Journal of Engineering and Technology

ISSN 2180-3811 eISSN 2289-814X https://jet.utem.edu.my/jet/index

UNCERTAINTY PROPAGATION IN PROBABILISTIC SAFETY ANALYSIS USING LOGNORMAL DISTRIBUTION

M. Boufenar*[1]

[1]Centre de Recherche Nucléaire de Draria, Algiers, Algeria

*corresponding: m-boufenar@crnd.dz

Article history:
Received Date: 16 March 2023
Revised Date: 4 June 2023
Accepted Date: 7 June 2023

Keywords: PSA Model, Uncertainty Propagation, Lognormal Distribution, RAE Approximation Monte Carlo,

Abstract— Uncertainty analysis, which is one of the major elements in Probabilistic Safety Assessments (PSA) of Nuclear Power Plants (NPP's), involves quantifying the uncertainties of the occurrence of accident scenarios. The traditional approximations used in current PSA models are limited, and normally conservative, based on not fully accounting for the dependence between Minimal Cut Sets (MCS). In this work, a mathematical development of an approximation method to propagate the uncertainty of lognormal distributions is carried out by modifying the approach suggested by Fenton and Wilkinson. When the uncertainties of basic events are modelled with lognormal random variables, the top event frequency or

Fenton-Wilkinson	probability is well approximated as the sum of the correlated lognormal random variables. This study focuses on how to minimize or eliminate the adverse effect of Rare Event Approximation (RAE) that induces the overestimated top event mean value. The probability distribution of the case study top event presented in this work is compared with analyses available in the literature using different approaches, such as Monte Carlo and Fenton-Wilkinson (FW) method. The application of this method to propagate uncertainty of lognormal distributions results in a better estimation of the top event probability distribution. It is shown how the cut set information for a model can be used together with the analytic expression to give closed-form approximation for the top event uncertainty distribution. This approach appears attractive and can provide a reasonable approximation, without incurring the computational expense.

I. Introduction

Uncertainties are present in PSA of Nuclear Power Plants, as incompleteness and imprecision, in probabilities of data which include, among others, components reliability, dependencies between plant systems, human interactions and common cause failures. These data are usually represented by uncertainty bounds or probability density functions (pdf), measuring the degree of knowledge or confidence in the available data. One of the key components of uncertainty analysis is the quantification of uncertainties in the system output performances propagated from uncertain inputs, named as uncertainty propagation. Many uncertainty propagation methods have been developed in various fields, such as simulation-based methods like Monte Carlo simulation (MC), which is widely used for evaluating the uncertainty distribution, but for such calculations, the computational cost becomes prohibitive as the number of input random variables increases and do not readily reveal the dominant contributors to the uncertainty. When selecting a method for uncertainty propagation various aspects should be considered such as the required level of uncertainty quantification, accuracy, or confidence level, as well as the computational cost or efficiency. The present work investigates how the uncertainties in input probabilities can be propagated through a PSA model to produce output uncertainties that can be communicated to decision-makers and how they can be accounted for in the ranking of the basic events or components. Since the performance of the uncertainty propagation methods are affected by the problem settings such as the type of input distribution, the number of input random variables and the required resolution, it would be beneficial to develop an appropriate method which fits for a specific situation [1]. In this work, an analytic method, to overcome the challenges of uncertainty propagation in a PSA model, which is applicable when the uncertainties in the basic events of the model are log-normally distributed is developed. The performances of the developed approach are compared in moment estimation and probability density function construction with results obtained by FW method based on order statistics and MC method based on sampling order and used as a reference. The comparison of the frameworks is explored by an application to a Level 1 NPP PSA benchmark problem.

