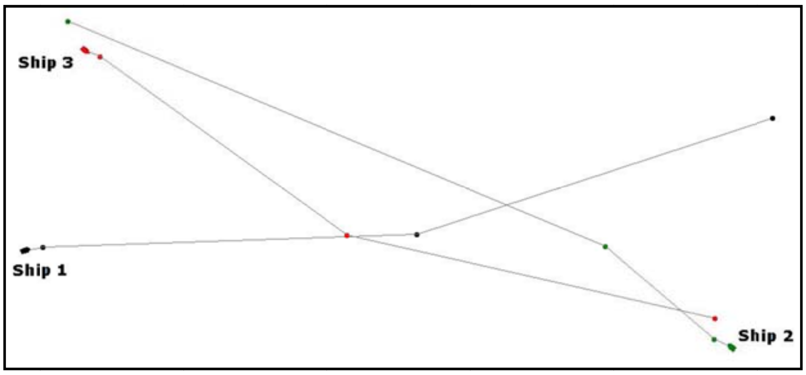


Figure 3. Open water basic scenario #2 – simulation starting screenshot

Table 2. Open water basic scenario #2 – ship positions & resulting fitness values

	Origin position	Destination position	V [kn]	Resulting trajectory fitness value [/]	Resulting general fitness value [/]
Ship 1	20° 20’ 45” E 58° 28’ 28” N	20° 44’ 57” E 58° 32’ 45” N	12.41	0.9806	0.9821
Ship 2	20° 43’ 35” E 58° 25’ 21” N	20° 22’ 08” E 58° 35’ 53” N	14.28	0.9856	
Ship 3	20° 22’ 41” E 58° 34’ 58” N	20° 43’ 05” E 58° 26’ 17” N	12.77	0.9591	

In the scenario presented in Figure 3 ship 2 & ship 3 have a head-on encounter while crossing with ship 1. Thus, ship 2 & ship 3 should alter their courses to starboard. Additionally ship 3 should give way to ship 1, while ship 1 should give way to ship 2. The resulting trajectories assure that all the restrictions are met and again there is no ahead crossing. In this situation (Table 2) ship 3 has to take a roundabout way resulting in the largest way loss (smallest fitness value).

4.3 Open water complex scenario

In the scenario presented in Figure 4 (with ship positions given in Table 3) there is a single ship (ship 1) crossing with two group of ships, namely:

- first group formed by ship 2, ship 3 and ship 4,
- second group formed by ship 5 and ship 6.

Ship 1 is a give-way vessel only to the first group of ships, thus it performs a substantial starboard course alteration to avoid ahead crossing. Ships 2, 3 & 4 are stand-on vessels (having no other vessels to their starboard) and due to that their courses remain unchanged until reaching their destination positions (maximum possible trajectory fitness value of 1.0). Unlike group 1, ships 5 & 6 from group 2 must give way to both ship 1 and group 1 ships. Due to mutual relation between origin and destination positions of ship 5 and ship 6 the former alters her course to port board, while the latter – to starboard. This way ship 5 reaches her destination safely bypassing ships 1, 2, 3 & 4 ahead with substantial distance to the ships. On the other hand, ship 6 avoids ahead crossing by her starboard maneuver. If both ship 5 & ship 6 changed courses to starboard, ship 6 would be forced to perform a larger alteration and the resulting way loss of the ships would be greater.

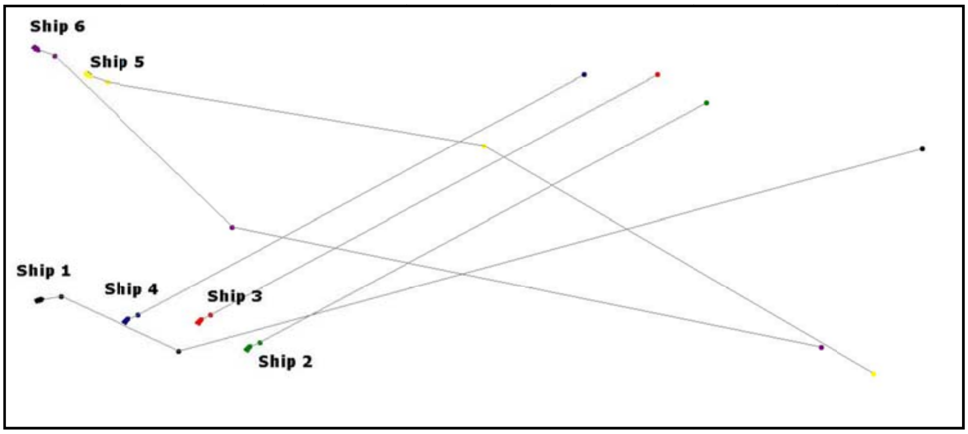


Figure 4. Open water complex scenario – simulation starting screenshot

Table 3. Open water complex scenario – ship positions & resulting fitness values

	Origin position	Destination position	V [kn]	Resulting trajectory fitness value [/]	Resulting general fitness value [/]
Ship 1	20° 18’ 29” E 58° 28’ 08” N	20° 47’ 17” E 58° 33’ 06” N	14.73	0.9275	0.9872
Ship 2	20° 25’ 17” E 58° 26’ 34” N	20° 40’ 14” E 58° 34’ 37” N	10.41	1.0000	
Ship 3	20° 23’ 42” E 58° 27’ 27” N	20° 38’ 39” E 58° 35’ 30” N	10.41	1.0000	
Ship 4	20° 21’ 20” E 58° 27’ 28” N	20° 36’ 16” E 58° 35’ 31” N	10.41	1.0000	
Ship 5	20° 20’ 04” E 58° 35’ 30” N	20° 45’ 41” E 58° 25’ 44” N	15.39	0.9575	
Ship 6	20° 18’ 21” E 58° 36’ 21” N	20° 43’ 59” E 58° 26’ 36” N	15.39	0.8984	

