Nuclear Engineering and Technology 53 (2021) 357–367

Contents lists available at ScienceDirect

Nuclear Engineering and Technology

journal homepage: www.elsevier.com/locate/net

Original Article

Copula-based common cause failure models with Bayesian inferences

Kyungho Jin, Kibeom Son, Gyunyoung Heo*

Kyung Hee University, 1732, Deogyeong-daero, Giheung-gu, Yongin-si, Gyeonggi-do, 17104, South Korea

ARTICLE INFO

Article history:
Received 1 March 2020
Received in revised form
15 July 2020
Accepted 16 August 2020
Available online 23 August 2020

Keywords:
Common cause failures
Asymmetric conditions
Copula
Bayesian inferences

ABSTRACT

In general, common cause failures (CCFs) have been modeled with the assumption that components within the same group are symmetric. This assumption reduces the number of parameters required for the CCF probability estimation and allows us to use a parametric model, such as the alpha factor model. Although there are various asymmetric conditions in nuclear power plants (NPPs) to be addressed, the traditional CCF models are limited to symmetric conditions. Therefore, this paper proposes the copula-based CCF model to deal with asymmetric as well as symmetric CCFs. Once a joint distribution between the components is constructed using copulas, the proposed model is able to provide the probability of common cause basic events (CCBEs) by formulating a system of equations without symmetry assumptions. In addition, Bayesian inferences for the parameters of the marginal and copula distributions are introduced and Markov Chain Monte Carlo (MCMC) algorithms are employed to sample from the posterior distribution. Three example cases using simulated data, including asymmetry conditions in total failure probabilities and/or dependencies, are illustrated. Consequently, the copula-based CCF model provides appropriate estimates of CCFs for asymmetric conditions. This paper also discusses the limitations and notes on the proposed method.

1. Introduction

Nuclear power plants (NPPs) employ the concept of redundancy through functional similarity or identity of components to improve safety. The failure probability of a redundant system with components A and B can be expressed as $P(A \cap B)$, and it can be quantified as $P(A) \cdot P(B)$ if they are independent. However, quite a few redundant components in NPPs have dependencies among them due to functional and environmental root causes or coupling mechanisms. To evaluate their dependency in probabilistic safety assessments (PSAs), common cause failures (CCFs) have been introduced.

A CCF that significantly affects the safety of a redundant system is a dependent failure in which two or more components fail at the same time or in a short time due to common causes. There are various methods to estimate the probability of CCFs. The most widely used CCF models are based on parametric approaches such as the alpha factor model (AFM) [1,2]. The alpha parameter for *k*-failure indicates the ratio of *k*-failures in total failures. The probability of common cause basic events (CCBEs) for *k*-failures can be evaluated using the alpha factor for *k*-failure and the total failure probability of the component. The parametric CCF model assumes that the components in a common cause component group (CCCG) are symmetric. In other words, the probability of CCBEs for *k*-failure within a CCCG is always the same. This assumption reduces the number of parameters required for the CCF probability estimation and allows use of the parametric model.

On the other hand, there are various asymmetric conditions in NPPs and this complex situation is extended as the scope of PSAs expands into multiple units of a site. For example, the dependency of emergency diesel generators (EDGs) in multiple units should be considered carefully when their symmetry cannot be guaranteed. In practice, there have been several studies of cases where a symmetry assumption is not satisfiable. IAEA TECDOC-648 [2] refers to the functional, operational, and environmental asymmetric conditions that may occur in NPPs and NUREG/CR-5485 [1] describes the CCF of component cooling water (CCW) pumps under different operation modes. Kang et Al [3]. proposed a way to estimate the probability of asymmetric CCFs by approximately formulating the CCF, which is decomposed into primary and secondary groups. Rasmuson et al. [4] illustrated the asymmetric conditions in total failure probability due to degradations, and O'Connor et al. [5] proposed the general dependency model (GDM) using a Bayesian network to overcome the limitations of traditional CCF models,

* Corresponding author.
E-mail address: gheo@khu.ac.kr (G. Heo).

https://doi.org/10.1016/j.net.2020.08.014

such as non-identical components.

Most research on asymmetric CCFs, as described above, is based on the parametric method. However, constraints may arise in dealing with asymmetry problems through the parametric approaches because they are basically developed based on the symmetry assumption. For example, it is difficult to handle asymmetries in both total failure probabilities and dependencies. Different components/operation modes cannot be grouped as the same CCCG. Therefore, this paper has tried to manage the symmetry and asymmetry in CCFs using the concept of the multivariate probability distribution of components using copulas. A copula is a multivariate probability distribution with uniform marginals to model dependencies among random variables [6–8]. It is widely used in many research areas that need to model the dependencies of multivariates, such as reliability engineering and hydrology [9–12]. If the joint distribution of CCFs can be constructed using copulas, the dependent failure probability will be logically estimated without the symmetry assumption.

However, the field data needed to construct the multivariate probability distribution are practically limited. It is difficult to secure plant-specific data because failure events are rare; therefore, data from various sources (industrial plants) have been collected and evaluated as generic data. This involves a qualitative and quantitative evaluation based on expert judgments or reference guidelines [1,2,5]. In addition, the failure data have been integrated through mapping-up/-down techniques [1] when the number of components in a redundant system varies. However, the data manipulated and merged from various sources cannot capture the specific features of components, such as asymmetry.

Therefore, this paper describes CCFs using copulas under the assumption that plant-specific data, including asymmetric conditions, can be collected and without considering the existing failure data. This assumption is applicable even when the existing data can be adjusted to build the probability distributions. Accordingly, this paper uses simulated data to validate the proposed method. In addition, Bayesian inference [11–14] for the parameters of marginal and copula distributions is employed to statistically back up their uncertainty, because not much data are available in practice. Markov Chain Monte Carlo (MCMC)-Metropolis Hastings (MH) algorithms are used to sample from the posterior distribution. Consequently, the multivariate probability distribution of the components is constructed using mean values of the posterior distributions for each parameter instead of prediction distributions for simplicity, and then the failure probability of each component is decomposed into CCBEs for *k*-failure [15,16].

2. Common cause failures

2.1. Symmetry assumptions in CCFs

The fault tree for two component failures, with one out of two success logic, is shown in Fig. 1. The failure of A (or B) is composed of an independent failure and a two-component failure due to common causes.

The total failure probability of A is given in Eq. (1):

$$P(A) = Q_T^A = Q_1^A + Q_2^{AB} \tag{1}$$

where Q_T^A is the total failure probability of A, Q_1^A is the independent failure probability, and Q_2^{AB} is the failure probability of both A and B due to common causes. As shown in Eq. (1), the total failure probability can be represented in CCBEs, and thus the analyst needs to know their probability rather than the total failure probability to quantify the system failure through minimal cut sets (MCSs). Furthermore, the symmetry assumption leads to simple formulations as follows:

$$Q_T^A = Q_T^B = Q_T \tag{2}$$

$$Q_1^A = Q_1^B = Q_1 \tag{3}$$

$$Q_2^{AB} = Q_2 \tag{4}$$

Consequently, the analyst requires only the probability of Q_1, Q_2 to quantify this redundant system. It has the advantage of reducing the number of parameters required for analysis as the number of components increases. The symmetric condition also allows the alpha factors to take α_1, α_2, which are not specific to each component, to estimate Q_k. As a result, the failure probability of *k*-component failures can be calculated by Eq. (5) in the case of a non-staggered testing scheme.