II. Methods of Uncertainty Quantification used in PSA

Various methods have been developed and applied to determine the pdf of the top event from the pdf's of basic events. In theoretical approaches, such as variable transformation [2], the pdf of the top event is synthesized by integrating the joint pdf's of basic events [3]. Hence, given a set of n inputs $\{X_i\}$, the top event probability of a specified fault tree can be expressed as a function, which returns an output random variable, denoted by Z below:

$$Z = F\,(X_1, X_1, \ldots, X_n) \qquad (1)$$

The n inputs $\{X_i\}$ are the basic events with defined failure probabilities. In a PSA model, the top event is expressed as a logical equation, involving the sum of cut sets or Minimal Cut Sets (MCSs) that represent combinations of basic events. Applying the rare event approximation, the function F can be expressed as a multi-linear function of the form:

$$F\,(X_1, \ldots, X_n) = \sum_{j=1}^{m} \prod_{i=1}^{n(j)} X_{ij} \qquad (2)$$

In equation (2), each term in the summation gives the probability of a cut set, where an MCS is defined as an irreducible combination of elementary failures that together result in the top event occurring. The integer m is the total number of cut sets in the problem, and $n(j)$ is the number of events in the j_{th} cut set. Among the approximate methods, such as: (i) analytical methods (Fenton-Wilkinson), (ii) fuzzy arithmetic, (iii) Dempster-Shafer Theory (evidence theory), (iv) and Monte Carlo simulations, the Monte Carlo simulation is most widely used in practice, owing to its ease of application to large-scale models of complex systems [4]. This technique provides an efficient and straightforward way for propagating uncertainties. However, Monte Carlo simulation has inherent limitations as a simulation-based approach, it can be inaccurate if the number of samples is not large enough, and it can be limited by computational costs. Furthermore, analytical methods consist in obtaining the distribution of the output of a model starting from probability distribution of input parameters. An analytical distribution of the output, however, can be derived only for specific models such as normal or log-normal distributions [5]. Moreover, analytic solutions provide deeper insights into an uncertainty analysis, thus providing better explanations of the numerical results [6]. This kind of closed form is helpful, when more detailed uncertainty analyses are required, for instance, in parametric studies involving uncertainty importance assessments, which require re-estimating the overall uncertainty distribution many times.

III. Top Event Distribution Approximation

The pdf of a lognormal variate, X, is given by:

$$P\left(X_{\mu,\sigma}\right) = \frac{1}{X\sigma\sqrt{2\pi}} \exp\left(\frac{-(\ln(X)-\mu)^2}{2\sigma^2}\right) \quad (3)$$

where μ and σ^2 are the mean and the variance of $\ln(X)$, respectively. A property of a lognormal distribution is that its logarithm is a Gaussian distribution, that is $Y = \ln(X)$ is a normal distribution with parameters μ and σ^2 denoted by $X \sim \ln(\mu, \sigma^2)$. Note that the parameters μ and σ^2 are not equal to the mean and variance of the random variable X. The latter are given by the following expressions:

$$E[X] = \exp\left(\mu + \frac{\sigma^2}{2}\right) \quad (4)$$

$$\mathrm{Var}[X] = \exp(2\mu + \sigma^2)\,[\exp(\sigma^2 - 1)] \quad (5)$$

Because the basic events in a fault tree are assumed to be independent of each other, the occurrence probability of a minimal cut set is the product of all the occurrence probabilities of the basic events included in the MCS, as follows:

$$Y_i = \prod_{j=1}^{n_i} X_{ij} \quad (6)$$

where Y_i is the random variable for the occurrence probability of the i_{th} MCS, X_{ij} is the random variable, which have lognormal parameters $\{\mu_{Xi}, \sigma_{Xi}\}$, for the occurrence probability of j_{th} basic event included in the i_{th} MCS, and n_i is the number of basic events. Then Y_i also has a lognormal distribution because the product of log-normal random variables is given as another log-normal random variable: $Y_i \sim \ln(\sum_{i=1}^{m} \mu_{Xi}, \sum_{i=1}^{m} \sigma_{Xi}^2)$. As described in equation (2), the top event frequency of a single top fault tree in Level 1 PSA, can be approximated with the sum of a finite number of high-ranking lognormal random variables, as follows:

