ROBUST PROBABILISTIC RISK ASSESSMENT WITH FAULT-TREES VIA AN ENTROPY-BASED AFFINE-INVARIANT STOCHASTIC MODEL UPDATING FRAMEWORK WITH PROBABILITY BOUNDS ANALYSIS

Adolphus Lye[1], Wasin Vechgama[2], Scott Ferson[3], Sicong Xiao[4]

[1] *Singapore Nuclear Research and Safety Institute, National University of Singapore: 16 Prince George's Park, #05-01 SNRSI Building, Singapore 118415, Singapore, Singapore. Email: snrltsa@nus.edu.sg*
[2] *Thailand Institute of Nuclear Technology: 9/9 Village no.7 Saimoon Sub-district, Ongkarak District, Nakhon Nayok Province 26120, Bangkok, Thailand. Email: wasin@tint.or.th*
[3] *Institute for Risk and Uncertainty, Department of Civil and Environmental Engineering, University of Liverpool, Chadwick Building, Liverpool L697ZF, Liverpool, United Kingdom. Email: ferson@liverpool.ac.uk*
[4] *Singapore Nuclear Research and Safety Institute, National University of Singapore: 16 Prince George's Park, #05-01 SNRSI Building, Singapore 118415, Singapore, Singapore. Email: snrxs@nus.edu.sg*

ABSTRACT

Fault-tree analysis is a widely adopted method for performing probabilistic risk assessment within the nuclear industry. However, its effectiveness is often constrained by the availability and quality of reliability data, especially concerning basic events. This limitation introduces significant uncertainty in the probability distributions used to model failure events. Furthermore, the assumption of independence among fault-tree events can result in an underestimation of the overall system risk. The paper proposes a two-step probabilistic framework aimed at improving fault-tree analysis under such uncertainties. The first step applies an entropy-based affine-invariant stochastic model updating scheme to construct probability-boxes over limited data. The second step propagates these probability-boxes through the fault-tree logic using probability bounds analysis, which explicitly incorporates dependency uncertainties. The result provides a more robust and conservative estimate of the top-event failure probability, enhancing the probabilistic risk assessment towards nuclear safety.

Keywords: Fault-tree, Probability-box, Bayesian model updating, Probability bounds analysis, Uncertainty analysis

I. INTRODUCTION

In recent years, nuclear energy has become an increasingly promising option on a global scale considering climate change and the need for de-carbonisation. Such an option, however, comes with significant risks as seen from the past severe nuclear accidents such as the Three Mile Island accident in 1979, the Chernobyl accident in 1986, and the Fukushima-Daiichi accident in 2011 [1]. This brings about the importance of nuclear safety, and therefore the need for probabilistic risk assessment. A standard tool for probabilistic risk assessment is fault-tree analysis. However, its effectiveness is often limited by two practical issues: 1) the polymorphic uncertainty associated with the limited availability of reliability data to characterise the failure probability distribution associated with a given basic event; and 2) the epistemic uncertainty over the dependency between the basic and the intermediate events of the fault-tree.

Hence, the research objectives and contributions of the paper are two-fold: 1) the proposal of a two-step approach towards performing a fault-tree analysis under limited data and uncertain event dependencies, which involves the entropy-based affine-invariant stochastic model updating framework along with probability bounds analysis. Such approach is yet to be presented within the existing literature; and 2) to implement and validate the proposed approach on a fault-tree model of a "reversed flow of the water inlet system" accident within the Thai Research Reactor-1/Modification 1 (TRR-1/M1) research reactor [2].

To realize the research objectives, the paper is structured as follows: Section II presents an overview of the proposed methodology. This includes a review of the Bayesian model updating theory, and the Boolean algebra under the independence assumption and uncertain dependencies between multiple events; Section III presents the application case study which introduces the TRR-1/M1 research reactor, the problem set-up, and the results and discussions; and Section IV provides a summary of the key findings and concludes the paper.