$$Q_k = \frac{k}{\binom{m-1}{k-1}} \frac{\alpha_k}{\alpha_T} Q_T \tag{5}$$

where $\binom{m-1}{k-1}$ is a binomial coefficient, α_k is the alpha factor for k-failure, and α_T is equal to $\sum_{k=1}^{m} k\alpha_k$. Accordingly, the CCBE probabilities for k-failure can be estimated when the alpha factor and the total failure probability are given. The alpha factors, regardless of test strategy, are evaluated from Eq. (6):

$$\alpha_k = \frac{n_k}{\sum_{i=1}^{n} n_i} \tag{6}$$

where n_k is the number of k-component failures and $\sum_{i=1}^{n} n_i$ is the summation of the total number of failures. The number of component failures n_k could be assessed through the impact vector involving a qualitative and quantitative analysis by experts because failure data should be classified into independent failure or dependent failure to count n_k. For example, there is a two-component failure data. If the expert concluded that this data was caused independently, then n_1 becomes two and n_2 is equal to zero. Otherwise, it the expert decided that the data was caused totally dependently, then n_1 is zero and n_2 is equal to one. In practice, the expert gives weights to the failure data to determine whether they are independent or dependent failure. The details of impact vector analysis are described in Ref. [1,2].

2.2. Asymmetry in CCFs

The symmetry assumption is quite useful and widely applicable, but the particular cases where an asymmetric condition occurs should also be addressed. These non-identical conditions in CCFs have already been studied, including three examples of asymmetric conditions that could occur in NPPs as described in Ref. [2]. One example is functional asymmetry, as when four EDGs are supported by three emergency service water (ESW) pumps, and these pumps are also supplied by one EDG. If three EDGs fail to start, the operator connects the last EDG to the ESW pumps to provide electrical power. In this case, the symmetry conditions cannot be guaranteed. The second example is environmental asymmetry. Fig. 2 shows two valves located inside containment and the others outside. Even if all the design and maintenance features are the same, asymmetry in dependencies or failure probabilities could occur. The last example case in Ref. [2] is operational asymmetry. In BWR, many safety relief valves are installed. One-third of the valves are disassembled and

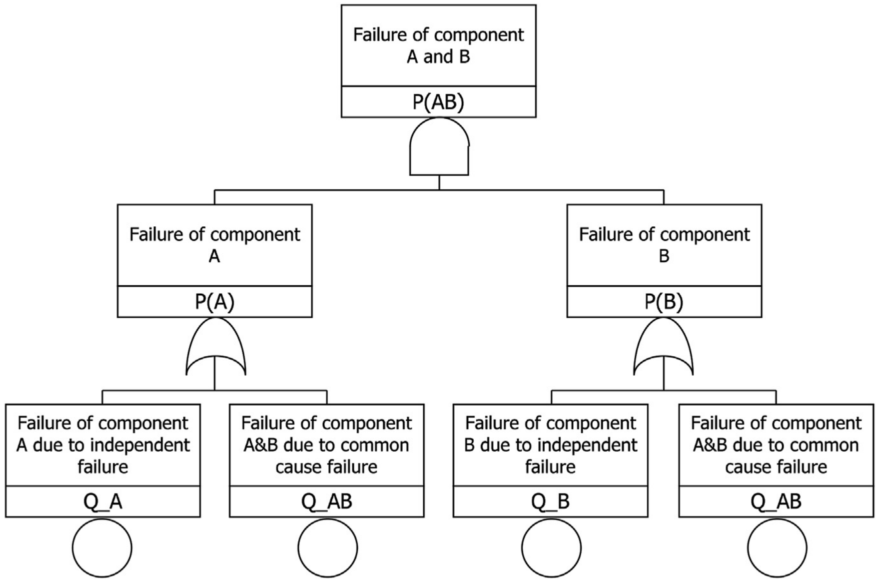


Fig. 1. The fault tree of two component failures, with one out of two success logic.

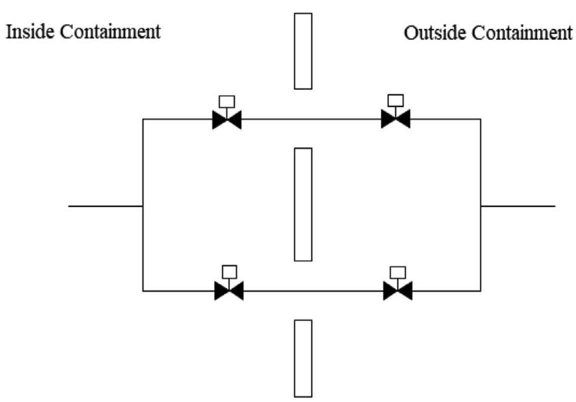


Fig. 2. An example of environmental asymmetry in CCFs [2].

reassembled at every refueling outage, hence, they might have different degradation levels or dependencies.

U.S. NRC [1] refers to the asymmetry of CCW pumps in different operation modes. Although they are regarded as the same CCCG, non-symmetric conditions could occur between the standby pump and the operation pump. It is also possible to have different total failure probabilities or dependencies. This asymmetry has been studied by adding an asymmetric basic event to the existing fault tree of the CCW pump. However, the method of estimating the probability of an asymmetric basic event is not explained in the reference.

Kang et al. [3] suggest approximate formulas to describe asymmetric components, such as EDGs and alternate AC DGs (AAC DGs), by the decomposition approach. The total failure events are decomposed into primary and secondary groups to deal with both symmetric and asymmetric conditions. The parameter estimation is done by approximating the formulas of the alpha factors and the basic parameter models.

Rasmuson et al. [4] explain unequal total failure probabilities. This asymmetry occurs when one component is degraded and this results in having different failure probabilities across the components. The problem can be solved by assuming the smallest failure probability between the components as a total failure probability based on the concept of the Fréchet-Hoeffding upper bound. In other words, it takes the assumption that the degradation impacts only the independent failures, not the CCFs.

O'Connor et al. [5] proposed the GDM using a Bayesian network to overcome the limitations of the traditional CCF models. GDM is based on the cause condition probability, the component fragility, and the coupling factor strength, which can be interpreted as physical features of the system. Using these parameters and the Bayesian network model, restricted conditions such as non-identical components can be considered.

All of these researches have tried to resolve asymmetric conditions in CCFs while keeping the parametric approach for estimating probabilities. However, restrictions or constraints for asymmetric CCFs still remain because the parametric models of CCFs are basically developed based on the symmetric assumption. For example, it is difficult to handle asymmetries in both total failure probabilities and dependencies. Different components or operation modes cannot belong to a same CCCG. Therefore, in this paper, copula-based CCF models are proposed to model the dependent failures of components without symmetry assumptions. The multivariate probability distribution of components in a CCCG, including their dependency, is constructed for estimating the CCF probabilities. The asymmetry in both total failure probabilities and dependencies can be measured since the proposed model is based on the probability distribution. In addition, different operation modes or components can be analyzed within this framework because a copula allows the uses of various marginal distributions.

