Reliability Engineering and System Safety 116 (2013) 20–27

Contents lists available at SciVerse ScienceDirect

Reliability Engineering and System Safety

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

α-Decomposition for estimating parameters in common cause failure modeling based on causal inference

Xiaoyu Zheng*, Akira Yamaguchi, Takashi Takata

Graduate School of Engineering, Osaka University, Japan

ARTICLE INFO

Article history:
Received 11 January 2012
Received in revised form
25 February 2013
Accepted 26 February 2013
Available online 22 March 2013

Keywords:
Common cause failure
α-factor model
Bayesian theory
Causal inference
α-decomposition
Probabilistic risk assessment

ABSTRACT

The traditional α-factor model has focused on the occurrence frequencies of common cause failure (CCF) events. Global α-factors in the α-factor model are defined as fractions of failure probability for particular groups of components. However, there are unknown uncertainties in the CCF parameters estimation for the scarcity of available failure data. Joint distributions of CCF parameters are actually determined by a set of possible causes, which are characterized by CCF-triggering abilities and occurrence frequencies. In the present paper, the process of α-decomposition (Kelly-CCF method) is developed to learn about sources of uncertainty in CCF parameter estimation. Moreover, it aims to evaluate CCF risk significances of different causes, which are named as decomposed α-factors. Firstly, a Hybrid Bayesian Network is adopted to reveal the relationship between potential causes and failures. Secondly, because all potential causes have different occurrence frequencies and abilities to trigger dependent failures or independent failures, a regression model is provided and proved by conditional probability. Global α-factors are expressed by explanatory variables (causes' occurrence frequencies) and parameters (decomposed α-factors). At last, an example is provided to illustrate the process of hierarchical Bayesian inference for the α-decomposition process. This study shows that the α-decomposition method can integrate failure information from cause, component and system level. It can parameterize the CCF risk significance of possible causes and can update probability distributions of global α-factors. Besides, it can provide a reliable way to evaluate uncertainty sources and reduce the uncertainty in probabilistic risk assessment. It is recommended to build databases including CCF parameters and corresponding causes' occurrence frequency of each targeted system.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

As a conclusion from probabilistic risk assessment (PRA) for nuclear power plants (NPPs), common cause failures (CCFs) are significant challenges to the availability of safety systems with redundant components. WASH-1400 defined common mode failure as multiple failures that result from a single event. The single event can be any one of a number of possibilities: a common property, process, environment, or external event [1]. NUREG/CR-4780 defined common cause events as a subset of dependent events in which two or more component fault states exist at the same time, or in short intervals, and are direct results of a shared cause [2].

In past years, great achievements are gained to understand and model the mathematical mechanism of CCF. Fleming [3] introduced the most widely used single parameter model to be applied to CCF analysis, which is known as the β-factor model. Thereafter, for a more accurate analysis of systems with higher level of redundancy, Fleming and Kalinowski [4] extended the β-factor model to Multiple Greek Letter model (MGL). The α-factor model was developed by Mosleh and Siu [5], which can be applied to multi-component system by using total component failure probability and the fractions of CCF probability. NUREG/CR-4780 [2] and NUREG/CR-5485 [6] provided basic principles, models and guidance for analysts performing CCF analysis. Noticeably, NUREG/CR-5485 proposed generic prior distributions of α-factors for various system sizes. Therefore, a question would be asked that how could analysts update the distributions of CCF parameters, since α-factors of every system differ from each other and from time to time.

Along with the development of CCF modeling methodology and database, the understanding of CCF occurring-mechanism is progressing. United States Nuclear Regulatory Commission (U.S. NRC) has been endeavoring to develop a database for CCF parameters estimation [7]. NUREG/CR-5497 provided the parameter estimation employing two quantitative models: MGL model and α-factor model [8]. Furthermore, since 21st century, the emphasis of CCF analysis has been converted from simple mathematical models to

* Corresponding author. Tel.: +81 6 6879 7895; fax: +81 6 6879 7891.
E-mail address: zheng_x@qe.see.eng.osaka-u.ac.jp (X. Zheng).

0951-8320/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ress.2013.02.025

more complicated event assessment. Based on U.S. commercial NPPs event data, NUREG/CR-6819 illustrated further understanding of CCF insights for emergency diesel generators, motor-operated valves, pumps, and circuit breakers [9]. Updated version of NUREG/CR-6268 presented the process of event data collection and grouped the hierarchy of proximate failure causes, which provided a way to gain further understanding of the CCF events' occurring [10].

Rasmuson et al. extended the previous work on the treatment of CCF in event assessment. Fully expanded fault trees were used, and it is specifically showed that all terms in the basic parameter model (BPM) for each failure model. They quantified the conditional probability of CCF, given independent failures or failures with common-cause potential, and the asymmetry within a common-cause component group (CCCG) is considered [11]. Kelly et al. proposed the preliminary framework of a causality-based model via Bayesian networks which has the potential to overcome limitations of BPM. This model aimed to tell the conditional failure probability of remaining equipment given observed equipment failures and associated causes. Furthermore, it aimed to provide cause-specific quantitative insights into likely causes of failures [12].

As one family of graphical representation of distributions, Bayesian network uses a directed graph (where the edges have a source and a target) to represent a set of independencies and to factorize a distributions [13,14]. This advantage of probabilistic graphical models can promote the visual analysis of CCF. Besides, Bayesian statistical inference provides a way of formalizing the process of learning from data to update beliefs in accord with recent notions of knowledge synthesis [15].

Based on the frequentist probability, the widely applied BPM (β-factor model, MGL model, α-factor model, etc.) enables the uncomplicated evaluation of CCF probability. However, there are unknown uncertainties in the parameter estimation, limitations to identify the risk of potential causes and it is impossible to arrive at posterior distribution as a result of sparse failure data or data-missing problem. Recent research reveals that CCF analysis is transferring from pure mathematical modeling to causality-based analysis. In order to reduce the uncertainty in CCF parameter estimation, Bayesian regression models can be applied to combine difference data sources of CCF failure events and cause occurrence. Therefore, posterior distributions of less uncertainty can be obtained.