$$Z = \sum_{i=1}^{n} Y_i \quad (7)$$

However, the lognormal random variables for minimal cut sets (Y_i's) may be dependent to each other. Thus, some MCSs may include the same initiating event or basic events. Since high-ranking MCSs may depend on each other, the top event frequency is given as the sum of correlated lognormal random variables. The probability density function of the sum of correlated log normally distributed random variables is a well-known challenging problem. For instance, an analytical closed-form expression of the lognormal sum distribution does not exist and is still an open problem. Ref. [7] gives an extension of the widely used iterative method known as Schwartz and Yeh (SY) method. Some other resources use an extended version of Fenton and Wilkinson methods [8]. In these methods, there is general agreement that a sum of lognormal distributions can be well approximated by another lognormal distribution. However, to find the parameters of the resultant lognormal distribution, different approaches have been proposed. Among these approaches are moments matching, cumulants matching, etc. The accuracy of each method relies highly on the region of the resulting distribution being examined, and the individual lognormal parameters. There is no such method which can provide the needed accuracy for all cases. For example, Schwartz and

Yeh's [7] method provide acceptable accuracy in low-precision region of the Cumulative Distribution Function (CDF) and the FW method offers high accuracy in the high-value region of the CDF. Both methods break down for high values of standard deviations.

IV. Proposed Lognormal Approximation Method

The lognormal approximation, developed in this work, for the top event uncertainty estimation, uses first and second moments of the input parameters (mean and variance) to estimate the mean and variance of the output function. Recall that for a set of MCS, the probability of the output event is given by the random variable Z which is approximated by:

$$P(Z) = P(\cup_i MCS) \tag{8}$$

In theory, the calculation of the probability of a sum of MCS can be performed thanks to the Higher-order method via the inclusion–exclusion principle, also known as the Sylvester-Poincaré development.

$$P(Z) = P\left(\bigcup_i MCS\right) = \sum_{i=1}^{n} P(MCS_i) - \sum_{j=1, i\neq j}^{n} P\left(MCS_i \cap MCS_j\right) + \cdots + (-1)^{n-1} P\left(\cap_{i=1}^{n} MCS_i\right) \tag{9}$$

It can be seen from equation (9) that the quantification of higher-order terms results in a combinatorial expansion of conjunctive terms even with a small number of cut sets. The computational cost of this calculation method is prohibitive. Owing to both the complexity of the calculation and the large size of the resulting equations for the real models, approximations are thus performed by the so-called Rare Event Approximation (RAE) that consists in considering only the first term of the development. The REA approximation ignores the possibility that two or more rare events can occur simultaneously. That is, this approximation only considers the most relevant terms of the expansion up to a given order; hence, the Sylvester-Poincaré development is often approximated up to the first order:

$$P_{RAE}(Z) = \sum_{i=1}^{n} P(MCS_i) \quad (10)$$

Accordingly, the first and second-order moments are matched to the true distribution moments, by matching the mean value and the variance:

$$E[Z] = \sum_{i=1}^{n} E[U_i] \quad (11)$$

$$\text{Var}[Z] = \sum_{i=1}^{n} \text{Var}[U_i] \quad (12)$$

Since we are forcing our approximate distribution for Z to be log-normal, this results in the following expressions (from the properties of log-normal parameters):

$$\exp\left(\mu_Z + \frac{\sigma_Z^2}{2}\right) = \sum_{i=1}^{n} \exp\left(\mu_i + \frac{\sigma_i^2}{2}\right) \quad (13)$$

$$\exp(2\mu_Z + \sigma_Z^2)\exp(\sigma_Z^2 - 1) = \sum_{i=1}^{n} \text{Var}[U_i] \quad (14)$$

The mean value and variance of the lognormal approximation for Z can be derived by solving these equations that gives the following results:

$$\sigma_Z^2 = \ln\left(1 + \frac{\sum_{i=1}^{n} \text{Var}[U_i]}{(E[Z])^2}\right) \quad (15)$$

$$\mu_Z = \ln\left(\sum_{i=1}^{n} \exp\left(\mu_i + \frac{\sigma_i^2}{2}\right)\right) - \frac{\sigma_Z^2}{2} \quad (16)$$