3. Copula and Bayesian approach

3.1. Copulas

A copula is a multivariate probability distribution with uniform marginal distributions, used to model dependency between random variables. It is widely used in many research areas [9–12]. By Sklar's theorem [6–8], a multivariate probability distribution is given in Eq. (7) through a copula and cumulative distribution functions of each random variable:

$$F(x_1, x_2, \ldots, x_k) = C(F(x_1), F(x_2), \ldots, F(x_k); \theta) \quad (7)$$

where x is a random variable, $F(x)$ is a cumulative probability distribution, and C_θ is a copula with a copula parameter θ. Each cumulative distribution function has a value from zero to one as if U (0,1). The advantage of using copulas is that it can be represented by various marginal probability distributions through the probability integral transform. Note that the expression in Eq. (7) is unique only when the random variables are continuous. If we denote $F(x_k) = u_k$ and differentiate Eq. (7) for each random variable, a multivariate density function is given in Eq. (8):

$$\begin{aligned} f(x_1, x_2, \ldots, x_k) &= \frac{\partial^2 C}{\partial u_1 \partial u_2 \ldots \partial u_n} \frac{dF_{x_1}}{dx_1} \frac{dF_{x_2}}{dx_2} \cdots \frac{dF_{x_k}}{dx_k} \\ &= c(u_1, u_2, \ldots u_k; \theta) f_1(x_1) f_2(x_2) \ldots f_k(x_k) \end{aligned} \quad (8)$$

One type of copula is the elliptical copula such as a normal and a t-copula. There is also the Archimedean copula family, which includes Frank, Gumbel, and Clayton copulas. The correlation between variables is determined through a copula parameter θ. Table 1 shows the typical copula distributions for bivariate variables and their scatter plots at certain copula parameters.

While normal and Frank copulas have symmetric structure in their distributions, Clayton and Gumbel copulas have an asymmetric structure as shown in Table 1. Although Table 1 represents copula functions for bivariate variables, it is not difficult to expand to n-random variables. For example, an n-variate Clayton copula can be written as $(u_1^{-\theta} + u_2^{-\theta} \ldots + u_n^{-\theta} - n - 1)^{-1/\theta}$. Since the Archimedean copula generally has a single parameter, a vine copula and the hierarchical Archimedean approach can be used [6–8] to assign various dependencies between variables.

3.2. Bayesian inferences for parameters

The basic concepts of the Bayesian approaches supporting copula-based CCFmodels are introduced here. The Bayesian inference treats a parameter as a random variable of a distribution (Parameters of this distribution is called hyper parameters). For this reason, it has the advantage of analyzing uncertainties of parameters with combining prior knowledge. In general, the alpha factors are updated through the posterior Dirichlet distribution [1]. Component unreliability data also employed the Bayesian statistics using their conjugate prior distribution [17]. Due to the feature of copula mentioned in Section 3.1, the prior knowledge of component unreliability data can be used as the prior of marginal distributions in a joint probability distribution. Therefore, the Bayesian approach was selected to infer copula parameters as well as parameters of marginal distributions. First, the posterior distribution can be defined using likelihoods and prior distributions given in Eq. (9) [6]:

$$\pi(\theta|x) = \frac{\pi(\theta) f(x|\theta)}{\int \pi(\theta) f(x|\theta) dx} \quad (9)$$

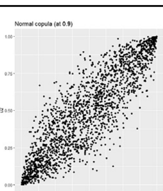
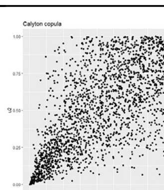
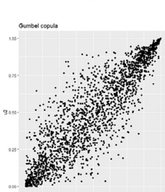
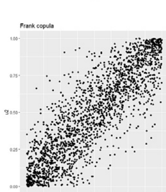
where θ is the parameter of the probability distribution and x is the observations. The posterior distribution $\pi(\theta|x)$ is determined using the prior distribution $\pi(\theta)$ and the likelihood function $f(x|\theta)$.

It is known that the use of conjugate prior distributions makes it easier to estimate a posterior distribution. For example, a Poisson distribution as a likelihood function for observations (n, t), and a gamma prior distribution $Gamma(\alpha, \beta)$ leads to a posterior distribution, $Gamma\ (\alpha + n,\ \beta + t)$.

If it is possible to sample random variables (i.e., independent and identically distributed, i.i.d) from a posterior distribution, Monte Carlo sampling instead of conjugate distributions can be used to infer all quantities of interest for a posterior distribution. When the sampling cannot be directly obtained from a posterior distribution, Markov Chain Monte Carlo (MCMC) is applied to draw random variables. The widely used sampling techniques for MCMC are of two types, Gibbs sampling [18] and Metropolis-Hastings (MH) sampling [19]. Gibbs sampling produces random samples using a full conditional distribution of each parameter rather than a posterior distribution. MH employs a proposal distribution, which can easily generate random numbers, instead of a posterior distribution. This paper uses MCMC-MH sampling algorithms; their simple descriptions and summaries of the algorithms are described below.

The algorithm starts from an arbitrary initial value of a parameter denoted by $\theta^{(0)}$. MCMC-MH proposes a new parameter value by a proposal distribution as a candidate. One of the proposal distributions is a normal distribution, which features a random walk chain and symmetrical structures around zero. The new value (candidate) generated from the proposal (t+1) is compared with the value in the previous step (t) through the acceptance probability (p) given in Eq. (10). This probability determines whether or not to accept the candidate value. Let us suppose that $\theta^{(t)}$ is a

Table 1
Typical copula distributions for bivariate random variables.

$C(u_1, u_2; \theta)$	Scatter plot	$C(u_1, u_2; \theta)$	Scatter plot
Normal copula $\int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{exp\left(\frac{2\theta sw - s^2 - w^2}{2(1-\theta^2)}\right)}{2\pi\sqrt{1-\theta^2}} dsdw$		**Clayton copula** $(u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}$ for $\theta > 0$	
Gumbel copula $\exp - [(-lnu_1)^\theta + (-lnu_2)^\theta]^{1/\theta}$		**Frank copula** $-\frac{1}{\theta}\ln[1 + (e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1) / (e^{-\theta} - 1)]$	

current state and $\theta^{(t+1)}$ is a candidate state. Then $\theta^{(t+1)}$ is generated from the proposal distribution around the current state $\theta^{(t)}$ with a standard deviation σ^2 in the case of a normal distribution. Then the acceptance probability (p) is given as:

$$p = min\left(\frac{\pi(\theta^{(t+1)})}{\pi(\theta^{(t)})}, 1\right) \tag{10}$$

where, in this paper, $\pi(\,)$ is a posterior distribution as a target distribution. The acceptance probability means which candidate has a higher possibility in the target distribution. If the acceptance probability is larger than 1 (i.e.$\pi(\theta^{(t+1)}) > \pi(\theta^{(t)})$), the candidate value is always accepted (See Fig. 3). In this case, the candidate is more likely to be a sample from the target distribution. If $\theta^{(t+1)}$ has a lower probability than $\theta^{(t)}$ (See Fig. 4), then $\theta^{(t+1)}$ is accepted or rejected randomly; if p is larger than a probability generated from U (0,1), then the candidate value $\theta^{(t+1)}$ is accepted. Otherwise, $\theta^{(t+1)}$ is rejected and the previous $\theta^{(t)}$ is held. Figs. 3 and 4 shows the simple example of a MCMC-MH procedure with an acceptance probability in case of a standard normal distribution, N (0,1) as a target distribution.

At first, the proposal distribution draws the sample ($\theta^{(1)}$) around $\theta^{(0)}$ with the variance 2.5. Then the acceptance probability in Eq. (10) can be calculated using the target distribution. In this case, $\theta^{(1)}$ has a higher possibility than $\theta^{(0)}$ in the target distribution (i.e. p is larger than 1). Thus, $\theta^{(1)}$ is accepted as a sample of the target distribution.