In this paper, the authors propose the α-decomposition process to estimate CCF parameters. The α-decomposition process is named as Kelly-CCF method since Dana Kelly firstly proposed the preliminary framework of a causality-based Bayesian network for CCF analysis: (a) based on the causal inference for CCF, global α-factors are decomposed as a regression model with explanatory variables (causes' occurrence frequencies) and parameters (decomposed α-factors); (b) Bayesian inference is applied to arrive at posterior distributions for global α-factors and decomposed α-factors, which quantitatively represent the CCF risk on component level and cause level, respectively; (c) a hypothetical database is constructed and recommended to be built for the α-decomposition analysis. This database must include causes occurrence frequencies and either α-factors or CCF event records for specific systems.

2. α-factor model for standard CCF analysis

Before the introduction of the α-decomposition process, a review of the standard α-factor model is necessary for the purpose of easy understanding of the notation. As shown in Fig. 1, let us consider a system of three identical components A, B, and C from the perspective of a two-out-of-three success logic [6]. The common-cause component group (CCCG) is A, B, and C. There, a group of components identified in the process of CCF analysis is called as a CCCG. The minimal cutsets of this system failure are:

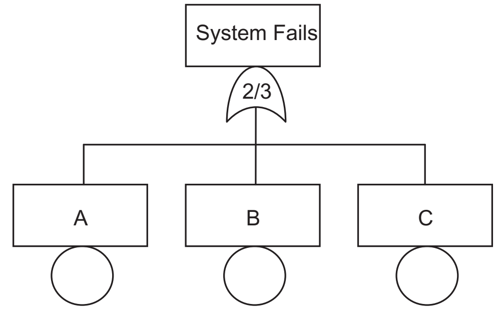

Fig. 1. Component-level fault tree of system.

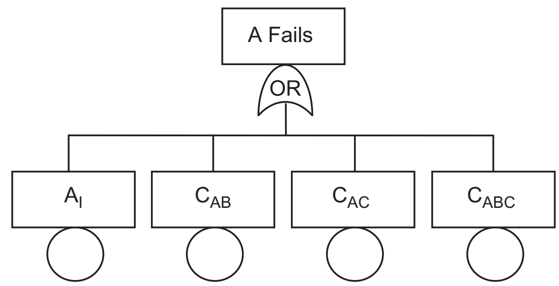

Fig. 2. Expanded CCBE fault tree for component A.

$\{A,B\}; \{A,C\}; \{B,C\}; \{A,B,C\}$

For the consideration of CCF, the fault tree is expanded to include corresponding common-cause basic event (CCBE). Take component A as an example as shown in Fig. 2.

The cutsets of component A are

$\{A_I\}; \{C_{AB}\}; \{C_{AC}\}; \{C_{ABC}\}$

Similarly, the cutsets of component B failure are

$\{B_I\}; \{C_{AB}\}; \{C_{BC}\}; \{C_{ABC}\}$

The cutsets of component C failure are

$\{C_I\}; \{C_{AC}\}; \{C_{BC}\}; \{C_{ABC}\}$

Here, A_I is the failures of component A from independent causes, B_I is the failures of component B from independent causes, C_I is the failures of component C from independent causes; C_{AB} is the failures of components A and B from common causes, C_{AC} is the failures of components A and C from common causes, C_{BC} is the failures of components B and C from common causes and C_{ABC} is the failures of components A, B and C from common causes.

Using the rare event approximation, the system failure probability of the two-out-of-three system is given by

$$P(S) \cong P(A_I)P(B_I)+P(A_I)P(C_I)+P(B_I)P(C_I)+P(C_{AB})+P(C_{AC})+P(C_{BC})+P(C_{ABC}) \tag{1}$$

In Eq. (1), the rare event approximation is defined as that the probability of the simultaneous occurrence of two independent failures is assumed as zero. It can be written as

$$P(a+b)=P(a)+P(b)-P(a \cdot b) \cong P(a)+P(b) \tag{2}$$

The failure probability of component A is decomposed as

$$P(A_t)=P(A_I)+P(C_{AB})+P(C_{AC})+P(C_{ABC}) \tag{3}$$

Here, A_t is the all failures of component A, and $P_{(X)}$ is the probability of event X.

Assume that

$$P(A_I)=P(B_I)=P(C_I)=Q_1 \tag{4}$$

$$P(C_{AB})=P(C_{AC})=P(C_{ABC})=Q_2 \tag{5}$$

$$P(C_{ABC})=Q_3 \tag{6}$$

The system failure probability is

$$Q_S=3(Q_1)^2+3Q_2+Q_3 \tag{7}$$

The component A failure probability

$$Q_t=Q_1+2Q_2+Q_3 \tag{8}$$

The testing scheme to which the CCCG is subjected has an impact on the mathematical form of models. In the case of staggered testing scheme, only one component is tested in a test episode. If there is a failure, the rest of components will be tested. In contrast, in the non-staggered testing scheme, all components in the group are tested in a test episode. The definitions of parameters of these two testing schemes are explained as follows.

The definition of α-factors (staggered testing scheme):

$$\alpha_1=\frac{Q_1}{Q_t} \tag{9}$$

$$\alpha_2=\frac{2Q_2}{Q_t} \tag{10}$$

$$\alpha_3=\frac{Q_3}{Q_t} \tag{11}$$

The definition of α-factors (non-staggered testing scheme):

$$\alpha_1=\frac{Q_1}{Q_t}\alpha_t \tag{12}$$

$$\alpha_2=\frac{Q_2}{Q_t}\,\frac{\alpha_t}{2} \tag{13}$$

$$\alpha_3=\frac{Q_3}{Q_t}\,\frac{\alpha_t}{3} \tag{14}$$

Here

$$\alpha_t=\alpha_1+2\alpha_2+3\alpha_3 \tag{15}$$

Note that these two estimates of the CCF probability differ by factors, but both models are reasonable. In the CCF modeling, it will affect estimates to make assumptions according to testing strategies. Such assumptions are usually necessary for the lack of success data in the database. In this paper, it is assumed as staggered testing scheme. Therefore, the system failure probability can be written as

$$Q_S=3(\alpha_1 Q_t)^2+\tfrac{3}{2}(\alpha_2 Q_t)+(\alpha_3 Q_t) \tag{16}$$

3. Proposed α-decomposition (Kelly-CCF method) with Hybrid Bayesian Network (HBN)

3.1. The definition of α-decomposition

For distinction, α-factors at the component level are named as global α-factors. After decomposition, α-factors at the cause level are named as decomposed α-factors. In PRA analysis, global α-factors are treated as distributions and uncertainty sources in distributions are always unknown. The uncertainty in the basic events analysis will be propagated through PRA models, e.g. Fault Tree and Event Tree. Distributions for top events will be flattened as a result of uncertainty propagation. The global α-factors in α-factor model can be affected by observable or unobservable variables. Information about such variables can be used in the Bayesian inference paradigm to obtain posterior distributions of less uncertainty.