4.4 Restricted water basic scenario *#1*

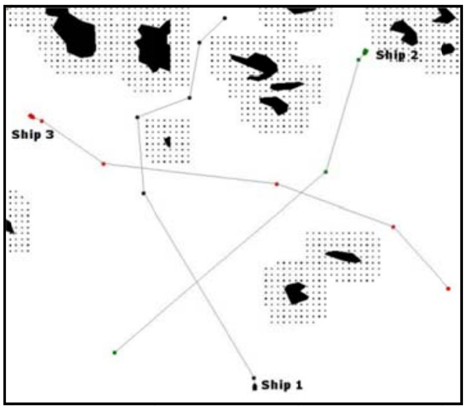


Figure 5. Restricted water basic scenario #1 – simulation starting screenshot (dotted areas depict non-approachable regions)

Table 4. Restricted water basic scenario #1– ship positions & resulting fitness values

	Origin position	Destination position	V [kn]	Resulting trajectory fitness value [/]	Resulting general fitness value [/]
Ship 1	21° 03' 33" E 60° 04' 35" N	21° 02' 18" E 60° 20' 05" N	14.39	0.9345	0.9481
Ship 2	21° 08' 11" E 60° 18' 39" N	20° 57' 40" E 60° 06' 02" N	12.78	0.9855	
Ship 3	20° 54' 10" E 60° 15' 57" N	21° 11' 41" E 60° 08' 44" N	10.82	0.9547	

In the scenario presented in Figure 5 (with ship positions given in Table 4) all the ships have one ship starboard and one port board, similar to open water scenario #1, but here ships also have to bypass obstacles (landmasses and areas limited by safety isobate). Ship 1 initially maneuvers to port board, securing safe bypassing of ship 2, ship 3 and obstacle being on her way. Later ship 1 has to change her course three more times to reach her destination hidden behind islands. Ship 2, although having ship 3 on her starboard requires only a small course alteration to port board to safely bypass the other ships. Possible collision threat between ship 2 and ship 3 is diminished also by initial starboard course change of ship 3, made originally due to obstacle bypassing.

4.5 Restricted water basic scenario *#2*

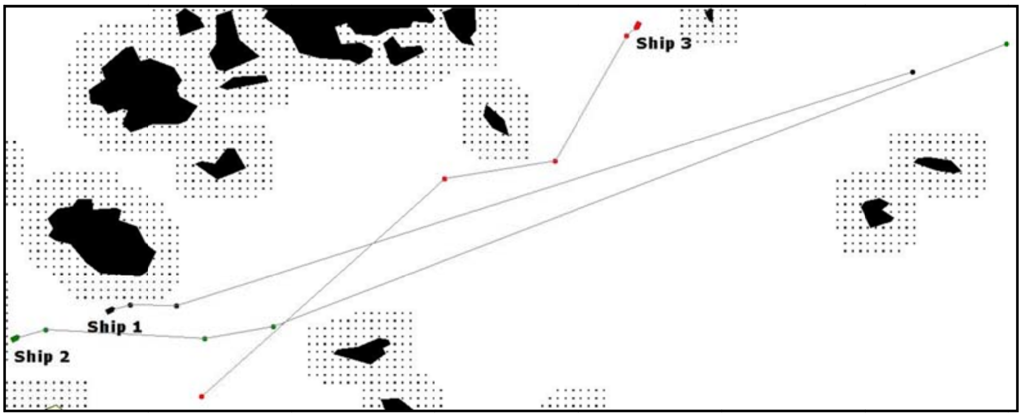


Figure 6. Restricted water basic scenario #2 – simulation starting screenshot (dotted areas depict non-approachable regions)

Table 5. Restricted water basic scenario #2 – ship positions & resulting fitness values

	Origin position	Destination position	V [kn]	Resulting trajectory fitness value [/]	Resulting general fitness value [/]
Ship 1	20° 40’ 34” E 60° 05’ 21” N	21° 06’ 28” E 60° 13’ 03” N	14.46	0.9930	0.9716
Ship 2	20° 37’ 32” E 60° 04’ 27” N	21° 09’ 30” E 60° 13’ 57” N	22.05	0.9735	
Ship 3	20° 57’ 36” E 60° 14’ 32” N	20° 43’ 33” E 60° 02’ 36” N	13.00	0.9587	

In the scenario presented in Figure 6 (with ship positions given in Table 5) a group of two ships (ship 1 & ship 2) crosses with ship 3, while in the group ship 1 is overtaken by ship 2. Ship 1 as the stand-on vessel in this case has to perform only a slight starboard alteration to avoid an obstacle and then keeps her course. Ship 2 as the overtaking vessel performs a substantial starboard alteration to safely bypass ship 1. Ship 3 must initially change her course to port board to avoid collision with an obstacle and then gets back to course towards her destination points, having ship 1 and ship 2 safely bypassed astern.

4.6 Restricted water complex scenario

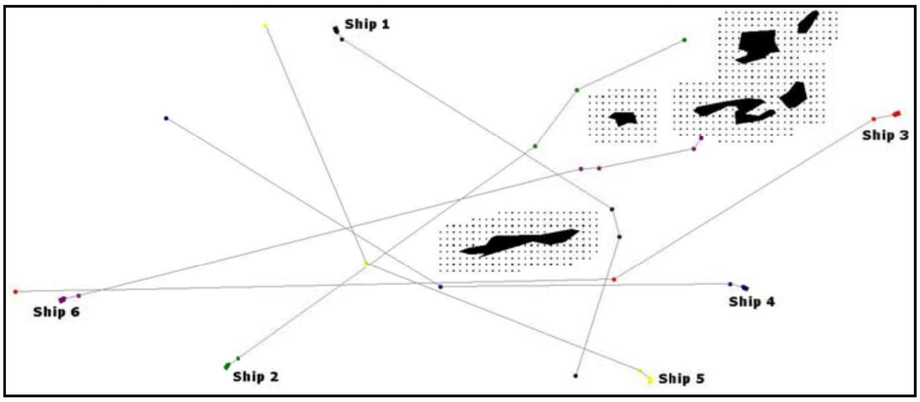


Figure 7. Restricted water complex scenario – simulation starting screenshot (dotted areas depict non-approachable regions)