In the next step, the proposal distribution generates the sample ($\theta^{(2)}$) around $\theta^{(1)}$. In this case, $\theta^{(2)}$ has a lower possibility than $\theta^{(1)}$ in the target distribution, therefore $\theta^{(2)}$ is accepted or rejected randomly. The chain will converge when there are enough Markov chains from the initial conditions. However, the variance of the proposal distribution should be carefully determined because it affects the jumping size to explore a sample space. If the variance (or jumping size) is too high or too low, the MCMC-MH explores the sample space of the target distribution inefficiently. In other words, this jumping size is significantly related with the convergence rate of MCMC-MH. One way of the convergence diagnosis for determining a proper jumping size is to check the acceptance rate. The acceptance rate is the fraction of the accepted candidate. The ideal acceptance rate in case of univariate normal distribution is about 50% [20]. Finally, the samples generated initially (called *burn-in*) will be eliminated to reduce the effects of the initial conditions. The overall algorithms of MCMC-MH are given in Table 2.

4. Copula-based CCF models

4.1. Posterior distributions of copula-based CCF models

In this section, the estimation scheme of CCF probabilities was proposed via copulas instead of the alpha factor model described in Section 2.1 to address the asymmetric conditions. The study of copulas with regards to CCFs has already been studied in Ref. [9]. This paper developed a practical CCF model using copulas to solve the asymmetric problem using the decomposition technique [16] for applying to the fault tree of the PSA model. In addition, the reliability data [17] can be used by employing the Bayesian approach as a prior knowledge. Assuming that the failure data under asymmetric conditions can be collected, the posterior distribution of the joint probability distribution for the components to be analyzed is firstly constructed. The likelihood function, which is the multivariate probability density function using the copula in Eq. (8), and the prior distribution $\pi(\theta_j)$ of each parameter θ_j, formulates

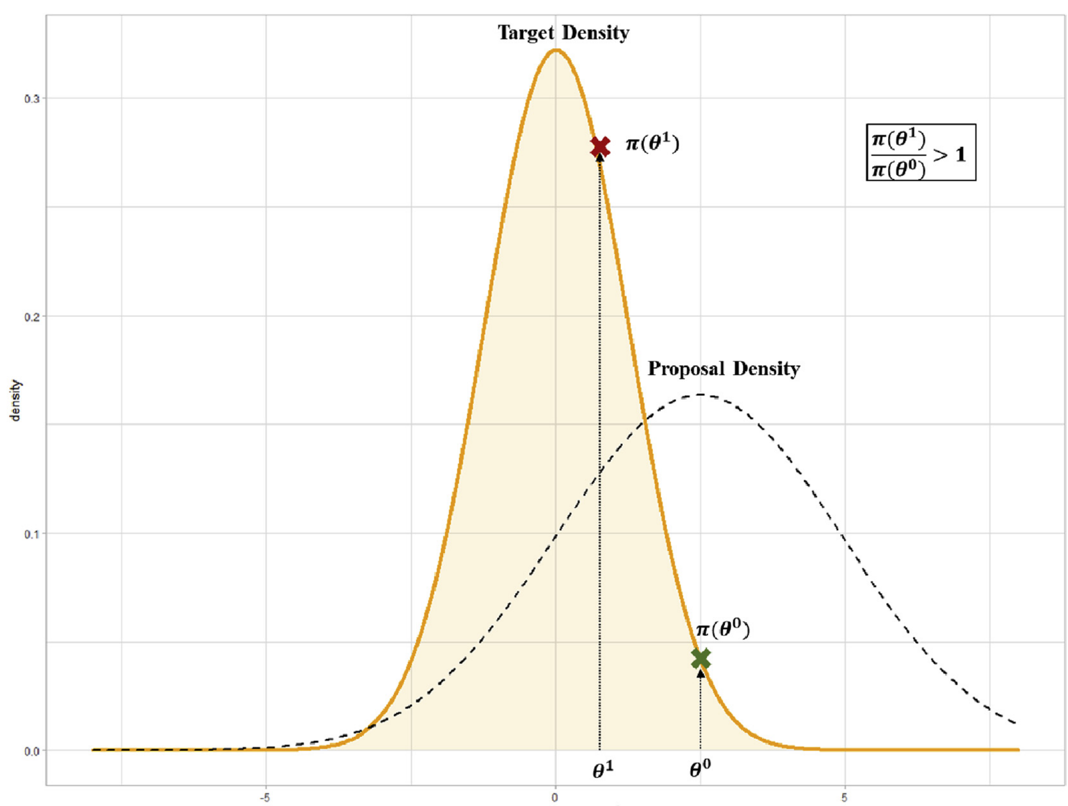


Fig. 3. The first candidate of a parameter ($\theta^{(1)}$) generated by the proposal distribution around the initial value ($\theta^{(0)}$).

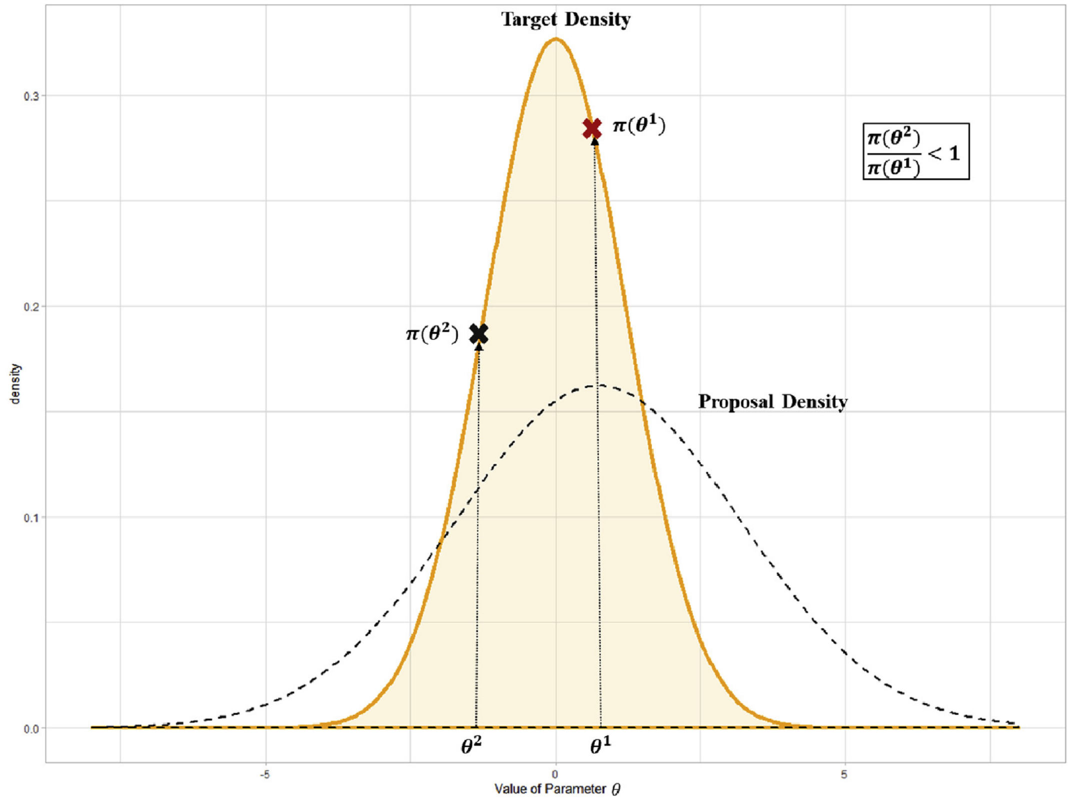


Fig. 4. The next candidate of a parameter ($\theta^{(2)}$) by the proposal distribution around the accepted parameter ($\theta^{(1)}$).