II. PROPOSED METHODLOGY

II.A. Stochastic Model Updating

II.A.1. Approximate Bayesian Computation

Bayesian model updating is a stochastic approach towards model updating to which the mathematical formalism follows:

$$P(\boldsymbol{\theta}|\boldsymbol{D}, M) = \frac{P(\boldsymbol{\theta}|M) \cdot P(\boldsymbol{D}|\boldsymbol{\theta}, M)}{P(\boldsymbol{D}|M)} \tag{1}$$

where $P(\boldsymbol{\theta}|M)$ is the prior distribution characterising the prior knowledge on the inferred parameter(s) $\boldsymbol{\theta}$ before collecting data $\boldsymbol{D}$, $P(\boldsymbol{D}|\boldsymbol{\theta}, M)$ is the likelihood function reflecting the degree of agreement between the observed data $\boldsymbol{D}$ and the model prediction from M given $\boldsymbol{\theta}$, and $P(\boldsymbol{D}|M)$ is the evidence which ensures that the posterior integrates to one. Details on the terms in Eq. (1) are found in [3] and [4]. As $P(\boldsymbol{D}|M)$ is a numerical constant, it can be neglected, and the posterior is re-expressed as:

$$P(\boldsymbol{\theta}|\boldsymbol{D}, M) \propto P(\boldsymbol{\theta}|M) \cdot P(\boldsymbol{D}|\boldsymbol{\theta}, M) \tag{2}$$

An important aspect in Bayesian model updating is the definition of the likelihood function $P(\boldsymbol{D}|\boldsymbol{\theta}, M)$, which at times may not be possible. In such case, the Approximate Bayesian Computation (ABC) approach is implemented as it provides a likelihood-free approach. For the work, the approximate Gaussian likelihood function is implemented [5]:

$$P(\boldsymbol{D}|\boldsymbol{\theta}, M) = \exp\left[-\frac{d^2}{\varepsilon^2}\right] \tag{3}$$

where d is the distance function, while ε is the width-factor acting as the pre-defined parameter controlling the centralisation degree of the posterior. It is proposed in [5] that the width-factor should lie within the interval of $[10^{-3}, 10^{-1}]$. For the work presented in the paper, the Jensen-Shannon divergence d_{JS} is implemented as the distance function which is defined as:

$$d_{\mathrm{JS}}(p_1, p_2) = \frac{1}{2} \cdot \left(d_{\mathrm{KL}}(p_1 \,||\, T) + d_{\mathrm{KL}}(p_2 \,||\, T)\right), \quad \text{for } \mathrm{T} = \frac{1}{2} \cdot (p_1 + p_2) \tag{4}$$

where:

$$d_{\mathrm{KL}}(p_1||p_2) = \sum_{x_d=1}^{N_{\mathrm{bin}}} \cdots \sum_{x_1=1}^{N_{\mathrm{bin}}} p_1\left(b_{x_1,\ldots,x_d}\right) \cdot \log\left[\frac{p_1\left(b_{x_1,\ldots,x_d}\right)}{p_2\left(b_{x_1,\ldots,x_d}\right)}\right] \tag{5}$$

for which N_{bin} denotes the total bin number used to approximate the distributions p_1 and p_2, and log is the natural logarithm (to the base e).
In the context of ABC, the interest would be to compute $d_{\mathrm{JS}}(p_M, p_{\boldsymbol{D}})$ where p_M is the distribution of the model prediction while $p_{\boldsymbol{D}}$ is the distribution of the observed data. Such distance function was first implemented for ABC in [7], and recently in [8] and [9].

An important aspect of the Jenson-Shannon divergence is the computation of the parameter N_{bin}. An approach to do so would be via the adaptive-binning algorithm, details on which the readers may refer to [10].

II.A.2. Transitional Ensemble Markov Chain Monte Carlo

An approach to sample from the posterior defined by Eq. (2) would be the Transitional Ensemble Markov Chain Monte Carlo (TEMCMC) method [11]. It is a variant of the Transitional Markov Chain Monte Carlo sampling technique proposed in

[12] which allows for the generation of samples from complex-shaped posteriors (e.g., very peaked or having multiple peaks) in an iterative manner. This is done using a series of intermediate functions known as transitional distributions P^j defined as:

$$P^j \propto P(\theta|M) \cdot P(D|\theta, M)^{\beta_j} \tag{11}$$

where $j \geq 0$ is the sampling iteration number, β_j is the tempering parameter such that $0 = \beta_0 < \beta_1 < \cdots < \beta_{m-1} < \beta_m = 1$, and m is the final iteration number. Readers may refer to [11] for further details on the algorithm and implementation of the TEMCMC sampler.