Table 6.Restricted water complex scenario – ship positions & resulting fitness values

	Origin position	Destination position	V [kn]	Resulting trajectory fitness value [/]	Resulting general fitness value [/]
Ship 1	21° 29’ 58” E 59° 58’ 05” N	21° 39’ 13” E 59° 44’ 44” N	13.18	0.9137	0.9565
Ship 2	21° 25’ 45” E 59° 45’ 05” N	21° 43’ 24” E 59° 57’ 44” N	14.54	0.9909	
Ship 3	21° 51’ 33” E 59° 54’ 51” N	21° 17’ 38” E 59° 47’ 58” N	17.67	0.9139	
Ship 4	21° 45’ 43” E 59° 48’ 07” N	21° 23’ 26” E 59° 54’ 42” N	12.43	0.9004	
Ship 5	21° 42’ 05” E 59° 44’ 35” N	21° 27’ 15” E 59° 58’ 17” N	14.61	0.9374	
Ship 6	21° 19’ 24” E 59° 47’ 39” N	21° 44’ 04” E 59° 53’ 56” N	13.32	0.9893	

To facilitate analysis of a scenario presented in Figure 7 (with ship positions given in Table 6) let's divide the ships as follows:

1. ship 3, ship 4 & ship 5, forming group 1, heading westbound,
2. ship 2 and ship 6, forming group 2, heading eastbound,
3. ship 1 heading southbound.

All group 1 ships must bypass an obstacle and perform this action by port board maneuvers assuring safe astern crossings. In the group 2 alone there a slight crossing threat and ship 2 & ship 6 are forced to minor course amendments. However, still group 2 ships have impact on ship 1 and ship 5 maneuverings. Ship 1 is in the worst situation here: she has to bypass a large obstacle (the same as group 1 & 2 but larger north-southbound than west-eastbound), give way to group 2 ships and make sure her maneuvering won't disturb group 1. Successfully ship 1 makes her so by severe port board course change and astern bypassing trajectories of ship 3, ship 4 and ship 5.

5. SUMMARY AND CONCLUSIONS

The paper presents a newly designed and implemented improvement to the evolutionary sets of safe trajectories method. The method finds the optimal or near optimal set of safe ship trajectories for given positions and motion parameters of all ships involved in an encounter situation. The method is a generalization of evolutionary trajectory determining.

A set of trajectories of all ships involved, instead of just the own trajectory, is determined. The method avoids violating ship domains and stationary constraints, while obeying the COLREGS and minimizing total way loss computed over all trajectories. Because of its low computational time the method can be applied to on-board collision-avoidance systems and VTS systems. In the former, in case of simple scenarios (where ship priorities are clearly described by COLREGS), the method is able to predict the most probable manoeuvre of a target and plan own ship manoeuvre in advance, so that own manoeuvre could be initiated as soon as the target's manoeuvre is executed. In the latter, due to central planning, it could successfully solve any given scenario involving multiple ships and stationary constraints. The improvement, which the paper focuses on, is a set of rules that update fitness function values by penalizing unlawful manoeuvres. The solution has been tested and its better compliance with COLREGS has been confirmed by the experiments, whose examples are given in section 4.

The current version of the method is therefore able to plan trajectories not only of minor way loss spent on collision avoidance manoeuvres but also of full compliance with regulations and therefore – much safer. The further research on the method is planned and it will focus on VTS-specific issues and on planning ship trajectories on Traffic Separation Schemes with high ship density.

ACKNOWLEDGEMNTS

The author thanks the Polish Ministry of Science and Higher Education for funding this research under grant no. N N516 186737.

REFERENCES

1. Cockroft A.N., Lameijer J.N.F.(1993): *A Guide to Collision Avoidance Rules*, Butterworth-Heinemann Ltd..
2. Coldwell T.G. (1982): Marine Traffic Behaviour in restricted Waters, *The Journal of Navigation*, 36, 431-444..

3. Davis P.V., Dove M.J., Stockel C.T. (1982). A Computer Simulation of multi-Ship Encounters. *The Journal of Navigation*, 35, 347-352.
4. Lisowski J. (2005): Dynamic games methods in navigator decision support system for safety navigation, *Proceedings of the European Safety and Reliability Conference*, vol. 2, 1285 – 1292.
5. Szlapczynski R. (2006a). A new method of ship routing on raster grids, with turn penalties and collision avoidance, *The Journal of Navigation*, 59, 27-42.
6. Szlapczynski R.(2006b). *A* unified measure of collision risk derived from the concept of a ship domain, *The Journal of Navigation*, 59, 477-490.
7. Szlapczynski R. (2008): A new method of planning collision avoidance manoeuvres for multi target encounter situations, *The Journal of Navigation*, 61, 307-321.
8. Szlapczynski R. (2009). Solving multi-ship encounter situations by evolutionary sets of cooperating trajectories, *Marine Navigation and Safety of Sea Transportation*, CRC Press / Taylor & Francis Group / Balkema, 437-442.
9. Smierzchalski, R., Michalewicz, Z. (2000). Modeling of ship trajectory in collision situations by an evolutionary algorithm. *IEEE Transactions On Evolutionary Computation*, 4, 227–241.
10. Tsou M.-C., Kao, S.-L., Su C.-M. (2010). Decision Support from Genetic Algorithms for Ship Collision Avoidance Route Planning and Alerts, *The Journal of Navigation*, 63, 167–182.
11. Zeng X. (2003). Evolution of the safe path for ship navigation. *Applied Artificial Intelligence*, 17, 87–104.