Table 2
MCMC-MH algorithms with a normal distribution as the proposal distribution.

Step 1	Set initial values for $\boldsymbol{\theta}^{(0)}$, $\boldsymbol{\sigma}^2$ and initialize $\boldsymbol{n^{accept}}$
Step 2	For $t = 1$ to N
	Generate $\theta^{(t+1)} \sim N(\theta^{(t)}, \sigma^2)$ and $u \sim U(0,1)$
	If $u < p$ then
	$\theta^{(t+1)} = \theta^{(t+1)}$
	$n^{accept} = n^{accept} + 1$
	$Acceptance\ rate = \frac{n^{accept}}{N}$
	Else
	$\theta^{(t+1)} = \theta^t$
Step 3	Repeat *Step 2* for iterations (N) or until convergences

the posterior distribution given in Eq. (11) [11,13,14]:

$$\pi(\boldsymbol{\Theta}|\boldsymbol{X}) \propto \prod_{j}^{m} \pi(\theta_j) \prod_{i=1}^{n} c(u_{1i}, u_{2i}, \ldots u_{ni}|\boldsymbol{\Theta}) f_1(x_{1i}|\boldsymbol{\Theta}) \times f_2(x_{2i}|\boldsymbol{\Theta}) \ldots f_k(x_{ki}|\boldsymbol{\Theta}) \quad (11)$$

where $\boldsymbol{\Theta}$ is the vector for m parameters ($\boldsymbol{\Theta} = \theta_1, \theta_2 \ldots \theta_m$), $\boldsymbol{X} = x_{1i}, x_{2i} \ldots x_{ki}$ for k random variables and i-th observations, and $u_{1i} = F_1(x_{1i}|\boldsymbol{\Theta})$. MCMC-MH allows us to draw random variables from the posterior distribution given in Eq. (11).

Because the actual failure data is not available, we generated 1000 sets of simulation data for the bivariate case, with exponential marginal distributions and a Clayton copula to verify how the copula-based CCF model with Bayesian inferences works. The multivariate data were generated using the R-programming package, *copula*; each exponential marginal had the parameter value of 0.005/hr (i.e., $\lambda_1 = \lambda_2 = 0.005$/hr) and the copula parameter (θ) was assumed to be 0.1. In this paper, it was also assumed that the prior distribution of all the parameters is a uniform distribution $\pi(\theta_j) \sim U(0,1)$.

Once the data are generated (or collected in practice), the type of marginal and copula in the likelihood function should be determined through various model selection methods [6,21]; however, this step was not considered in this paper because the simulation data were used. Accordingly, the same marginal and copula distribution as the generated data (i.e., exponential and Clayton copula) were selected to construct the likelihood function in Eq. (11). A total of 50,000 iterations of MCMC-MH were carried out and the initial 20,000 iterations were trimmed off. Fig. 5 shows the results of sampling for three parameters.

The first row in Fig. 5 presents the histograms of the posterior distribution of each parameter. The dotted lines in Fig. 5 indicate the mean value of parameters. The second row is the trace plots that can check the convergence of the MCMC. All the parameters are converged around their respective values; this means the copula-based CCF model with Bayesian inferences worked well. The various example cases with the simulation data will be described in Section 5.

4.2. Probability decomposition into CCBEs

If the multivariate probability distribution is constructed using the parameters after Bayesian inferences, it is possible to calculate the failure probability of each component, such as P(A), P (AB). However, the ultimate probability that we want to know is the probability of CCBEs (Q_1^A, Q_2^{AB}) rather than P(A) or P (AB). Therefore, the probability of each component should be decomposed into the CCBEs. A method that allocates the probabilities of correlated failures into CCF probabilities is well explained in Ref. [16]. For the bivariate cases in this paper, the system of equations of each failure

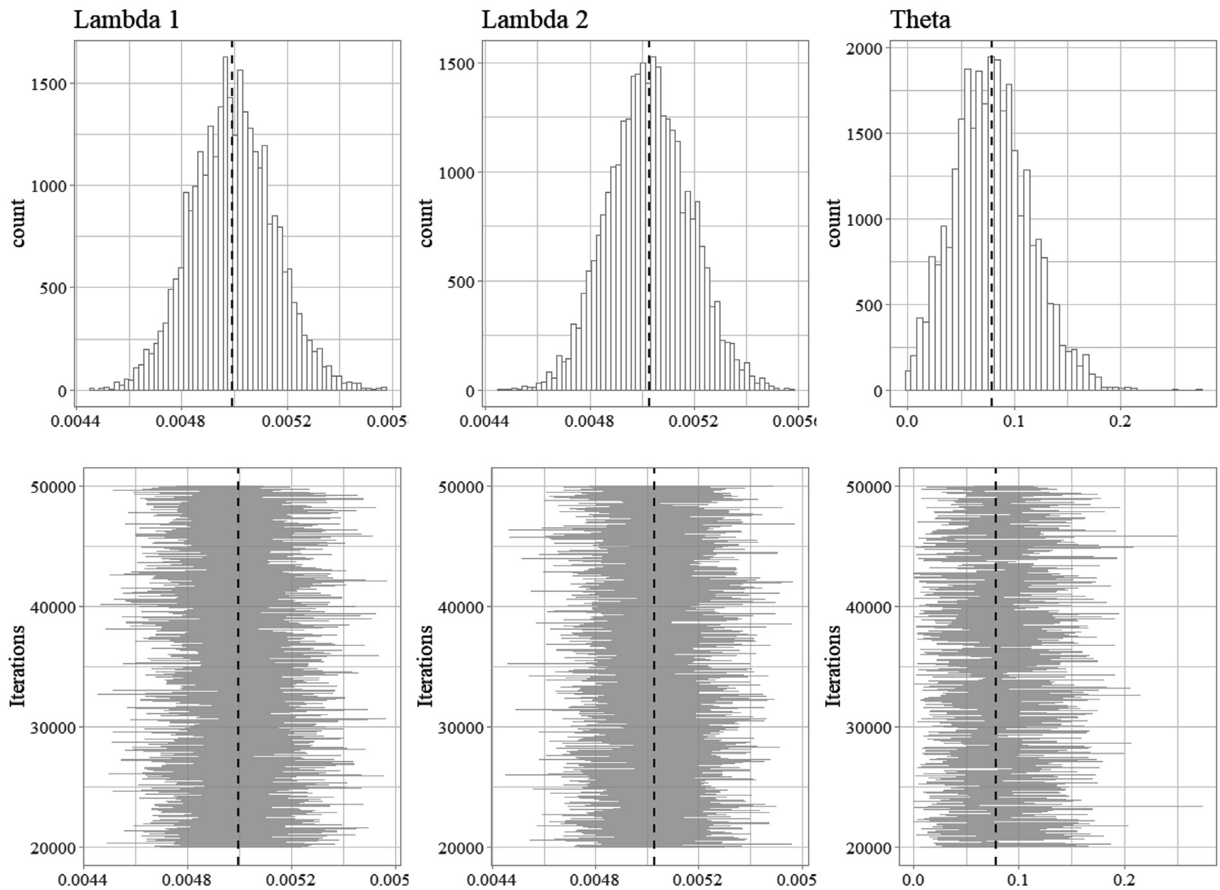


Fig. 5. The sampling results from the posterior distribution using MCMC-MH algorithms.

probability with respect to the CCBE representations is as follows:

$$\mathrm{P(B)} = Q_1^B + Q_2^{AB} \tag{13}$$

$$\mathrm{P(AB)} = Q_1^A Q_1^B + Q_2^{AB} \tag{14}$$

Thus, the CCBE probabilities without a symmetry assumption can be calculated by solving the system of equations using Eqs. (1), (13) and (14) because there are three equations and three unknowns for the bivariate cases. The system of equations can be established regardless of the number of components. Once the probability of CCBEs is given, it is available to put into the fault tree, as shown in Fig. 1. Fig. 6 shows the stepwise procedure of the copula-based CCF models described in Sections 3 and 4.