II.B. Probability Bounds Analysis With Uncertain Boolean Logic

For the research work, only the Boolean "AND" and "OR" logical operations of the fault-tree are of interest and, hence, discussed in the section. Consider n distinct events of the fault-tree denoted as x_i with an associated probability $P(x_i) = p_i$, for $i = 1, \dots, n$. In the interest of the research undertaken here, the probabilities p_i are characterised by a probability-box (p-box) [13] to express its associated polymorphic uncertainty such that: $p_i \sim [p_i^L, p_i^R]$, where p_i^L and p_i^R are the left and right-bounding distributional envelope of the p-box respectively.

Under the independence assumption, the resulting associated probability of the event defined by the Boolean "AND" logical operation between the events x_i follows [14]:

$$P(\wedge_{i=1}^n x_i) = \prod_{i=1}^n p_i = \left[\prod_{i=1}^n p_i^L, \prod_{i=1}^n p_i^R\right] \tag{12}$$

while the resulting associated probability of the event by the Boolean "OR" logical operation between the events x_i follows [14]:

$$P(\vee_{i=1}^n x_i) = 1 - \prod_{i=1}^n (1 - p_i) = \left[1 - \prod_{i=1}^n \left(1 - p_i^L\right), 1 - \prod_{i=1}^n (1 - p_i^R)\right] \tag{13}$$

In general, the individual events x_i may not be independent from one another. Instead, some form of dependency may exist between them. To provide for a relatively conservative yet robust risk analysis, the element of event dependency can be treated as an epistemic entity. Under the uncertain dependency between the n input p-boxes $P_i(x) \sim [P_i^L(x), P_i^R(x)]$ for $i = 1, \dots, n$, the output imprecise distribution is defined as:

$$P(x) = [P^L(x), P^R(x)] \tag{14}$$

Such that in the case of the Boolean "AND" logical operation between the events x_i, the bounds are [15]:

$$P^L(x) = \sup_{x = \wedge_{i=1}^n z_i} \left[\max\left(0, \sum_{i=1}^n P_i^L(z_i) - (n-1)\right)\right] \tag{15a}$$
$$P^R(x) = \inf_{x = \wedge_{i=1}^n z_i} \left[\sum_{i=1}^n P_i^R(z_i) - \max\left(0, \sum_{i=1}^n P_i^R(z_i) - (n-1)\right)\right] \tag{15b}$$

whereas in the case of the Boolean "OR" logical operation between the events x_i, the bounds follow [15]:

$$P^L(x) = \sup_{x = \vee_{i=1}^n z_i} \left[\sum_{i=1}^n P_i^L(z_i) - \min\left(1, \sum_{i=1}^n P_i^L(z_i)\right)\right] \tag{16a}$$
$$P^R(x) = \inf_{x = \vee_{i=1}^n z_i} \left[\min\left(1, \sum_{i=1}^n P_i^R(z_i)\right)\right] \tag{16b}$$

It is to be highlighted that the proposed conservative risk analysis under uncertain event dependencies is currently applicable when there are no repeated variables (i.e., repeated basic events of a fault-tree). Such challenge is discussed in [14] and remains an open research question.

The proposed methodology is outlined as follows: The entropy-based affine-invariant stochastic model updating framework is implemented to probabilistically update a given distribution model over an Empirical Cumulative Distribution

Function (ECDF) of p_i based on the limited experimental data. This yields a p-box over p_i to account for the epistemic uncertainty over the distribution shape parameters. This is done for each root event. From there, the resulting p-boxes across all the root events are propagated through the Boolean logic, while accounting for the uncertain event dependencies, to yield the output p-box for the Top event. In summary, a flow-chart is presented in Fig. 1.