Definition 1 (α-decomposition). The global α-factors are decomposed into explanatory variables, occurrence frequencies and CCF triggering abilities (decomposed α-factors) of potential causes. The Bayesian network is applied to represent the relationship of global α-factors and decomposed α-factors.

Potential causes have different occurrence frequencies. For instance, as depicted in NUREG/CR-6819, there are totally 41 EDG CCF events happened. 15 events (36.6%) are caused by Design/Construction/Installation/Manufacture proximate cause group; 9 events (22.0%) are caused by Internal to Component proximate cause group; 9 events (22.0%) are caused by Operational/Human Error proximate cause group; 5 events (12.2%) are caused by External Environment proximate cause group; 3 events (7.3%) are caused by Other proximate cause group [9]. According to the research of seismic probabilistic safety assessment (SPSA), in the low peak ground acceleration (PGA) range, the β-factor is quite small, but in the high PGA range the β-factor is quite large [16]. Here, the β-factor model is functionally the same as α-factor model, but it is simpler for using one single parameter (β) to represent fraction of CCF probability. Obviously, the higher PGA tends to generate CCF much easier than the lower PGA. Therefore, the same type of CCF events is the result of that the same cause frequently happens or the cause is good at CCF triggering.

The previous two examples show that global α-factors are affected by two variables (causes' occurrence frequency and decomposed α-factors). It is important to infer the mathematical model which can explain the relationship of global α-factors and two explanatory variables.

3.2. The HBN for the α-decomposition process

In order to infer the mathematical model of the α-decomposition process, two kinds of relationships are considered: the relationship between system failures and component failures; the relationship between component failures and causes. Hybrid Bayesian network (HBN) is applied to express both relationships. A hybrid network means to use a combination of two or more topologies. In the current article, the HBN is a combination of Fault Tree (FT) and Bayesian network. Let us consider a system of three components A, B, and C. In a standard α-factor model, all components of one system are assumed identical. For the simplest consideration, it is also assumed that all of three components A, B, and C are identical. There are three potential causes C_1, C_2, and C_3 that will possibly lead to a component failure (as shown in Fig. 3).

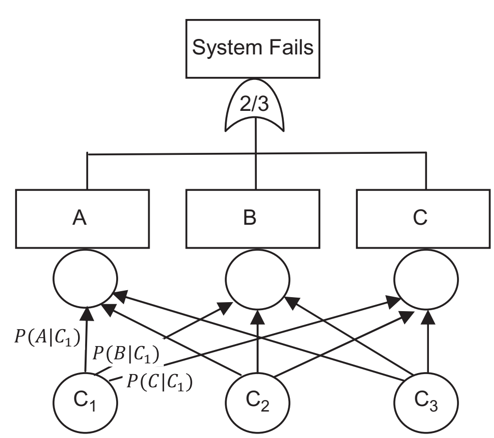

Fig. 3. Hybrid Bayesian network for CCF analysis.

The conditional probabilities of component failures and system failure are expressed in Fig. 3. Take component A as an example:

$$P(A) = P(A|C_1)P(C_1) + P(A|C_2)P(C_2) + P(A|C_3)P(C_3) \quad (17)$$

The process of α-decomposition is illustrated in Fig. 4. Cause 1 results in three types of failure: independent failure, partial CCF and complete CCF.

$$P(A|C_1) = P(A_I|C_1) + P(AB|C_1) + P(AC|C_1) + P(ABC|C_1) \quad (18)$$

Both sides of Eq. (18) are divided by $P(A|C_1)$.

$$1 = \alpha_1^{C_1} + \alpha_2^{C_1} + \alpha_3^{C_1} \quad (19)$$

Here,

$$\alpha_1^{C_1} = \frac{P(A_I|C_1)}{P(A|C_1)} \quad (20)$$

$$\alpha_2^{C_1} = \frac{P(AB|C_1)}{P(A|C_1)} + \frac{P(AC|C_1)}{P(A|C_1)} \quad (21)$$

$$\alpha_3^{C_1} = \frac{P(ABC|C_1)}{P(A|C_1)} \quad (22)$$

The $\alpha_j (j=1,2,3)$ are global α-factors, and the $\alpha_j^{C_i} (i,j=1,2,3)$ are decomposed α-factors. Equivalently, for Cause 2 and Cause 3, the conditional probability and decomposed α-factors can be given by

$$P(A|C_2) = P(A_I|C_2) + P(AB|C_2) + P(AC|C_2) + P(ABC|C_2) \quad (23)$$

$$1 = \alpha_1^{C_2} + \alpha_2^{C_2} + \alpha_3^{C_2} \quad (24)$$

$$P(A|C_3) = P(A_I|C_3) + P(AB|C_3) + P(AC|C_3) + P(ABC|C_3) \quad (25)$$

$$1 = \alpha_1^{C_3} + \alpha_2^{C_3} + \alpha_3^{C_3} \quad (26)$$

The independent failure of each component is affected by Cause 1, Cause 2 and Cause 3. According to Fig. 4, the relationship is expressed as

$$P(A_I) = P(A_I|C_1)P(C_1) + P(A_I|C_2)P(C_2) + P(A_I|C_3)P(C_3) \quad (27)$$

Replace the independent part of probability with $\alpha_1^{C_i}$ elements as follows:

$$P(A_I) = \alpha_1^{C_1} P(A|C_1)P(C_1) + \alpha_1^{C_2} P(A|C_2)P(C_2) + \alpha_1^{C_3} P(A|C_3)P(C_3) \quad (28)$$