RAE is limited and normally conservative, based on not fully accounting for the dependence between MCSs. The amount of conservatism is generally unknown, and if the basic event probabilities get too large or the dependence among cut sets gets large, the REA overestimation can be quite noticeable. Since the REA tends to overestimate the top event probability, Min Cut Upper Bound (MCUB) approximation is introduced in this work to minimize or eliminate the high overestimation in the calculation. This takes advantage of the fact that is a better estimate of the mean of the top event probability than the rare event approximation (simple sum of MCSs). The Min Cut Upper Bound approximation is thus given by the following expression:

$$P_{\text{MCUB}}(Z) = 1 - \prod_{i=1}^{n}\left(1 - P(MCS_i)\right) = 1 - \prod_{i=1}^{n}\left(1 - \prod_{j=1}^{m} P(X_j)\right) \quad (17)$$

If, all the X_j have a lognormal distribution, then using the results of equation (17), the distribution of Z can be approximated in closed form as a lognormal distribution. The original estimate of the sigma parameter in equation (15), of the assumed log-normal distribution via the Rare Event approximation is still used, since equation (17) cannot be used to derive a better estimate of the variance. These two estimates together define a unique lognormal approximation distribution.

V. Application to a Benchmark Problem

In this section, the accuracy of the proposed method is highlighted through the example presented in Ref. [9], which is associated with the uncertainty analysis of a Level 1 PSA of a nuclear power plant (NPP) in the United Kingdom. The fault tree shown in Figure 1 was constructed for the top event 'Reactor Core Damage' caused by a failure of Core Residual Heat Removal due to a short-term loss of offsite power, which is a small part of the full PSA model. Tables 1 and 2 present the relevant basic events and minimal cut sets parameters, respectively. As shown in Table 2, RiskSpectrum software was used to derive the MCSs and their associated probabilities that quantitatively describe the illustrative example analyzed in the literature. A comparison of the proposed method with MC simulations and FW method was done, considering the resultant probability density function, cumulative distribution function, and percentiles. A Matlab code is developed to implement the analytic approach shown above and the procedure for obtaining the pdf or the cumulative distribution function for the event of interest.

Next, Monte Carlo simulations are performed with 100,000 samples using RiskSpectrum software to estimate the top event uncertainty distribution. Figure 2 shows the probability density functions of the core damage frequency obtained using Monte Carlo simulations, Fenton-Wilkinson's method, and the proposed method. As can be seen in Figure 2, FW method shows a relatively high accuracy in the right tail of the probability distribution function, whereas the method shows poor accuracy in the left tail of the distribution. Indeed, Fenton - Wilkinson method relies on matching the first and second moments using REA approximation of the mean value. However, the proposed method based on MCUB approximation of the mean provides a theoretical probability density function which is in good agreement with the MC simulation and being slightly lower as can be shown in Figure 2. Table 3 presents a comparison of the percentiles estimated using the three methods. As can be seen in this table, the proposed method provides theoretically more accurate results since the median values of the approximation method show a good agreement with MC simulation.

Table 1: Relevant Basic Events and Parameters (E(x) and σ^2 values were taken from Ref. [9])

Event	Designation	Parameters	
		$E[X]$	σ^2
A	Short-Term Loss of Offsite Power	6.00E-2	7.60E-2
B	CCF of Batteries for Short-Term Start of EDG	6.60E-6	8.36E-6
C	Operator fails to Start the Backup DG by Local Action	1.00E-2	1.27E-2
D	Operator fails to Start the Backup DG or Close Breakers	2.13E-3	2.70E-3
E	CCF to Run Backup Diesels	8.33E-4	1.06E-3
F	CCF of Batteries Via Two-Hour Discharge	5.20E-5	6.59E-5
G	CCF to Start Emergency Diesel Generators	6.10E-5	7.73E-5
H	CCF to Run Emergency Diesel Generators	4.20E-5	5.32E-5
I	Fail to Run Individual Backup DG	1.58E-3	2.00E-3
J	CCF to Start Backup Diesels	1.00E-4	1.27E-4
K	Damage to O-ring Seals	9.00E-2	1.14E-1
L	Severe Seal Damage on Reactor Cool. Pumps	1.00E-1	1.27E-1
M	CCF to Run High-Pressure Injection Pumps	1.20E-4	1.52E-4