4.3. Comparisons with the alpha factor model

In this section, the results of copula-based CCF model were compared with the alpha factor model to validate the proposed method using the given simulation data (Section 4.1). The copula-based CCF model used the mean value of the sampling results in Fig. 5 to calculate the failure probability of each component (e.g. P(A), P(B), P (AB)). The failure probabilities were calculated through a bivariate Clayton copula in Table 1. For example, the mean value of $\overline{\lambda}_A$, $\overline{\lambda}_B$ is 5.0E-03 and $\overline{\theta}$ is 7.7E-02 as shown in Fig. 5. Thus, P (AB) can be calculated by the following equation:

$$F(t_A < T, t_B < T) = \left(\left(1 - e^{-\overline{\lambda}_A T}\right)^{-\overline{\theta}} + \left(1 - e^{-\overline{\lambda}_B T}\right)^{-\overline{\theta}} - 1\right)^{-1/\overline{\theta}} \tag{15}$$

In this case, T is assumed to be 24 h. Finally, the failure probability of each component was decomposed into CCBEs as shown in Table 3.

On the other hand, the number of component failures for the alpha factor model was calculated depending on the cases. The first case (denoted by *Alpha*_min in Table 3) included the assumption that all failures of two components are independent and the second case (denoted by *Alpha*_max in Table 3) assumed that all failures of two components are dependent. For example, the number of failures of A (n_A) in the simulation data is 112, n_B is 113 and $n_{AB} = 15$. The first case (*Alpha_min*) assumes that n_{AB} is caused independently, therefore $n_2 = 0$. On the contrary, the second case (*Alpha_max*) assumes that n_{AB} are totally dependent, which results in $n_2 = 15$. The alpha factor was calculated using Eq. (6) and the CCBE probability through the alpha factor model was estimated using Eq. (5). The comparison results are summarized in Table 3.

It can be found that the results of the copula-CCF model are properly located between Alpha_min and Alpha_max. Since the actual results using the alpha factor are also located between min and max in practice, it is confirmed that the proposed method can make appropriate results for the dependent failures.

As mentioned in Section 3.1, since a copula can construct a joint probability distribution with various marginal distributions, the proposed method has the advantage of dealing with asymmetric CCFs as well as symmetric conditions. To highlight the strengths of using copula, we simulated 10 sets of failure data under asymmetric conditions. It was assumed that the asymmetry occurs in total failure probability of component A (The probability of failure of A is about 2 times higher than that of B). The results of the alpha factor model were compared with the results of the copula-based CCF

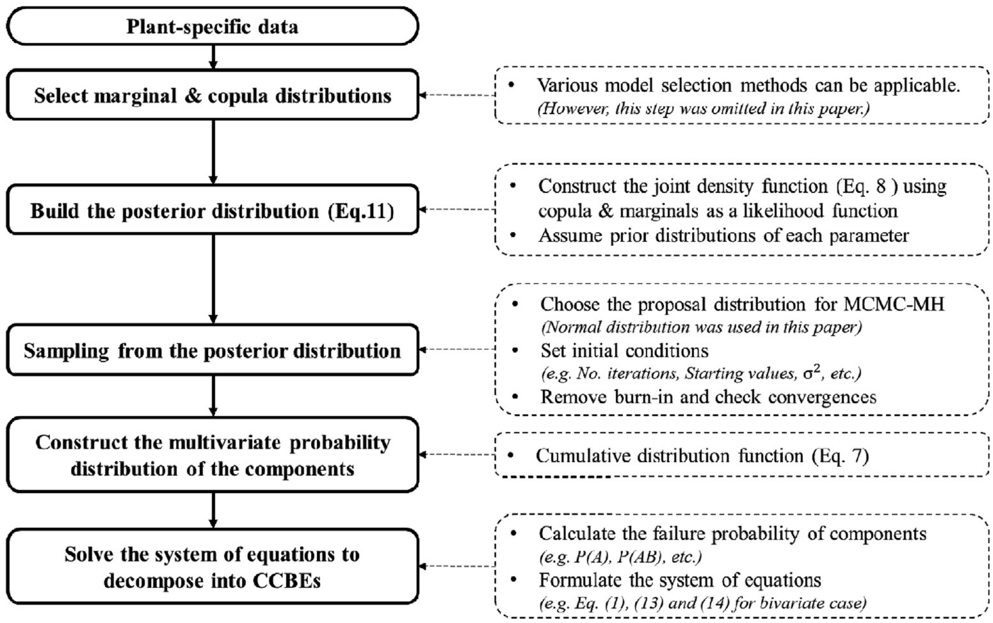


Fig. 6. Flowchart of a copula-based CCF model with the Bayesian approach.

Table 3
Comparisons between the copula-based CCF and alpha factor model.

CCBE	Q_1^A	Q_1^B	Q_2^{AB}
Alpha_min	1.13E-01	1.13E-01	0
Copula-CCF	1.07E-01	1.08E-01	6.06E-03
Alpha_max	8.60E-02	8.60E-02	2.65E-02

model using the same simulated data. Fig. 7 shows the results of independent failure probability between AFM and Copula-CCF using the same asymmetric conditions.

In Fig. 7, the results using AFM are represented for only A component because it assumes symmetry of components ($Q_1^A = Q_1^B$). Since the AFM does not consider the difference between A and B, the results indicate the average of the differences of components. However, the copula-based CCF model can capture such differences. It can be found that Q_1^A is almost twice of Q_1^B in the proposed method. Therefore, the proposed method has strengths and flexibility in estimating CCF probability under general conditions without special restrictions. This is not common in risk analysis, but it can work for particular issues aforementioned. Chapter 5 includes more example studies using simulated data under various asymmetric conditions.

5. Example studies using simulated data

This section describes asymmetric CCFs based on the proposed method using simulation data including asymmetric conditions. As mentioned in the Introduction, this paper assumes there are no existing failure data available to construct a probability distribution. If existing data are available or are used as prior information, then it is obvious that the results will be much better. For example, in the generic data in Ref. [17], the failure rate of an EDG is known to be 8.48E-04/hr. Thus, the gamma prior, Gamma (2.01, 2.37E+03) [17] can be used. The copula parameter can also be roughly estimated by the alpha factors in CCFs [22]. We can use this information as a prior distribution for a copula parameter. However, in this paper the prior distributions for all parameters are assumed to be uniform distributions, U (0,1) to describe more general conditions.

While a prediction distribution should be evaluated to estimate probabilities using a posterior distribution, this paper used the mean value of posterior distribution as a simple way. Note that the results of sampling or convergences will change at every calculation because simulation data is used. All calculations were carried out using R-programming packages (*copula* and *rootSolve* for solving systems of equations and *ggplot2*, *gridExtra* for visualizations).

5.1. CCCG 2: Asymmetry in total failure probabilities

Let us suppose that there are EDG failure data available for a certain NPP. In particular, EDG B is more likely to fail than EDG A due to degradations. (In this example, $\lambda_{EDG\ B} = 10 \times \lambda_{EDG\ A}$ and the others are the same as the conditions described in Section 4; we generated 100 sets of simulation data using an exponential marginal and a Clayton copula.) Using the posterior distribution given in Eq. (11) and MCMC-MH sampling, the posterior means of each parameter and the CCBEs are given in Table 4.