Both sides of Eq. (28) are divided by $P(A)$ and because $P(A_I)/P(A) = \alpha_1$, so it can be written as

$$\alpha_1 = \alpha_1^{C_1} \frac{P(A|C_1)P(C_1)}{P(A)} + \alpha_1^{C_2} \frac{P(A|C_2)P(C_2)}{P(A)} + \alpha_1^{C_3} \frac{P(A|C_3)P(C_3)}{P(A)} \quad (29)$$

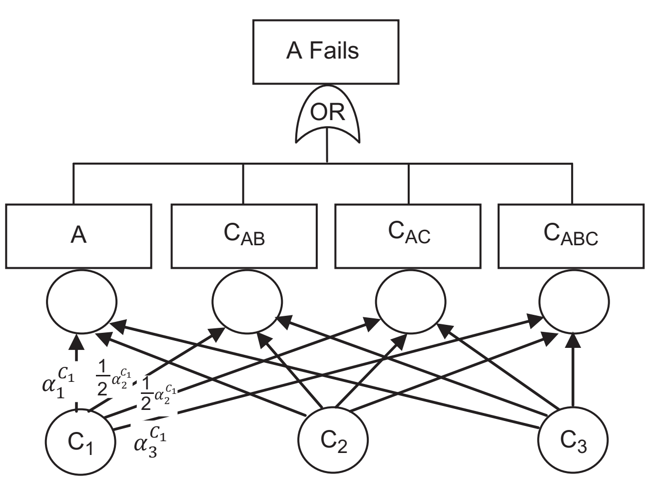

Fig. 4. α-decomposition for CCF analysis with HBN.

Deduced from Fig. 3, it is known that failure of A is generated by three parts. These three parts are notated as occurrence rates

$$r_1 = \frac{P(A|C_1)P(C_1)}{P(A)} = P(C_1|A) \quad (30)$$

$$r_2 = \frac{P(A|C_2)P(C_2)}{P(A)} = P(C_2|A) \quad (31)$$

$$r_3 = \frac{P(A|C_3)P(C_3)}{P(A)} = P(C_3|A) \quad (32)$$

$$r_1 + r_2 + r_3 = 1 \quad (33)$$

A simple form of α_1-decomposition is obtained.

$$\alpha_1 = \alpha_1^{C_1} r_1 + \alpha_1^{C_2} r_2 + \alpha_1^{C_3} r_3 \quad (34)$$

The simple forms of α_2-decomposition and α_3-decomposition are given by

$$\alpha_2 = \alpha_2^{C_1} r_1 + \alpha_2^{C_2} r_2 + \alpha_2^{C_3} r_3 \quad (35)$$

$$\alpha_3 = \alpha_3^{C_1} r_1 + \alpha_3^{C_2} r_2 + \alpha_3^{C_3} r_3 \quad (36)$$

Here, $\alpha_j (j=1,2,3)$ is the global α-factors for j components failure; $\alpha_j^{C_i} (i,j=1,2,3)$ is the decomposed α-factors for j components faiure as a result of Cause i; and $r_i (i=1,2,3)$ is the occurrence rate for Cause i.

The $\alpha_j^{C_i}, r_i (i,j=1,2,3)$ have practical engineering meanings. The $\alpha_j^{C_i}$ means the ability of Cause i to lead to the failure involving j components. The r_i means that among all failures, so totally there are $r_i \times 100\%$ failures generated by Cause i. In other words, r_i is the occurrence rate over the occurrence of all causes. If Cause 1 happens much more frequently than other causes, the global QUOTE will tend to be $\alpha_j^{C_i}$. However, if Cause 2 happens rarely with low $\alpha_j^{C_i}$, it means that risk-significance of Cause 2 is negligible.

4. Bayesian inference for α-decomposition process

4.1. Hierarchical Bayesian modeling

In the traditional CCF parameter estimation, the Dirichlet distribution is chosen as the prior distribution for global α-factors. The Dirichlet distribution is conjugate to the multinomial likelihood. The failure events of a CCCG are always assumed as Multinomial distributions, so the posterior distribution for global α-factors is still a Dirichlet distribution. Usually, in the Bayesian inference for CCF parameters, parameters in prior distributions are set as constant. For example, the noninformative prior for global α-factors, which is a Dirichlet distribution with all parameters equal to 1 [17–20].

When the CCF parameter estimation is considered based on the information from cause level, the prior in the previous model should be specified in multiple stages. The graphical representation of a multi-stage hierarchical Bayesian modeling is shown in Fig. 5. Oval nodes refer to stochastic components of the model, and squared nodes refer to constant parameters. Solid arrows indicate stochastic dependence while hollow arrows indicate logical functions. The node (t[1:3]) is the data of CCF events, which is assumed as a Multinomial distribution. This value is affected by the node global α-factors (α[1:3]) and the distribution of global α-factors is assumed as a Dirichlet distribution. The node (θ[1:3]) represses parameters in the prior distribution of node (α[1:3]), which is affected by hyper variables. Hyper variables are decomposed α-factors and causes' occurrence frequencies. The distribution of global α-factors is affected by these two sets of variables. Decomposed α-factors are called hyper variables since they are uncertain variables. Other constant values are called as hyper parameters.

Usually, $p(\boldsymbol{t}|\boldsymbol{\alpha})$ is called likelihood, $p(\boldsymbol{\alpha}|\boldsymbol{\theta})$ is called prior distribution (1st level), and $p(\boldsymbol{\theta}|\boldsymbol{\alpha}^{C_1}_{[]},\boldsymbol{\alpha}^{C_2}_{[]},\boldsymbol{\alpha}^{C_3}_{[]},\boldsymbol{r})$ is called hyper prior (2nd level). It is a two-stage hierarchical Bayesian model.