FUZZY FAILURE RISK ANALYSIS IN DRINKING WATER TECHNICAL SYSTEM

B. Tchórzewska-Cieślak
Rzeszow University of Technology, Rzeszow, Poland
•
e-mail: cbarbara@prz.rzeszow.pl

ABSTRACT

Drinking water technical system is an essential element of urban infrastructure. The operation of this system is inseparably connected with a risk of failure. The main problem in the risk of failure analysis of water pipe network is the uncertainty of the information collected on the description of failure. In order to consider the uncertainty of information, the theory of fuzzy sets was used. The fuzzification of frequency, severity and the consequences of the incident scenario is basic input for fuzzy risk analysis. The presented model is part of a complex model of risk management of failures in drinking water technical system manly in water pipe network and can be used in practice in system operator's decision-making process. An adaptation of the fuzzy set theory to analyse risk of failure of water mains is not a standard approach for water works. An effect of the analysis of different sources of risk can be used for the design of a more reliable safety system assurance.

1 INTRODUCTION

A drinking water technical system (DWS) belongs to the so called critical infrastructure of cities. That is why it should be a priority task for waterworks and even for the local authorities to ensure the suitable level of its safety. Its aim is to supply consumers with a required amount of water, with a specific pressure and a specific quality, according to binding standards. Disturbances in its operation may be the result of random events, the forces of nature, material defects, mechanical damage or deliberate actions of third parties, including terrorist activities. At the same time this system plays a priority role in the functioning of urban areas, even a short lack of suitable drinking water always causes anxiety or panic among people. In many countries in the world, including the EU countries, the regulations obligate waterworks to prepare accurate water supply management procedures, including an analysis of risk of failure. Failure in water network is one of the most common failure of DWS(Kleiner 2004, Sadig et.al 2007, Tchorzewska-Cieślak 2007).

Modelling the risk of failure in water network consists of three main tasks:

- assessment (estimation) of the frequency/ probability of emergency scenarios (undesirable events),
- assessment (estimation) of various consequences of emergency scenarios (undesirable events),
- estimation of water mains protection level and the various types of protection minimizing the possible consequences of emergency scenarios (undesirable events).

The case that occurs most frequently in the risk analysis is a statistical uncertainty caused by the random nature of the studied phenomenon, the influence of external factors, as well as the time factor that determines a change of analysed undesirable event (failure) (Tchorzewkusa-Cieślak 2009) .In many cases, data on failures of water pipe network are obtained from experts (water supply system operators, engineers or researchers).

These data are often imprecise and incomplete. The following data, among others, are necessary to perform risk analysis in the DWS (Ezell et al. 2000, Hubbard 2009, Tchorzewska-Cieślak 2010) :

- data identifying the analysed object (e.g. water treatment station, distribution pipeline), the name and type of the object and its basic technical data,
- data about failures (undesirable events), repairs and other breaks in the DWS's operation (information about the date, time and duration of failure, and a description of the failure),
- data relating to the reasons behind the occurrence of undesirable events,
- data relating to the consequences of these events.

The main aim of this study is to present a risk analysis model using fuzzy set theory and the application of this theory in the risk management process in water network.

2. The risk of failure of water pipe network

Risk assessment includes the so called risk analysis, which is the process aimed at the determination of the consequences of failures (undesirable events) in the DWS, their extend, sources of their occurrence and the assessment of the risk levels. Haimes (Haimes 1998, Heimes 2009) suggests that risk assessment concerns its reasons, as well as its likelihood and consequences.

Drinking water infrastructure system uncertainty or risk is defined as the likelihood or probability that the drinking water service fails to provide water on-demand to its customers (Kleiner et.al 2006, Sadig et. al. 2009).

For purposes of this paper, operational reliability of the DWS is defined as the ability to supply a constant flow of water for various groups of consumers, with a specific quality and a specific pressure, according to consumer demands, in specific operational conditions, at any or at a specific time.

Failure is defined as the event in which the system fails to function with respect to its desired objectives.

Safety of the DWS means the ability of the system to safely execute its functions in a given environment. The measure of DWS safety is risk.

Risk (r) is a function of the parameters: the probability or frequency (f_i) that representative emergency scenario occurs (S), the magnitude of losses (C_j) caused by RES and the degree of sensitivity (E_k) to RES, according to equation (1) (Rak 2009,Rak et al 2006).

$$r = \sum_{S=1}^{N} (P_i \quad C_j \quad E_k) \tag{1}$$

where:
S- a series of the successive undesirable events (failures),
P_i - a point value depending on the frequency of RES or a single failure,
i- a number of the scale for the frequency,
C_j--a point value of losses caused by RES or a single failure,
j- a number of the scale for the losses,
E_k–a point value for the parameter of exposure (sensitivity of water mains) associated with RES or a single failure,
k- a number of the scale for the sensitivity,
N-number of *RES*.

To analyse risk defined in this way the matrix methods can be used (Markowski et.al 2008, Rak et al 2006). According to equation (1) the qualitative risk matrix was developed, assuming a descriptive point scale for the particular risk parameters (Tchorzewska-Cieślak 2010). Depending on the frequency of a given failure the point weights for the parameter *f* are presented in Table 1.

Table1. Criteria for a descriptive point scale for the parameter $P_i.(i=1,2,3,4,5)$

P_i	Probability of failure
P_1	very low probability once in 10 years and less often
P_2	low probability, once in (10÷5) years
P_3	medium probability, once in (5÷1)years
P_4	probability, once in (1÷0.5) year
P_5	once a month and more often,

The criteria and the point weights for the assumed descriptive point scale for the parameter of losses C_j and sensitivity E_k are presented in Tables 2 and 3.