The quantities of interest, such as quantiles or means of the posterior parameters, are presented in Table 3. The posterior mean of $\lambda_{EDG\ B}$ is much greater than $\lambda_{EDG\ A}$. As a result, the independent failure probability of EDG B is higher than EDG A. These results were as we expected.

5.2. CCCG 3: Asymmetry in dependencies

This example case describes the asymmetry in dependencies of three EDGs using the normal copula. For EDG A, B, and C, it was assumed that the failures between EDG B and C occur more frequently than the other combinations. (The normal copula parameter θ_{BC} is assumed to be three times that of $\theta_{AB} (= \theta_{AC})$, and the others are the same as the condition described in Section 4; One hundred sets of simulation data were generated using exponential marginals and a normal copula.) The results are shown in Table 5.

The posterior mean of the normal copula parameter θ_{BC} cannot be calculated as three times that of θ_{AB} *or* θ_{AC} because there were not enough simulated data to describe the parameters exactly, as mentioned previously. Nevertheless, this proves that the failure

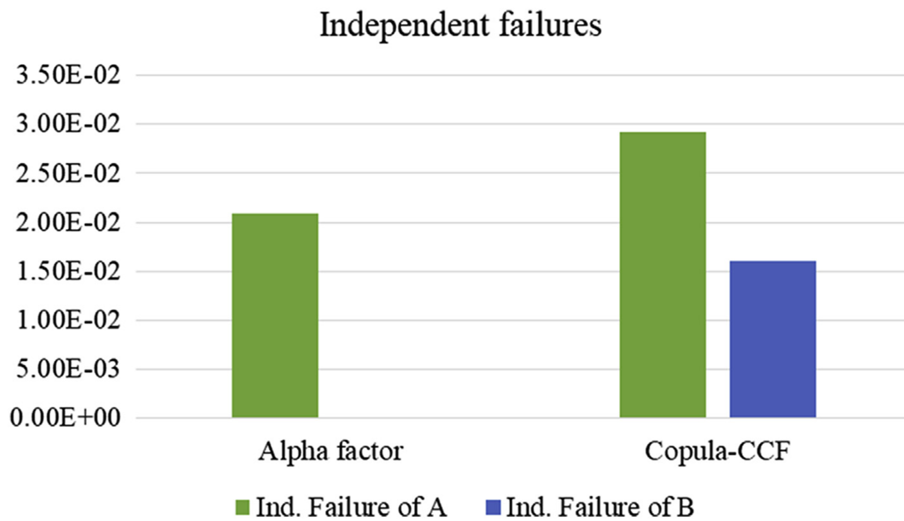


Fig. 7. Comparison of the results between AFM and Copula-CCF using the same asymmetric data.

Table 4
Example case 1: Asymmetries in total failure probabilities for CCCG 2 of EDGs.

	Posterior 2.5% Quantile	Posterior Mean	Posterior 97.5% Quantile	CCBEsProbability
$\lambda_{EDG\ A}$	7.69E-04	9.50E-04	1.14E-03	Q_1^A 1.69E-02
$\lambda_{EDG\ B}$	8.39E-03	1.24E-02	1.11E-02	Q_1^B 2.10E-01
$\theta_{clayton}$	1.17E-01	1.53E-01	3.80E-01	Q_2^{AB} 5.59E-03

probability of Q_2^{BC} is higher than that of Q_2^{AB} or Q_2^{AC}. The results can be improved when more data are simulated (or collected in practice).

5.3. *CCCG 3: Asymmetries in both total failure probabilities and dependencies*

The final example is the asymmetries in both the total failure probability and the dependency. A total of 200 sets of data were generated and evaluated using copula-based CCF models, assuming that the normal copula was used as described in Section 5.2. In addition, it was assumed that EDG B's failure rate is 10 times higher than EDG A or C, and the dependency between B and C is three times higher than AB or AC. Figs. 8 and 9 show the probabilities of CCBEs of independent failures and dependent failures, respectively.

The independent failure of EDG B is greater than the others, and the dependent failure of EDG B–C is also higher than A-B or A-C. Especially, the failure associated with EDG B is likely to occur because it has a higher total failure probability and also a greater dependency. This demonstrates that the copula-based CCF model with Bayesian approaches can provide the probability of CCBEs under asymmetrical conditions.

6. Conclusions

This paper proposed the copula-based CCF model for estimating asymmetric CCFs in situations where plant-specific failure data can be collected. The multivariate probability distribution across the components in the redundant systems was constructed using the marginal and copula distributions, and Bayesian inferences were also applied to the copula-CCF model. The posterior distribution of the copula-based CCF model allowed us to derive the asymmetric CCBEs by solving systems of equations without symmetry assumptions. Three example cases using simulation data are described, which confirmed that the CCBEs were appropriately derived by reflecting the characteristics of the data, such as asymmetric conditions.

Although the proposed method is shown to be valid for estimating CCF probabilities under asymmetric conditions, there are certain limitations and notes to be addressed as follows. In terms of the use of copula, while this paper used a normal copula to allocate different dependencies onto the variables, the vine copula

Table 5
Example case 2: Asymmetries in dependencies for CCCG 3 of EDGs.

	Posterior 2.5% Quantile	Posterior Mean	Posterior 97.5% Quantile	CCBEs	Probability
$\lambda_{EDG\ A}$	7.71E-04	9.47E-04	1.14E-03	Q_1^A	2.15E-02
$\lambda_{EDG\ B}$	7.17E-04	8.80E-04	1.05E-03	Q_1^B	1.95E-02
$\lambda_{EDG\ C}$	8.22E-04	1.00E-03	1.21E-03	Q_1^C	2.21E-02
θ_{AB}	4.03E-03	9.29E-02	2.54E-01	Q_2^{AB}	3.07E-04
θ_{AC}	1.79E-02	1.56E-01	3.16E-01	Q_2^{AC}	6.48E-04
θ_{BC}	**4.63E-02**	**2.28E-01**	**3.97E-01**	Q_2^{BC}	**1.02E-03**
-	-	-	-	Q_3^{ABC}	4.30E-05

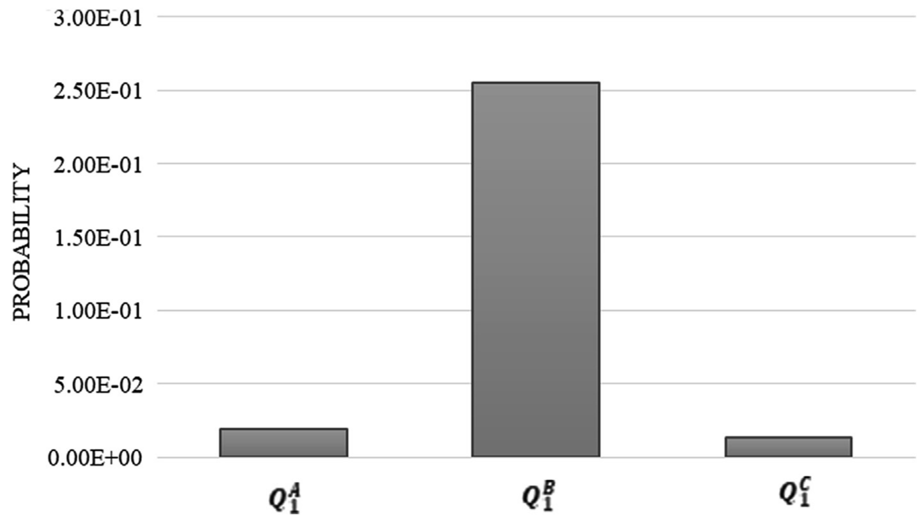


Fig. 8. The probability of CCBEs: Independent failures.