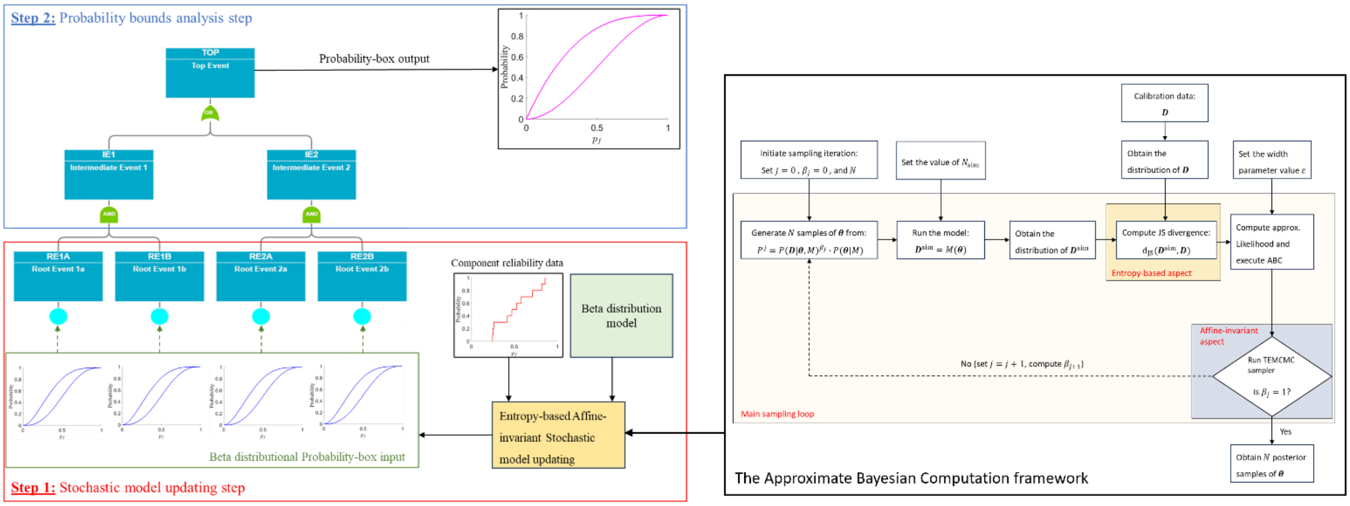

FIGURE 1. Flow-chart Of The Proposed Methodology

III. CASE STUDY - THAI RESEARCH REACTOR-1/MODIFICATION 1

The case study is based on the TRR-1/M1, which is a Training, Research, Isotopes, General Atomics nuclear research reactor located within Thailand, and is operated by the Thailand Institute of Nuclear Technology [2]. Since reaching criticality on 7-NOV-1977, the TRR-1/M1 nuclear research reactor operates at a normal operating power of 1MW with a maximum licensed power at 1.3 MW [2]. During normal operations, the research reactor generates radioisotopes for industrial, medical, and agricultural purposes. On top of that, these radioisotopes have also been used to conduct various beam experiments, neutron radiography, and the prompt-gamma neutron activation analysis. The schematic diagram of the research reactor is provided in Fig. 2.

III.A. Problem Set-up

The Top event of interest for the fault-tree analysis is the "reversed flow of the water within the inlet system", for which the imprecise probability distribution is to be obtained via the proposed methodology. The corresponding fault-tree is illustrated in Fig. 3, and details on the respective intermediate and basic events are presented in Table 1. The failure probability data for the associated basic events of the fault-tree is obtained from Vechgama et al. (2021) [2], and presented in Table 2.

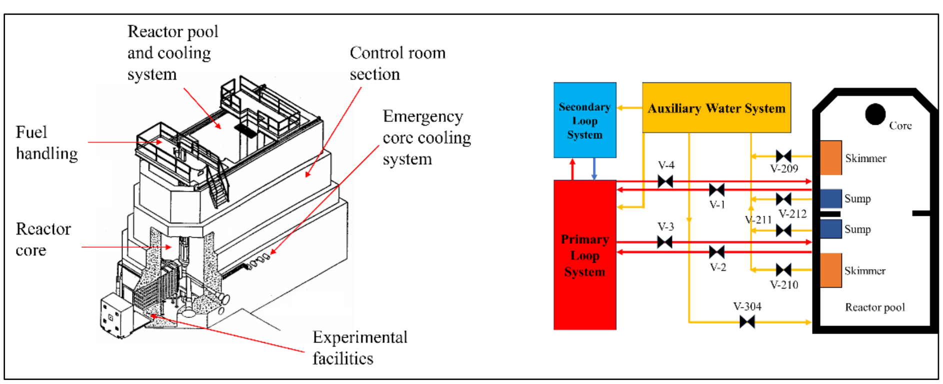

FIGURE 2. Schematic Diagram Of The TRR-1/M1 Research Reactor Adopted From Vechgama et al. (2021) [2].