Table 2. Criteria for a descriptive point scale for the parameter C_j, (j=1,2,3)

C_j	Description
$C_1=1$	small losses : • perceptible organoleptic changes in water, • isolated consumer complaints, • financial losses up to $5 \cdot 10^3$EUR
$C_2=2$	medium losses: • considerable organoleptic problems (odour, changed colour and turbidity), • consumer health problems, numerous complaints, • information in local public media, financial losses up to 10^5 EUR
$C_3=3$	large losses: • endangered people require hospitalisation, • professional rescue teams involved, serious toxic effects in test organisms, • information in nationwide media, financial losses over 10^5 EUR

Table 3. Criteria for a descriptive point scale for the parameter E_k, (k={1,2,3})

E_k	Description
$E_1=1$	small sensitivity to failure (high resistance): • the network in the ring system, • the ability to cut off the damaged section of the network by means of gates (for repair), • the ability to avoid interruptions in water supply to customers, • full monitoring of water mains, continuous measurements of pressure and flow rate at strategic points of the network covering the entire area of water supply, utilising SCADA and GIS software, the possibility to remote control network hydraulic parameters
$E_2=2$	medium sensitivity to failure: • the network in the radial or mixed system, • the possibility to cut off the damaged section of the network by means of gates, but the network capacity limits water supply to customers, • water mains standard monitoring, measurements of pressure and flow rate
$E_3=3$	large sensitivity to failure (low resistance): • the network in the radial system, • the inability to cut off the damaged section of the network by means of gates (for repair) without interrupting water supply to customers, • limited water mains monitoring

The use of fuzzy set theory in the analysis of risk of water mains failure

The notion of fuzzy sets was introduced in 1965 by L.A Zadeh of the University of Berkeley. Unlike in the classical set theory, the limit of the fuzzy set is not precisely determined, but there is a gradual transition from non-membership of elements in a set, through their partial membership, to membership (Dubois et.al 1980, Kluska 2009, lee 1999). This gradual transition is described by the so called membership function $\mu_{A,}$ where A is a set of fuzzy numbers. Risk analysis is based largely on expert opinions and uses such linguistic terms as small losses, high risk and can be described by means of fuzzy sets (Braglia et.al 2003a, Karwowski 1986,Kleiner 2004,markowski 2008, Sadig 2009, Shang-Lien 2002, Tchórzewska-Cieślak 2010). For risk analysis of water mains failure the membership function class type *t* (a triangular function) according to equation (2), the membership function class type γ according to equation (3) and the membership function class type *L* according to equation (4), were proposed.

$$\mu_A(x,a,b,c)=\begin{cases} 0 & for \quad x \le a \\ \dfrac{x-a}{b-a} & for \ a \le x \le b \\ \dfrac{c-x}{c-b} & for \ b \le x \le c \\ 0 & for \quad x \ge c \end{cases} \tag{2}$$

$$\mu_A(x,a,b)=\begin{cases} 0 & for \quad x \le a \\ \dfrac{x-a}{b-a} & for \ a \le x \le b \\ 1 & for \quad x > b \end{cases} \tag{3}$$

$$\mu_A(x,a,b)=\begin{cases} 1 & for \quad x \le a \\ \dfrac{b-x}{b-a} & for \ a < x \le b \\ 0 & for \quad x > b \end{cases} \tag{4}$$

where:
x- variable, parameter value,
μ_A- the membership function of variable x in the fuzzy set A ,
a, b, c -the membership function parameters (minimal, median (central) and maximum value of fuzzy number),

For the probability parameter the set of possible linguistic characterization is defined as:

$\overline{P}=\{P_i\}$, i={1,2,3,4,5}.

Table 4 shows the linguistic characterization, type and parameters of membership function.

Table 4. The linguistic characterization, type and parameters of membership function, for P parameter, $\overline{P}=\{P_i\}$, i={1,2,3,4,5}

Fuzzy set	linguistic characterization	type of membership function	membership function parameters		
			a	b	c
P_1	very low probability	type *L*, *acc.to eq.(4)*	0.125	0.25	-
P_2	low probability	triangular *t, acc.to eq.(2)*	0.125	0.25	0.375
P_3	medium probability	triangular *t, acc.to eq.(2)*	0.25	0.375	0.5
P_4	probability	triangular *t, acc.to eq.(2)*	0.375	0.5	0.625
P_5	high probability	type γ, *acc.to eq.(3)*	0.5	0.625	-

Figure 1 shows forms of membership function for the *P* parameter.

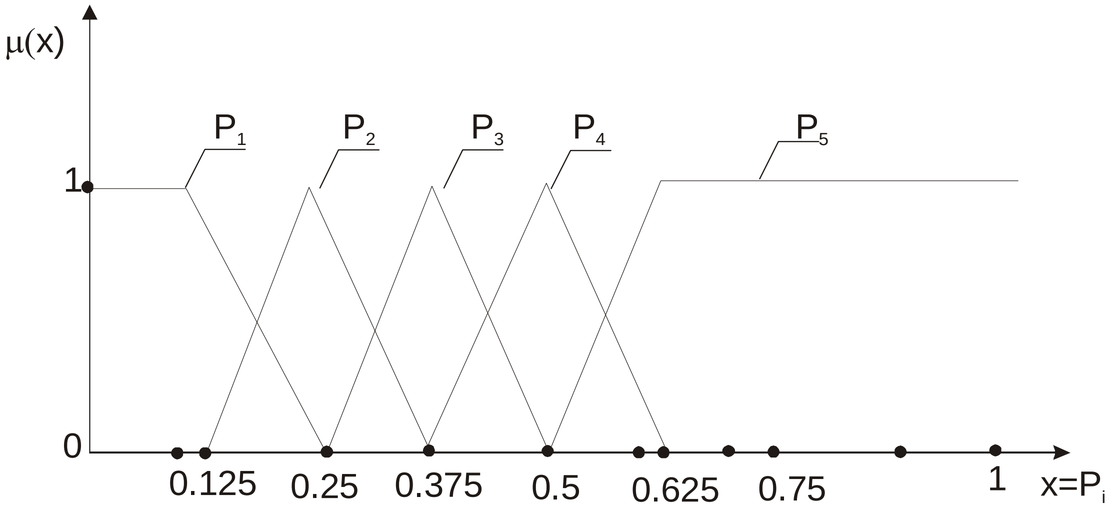


Figure1. Form of membership function for the parameter P.

For the losses parameter the set of possible linguistic characterization is defined as:

$$\overline{C}=\{C_j\},\ j=\{1,2,3\}.$$

Table 5 shows the linguistic characterization, type and parameters of membership function for C parameter.

