Atom Indonesia

Journal homepage: http://aij.batan.go.id

A Bayesian Network Approach to Estimating Software Reliability of RSG-GAS Reactor Protection System

S. Santoso[*], S. Bakhri and J. Situmorang

Center for Nuclear Reactor Technology and Safety, National Nuclear Energy Agency,
Puspiptek Area, Serpong, Tangerang Selatan 15314, Indonesia.

ARTICLE INFO

Article history:
Received 14 December 2017
Received in revised form 29 November 2018
Accepted 17 December 2018

Keywords:
Reliability
Software
RSG-GAS
Bayesian network
Protection system

ABSTRACT

Reliability represents one of the most important attributes of software quality. Assessing the reliability of software embedded in the safety of highly critical systems is essential. Unfortunately, there are many factors influencing software reliability that cannot be measured directly. Furthermore, the existing models and approaches for assessing software reliability have assumptions and limitations which are not directly acceptable for all systems, such as reactor protection systems. This paper presents the result of a study which aims to conduct quantitative assessment of the software reliability at the reactor protection system (RPS) of RSG-GAS based on software development life cycle. A Bayesian network (BN) is applied in this research and used to predict the software defect in the operation which represents the software reliability. The availability of operation failure data, characteristics of the RPS components and their operation features, prior knowledge on the software development and system reliability, as well as relevant finding from references were considered in the assessment and the construction of nodes on causal network model. The structure of causal model consists of eight nodes including design quality, problem complexity, and defect inserted in the software. The calculation result using Agenarisk software revealed that software defect in the operation of RPS follows binomial statistic distribution with the mean of 1.393. This number indicated the high software maturity level and high capability of the organization. The improvement of software defect concentration range on the posterior distribution compared with the prior’s is also identified. The result achieved is valuable for further reliability estimation by introducing new evidence and experience data, and by setting up an appropriate plan in order to enhance software reliability in the RPS.

INTRODUCTION

In the past few years, the implementation of digital technology for monitoring and controlling critical systems, such as a nuclear power plant, has become essential towards the increasing of plant complexity. G.A. Siwabessy Reactor (RSG-GAS) which is facing the problem of ageing also has applied digital technology on its reactor protection system, particularly in the monitoring function. It is reasonable since a digital system has much better technical features, i.e. better accuracy of monitor and control, friendlier human-machine interface, easier configurability and maintainability, and higher automation level over analog I&C systems [1]. Moreover, digital I&C systems seem much easier to incorporate with new features [2,3].

Extensive research has been carried out to develop a method for estimating the average probability of failure on demand of a safety instrument system, including the digital system platform [4,5]. Referring to IEC (International Electrotechnical Commission) standard, the probability of a failure on demand of a safety critical system should be on the order of 10^{-5}. Therefore, a new approach should be developed since it is not practical to perform millions of tests to assess the reliability of such a software. Usually,

[*] Corresponding author.
E-mail address: sigitsan@batan.go.id
DOI: https://doi.org/10.17146/aij.2019.775

hardware faults occur due to aging or wearing out, but software faults causing a digital system failure are design fault introduced during the software development or maintenance phase [5]. A root cause analysis of the software defects based on the orthogonal defects classification in software development has been proposed [6]. It also has been identified that most of software failures are due to a design error and system loop holes which are not independent to human performance shaping factor [7,8].

The probabilistic safety assessment of digital reactor protection systems (RPSs) has been an interest of many researches due to the ambiguity over the mechanism of the software failure and quantification of the failure probability [1,9]. There are many approaches and methods developed for qualitative and quantitative software reliability assessment. The software reliability growth method (SRGM), Bayesian belief network (BBN) method, and test-based methods are well known methods for a software reliability analysis [2]. Traditional prediction approaches are based on the stochastic process models, such as Jelinki-Moranda model and the Goel-Okumoto non-homogeneous Poisson process (NHPP) model [10]. The current model suggests that a stochastic process, especially a non-homogeneous Poisson process (NHPP) can describe the failure process of a software. Non-probability models, e.g. neural networks, have also been applied in the assessment of software reliability [11,12]. Another approach is software metrics assessment based on the defect per thousand line of codes (kilo line of code - KLOC). The existing models and approaches have assumptions and limitations which are not directly acceptable for all systems, such as reactor protection systems. Moreover, many of the models apply unrealistic restrictive assumptions to ensure tractability and solvability, such as removing defects completely (no new defects introduced when removing the one detected) or data completeness [13,14].