In the case of the basic events 1, 3, 5, and 6, the distribution on the failure probability p_f refers to the stochastic variability (i.e., aleatory uncertainty) of the component's reliability [14]. To simulate the case where there is limited component reliability data due to the limited reliability/life tests, the work shall consider $N_{\text{obs}} = 20$ observations on the p_f for the aforementioned basic events. The resulting ECDF on the p_f for the corresponding basic events are presented in Fig. 5.

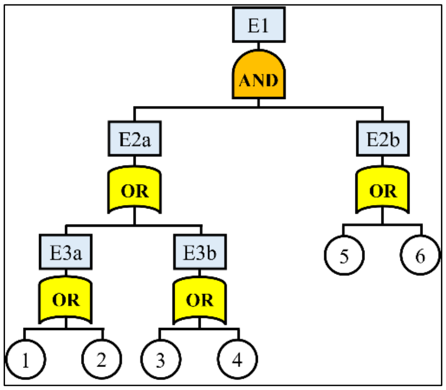

FIGURE 3. Fault-tree Diagram For The Top Event – Reversed Flow Of The Water Inlet System

TABLE 1. Details On The Respective Events Of The Fault-Tree Presented In Fig. 3

Intermediate events		Basic events	
Symbol	**Description**	**Symbol**	**Description**
E1	Reversed flow of the water within the inlet system.	1	Valve V-3 failed to open.
E2a	Inlet valve failed to open.	2	Operator failed to open valve V-3.
E2b	Inlet pipe breaks.	3	Valve V-4 failed to open.
E3a	Valve V-3 failed.	4	Operator failed to open valve V-4.
E3b	Valve V-4 failed.	5	Inlet pipe of V-3 breaks.
		6	Inlet pipe of V-4 breaks.

To calibrate a probability distribution over the component reliability data ECDF for basic events 1, 3, 5, and 6, a Beta distribution is chosen due to such distribution having defined bounds between 0 and 1, and having sufficient degrees of freedom in characterising different distributional shapes. Next, the entropy-based affine-invariant stochastic model updating framework is implemented to update the Beta distribution over the respective ECDF where the inferred parameters are as follows: $\boldsymbol{\theta} = \{\alpha_i, \beta_i\}$, for $i = 1, 3, 5, 6$. For each inferred parameter, the prior is defined by a Uniform distribution with the corresponding bounds defined in Table 3. The likelihood is defined by Eq. (3) with the corresponding width parameter ε defined in Table 3. The choice of ε is to ensure that the TEMCMC sampler samples from the posterior over 7 sampling iterations to achieve convergence over the posterior sample distribution.

TABLE 2. Failure Probability Data For The Associated Basic Events Of The Fault-tree Presented In Fig. 3

Symbol	Description	Distribution	Shape parameters
1	Valve V-3 failed to open.	Beta	$\alpha_1 = 1.5,\ \beta_1 = 43.00$
2	Operator failed to open valve V-3.	Fixed value	3.45×10^{-5} per reactor year
3	Valve V-4 failed to open.	Beta	$\alpha_3 = 1.5,\ \beta_3 = 43.00$
4	Operator failed to open valve V-4.	Fixed value	3.45×10^{-5} per reactor year
5	Inlet pipe of V-3 breaks.	Beta	$\alpha_5 = 1.5,\ \beta_5 = 62.40$
6	Inlet pipe of V-4 breaks.	Beta	$\alpha_6 = 1.5,\ \beta_6 = 62.40$

TABLE 3. Parametric Settings Implemented For The Stochastic Model Updating Step

Basic event	1	3	5	6
Prior bounds α	$[0.01, 100]$	$[0.01, 100]$	$[0.01, 100]$	$[0.01, 100]$
Prior bounds β	$[0.01, 100]$	$[0.01, 100]$	$[0.01, 100]$	$[0.01, 100]$

Width parameter, ε	0.015	0.015	0.015	0.010

III.B. Results And Discussions

The resulting posterior distribution for the respective inferred parameter is illustrated in Fig. 4, which is also interpreted as a fuzzy set [17]. From each fuzzy set, the resulting updated epistemic bound over the inferred parameter is obtained at an alpha-cut level of 0.8 to ensure a non-conservative coverage which encloses the true value. The resulting updated epistemic bound for the respective inferred parameter is presented in Table 4.