Table 5. The linguistic characterization, type and parameters of membership function, for C parameter, $\overline{C}=\{C_1,C_2,C_3,\}$

Fuzzy set	linguistic characterization	type of membership function	membership function parameters		
			a	b	c
C_1	small	triangular *t, acc. to eq.(2)*	0.0	0.0	1.5
C_2	medium	triangular *t, acc. to eq.(2)*	0.5	1.5	2.5
C_3	large	triangular *t, acc. to eq.(2)*	1.5	3.0	3.0

For the sensitivity parameter the set of possible linguistic characterization is defined as: $\overline{E} = \{E_j\}$, $j=\{1,2,3\}$.

Table 6 shows the possible linguistic characterization for the sensitivity parameter E, type and parameters of membership function.

Table 6. The linguistic characterization, type and parameters of membership function, for E parameter, $\overline{E} = \{E_1, E_2, E_3,\}$

Fuzzy set	linguistic characterization	type of membership function	membership function parameters		
			a	b	c
E_1	small	triangular *t, acc. to eq.(2))*	0.0	0.0	1.5
E_2	medium	triangular *t, acc. to eq.(2)*	0.5	1.5	2.5
E_3	large	triangular *t, acc. to eq.(2)*	1.5	3.0	3.0

Figure 2 shows forms of membership function for the parameters *C* and *E*.

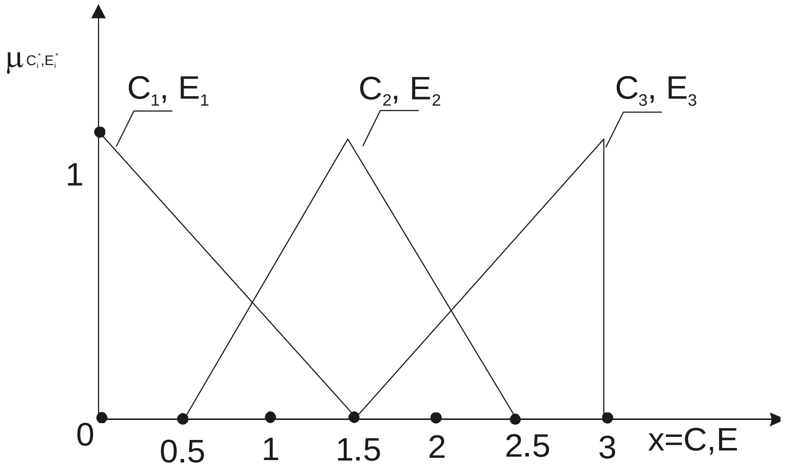


Figure 2. Form of membership function for the parameters E and C.

For risk the set of possible linguistic characterization is defined as: $\overline{R} = \{R_l\}$, $l=\{1,2,3\}$.

Table 7 shows the linguistic characterization of risk, type and parameters of membership function.

Table 7. The linguistic characterization, type and parameters of membership function, for risk, $\overline{R} = \{R_1, R_2, R_3\}$

Fuzzy set	linguistic characterization	type of membership function	membership function parameters		
			a	b	c
R_1	tolerable risk (TR)	triangular *t, acc eq.(2)*	0.0	0.0	18
R_2	controlled risk (CR)	triangular *t, acc eq.(2)*	5	22.5	40.5
R_3	unacceptable risk(Ur)	triangular *t, acc eq.(2)*	27	45	45

4. THE DECISION MODEL

Decision-making tools help in the selection of prudent, technically feasible, and scientifically justifiable actions to protect the environment and human health in a cost-effective way. Processing of information obtained from experts is carried out by the following steps: fuzzification, inference with the use of the rules, and, in case of Mamdani-type inference (Mamdani 1977), defuzification, to get the result in a discrete form. For each linguistic variable a set of membership functions that correspond to the values of this variable is defined.

The Mamdani – Zadeh type decision model was proposed:

- The input base of the proposed model consists of three values of risk parameters:
  - the probability of failure $x_1(P_i)$: five possible fuzzy sets: P={P_1, P_2 ,P_3, P_4, P_5},
  - losses associated with the occurrence of failure $x_2(C_j)$, three possible fuzzy sets: C={C_1, C_2 ,C_3}
  - and a degree of exposure (resistance) to failure $x_3(E_k)$, three possible fuzzy sets E={E_1, E_2 ,E_3}
- The output of the model, which allows making an operational decision, is the index risk value for water mains failure $y\ (R_l)$, three possible fuzzy sets: $R=\{R_1,R_2,R_3\}$.
- The fuzzification, which converts a vector of numbers (the crisp input values of risk parameters) into a vector of degrees of membership (a singleton method was used).
- The inference – the determination of a fuzzy conclusion model in form of the resulting membership function. In this block all rules whose premises are satisfied, are activated. At this moment the following steps are used:
  - determination of the so called rule base, which provides a knowledge base for qualitative knowledge and consists in determining the relationship between the particular parameters of the model. A rule base determines the relationships between the inputs and outputs of a system using linguistic *antecedent* and *consequent* propositions in a set of IF-THEN rules, where if – premise, then- conclusion.
  - the rule base of a complex system usually requires a large number of rules to describe the behaviour of a system for all possible values of the input variables. The base of rules contains the logical rules which determine cause and effect relationship between the particular risk parameters in water mains. Based on the risk matrix shown in Table 4, the base of rules was determined. It is a set of rules: $R_M = \{R_{M1}, R_{M2}, \dots R_{M45}\}$, in a general form:

    If probability is P_i and possible consequences are C_j and sensitivity is E_k then risk is R_l
  - Aggregation of rules - in the process of aggregation a degree of fulfilment of each rule is calculated based on the degree of fulfilment of its premises. For this purpose, the fuzzy logic operations: (*and, or*), are used. Based on the presented base of rules, the inference min-max, using the operator S-norms and T-norms, was proposed (Dubois et al. 1980, Mamdani 1977). The aggregating output membership function of a resultant output fuzzy risk category is expressed as:

$$\mu_R(R_l) = \max_m\{\min \mu_F^m(f_i), \mu_C^m(C_j), \mu_E^m(E_k), \mu_R^m(r_l)\ )$$

where *m* is the number of rules, *i* the number of fuzzy frequency sets, *j* the number of loss parameter sets, *k* the number of sensitivity and *l* is the number of fuzzy risk sets.