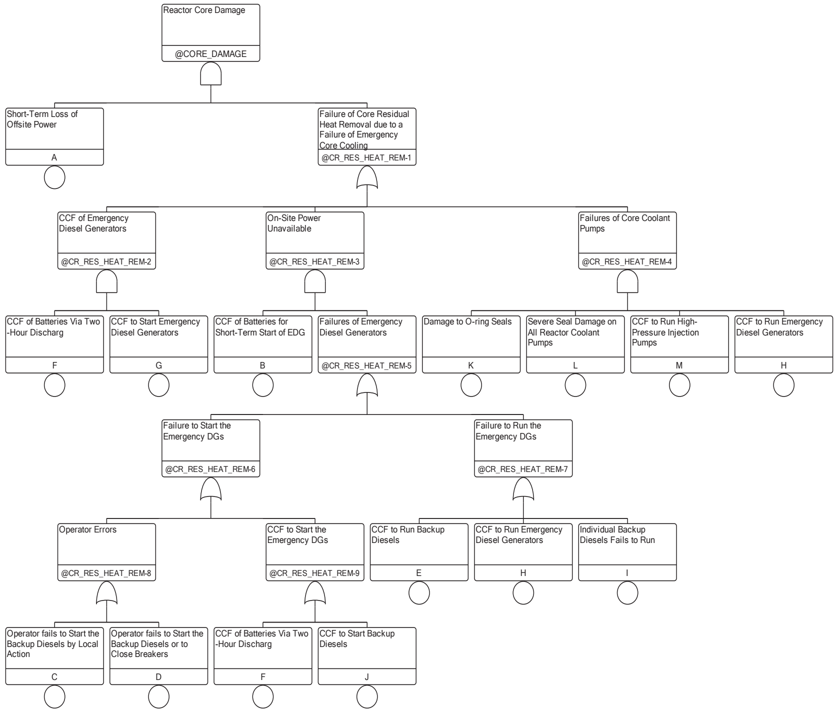

Figure 1: FT analysis for the Core Damage caused by the Core Residual Heat Removal Failure

Table 2: Minimal Cut Sets probabilities for Reactor Core Damage frequency due to short-term loss of offsite power

	Minimal Cut Sets (MCS)	Probability	Percentage (%)
1	ABC	1.130E-09	68.10
2	ABD	1.235E-10	7.22
3	ABI	1.147E-10	6.62
4	ABE	1.119E-11	1.19
5	AFG	1.103E-11	1.07
6	ABJ	3.960E-11	0.66
7	ABF	2.059E-11	0.34
8	ABH	1.663E-11	0.28
9	AHKLM	2.722E-12	0.05

This is also illustrated in Figure 3, where the cumulative distribution function obtained by lognormal approximation method is compared with results obtained by MC method used as a reference and the data from the mentioned literature using FW method. Figures 2 and 3 and the comparison of percentiles in Table 3, show that the proposed method gives reasonable overall results for this problem, and capture the broad shape of the top event distribution quite well. The use of MCUB approximation in uncertainty analysis with the analytic solution is expected to significantly reduce the computational burden in calculating the pdf's of top events in a fault trees with limited errors. However, for more complex systems and large fault and event trees, computer implementation of the described approach can be performed and examined in depth to understand the characteristics and limitations, if any.

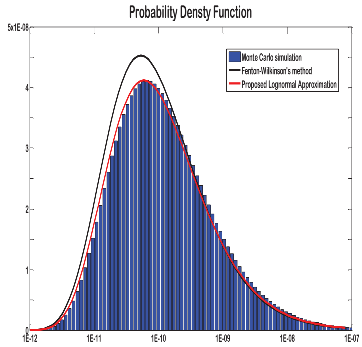

Figure 2: Top event pdf, via MC, FW and the Proposed method

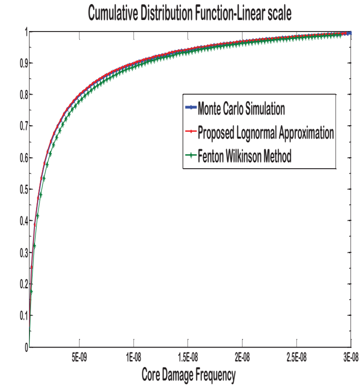

Figure 3: Top event CDF, via (i) MC sampling, (ii) FW method and (iii) Proposed Approximation