To address these challenges, Bayesian networks (BN) had been proposed in some studies as an alternative to traditional reliability estimation approaches [4,9,15]. The BN method evaluates the software development life cycle activities for estimating the potential number of remaining faults in the software [9,15]. The BN has been increasingly recognized as a potentially powerful solution to complex risk assessment problems. Historical background and recent applications of the BN methodology for software reliability and system reliability have been discussed in several studies [15-17].

The BN method was elaborated for assessing software reliability of this study. The objective of the study was to conduct the assessment of software reliability of reactor protection system (RPS) of RSG-GAS. The reliability calculation was performed by predicting the amount of defects (faults) on the operation of RPS software through Bayesian network causal model approach. The approach evaluated the software development life cycle activities experienced in the RPS software for predicting the potential number of remaining faults in the software. The Bayesian inference program tools developed by Agena were used for conducting statistic calculation [18]. The result achieved is valuable to set up a plan for improving the software reliability of RPS. Moreover, the result can be further developed in backward analysis in order to define process criteria, such as the level of process quality or testing quality when some requirements for the software reliability should be met.

THEORY

Software reliability assessment and bayesian network approach

Software reliability can be defined as the probability of failure free operation of a computer program in a specified environment for a specified time [19]. The reliability of the software in this study was associated to the reliability of the RSG-GAS reactor protection system in generating the valid scram signal and in performing the required scram function of the reactor correctly. Bayesian Network is a probabilistic approach that can be used to model and to predict the behavior of such a system based on observed stochastic events [20]. The application of a BN to quantify software residual faults shows that a BN can be an attractive approach to estimate the number of remaining faults after verification and validation of the software through an evaluation of the SDLC activities. The sample of BN modelling approach is given in Fig. 1 [15].

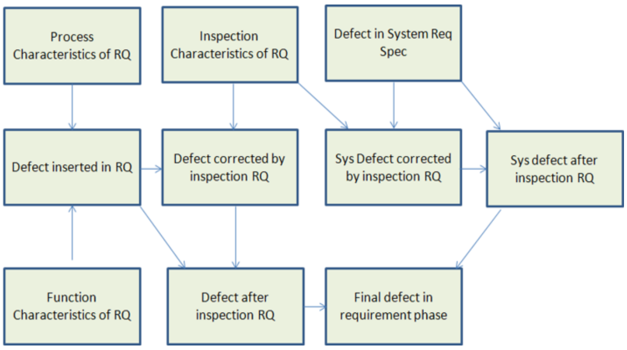
Process Characteristics of RQ
Inspection Characteristics of RQ
Defect in System Req Spec
Defect inserted in RQ
Defect corrected by inspection RQ
Sys Defect corrected by inspection RQ
Sys defect after inspection RQ
Function Characteristics of RQ
Defect after inspection RQ
Final defect in requirement phase

Fig. 1. Software requirement (RQ) specification phase in Bayesian network [15].

There are two main components of Bayesian network; they are Bayesian network structure graph which is also known as directed acyclic graph (DAG) and quantitative component represented by nodes. In this BN framework, a directed acyclic graph defines characteristics of the independent structure in a joint distribution, while the nodes represent system components as random variables, and directed arcs represent relationships among variables [16,21]. Another component of BN, the conditional probability distribution (CPD) is associated with each node in the graph. It determines quantitatively the relationship between a parent node and a child node. The Bayesian network's probability functions are usually represented as node probability tables (NPTs), which are used to represent CPD particularly when the variables are discrete. NPT shows the list of probability that the child node takes on each of its different values for each combination of values of its parents.