- The defuzzification, whose aim is to obtain a specific value of risk.

This process is the final stage of the model and provides the basis for the water supply system operator's decision-making process. For example, if the risk value corresponds to the category of unacceptable risk an operator undertakes some measures to reduce risk of failure (water mains modernization). The transformation of fuzzy set into not fuzzy value (determined) can be performed by various methods. For the proposed model the centre of gravity method was used:

$$R = \frac{\sum_{l=1}^{3} R_l \quad \mu_l(r_l)}{\sum_{l=1}^{3} \mu_l(r_l)} \tag{5}$$

Using the operating data of water supply system in a city with population of 200 thousand, the risk analysis, using the fuzzy software, was performed. Data on water main failures for five years of water supply network operation were collected and analysed in terms of frequency of failures and their consequences.

Risk assessment was performed for three diameter ranges:

- up to ϕ150mm,
- ϕ(150-400)mm,
- ϕ> 400 mm.

The result of the analysis for defined membership functions according to tables 1,2,3 and 45 rules of interference is presented in figure 3(using Matlab fuzzy toolbox). Defined rules are in the form:

1. If (P is P1) and (C is C1) and (E is E1) then (R is R1) (1)
2. If (P is P1) and (C is C1) and (E is E2) then (R is R1) (1)
3. If (P is P1) and (C is C1) and (E is E3) then (R is R2) (1)
4. If (P is P1) and (C is C2) and (E is E1) then (R is R1) (1)
5. If (P is P1) and (C is C2) and (E is E2) then (R is R2) (1)
6. If (P is P1) and (C is C3) and (E is E3) then (R is R2) (1)
7. If (P is P1) and (C is C3) and (E is E1) then (R is R1) (1)
8. If (P is P1) and (C is C3) and (E is E2) then (R is R2) (1)
9. If (P is P1) and (C is C3) and (E is E3) then (R is R2) (1)
10. If (P is P2) and (C is C1) and (E is E1) then (R is R1) (1)
11. If (P is P2) and (C is C1) and (E is E2) then (R is R2) (1)
12. If (P is P2) and (C is C1) and (E is E3) then (R is R2) (1)
13. If (P is P2) and (C is C2) and (E is E1) then (R is R1) (1)
14. If (P is P2) and (C is C2) and (E is E2) then (R is R2) (1)
15. If (P is P2) and (C is C2) and (E is E3) then (R is R2) (1)
16. If (P is P2) and (C is C3) and (E is E1) then (R is R2) (1)
17. If (P is P2) and (C is C3) and (E is E2) then (R is R3) (1)
18. If (P is P2) and (C is C3) and (E is E3) then (R is R2) (1)
19. If (P is P3) and (C is C1) and (E is E1) then (R is R1) (1)
20. If (P is P3) and (C is C1) and (E is E2) then (R is R2) (1)
21. If (P is P3) and (C is C1) and (E is E3) then (R is R2) (1)
22. If (P is P3) and (C is C2) and (E is E1) then (R is R2) (1)
23. If (P is P3) and (C is C2) and (E is E2) then (R is R2) (1)
24. If (P is P3) and (C is C2) and (E is E3) then (R is R2) (1)
25. If (P is P3) and (C is C3) and (E is E1) then (R is R2) (1)
26. If (P is P3) and (C is C3) and (E is E2) then (R is R2) (1)
27. If (P is P3) and (C is C3) and (E is E3) then (R is R3) (1)
28. If (P is P4) and (C is C1) and (E is E1) then (R is R2) (1)
29. If (P is P4) and (C is C1) and (E is E2) then (R is R2) (1)
30. If (P is P4) and (C is C1) and (E is E3) then (R is R2) (1)
31. If (P is P4) and (C is C2) and (E is E1) then (R is R2) (1)
32. If (P is P4) and (C is C2) and (E is E2) then (R is R2) (1)
33. If (P is P4) and (C is C2) and (E is E3) then (R is R3) (1)
34. If (P is P4) and (C is C3) and (E is E1) then (R is R2) (1)
35. If (P is P4) and (C is C3) and (E is E2) then (R is R2) (1)
36. If (P is P4) and (C is C3) and (E is E3) then (R is R3) (1)
37. If (P is P5) and (C is C1) and (E is E1) then (R is R2) (1)
38. If (P is P5) and (C is C1) and (E is E2) then (R is R2) (1)
39. If (P is P5) and (C is C1) and (E is E3) then (R is R2) (1)
40. If (P is P5) and (C is C2) and (E is E1) then (R is R2) (1)

41. If (P is P5) and (C is C2) and (E is E2) then (R is R2) (1)
42. If (P is P5) and (C is C2) and (E is E3) then (R is R3) (1)
43. If (P is P5) and (C is C3) and (E is E1) then (R is R2) (1)
44. If (P is P5) and (C is C3) and (E is E2) then (R is R3) (1)
45. If (P is P5) and (C is C3) and (E is E3) then (R is R3) (1)

In Table 8 the results of the analysis are presented.

Table 8. Risk assessment for the analysed water mains.

ϕ [mm]	P	C	E	Risk value	Risk category
up to 150	0.223	1.97	2.0	24.4	CR
150-400	0.173	2.08	2.15	23.8	CR
≥ 400	0.114	2.66	1.07	20.4	TR