Bayes' theorem for the α-decomposition process can be given by

$$P(\alpha_j^{C_i}|\mathrm{r},\theta,\alpha,\mathrm{t}) = \frac{P(\alpha_j^{C_i},\mathrm{r},\theta,\alpha,\mathrm{t})}{P(\mathrm{r},\theta,\alpha,\mathrm{t})} = \frac{P(\mathrm{r},\theta,\alpha,\mathrm{t}|\alpha_j^{C_i})P(\alpha_j^{C_i})}{P(\mathrm{r},\theta,\alpha,\mathrm{t})} \tag{37}$$

In Eq. (37), $P(\alpha_j^{C_i})$ is the prior distribution for decomposed α-factors, and $P(\mathrm{r},\theta,\alpha,\mathrm{t}|\alpha_j^{C_i})$ is defined as the likelihood of parameters. Therefore, according to the chain rule of conditional probability and the independence among parameters, the posterior distribution can be written as

$$\pi_2(\alpha_j^{C_i}|\mathrm{r},\theta,\alpha,\mathrm{t}) \propto f_t(\mathrm{t}|\alpha) f_\alpha(\alpha|\theta) f_\theta(\theta|\alpha_j^{C_i},r)\pi_1(\alpha_j^{C_i}) \tag{38}$$

Here, all bold letters mean vectors of respective parameters. Therefore, the posterior distribution $\pi_2(\alpha_j^{C_i}|\mathrm{r},\theta,\alpha,\mathrm{t})$ can be obtained according to the prior distribution $\pi_1(\alpha_j^{C_i})$ and likelihood functions. Moreover, the estimation of posterior global α-factors can be obtained as well.

4.2. *An example for α-decomposition's Bayesian inference*

The following example is proposed by authors to explain how to do Bayesian inference for the α-decomposition process [21]. We

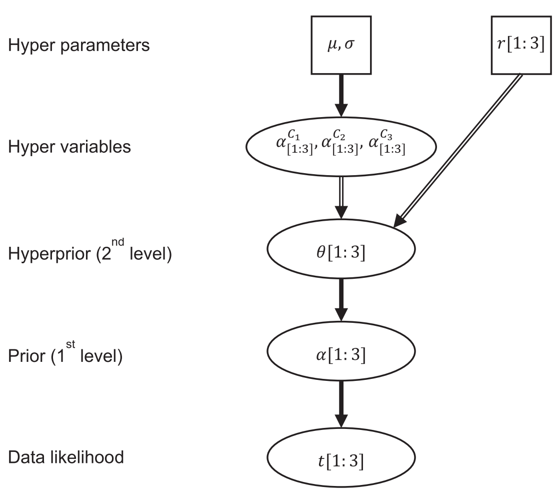

Fig. 5. Directed acyclic graph for Bayesian inference.

would like to caution that all the CCF data used in this paper is for illustration only and not from any real database. All the calculation of Bayesian inference is conducted by OpenBUGS version 3.2.2, and the graphs of posterior distributions are drawn by R version 2.13.1.

Table 1 lists the hypothetical database including the occurrence of causes and CCF events. CCF data of 16 systems is assumed. There are occurrence record of causes and single failure (1/3), partially dependent failure (2/3), and complete failure (3/3). It is remarked that all the recorded causes have triggered failures and each failure has a single cause. Take the System 1 as an example. There are totally 127 times of failures, of which 32 failures are generated by Cause 1, 28 by Cause 2, and 67 by Cause 3. There are also 113 single failure events, 11 partial CCF and 3 complete CCF.

Additionally, it is assumed that all the occurrences of causes are independent. From the perspective of conditional probability, it is

$$P(\mathit{Cause}\ j|\mathit{Cause}\ k) = P(\mathit{Cause}\ j), j \neq k \tag{39}$$

Common causes' occurrence rates are defined in Eq. (40). For System 1, occurrence rates for Cause 1, Cause 2, and Cause 3 are 25.20%, 22.00%, and 52.80%, respectively.

$$r_i = \frac{\mathit{Cause}\ i\text{'s}\ \mathit{occurrence\ frequency}}{\mathit{Total}} \quad (i = 1,2,3) \tag{40}$$

Therefore, based on the information of failure and cause occurrence, it is able to evaluate the posterior distributions of decomposed α-factors and global α-factors. As introduced before, the Multinomial distribution serves as the aleatory model for CCF events, and the prior distribution for global α-factors is assumed as a Dirichlet distribution.

$$t[k,1:m] \sim dmulti(\alpha[k,1:m], T[k]) \tag{41}$$

$$\alpha[k,1:m] \sim ddirich(\theta[k,1:m]) \tag{42}$$

Similarly, prior distributions for decomposed α-factors are assumed as Dirichlet distributions. A noninformative prior is the Dirichlet distribution with each parameter $\delta_i[\,] = 1$.

$$\alpha^{C_i}[1:m] \sim ddirich(\delta_i[1:m]) \tag{43}$$

On the basis of Eqs. (34)–(36), parameters in the hyper prior distribution can be expressed by decomposed α-factors and occurrence rates:

$$\theta[k,j] = \left(\sum_{i=1}^{3} \alpha^{C_i}[j] \times r[k,i]\right) T[k] \tag{44}$$

Table 1
Hypothetical database of common cause failure events.

	Common causes' occurrence				CCF event		
	Cause 1	Cause 2	Cause 3	Total	Single (1/3)	Partial (2/3)	Complete (3/3)
System 1	32 (25.20%)	28 (22.00%)	67 (52.80%)	127	113	11	3
System 2	17 (16.00%)	78 (73.60%)	11 (10.40%)	106	98	7	1
System 3	18 (20.70%)	19 (21.80%)	50 (57.50%)	87	73	9	5
System 4	29 (43.90%)	6 (9.10%)	31 (47.00%)	66	53	5	8
System 5	7 (14.00%)	33 (66.00%)	10 (20.00%)	50	45	4	1
System 6	15 (36.60%)	9 (22.00%)	17 (41.40%)	41	33	3	5
System 7	12 (35.30%)	15 (44.10%)	7 (20.60%)	34	32	2	0
System 8	2 (6.50%)	22 (71.00%)	7 (22.50%)	31	29	2	0
System 9	7 (31.80%)	4 (18.20%)	11 (50.00%)	22	20	2	0
System 10	10 (47.60%)	8 (38.10%)	3 (14.30%)	21	20	1	0
System 11	3 (15.80%)	6 (31.60%)	10 (52.60%)	19	16	2	1
System 12	7 (43.80%)	3 (18.80%)	6 (37.40%)	16	14	1	1
System 13	3 (20.00%)	5 (33.33%)	7 (46.67%)	15	13	1	1
System 14	5 (33.30%)	3 (20.00%)	7 (46.70%)	15	12	1	2
System 15	4 (36.40%)	5 (45.50%)	2 (18.10%)	11	9	1	1
System 16	1 (11.10%)	6 (66.70%)	2 (22.20%)	9	7	1	1

Thus,

$$\text{Posterior } \alpha_j^{C_i} \propto \text{Prior } \alpha_j^{C_i} \times \text{Likelihood } (\mathrm{r,t}|\alpha_j^{C_i}), \quad i,j=(1,2,3) \tag{45}$$

The graphical model for the example is provided in Fig. 6. m is the size of the CCCG which equals three in current article. The *case number* is 16 as there are 16targeted systems in the database. T[] is the total failure events happening in every system, which is used in Eq. (41) as a parameter in Multinomial distribution. The meaning of symbols are the same to Fig. 5. Two plates represent two repeated loops, which are shown with the range (j in 1: m) and the range (k in 1: case number).