The approach of Bayesian can be presented in equations (1) to (3). The probability of an event or hypothesis is conditional on the available evidence on current context. This can be made explicit by the notation *P(h|D)* which is read as the probability of event h given the evidence *D*. *P(h)* represents prior probability of event or hypothesis where *P(h|D)* represents as posterior. In probability theory, a joint distribution of *n* random variables in a graph always can be developed by using the chain rule as shown in equation (2). Each node of a BN can be described by a local conditional probability distribution (CPD) function given its parents in the graph, that is *P(*V_i*| parents(*V_i*))*, where *parents(*V_i*)* indicates the parent nodes of node V_i. The joint distribution of variables in Bayesian network is given in equation (3).

$$(P(h|D) = \frac{P(D|h)P(h)}{P(D)} \qquad (1)$$

$$P(V_1, V_2, \ldots, V_n) = P(V_1)P(V_2|V_1)P(V_3|V_2V_1) \ldots (V_n|V_{n-1,} \ldots V_1) \qquad (2)$$

$$P(V_1, V_2, \ldots, V_n) = \prod_{i=1}^{n} P\,(V_1|parents(V_1)) \qquad (3)$$

Equation 3 shows the chain rule to construct Bayesian inferences for BN. A Bayesian network is defined by specifying every term on the right-hand side of this equation. In general, if there are *n* binary nodes, the number of parameters for constructing the full joint distribution $P(V_1, V_2, \ldots V_n)$ is on the order of 2^n. The concept of Bayesian approach and BN was described in some literatures.

Bayesian statistic inference is powerful since the method enables combination of common-sense knowledge and observational evidence. The BN approach is able to cope with the known limitations of data required in the traditional assessment model. Table 1 shows the comparison between software reliability growth modelling (SRGM) and Bayesian belief network (BBN) i.e. BNs evaluation based on desired characteristics [22]. Further improvement to the BNs approach is being developed such as the problem dealing with the limited prior data, discretization, and others. Zwirglmaier *et al*. [23] proposed parametric formulations for efficient discretization of random variables in BNs for reliability problems based on numerical investigations. Based on the sufficient data, the challenge for applying BN in a software reliability assessment mostly corresponds to a scenario that a network structure is known, while some variables are hidden.

The calculation of software reliability can be performed based on the software development process and the operation experienced in the software system. For the causal BN model employed in this study, the defect remained to exist in the RPS software, depending on the residual defect which was introduced in the initial program design (*defect inserted*) as well as on the defect that could be identified and fixed during the test or operation (*defect found and fixed*).

Table 1. Evaluation of quantitative software reliability methods against their characteristics [22].

Characteristics	Methods	
	SRGM	BBN
1. Comprehensive method description	Yes	Yes
2. Reasonable assumptions	Maybe	Maybe
3. Consideration of operating condition	No	No
4. Consideration of SDLC quality	No	Yes
5. Use of test and operation experience data	Yes	Yes
6. Uncertainty addressing	Yes	Yes
7. V&V of the method	Yes	Maybe
8. Demonstrating high system reliability	No	Maybe
9. Evaluation of CCF parameters	No	No
10. Data availability	No	Maybe

RSG-GAS reactor protection system

Reactor Protection System of RSG-GAS is the essential part for securing the safety in reactor operation. The main components of the RSG-GAS reactor protection system (RPS) consist of data acquisition systems, analog systems, logic systems, and six contact systems (Fig. 2) [24]. From the diagram it is understood that reactor scram can be activated based on established reactor parameter measurements considered as safety critical. These parameters include gamma dose rate in the reactor pool ventilation system, surface elevation of reactor pool water, primary isolation valve position, reactor core neutron flux density, and gamma dose rate in the primary system. Otherwise, the scram can also be performed manually by the operator in the control room.