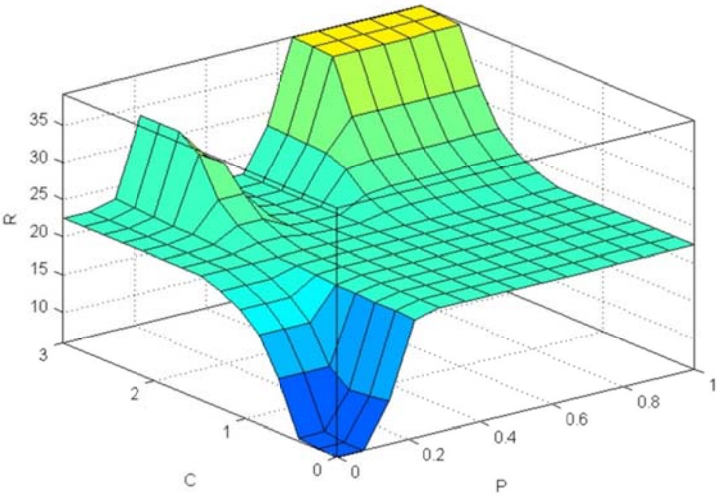


Figure 3. The fuzzy risk graph

6. CONCLUSIONS

- Technical systems for drinking water supply require specific procedures for supervision and monitoring. Safety analyses of such systems are based on risk analyses and assessments primarily for the health safety of urban residents.
- Now, however, new trends in the technical management of water supply systems are heavily focused on security issues related to the design, operation and management of these systems. Of particular interest in the future will be the evaluation of risk and reliability issues of the various components, subsystems and the systems as a whole, from the viewpoint of their susceptibility to terrorism.
- Effective risk management requires operators to monitor actively the entire water supply system, to develop emergency plans, to response fast in emergency, and to be able to analyse and assess risk.
- One of the methods to deal with uncertainties in risk assessment is a fuzzy logic where fuzzy sets are a fundamental issue.

- In the study the application of fuzzy logic theory to analyse risk of failure of DWS was proposed. In case of having inaccurate or various (eg from different experts) data on particular risk parameters, there is the possibility to describe them by a linguistic variables.
- In contrast to the traditional risk analysis, all variables of the risk parameters (according to equation (1)) are expressed in fuzzy sets defined by appropriate membership functions
- The probability or frequency of failures and their possible consequences can be defined as fuzzy values, particularly when they are estimated and not precisely determined, which often occurs at the analysis of failures in water supply network.
- The decision model presented in the study, based on assumptions of Mamdani's fuzzy modelling, may be used in practice in water mains as an element of a complex management of risk of failures of water mains.
- A certain limitation of the proposed method is the need to develop a database of the rules, based on the knowledge of experts, whose opinions on the assumed criteria may differ from each other.
- In order to develop the complete and most reliable database of the rules (the knowledge base), as much information as possible about failures of water mains, their possible consequences and causes, should be collected.
- Proposed method provides an alternative to other methods for assessing and managing risk of water network failure (subjective probability theory, mathematical theory of records) and its use is justified if you have a subjective assessment of risk parameters.

REFERENCES

1. Braglia, M. & Frosolini, M., & Montanari, R. (2003a). Fuzzy criticality assessment model for failure modes and effects analysis. *International Journal of Quality & Reliability Management*, 20(4), 503–524
2. Dubois, D., & Prade, H. (1980). *Fuzzy sets and systems: Theory and application*. New York: Academic Press.
3. Ezell B. & Farr J. & Wiese I.(2000) Infrastructure risk analysis of municipal water distribution system. *Journal of Infrastructure Systems*, ASCE. 6(3) ,118–122.
4. Haimes, Y. Y., (1998) *Risk Modeling, Assessment and Management*. Wiley, New York.
5. Haimes. Y.Y. (2009). On the Complex definition of risk: a systems-based approach. *Risk Analisys*.29(12), 1647-1654.
6. Hubbard D.W. (2009). *The failure of risk management*. Wiley. New York.2009
7. Karwowski W, Mital A (1986). Potential applications of fuzzy sets in industrial safety engineering, *Fuzzy Sets System*. 19, 105–120.
8. Kleiner, Y. & Rajani, B.B. & Sadiq, R. (2006). Failure risk management of buried infrastructure using fuzzy-based techniques, *Journal of Water Supply Research and Technology: Aqua*, 55 (2), 81-94.
9. Kleiner, Y. (2004). A fuzzy based method of soil corrosivity evaluation for predicting water main deterioration. *Journal of Infrastructure Systems, ASCE.*10(4),149-156
10. Kluska J. 2009. *Analytical Methods in Fuzzy Modelling and Control*. Springer-Verlag GMbH.
11. Lee, H.-M. (1996), Applying fuzzy set theory to evaluate the rate of aggregative risk in software development. *Fuzzy Sets and Systems*, 79, 323-336
12. Mamdani, E.H. 1977. Application of fuzzy logic to approximate reasoning using linguistic systems, *Fuzzy Sets and Systems*, 26,1182-1191
13. Markowski A., Mannan S. (2008). Fuzzy risk matrix. *Journal of Hazardous Materials*, *59(1,)152-156.*
14. Rak J.(2009). Selected problems of water supply safety. *Environmental Protection Engineering.* 35,29-35

15. Rak J. & Tchórzewska-Cieślak B.(2006) Review of matrix methods for risk assessment in water supply system. *Journal of Konbin*, 1(1), 67-76.
16. Sadig R. & Kleiner Y., & Rajani B.(2007) Water quality failures in distribution networks-risk analysis using fuzzy logic and evidential reasoning. *Risk Analysis*.27(5),1381-1394
17. Sadiq, R., & Tesfamariam, S. (2009). Environmental decision-making under uncertainty using intuitionist fuzzy analytic hierarchy process (IF-AHP), *Stochastic Environmental Research and Risk Assessment*, 23(1),75-91
18. Shang-Lien L. & Ruei-Shan L. (2002). Diagnosing reservoir water quality using self-organizing maps and fuzzy theory. *Water Reserch*.36 ,265-2274.
19. Tchórzewska-Cieślak B.(2007) Method of assessing of risk of failure in water supply system. European safety and reliability conference ESREL. *Risk, reliability and societal safety*. Taylor & Francis, 2.1535-1539, Norway, Stavanger
20. Tchórzewska-Cieślak B. (2010). Model of risk of water mains failure using fuzzy logic. *Journal of Polish Safety and Reliability Association.(Summer Safety &Reliability Seminars). Polish Safety and Reliability Association*. vol 1. 255-265.
21. Tchórzewska-Cieślak B, (2009) Water supply system reliability management. *Environmental Protection Engineering*. 35 ,29-35

ISSN 1932-2321