The summary of posterior distributions for decomposed α-factors is shown in Table 2. Causes have different abilities to trigger CCF events, so the posterior distributions for decomposed α-factors differ from each other. If the best point estimate for posterior distributions is applied, the risk significances of decomposed α-factors are compared in the order shown below. Cause 2 is of least CCF risk as the $\alpha_2^{C_1}$ is of greatest value. Cause 3 is best at triggering partial CCF and Cause 1 is best at triggering complete CCF. However, the difference is not significant considering the range of 90% interval. It is practically meaningful to recognize the ability of CCF triggering. The mechanism through which failures of multiple components are coupled is termed as the coupling factor. In other words, it is the coupling factor of multiple components that identifies them as susceptible to the same causal mechanism of failure [13]. Therefore, based on the α-decomposition process, it is possible to recognize the most hazardous causes. Thus, if specific measures could be taken to disrupt the coupling factor of multiple components, the occurrence probability of CCF events would be reduced.

$$\alpha_1^{C_2} < \alpha_1^{C_3} < \alpha_1^{C_2}$$

$$\alpha_2^{C_2} < \alpha_2^{C_1} < \alpha_2^{C_3}$$

$$\alpha_3^{C_2} < \alpha_3^{C_3} < \alpha_3^{C_1}$$

The specific posterior probability density functions for $\alpha_1^{C_i}$,$\alpha_2^{C_i}$ and $\alpha_3^{C_i}(i=1,2,3)$ are shown in Figs. 7–9, respectively. Fig. 7 shows the ability of each cause to result in a single failure. Figs. 8 and 9 show the ability of each cause to result in partial CCF and complete CCF, respectively. The positions of curves represent risk significances of possible causes and ranges represent uncertainties in estimates. One important problem in PRA parameter estimation is the evaluation of uncertainty. According to the probability density functions of decomposed α-factors, uncertainty sources in global α-factors estimation can be judged.

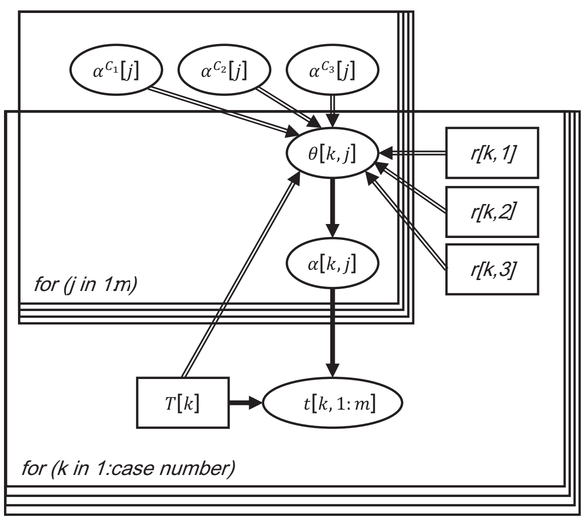

Fig. 6. Graph model for specific α-decomposition analysis.

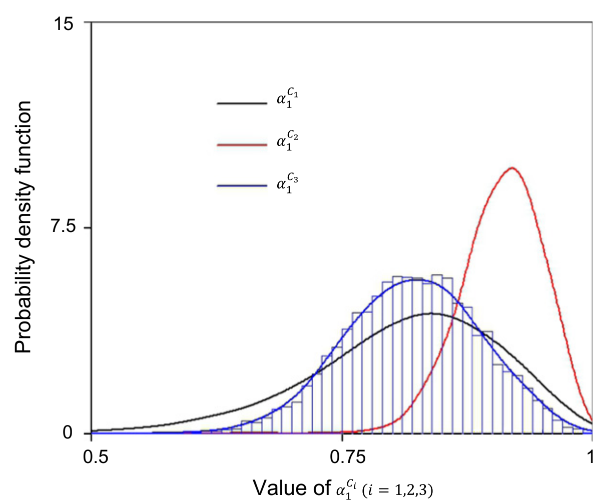

Fig. 7. Posterior distribution of decomposed α_1.

Table 2
Summary of posterior distributions for decomposed α-factors.

Posterior $\alpha_j^{C_i}$ $(i,j=1,2,3)$	Mean	SD	2.5%	Median	97.5%
α_1-decomposition					
$\alpha_1^{C_1}$	8.13E−01	4.73E−03	6.03E−01	8.24E−01	9.61E−01
$\alpha_1^{C_2}$	9.10E−01	1.54E−03	8.26E−01	9.13E−01	9.78E−01
$\alpha_1^{C_3}$	8.20E−01	3.10E−03	6.87E−01	8.21E−01	9.44E−01
α_2-decomposition					
$\alpha_2^{C_1}$	9.88E−02	3.90E−03	3.89E−03	8.34E−02	2.77E−01
$\alpha_2^{C_2}$	7.29E−02	1.47E−03	9.79E−03	6.99E−02	1.53E−01
$\alpha_2^{C_3}$	1.12E−01	2.72E−03	1.20E−02	1.12E−01	2.21E−01
α_3-decomposition					
$\alpha_3^{C_1}$	8.79E−02	1.76E−03	4.56E−03	8.10E−02	2.18E−01
$\alpha_3^{C_2}$	1.74E−02	2.68E−04	5.21E−04	1.34E−02	5.61E−02
$\alpha_3^{C_3}$	6.79E−02	1.21E−03	5.40E−03	6.59E−02	1.51E−01