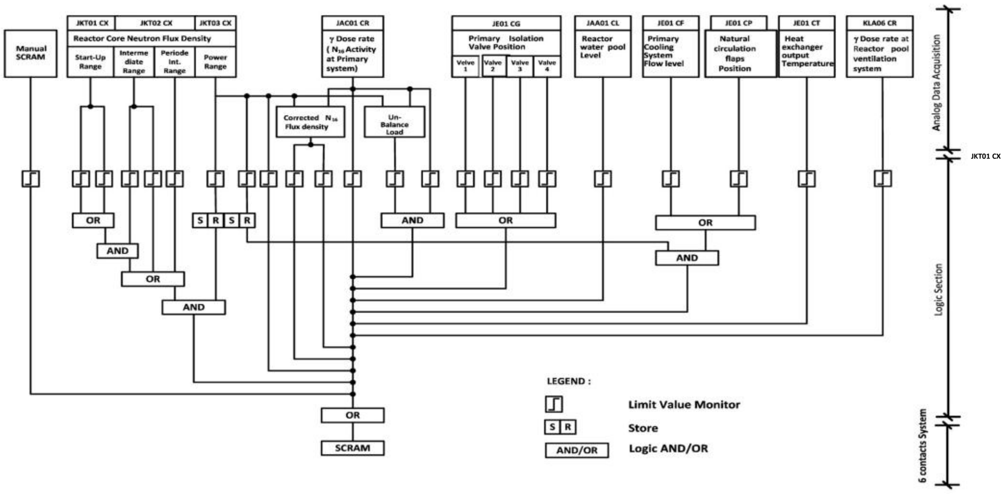

Fig. 2. Scram logic diagram for the RSG-GAS reactor protection system [24].

The RPS software is embedded in the reactor protection system in order to perform reactor condition monitoring task and to support control task. According to the control function, the initial step is to perform the signal conversion process on a selected calculation parameter. Furthermore, the logic sections will conduct the comparison of the parameter with corresponding scram boundary values. The software program intended for monitoring task is integrated on the display system that helps operator in conducting reactor operation. The embedded program for conducting RPS function in controlling and monitoring was developed by reactor vendor and integrated into the reactor system and operation. However, during its operation since the first critical in 1987, some modifications have been applied to the system specifically on the RPS.

CALCULATION METHODS

The software reliability assessment was performed by predicting the defect which exists on the RPS software using Bayesian network causal model approach. The steps for conducting assessment were as follows. The first one was to identify the software applied in the digital protection system of RSG-GAS, including the software process design and operation. The second step was the construction of Bayesian network graph and causal modelling based on software life cycle information data. Activities in the software development life cycle were represented by a node of the causal graph. Node probability table (NPT) or conditional probability distribution (CPD) for each node was prepared to set up the relationship between child nodes and parent nodes of the BN graph. The final step was to estimate the remaining software defects based on given prior inputs and evidence. The steps of BN approach for assessment of RPS software are presented in Fig. 3.

Bayesian tool from AgenaRisk was used for conducting the Bayesian modelling and calculation. This tool is efficient for calculation since it applied dynamic discretization on the continuous variables contained in the network. In this study, the input data for prior distribution of nodes in the causal networks would be generated based on the RPS software data of the RSG-GAS and data from the program module.

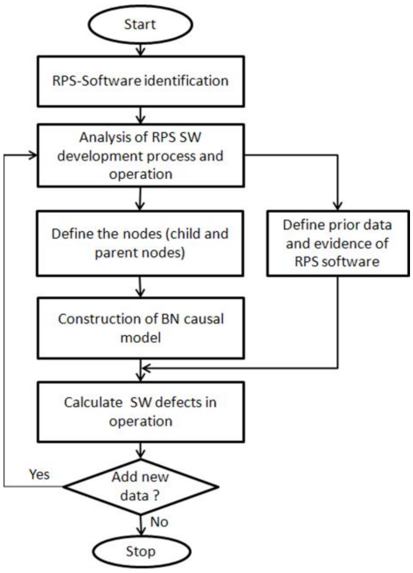
Start
RPS-Software identification
Analysis of RPS SW development process and operation
Define the nodes (child and parent nodes)
Define prior data and evidence of RPS software
Construction of BN causal model
Calculate SW defects in operation
Yes
Add new data ?
No
Stop

Fig. 3. Flowchart for the calculation of defects in the operation of RSG-GAS RPS software.

RESULTS AND DISCUSSION

Identification and analysis on the process development of RPS software for RSG-GAS showed its conformity to the required development procedure which includes the identification of scope and function, software development, testing and validation, and integration of the released software as a final product on the system. Considering the process development and characteristics, the causal Bayesian network model for software defects and reliability prediction developed by Fenton [25] seems to be the most appropriate for conducting RPS software assessment. The Bayesian causal network of RPS software consists of eight nodes with the causal diagram presented in Fig. 4.