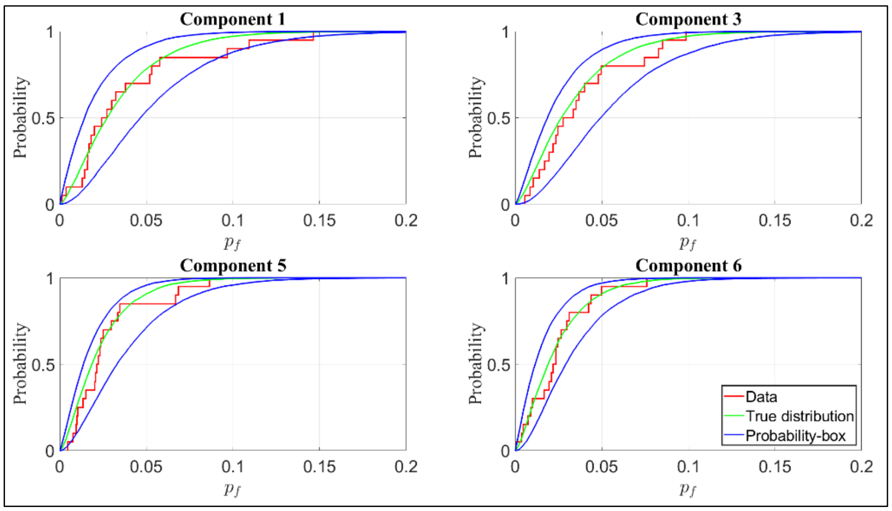

FIGURE 4. Resulting Beta Distribution P-box Over The ECDF Of The Component Reliability Data

TABLE 4. Results From The Stochastic Model Updating Step

Basic event	1	3	5	6
Updated bounds α	$[0.992, 1.950]$	$[1.326, 2.228]$	$[1.283, 1.843]$	$[1.129, 1.960]$
Updated bounds β	$[32.764, 49.039]$	$[36.713, 54.261]$	$[43.715, 72.931]$	$[54.029, 72.905]$

Based on the results in Table 4, a p-box is constructed over the failure probability p_f for the basic events 1, 3, 5, and 6. This is done via a Double-loop Monte Carlo approach where the outer loop generates $N_e = 1000$ samples from the epistemic bounds, while the inner loop generates $N_a = 10000$ samples from the resulting Beta distribution given the input realization from the outer loop. The creates a family of Beta distributions from which the distributional bounds are obtained and presented in Fig. 5. From there, the resulting p-box is propagated through the fault-tree in Fig. 3 under: 1) the independence assumption between the events (i.e., see Eq. (12) to (13)); and 2) under uncertain event dependency (i.e., see Eq. (15a) to (16b)). The results are illustrated in Fig. 6 which shows that under the independence assumption between events, the risk estimates of the Top event could be significantly underestimated, whereas that under the uncertain event dependency provides a conservative imprecise risk estimates but one that accounts for the worst-case scenario and the true risk of the Top event. Both resulting p-boxes enclose the true distribution (in black) as seen in Fig. 6 which verifies the applicability of the proposed methodology. The latter is obtained by propagating the true probability distribution and probability values (i.e., presented in Table 2) through the fault-tree under the independence assumption.