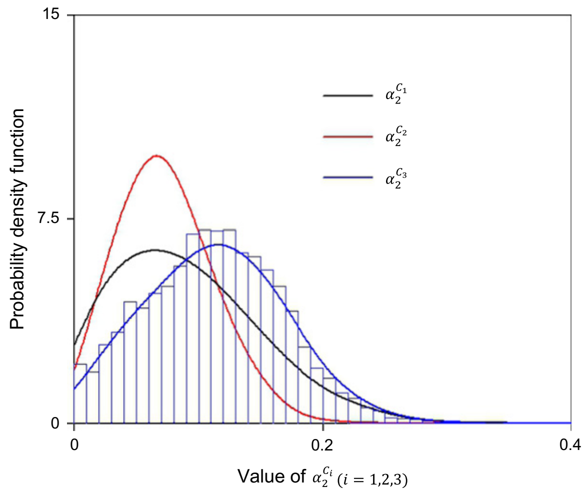

Fig. 8. Posterior distribution of decomposed α_2.

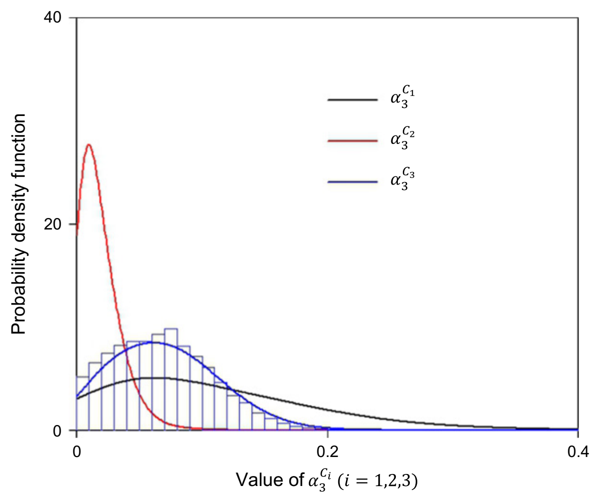

Fig. 9. Posterior distribution of decomposed α_3.

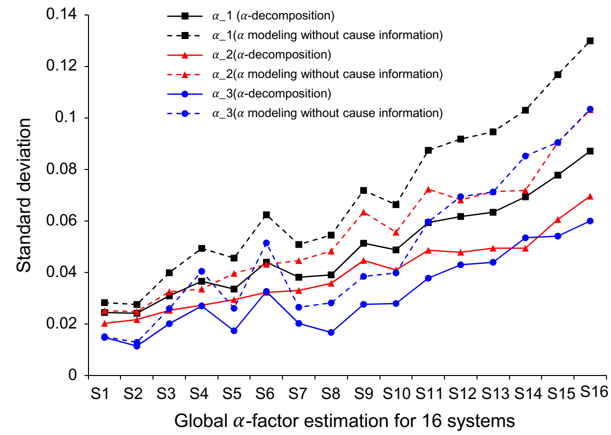

Fig. 10. Comparison of uncertainty in α-decomposition process and previous α-factor model.

Because the standard deviations for $\alpha_j^{C_2}$ are relatively less than other two causes, it can be decided that most of uncertainty comes from the innate uncertainty of Cause 1 and Cause 3. Such uncertainty will flatten the joint distributions for global α-factors by the means of uncertainty propagation. Under the current situation, the learning of Cause 1 and Cause 3 is important for PRA analysis. Other potential valuable data from different data sources and non-data information should be mined to support Bayesian inference and then to reduce the uncertainty in calculation [22–23].

Fig. 10 shows the uncertainty analysis for global α-factors estimation. Three important results should be remarked. Firstly, comparing the α-decomposition process with the previous α-factor modeling, the uncertainty in posterior distributions for global α-factors is reduced. Each dashed line is higher than the corresponding solid line. This result proves that the regression model for global α-factors, decomposed α-factors and occurrence rates can combine valuable information together to reduce the uncertainty in the CCF parameter estimation. Secondly, in the hypothetical database, systems are listed according to the number of failures from largest to smallest. In the CCF analysis, the scarcity of CCF data is a major source of uncertainty which can be observed from Fig. 10. All lines have the same trend that the uncertainty increases with the decreasing of available data. In the α-decomposition process, the estimation of systems with limited failure data is of greater uncertainty than systems with abundant data. Finally, since all causes and failure events have innate uncertainty as introduced before, there are some singular points in the trend. Therefore, the α-decomposition process can reduce the uncertainty in the parameter estimation by introducing information from cause level but this process cannot diminish all uncertainty. The posterior distributions for all global α-factors are omitted to avoid a lengthy explanation.

5. Conclusions and discussion

Based on the α-factor model and causal inference, the α-decomposition process (Kelly-CCF method) has been introduced to analyze common cause failures. Firstly, in order to easily understand the notation of α-decomposition process, the α-factor model is illustrated. Secondly, based on causes of different risk-significance, the CCF occurring mechanism is recognized as hierarchical network. The first level is system failure, the second level is components failure and the third level is possible causes. The theory of hybrid Bayesian network and conditional probability are applied to prove the correlation among the three levels. A regression model is proposed. Finally, based on the reasonable annotation of CCF and causes, the theory of Bayesian inference with MCMC method is introduced to evaluate all parameters of α-decomposition process. An example is illustrated in order to show the calculation procedure of α-decomposition process for a simple 3-component system. Global α-factors are lumped parameters, whose joint distributions are integrated reflection of all cause and failure information. From the viewpoint of data analysis, global α-factors could be considered as weighted averages of all decomposed α-factors. Decomposed α-factors proposed in the current article are risk characteristics of possible causes, which represent different CCF triggering abilities. This analysis is useful to assist analysts to rank causes according to different CCF risk significance to systems. The uncertainty analysis in CCF parameter estimation is conducted. Uncertainty sources are able to be judged and the uncertainty in global α-factors estimation can be reduced. Furthermore, a new type of database is recommended to be built, in which the record of CCF events and causes occurrence information should be combined.