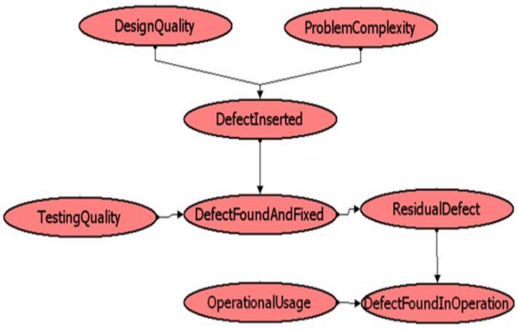

Fig. 4. A Bayesian network model for RPS software defects and reliability prediction.

The prior distribution for each node in the causality diagram was constructed based on design and operation information in the software development life cycle (SDLC). Generic software assessment data, result data from similar study, and characteristics of the model were considered in the establishment of assumptions and values on every node in the BN causality graph. Prior distribution data adopted from the available module in the Agena software were selected to represent the initial prior of the RPS software. The prior data generated by the module and evidence from observation applied in the calculation are summarized below: (i) Prior for the causality diagram root nodes such as design process quality node, complexity, operation level, and quality testing node were set as uniform distribution with a mean and a median value of 0.5 respectively. At this stage, the probability was considered as equal for all five classifications/discretization which were ranged from very low to very high probability categories, (ii) For the intermediate nodes and leaf nodes such as residual defect, defect found and fixed, the initial prior was approximated by binomial distribution while the prior distribution for defect contained in initial design of device software (defect inserted) was given in truncated normal. The statistic distribution of every node of BN graph is presented in Table 2, (iii) For subsequent calculations of RPS software, data characteristics and software operations were assumed in a medium complexity level while the process quality level was very high. Meanwhile, quality in testing was considered high and the level of operational usage was low. It was because the RPS will only be actuated when some reactor critical parameters achieve the operation limit condition.

Table 2. Selected node probability distributions for the nodes of the BN model in Fig. 4, adapted from Fenton [25].

Node	Node name	CPD
1-4	Design process quality, problem complexity, testing quality, operational usage	Ranked, with uniform distribution
5	Defect found in operation	Binomial(n,p) where n = residual defect and p = operational usage
6	Residual defect	Max (0, Defect inserted – defect found and fixed)
7	Defect found in testing	Binomial (n,p) where n = defect inserted and p = testing quality
8	Defect inserted	Truncated normal in the range of 0 to 500, complexity x (1-design process quality)x90; variance 300

The calculation result for the software defects prediction found in operation using BN model is presented in Fig. 5.

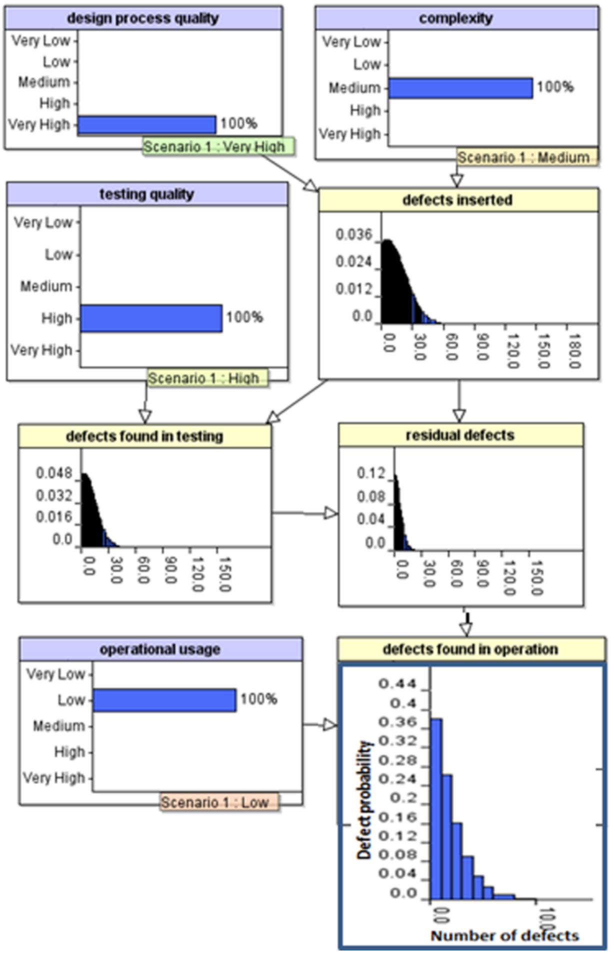

Fig. 5. Bayesian network model for RPS software reliability and prediction results.