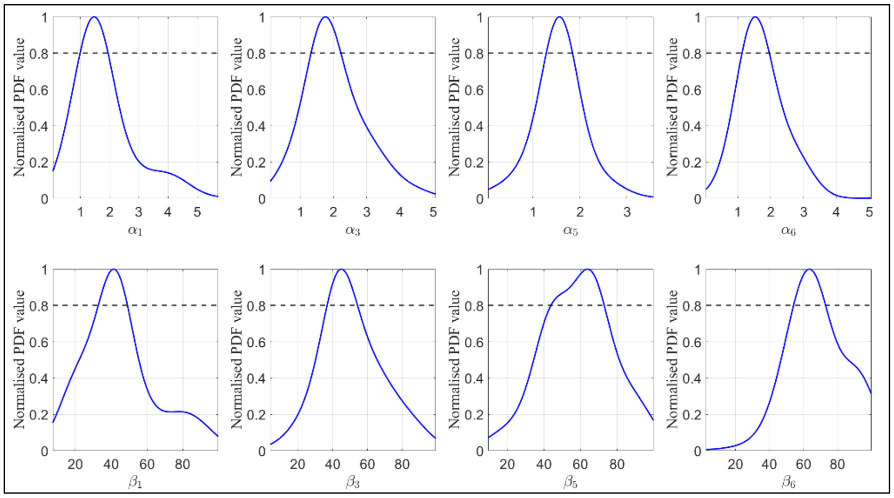

FIGURE 5. Resulting Posterior Distributions Over The Inferred Parameters

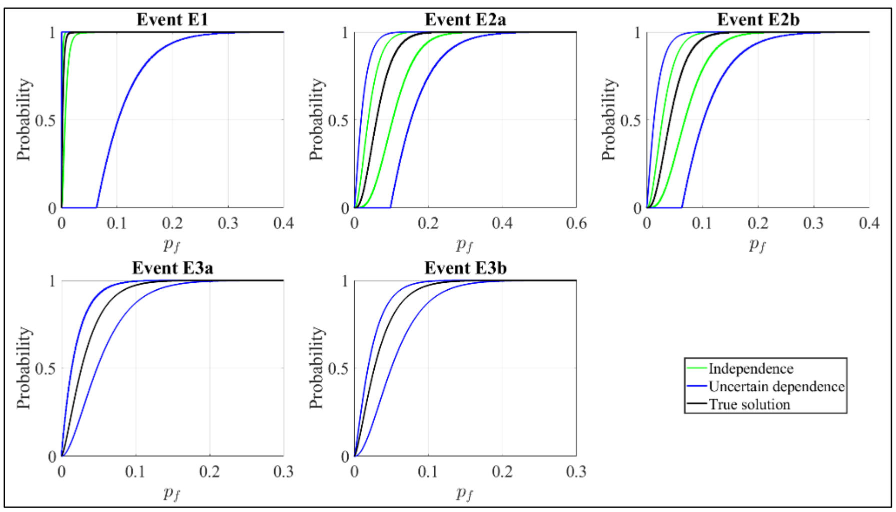

FIGURE 6. Resulting P-box Over The Top Event Under The Different Event Dependency Assumptions

IV. CONCLUSIONS

The paper proposed a two-step approach towards performing a robust fault-tree analysis under limited data and uncertain event dependence. The validity and feasibility of the proposed framework is demonstrated in computing the probability distribution for the reversed flow of the water within the inlet system within the TRIGA nuclear research reactor. The results support the hypothesis, and is shown to be consistent with what is expected in reality. A key selling point of the proposed methodology is that it is not reactor specific and can be applied towards any reactor design of interest. In fact, large conventional reactors such as pressurized water reactors can benefit from the proposed methodology given its relative system configuration complexity which presents significant epistemic uncertainty over the failure dependencies between its components. Future research efforts can be invested towards investigating the following: 1) a distribution-free approach towards performing such analysis with distribution-free probability boxes, aimed at eliminating the element of model form uncertainty; and 2) to investigate the simultaneous use of both confidence box and probability box to perform a fault-tree analysis, and provide a statistical interpretation over the resulting statistical structure associated with the Top event probability.

The MATLAB and R codes used to perform the analysis in the paper are made publicly available on GitHub via: https://github.com/Adolphus8/stochastic-model-updating.git

ACKNOWLEDGMENTS

The authors of the work would like to express their utmost gratitude to the University of Liverpool, the Thailand Institute of Nuclear Technology, and the Singapore Nuclear Research and Safety Institute for the support required to perform the necessary studies presented in the paper. In addition, the authors are grateful to the National Research Foundation Singapore (Funding no.: A-0001360-06-00) for their gracious funding, without which the research would not have been possible.

REFERENCES

[1] A. LYE, J. CHANG, S. XIAO, and K. Y. CHUNG, "An Overview of Probabilistic Safety Assessment for Nuclear Safety: What Has Been Done, and Where Do We Go from Here?", *Journal of Nuclear Engineering*, **5**, 456-485 (2024).