Appendix

A.1. OpenBUGS script for hierarchical Bayesian inference of α-decomposition process

```
model {
# model's likelihood
  for(k in 1:case.number) {
  t[k, 1:m]~dmulti(alpha[k, 1:m], T[k]) # Stochastic model with
  multinomial likelihood function
  T[k] < -sum(t[k, 1:m]) # T is the total number of "group failure
  events"
  alpha[k, 1:m]~ddirich(theta[k, ]) #Transition variable
  for (j in 1:m) {
  theta[k, j] < -(alpha.c1[j]*r[k, 1]+alpha.c2[j]*r[k, 2]+alpha.c3
  [j]*r[k, 3])*T[k]#Alpha decomposition predictor function
  }
  }
# A noninformative prior distributions for decomposed alpha-
  factors
# Each parameter in dirichlet distributions is 1
alpha.c1[1:3]~ddirich (delta.1[])
alpha.c2[1:3]~ddirich (delta.2[])
alpha.c3[1:3]~ddirich (delta.3[])
delta.1[1] < -1
delta.1[2] < -1
delta.1[3] < -1
delta.2[1] < -1
delta.2[2] < -1
delta.2[3] < -1
delta.3[1] < -1
delta.3[2] < -1
delta.3[3] < -1
}
```

A.2. DATA

list(*t*=structure(.Data=c(113, 11, 3, 98, 7, 1, 73, 9, 5, 53, 5, 8, 45, 4, 1, 33, 3, 5, 32, 2, 0, 29, 2, 0, 20, 2, 0, 20, 1, 0, 16, 2, 1, 14, 1, 1, 13, 1, 1, 12, 1, 2, 9, 1, 1, 7, 1, 1),.Dim=c(16,3)), *r*=structure(.Data=c(0.2520, 0.2200, 0.5280, 0.1600, 0.7360, 0.1040, 0.2070, 0.2180, 0.5750, 0.4390, 0.0910, 0.4700, 0.1400, 0.6600, 0.2000, 0.3660, 0.2200, 0.4140, 0.3530, 0.4410, 0.2060, 0.0650, 0.7100, 0.2250, 0.3180, 0.1820, 0.5000, 0.4760, 0.3810, 0.1430, 0.1580, 0.3160, 0.5260, 0.4380, 0.1880, 0.3740, 0.2000, 0.3333, 0.4667, 0.3330, 0.2000, 0.4670, 0.3640, 0.4550, 0.1810, 0.1110, 0.6670, 0.2220),.Dim=c (16,3)), *m*=3, case.number=16)

References

[1] WASH-1400, NUREG-75/014. Reactor safety study: an assessment of accident risks in U.S. commercial nuclear power plants. Washington, DC: U.S. Nuclear Regulatory Commission; 1975.
[2] NUREG/CR-4780, vols. 1 and 2. Procedures for treating common-cause failure in safety and reliability studies: procedural framework and examples. Washington, DC: U.S. Nuclear Regulatory Commission; 1989.
[3] Fleming KN. A reliability model for common mode failure in redundant safety systems. In: Proceedings of the sixth annual Pittsburgh conference on modeling and simulation; 1975.
[4] Fleming KN, Kalinowski AM. An extension of the beta factor method to systems with high levels of redundancy. Pickard, Lowe and Garrick, Inc.; 1983 PLG-0289.
[5] Mosleh A, Siu NO. A multi-parameter, event-based common-cause failure model. In: Proceedings of the ninth international conference on structural mechanics in reactor technology; 1987.
[6] NUREG/CR-5485. Guidelines on modeling common-cause failures in probabilistic risk assessment. Washington, DC: U.S. Nuclear Regulatory Commission; 1998.
[7] CCF parameter estimations, 2009 update. U.S. Nuclear Regulatory Commission; 2011.
[8] NUREG/CR-5497. Common-cause failure parameter estimations. Washington, DC: U.S. Nuclear Regulatory Commission; 1998.
[9] NUREG/CR-6819, Vol. 1–Vol. 4. Common-cause failure event insights. Washington, DC: U.S. Nuclear Regulatory Commission; 2003.
[10] NUREG/CR-6268. Common-cause failure database and analysis system: event data collection, classification, and coding. Washington, DC: U.S. Nuclear Regulatory Commission; 2007.
[11] Rasmuson DM, Kelly DL. Common-cause failure analysis in event assessment. Journal of Risk and Reliability 2008;222:521–32.
[12] Kelly DL, et al. Common-cause failure treatment in event assessment: basis for a proposed new model. In: Proceedings of Probabilistic Safety Assessment and Management (PSAM) 10; 2010.
[13] Koller D, Friedman N. Probabilistic graphical models. Cambridge, MA: The MIT Press; 2009.
[14] Pearl J. Causality: models, reasoning, and inference. 2nd edCambridge University Press; 2009.
[15] Congdon P. Bayesian statistical modeling. John Wiley & Sons, Ltd; 2006.
[16] Yamaguchi A. Seismic fragility analysis of the heat transport system of LMFBR considering partial correlation of multiple failure modes. SMiRT 11 Transaction M 1991.
[17] Kelly DL, Smith CL. Bayesian inference in probabilistic risk assessment—the current state of the art. Reliability Engineering and System Safety 2009;94:628–43.
[18] Atwood CL, Kelly DL. The binomial failure rate common-cause model with WinBUGS. Reliability Engineering and System Safety 2009;94:990–9.
[19] Kelly DL, Atwood CL. Finding a minimally informative Dirichlet prior distribution using least squares. Reliability Engineering and System Safety 2011;96:398–402.
[20] Siu NO, Kelly DL. Bayesian parameter estimation in probabilistic risk assessment. Reliability Engineering and System Safety 1998;62:89–116.
[21] Zheng X, Yamaguchi A, Takata T. α-Decomposition Method: a new approach to the analysis of common cause failure. In: Proceedings of Probabilistic Safety Assessment and Management (PSAM) 11 & European Safety and Reliability Association (ESREL) 2012; Helsinki, Finland, 25–29 June 2012.
[22] Mosleh A. Hidden sources of uncertainty: judgment in the collection and analysis of data. Nuclear Engineering and Design 1986;93:187–98.
[23] Mosleh A, Parry GW, Zikria AF. An approach to the analysis of common cause failure data for plant-specific application. Nuclear Engineering and Design 1994;150:25–47.