The result achieved showed that the mean of the number of defects in the posterior distribution was equal to 1.393 with a standard deviation of

1.682. These values are much smaller compared to the same variables in the prior. The prior mean of such defects was 7.150 and its standard deviation was 14.411. The result indicated a trend that a lower operational usage can significantly reduce defects. However, it does not indeed lead to the belief that the RPS software reliability is definitely high. Residual defects which depend on the other nodes of the network also contributed to the result. Moreover, adding new evidence and observation data would become necessary to achieve more accurate results.

It was indicated in the result that the probability distribution of other nodes was also updated as new evidence was entered. It showed that BN model was doing both forward and backward inferences to the variables. The result showed the mean value of defects inserted in the design was equal to 15.397. Process quality and system complexity level had reduced the defect inserted. On the other hand, the defects found and repaired in testing and the residual defects were 10.790 and 4.619 respectively. The estimation results can be combined with new evidence/observation data from the next operation experience of the RPS software to generate better predictive result.

The defects found in operation of the posterior followed a binomial statistic distribution. Figure 6 shows that the probability distribution of posterior tended to be concentrated in smaller range in comparison with the prior distribution. In the posterior distribution, 50 % defects in operation values existed in the range of between 0.665 (lower percentile) and 2.662 (upper percentile). For the prior distribution, the range was from 1.176 to 9.980. It means the observed data applied in the calculation decreased significantly the uncertainty of the obtained result. The result and the model achieved in this study can be further developed for backward inferences in order to define requirements, such as the level of processs quality or testing quality when some specific criteria of the system software reliability should be met.

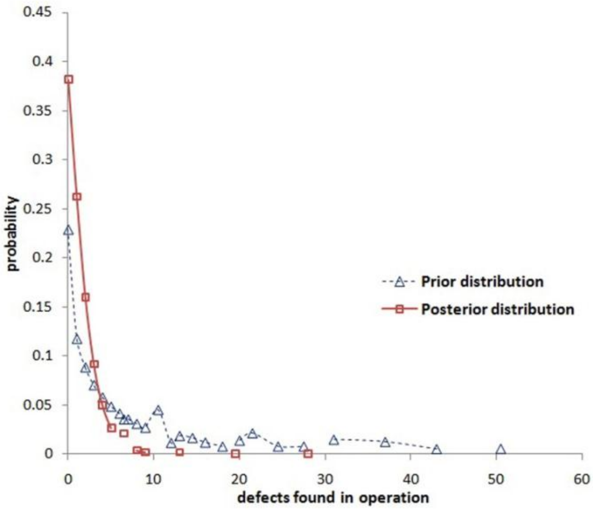

Fig. 6. Comparison of prior and posterior distributions on the defects found in operation node

CONCLUSION

The assessment on software reliability of RPS had been conducted in this research through a Bayesian network approach. The result and analysis presented in this paper demonstrated the capability of the BN for estimating RPS software reliability based on the software development life cycles. The calculation results of software defect in operation using Bayesian network revealed that defects found in the operation followed a binomial distribution with an average of 1.393 and the standard deviation of 1.682. From the perspective of calculation process in Bayesian network, this result indicated the improvement of accuracy of the software defect prediction. These numbers also indicated the high software maturity level and high capability of the organization. The result on relationship among nodes presented in the BN model would be helpful in understanding the dynamics behind the software reliability of RPS.

The result achieved in this study is essential for developing a plan aimed at enhancing RPS software reliability. Moreover, the developed BN model can be used in the backward analysis in order to define requirements in the software development process when certain criteria of the software reliability should be met.

ACKNOWLEDGMENT

This research was a part of a research project supported and funded by Center for Nuclear Reactor Technology and Safety of the National Nuclear Energy Agency of Indonesia and the Ministry of Research, Technology and Higher Education of the Republic of Indonesia.