[2] W. VECHGAMA, K. SILVA, A. PECHRAK, and S. WETCHAGARUN, "Application of Hazard and Operability Technique to Level 1 Probabilistic Safety Assessment of Thai Research Reactor-1/Modification 1: Internal Events and Human Errors", *Progress in Nuclear Energy*, **138**, 103838 (2021).

[3] A. LYE, "Robust and Efficient Probabilistic Approaches towards Parameter Identification and Model Updating", *University of Liverpool Ph.D. Thesis Repository* (2023).

[4] A. LYE, S. FERSON, and S. XIAO, "Comparison between distance functions for Approximate Bayesian Computation towards Stochastic model updating and Model validation under limited data", *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems Part A: Civil Engineering*, **10**, 03124001 (2024).

[5] S. BI, M. BROGGI, and M. BEER, "The role of the Bhattacharyya distance in stochastic model updating", *Mechanical Systems and Signal Processing*, **117**, 437-452 (2019).

[6] F. NIELSEN, "On the Jensen–Shannon Symmetrization of Distances Relying on Abstract Means", *Entropy*, **21**, 485 (2019).

[7] L. YANG, S. BI, M. G. R. FAES, M. BROGGI, and M. BEER, "Bayesian inversion for imprecise probabilistic models using a novel entropy-based uncertainty quantification metric", *Mechanical Systems and Signal Processing*, **162**, 107954 (2022).

[8] A. LYE, S. FERSON, and S. XIAO, "Stochastic Model Updating Using The Jenson-Shannon Divergence For Calibration and Validation Under Limited Data", *34th European Safety and Reliability Conference: Advances in Reliability, Safety and Security*, Krakow, 23-27 June 2024, 1, 173-182 (2024).

[9] A. LYE, S. FERSON, and S. XIAO, "Distribution-free stochastic model updating for the Physics-guided reliability analysis of a material thermal property under limited data", *17th International Conference on Probabilistic Safety Assessment and Management & Asian Symposium on Risk Assessment and Management*, Sendai, 7-11 October 2024, 1 (2024).

[10] W. ZHAO, L. YANG, C. DANG, R. ROCCHETTA, M. VALDEBENITO, and D. MOENS, "Enriching stochastic model updating metrics: An efficient bayesian approach using Bray-curtis distance and an adaptive binning algorithm", *Mechanical Systems and Signal Processing*, **171**, 108889 (2022).

[11] A. LYE, A. CICIRELLO, and E. PATELLI, "An efficient and robust sampler for Bayesian inference: Transitional Ensemble Markov Chain Monte Carlo", *Mechanical Systems and Signal Processing*, **167**, 108471 (2022).

[12] J. CHING, and Y. C. CHEN, "Transitional Markov Chain Monte Carlo Method for Bayesian Model Updating, Model Class Selection, and Model Averaging", *Journal of Engineering Mechanics*, **133**, 816-832 (2007).

[13] S. FERSON, V. KREINOVICH, L. GRINZBURG, D. MYERS, and K. SENTZ, "Constructing probability boxes and Dempster-Shafer structures", *U.S. Department of Energy Office of Scientific and Technical Information*, 1 (2003).

[14] A. LYE, W. VECHGAMA, M. SALLAK, S. DESTERCKE, S. FERSON, and S. XIAO, "Advances in the Reliability Analysis of Coherent Systems under Limited Data with Confidence Boxes", *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems Part A: Civil Engineering*, **11**, 04024074 (2025).

[15] M. FRECHET, "Généralisation du théorème des probabilités totals", *Fundamenta Mathematicae*, **25**, 379-387 (1935).

[16] A. GRAY, A. WIMBUSH, M. DE-ANGELIS, P. O. HRISTOV, D. CALLEJA, E. MIRALLES-DOLZ, and R. ROCCHETTA, "From inference to design: A comprehensive framework for uncertainty quantification in engineering with limited information", *Mechanical Systems and Signal Processing*, **165**, 108210 (2022).

[17] A. LYE, M. KITAHARA, M. BROGGI, and E. PATELLI, "Robust optimization of a dynamic Black-box system under severe uncertainty: A distribution-free framework", *Mechanical Systems and Signal Processing*, **167**, 108522 (2022).