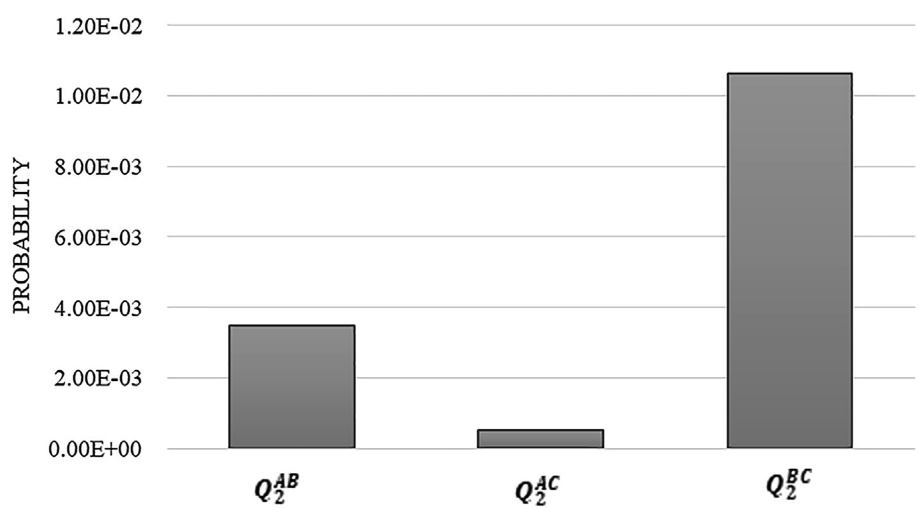


Fig. 9. The probability of CCBEs: CCFs of two component failures.

approach can be used to enable an Archimedean copula to assign various dependencies. The copula model selection should be also performed to fit failure data. In addition, this study addressed only continuous random variables, so a copula study for discrete random variables in CCFs should be also carried out.

In the Bayesian framework, assuming a prior distribution has to be carefully considered, i.e., a sensitivity analysis using various prior distributions should be carried out because there might be not enough data in practice. The performance of the MCMC-MH method used in this paper was significantly affected by our choice of proposed distributions, their variances, and the initial conditions. Therefore, empirical adjustments to MCMC-MH (e.g., the best acceptance probability or variance for this case) should also be considered. Finally, a prediction distribution rather than the use of the posterior mean should be considered.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the Nuclear Safety Research Program through the Korea Foundation Of Nuclear Safety (KOFONS), using the financial resource granted by the Nuclear Safety and Security Commission(NSSC) of the Republic of Korea (No. 1705001).

References

[1] A. Mosleh, D. Rasmuson, F. Marshall, Guidelines on modeling common-cause failures in probabilistic risk assessment, Idaho Natl. Eng. Environ. Lab. (1999).
[2] IAEA, Procedures for Conducting Common Cause Failure Analysis in Probabilistic Safety Assessment: IAEA TECDOC Series No, vol. 648, 1992.
[3] D. Il Kang, M.J. Hwang, S.H. Han, J.E. Yang, Approximate formulas for treating asymmetrical common cause failure events, Nucl. Eng. Des. 239 (2009) 346–352, https://doi.org/10.1016/j.nucengdes.2008.10.004.
[4] D.L. Kelly, D.M. Rasmuson, Common-cause failure analysis in event assessment, in: Proc. Inst. Mech. Eng. Part O J. Risk Reliab., 2008, pp. 521–532, https://doi.org/10.1243/1748006XJRR121.
[5] A. O'Connor, A. Mosleh, A general cause based methodology for analysis of common cause and dependent failures in system risk and reliability assessments, Reliab. Eng. Syst. Saf. 145 (2016) 341–350, https://doi.org/10.1016/j.ress.2015.06.007.
[6] A. Shemyakin, A. Kniazev, Introduction to Bayesian Estimation and Copula Models of Dependence, 2017, https://doi.org/10.1002/9781118959046.

[7] C. Czado, Analyzing Dependent Data with Vine Copulas, Springer International Publishing, Cham, 2019, https://doi.org/10.1007/978-3-030-13785-4.
[8] R.B. Nelsen, An Introduction to Copulas, 2006, https://doi.org/10.1007/0-387-28678-0.
[9] D.L. Kelly, Using copulas to model dependence in simulation risk assessment, ASME Int. Mech. Eng. Congr. Expo. Proc. 14 (2008) 81–89, https://doi.org/10.1115/IMECE2007-41284.
[10] X. Jia, L. Wang, C. Wei, Reliability research of dependent failure systems using copula, Commun. Stat. Simulat. Comput. 43 (2014) 1838–1851, https://doi.org/10.1080/03610918.2013.800879.
[11] H.H. Kwon, A copula-based nonstationary frequency analysis for the 2012–2015 drought in California, Water Resour. Res. 52 (2016) 1–20, https://doi.org/10.1002/2014WR015716.
[12] J.-Y. Kim, J.-G. Kim, Y.-H. Cho, H.-H. Kwon, A development of Bayesian Copula model for a bivariate drought frequency analysis, J. Korea Water Resour. Assoc. 50 (2017) 745–758, https://doi.org/10.3741/JKWRA.2017.50.11.745.
[13] R.D.S. Silva, H.F. Lopes, Copula, marginal distributions and model selection: a Bayesian note, Stat. Comput. 18 (2008) 313–320, https://doi.org/10.1007/s11222-008-9058-y.
[14] E.F. Saraiva, A.K. Suzuki, L.A. Milan, Bayesian computational methods for sampling from the posterior distribution of a bivariate survival model, based on AMH copula in the presence of right-censored data, Entropy 20 (2018) 1–21, https://doi.org/10.3390/e20090642.
[15] M.F. Pellissetti, U. Klapp, Integration of Correlation Models for Seismic Failures into Fault Tree Based Seismic Psa, 2011, pp. 6–11.
[16] W.S. Jung, K. Hwang, S.K. Park, A new method to allocate combination probabilities of correlated seismic failures into CCF probabilities, in: PSA 2019 - Int. Top. Meet. Probabilistic Saf. Assess. Anal., 2019, pp. 369–373.
[17] U.S.NRC, Industry-average performance for components and initiating events at U.S. Commercial, Nuclear Power Plants (2007).
[18] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images, IEEE Trans. Pattern Anal. Mach. Intell. PAMI- 6 (1984) 721–741, https://doi.org/10.1109/TPAMI.1984.4767596.
[19] W.K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika 57 (1970) 97–109, https://doi.org/10.2307/2334940.
[20] G.O. Roberts, A. Gelman, W.R. Gilks, Weak convergence and optimal scaling of random walk Metropolis algorithms, Ann. Appl. Probab. 7 (1997) 110–120.
[21] D. Huard, G. Évin, A.C. Favre, Bayesian copula selection, Comput. Stat. Data Anal. 51 (2006) 809–822, https://doi.org/10.1016/j.csda.2005.08.010.
[22] K. Jin, G. Heo, Copula study for common cause failures under asymmetric conditions, in: Trans. Korean Nucl. Soc. Spring Meet. Jeju, Korea, 2019. May 23-24.