REFERENCES

1. C. Guo, H. Xiong, X. Huang *et al.*, Sci. Technol. Nucl. Install. **2017** (2017) 1. https://doi.org/10.1155/2017/2981943
2. M. Khalaquzzaman, S.J. Lee, M.K. Ki *et al.*, J. Nucl. Eng. Des. **280** (2014) 201. https://doi.org/10.1016/j.nucengdes.2014.09.008
3. J. Cho, S.J. Lee, W. Jung, J. Nucl. Eng. Des. **316** (2017) 198. https://doi.org/10.1016/j.nucengdes.2017.03.016

4. B. Zou, M. Yang, E.R. Benjamin *et al.*, Progress in Nuclear Energy **98** (2017) **85**. https://doi.org/10.1016/j.pnucene.2017.03.006
5. G.Y. Park and S.C. Jang, J. Nucl. Eng. Technol. **46** (2014) 55. https://doi.org/10.5516/NET.04.2012.067
6. N. Silva, J. Cunha, M. Vieira, J. Reliability Engineering and System Safety **158** (2017) 213. http://dx.doi.org/10.1016/j.ress.2016.08.016
7. F. Huang and B. Liu, Chinese Journal of Aeronautics. **30** (2017) 1054. http://dx.doi.org/10.1016/j.cja.2017.03.005
8. S. Santoso, Journal of Nuclear Reactor Technology Tri Dasa Mega **18** (2016) 135. http://dx.doi.org/10.17146/tdm.2016.18.3.3017
9. F. Zare, H.K. Zare and M.S. Fallahnezhad, Applied Soft Computing **49** (2016) 968. http://dx.doi.org/10.1016/j.asoc.2016.08.004
10. A.L. Goel and K. Okumoto, IEEE Transactions on Reliability **28** (1979) 206.
11. C. Lopez-Martin, Appl. Soft Comput. **27** (2015) 434. http://dx.doi.org/10.1016/j.asoc.2014.10.033
12. D. Miholca, G. Czibula, I. Czibula., J. Information Sciences. **441** (2018) 152. https://doi.org/10.1016/j.ins.2018.02.027
13. M. Shepperd, D. Bowes and T. Hall, IEEE Transactions on Software Engineering, **40** (2014) 603.
14. O. Yazdanbakhsh, S. Dick, I. Reay *et al.*, Applied Soft Computing **49** (2016) 1256. http://dx.doi.org/10.1016/j.asoc.2016.08.006
15. H.S. Eom, G.Y. Park, S.C. Jang *et al.*, Ann. Nucl. Energy. **51** (2013) 38. http://dx.doi.org/10.1016/j.anucene.2012.06.030
16. M. Perkusich, G. Soares, H. Almeida *et al.*, J. Expert Systems With Applications **42** (2015) 437. http://dx.doi.org/10.1016/j.eswa.2014.08.015
17. R. Rana, M. Staron, C. Berger *et al.*, J. Syst. and Software. **1** (2016) 229. https://doi.org/10.1016/j.nucengdes.2017.03.0
18. Anonymous, AgenaRisk: Advanced Risk Analysis for Important Decisions, Agena (2007).
19. H. Okamura, T. Dohi and S. Osaki, Reliability Eng. and Syst. Safety **116** (2013) 135. https://doi.org/10.1016/j.ress.2012.02.002
20. Y. Zhao, J. Tong, L. Zhang *et al.*, Nucl. Eng. Des. **291** (2015) 154. http://dx.doi.org/10.1016 /j.nucengdes.2015.05. 010
21. I. Tien and A.D. Kiureghian, Reliability Eng. and Syst. Safety **156** (2016) 134. http://dx.doi.org/10.1016/j.ress.2016.07.022
22. T.L. Chu, M.Yue, Guridi *et al.*, *Development of Quantitative Software Reliability Models for Digital Protection Systems of Nuclear PowerPlants*. NUREG/CR-7044, BNL-NUREG-99068 (2013).
23. K. Zwirglmaier and D. Straub, J. Reliability Eng. and Syst. Safety **153** (2016) 96. http://dx.doi.org/10.1016/j.ress.2016.04.008
24. Anonymous, Safety Analysis Report of RSG-GAS, Rev 10, **2** (2008).
25. N. Fenton, M. Neil and D. Marquez, J. Risk and Reliability **222** (2008) 701.