Table 3: Comparison of core damage frequency obtained by proposed lognormal approximation with data from literature

Method	5^{th} Percentile	Median	95^{th} Percentile	% difference of median
Proposed Approximation	1.17×10^{-10}	1.57×10^{-9}	2.13×10^{-8}	3.18%
MC simulation	1.13×10^{-10}	1.59×10^{-9}	2.27×10^{-8}	1.89%
FW method	1.26×10^{-10}	1.62×10^{-9}	2.08×10^{-8}	--

VI. Conclusion

In this study, a lognormal approximation approach is proposed for analyzing the uncertainty in the top event frequency or probability in PSA when basic events are given with lognormal random variables. The effectiveness of the method developed in this work has been tested using a benchmark problem of a PSA part of a nuclear power plant. The value of the proposed log-normal approximation lies in reducing the excessive computational load involved in conventional Monte Carlo methods of uncertainty estimation and improving confidence in the values of important percentiles. An additional area in which the approximate method may be of value is in understanding of the contributions to the uncertainty distribution; mainly, in importance analysis and sensitivity studies related to basic events which require re-estimating the uncertainty distribution several times (once for every input parameter) which is hardly achievable with Monte Carlo simulation. Furthermore, it is worth noting that decreasing of the conservatism that is caused by the REA approximation of the top event mean value allows an improvement of importance measures calculations and provides a more accurate overview of uncertainty contributors. In summary, the proposed method proves to be an efficient approach providing the probability density function of the top event frequency or probability with high accuracy. As a theoretical approach, the proposed method provides deeper insights into the uncertainty analysis. The mathematical formulation and analytic solutions provided in this study are expected to serve as a basis for future studies on theoretical approaches for the uncertainty analysis in PSA.

VII. References

[1] G. S. Song and M. C. Kim, “Mathematical Formulation and Analytic Solutions for Uncertainty Analysis in Probabilistic Safety Assessment of Nuclear Power Plants”, *Energies* (2021), 14, 929.

[2] A. S. Cook and T. Downs, “Estimating Manufacturing Yield by Means of Jacobian of

Transformation". *IEEE Proc. G* 1982, vol. 129, pp. 127–133.

[3] A. Papoulis and H. Saunders, "Probability, Random Variables and Stochastic Processes (2nd Edition)", *J. Vib. Acoust*. 1989, vol. 111, pp. 123–125.

[4] K. Durga Rao, H.S.Kushwaha, A.K. Verma and A. Srividya, "Quantification of Epistemic and Aleatory Uncertainties in Probabilistic Safety Assessment Studies". *Rel. Eng. Syst. Safety* 2007, vol. 92, pp. 947–956.

[5] M. Yazdi, S. Kabirb and M. Walker, "Uncertainty handling in fault tree based risk assessment: State of the art and future perspectives". *Process Safety and Environmental Protection*. 131 (2019), pp. 89–104.

[6] M. Boufenar, "Treatment of uncertainties propagation in probabilistic safety analysis", *3rd International Conference on Radiations and Applications*, Algiers, (2022).

[7] S. C. Schwartz and Y. S . Yeh, "On the distribution function and moments of power sums with lognormal components", *Bell Syst. Tech. J.*, vol. 61, pp. 1441-1462, Sept. 1982.

[8] N. C. Beaulieu, A. A. Abu-Dayya and P. J. McLane, "Estimating the Distribution of a Sum of Independent Lognormal Random Variables", *IEEE Transactions On Communications*, Vol. 43, No. 12, Dec. 1995- 2869.

[9] A. B. El-Shanawany, K.H. Ardron and S.P. Walker, "Lognormal approximations of fault tree uncertainty distributions", *Risk Analysis* vol. 38 (2018) 1576-1584.