Dependence assessment in human reliability analysis based on cloud model and best-worst method

Changcheng Ji[a], Fei Gao[a,*], Wenjiang Liu[a]

[a]*School of Aeronautics, Shandong Jiaotong University, Jinan, Shandong, China*

Abstract

Dependence assessment, which is to assess the dependence level of human errors, is an essential part of human reliability analysis, which could be affected by the complexity and uncertainty of the real world. In this paper, a novel dependence assessment method based on cloud model and best-worst method (BWM) is proposed. Firstly, the influential factors used to measure the dependence level are identified. Then, the social network trust graph of different experts is constructed, and the weights of different experts are determined. Next, the cloud model is adopted to represent the linguistic judgments of experts, where the linguistic judgments are transferred into cloud models, and the assessments of different experts are combined. Finally, based on the dependence level of each factor, the final dependence assessment result is obtained. Two numerical examples are presented to show that the proposed method can effectively provide reliable assessment results under uncertainty. In conclusion, the proposed method provides a novel and effective way for dependence assessment in human reliability analysis.

Keywords: Dependence assessment, Human reliability analysis, Decision-making, Cloud model, Best-worst method

1. Introduction

In complex human-machine systems, reliability is one of the most important concerns, and human operation often plays a significant role, especially regarding the probability safety assessment (PSA) of the system as human

*Corresponding Author
Email address: gaofei1995@hotmail.com (Fei Gao)

errors could often be the cause of engineering accidents and the contribution of human operation to the reliability and safety of the system is often large [1–4]. Human reliability analysis (HRA), which is to quantify human contribution to the system risk for a given task, has drawn significant attention as most engineering systems could be regarded as human-machine systems, and there have been various studies focusing on the human reliability aspect in reliability engineering. For instance, Sezer et al. [3] proposed an extended human error assessment and reduction technique (HEART) evidence theory approach for assessing human reliability systematically during the gas freeing process on a chemical tanker ship. Catelani et al. [5] proposed a human reliability analysis technique based on an enhanced simulator for human error probability analysis (SHERPA) method for railway applications to simulate a time-dependent model of the human error probability during the work shift. Liu et al. [6] developed a large group success likelihood index method (SLIM) model to calculate the human error probabilities of operational tasks considering experts' noncooperative behaviors and social relations. Abreu et al. [7] proposed a HRA approach based on the prospective technique for early consideration of human reliability (TECHR) combined with Bayesian networks (BNs) to better understand the contribution of the human and organizational factors in maritime pilotage operations. Ahn and Kurt [8] introduced a new approach based on cognitive reliability and error analysis method (CREAM) for human reliability assessment in the maritime domain. By conducting HRA, it is possible to evaluate and determine the operator's contribution to system reliability by predicting human error rates and evaluating the degradation in human-machine systems likely caused by human errors in association with factors that could influence the system behavior [9–11].

In HRA, dependence assessment, which is to evaluate the dependence among human failure events (HFEs) and the effect of the dependence on the final human error probability (HEP), is an important issue [12–14]. Normally, when there is dependence among tasks, the failure probability of a task would be higher given the failure of its preceding task. Hence, it is important to properly assess the dependence among tasks to more accurately evaluate system risk. The result of dependence assessment is the conditional human error probability (CHEP), given the failure of the preceding task [15]. Dependence assessment is often considered in most of human reliability methods, including technique for human error rate prediction (THERP) [16–18], CREAM [19–21], accident sequence evaluation program (ASEP) [22, 23], standardized plant analysis risk-human reliability analysis method (SPAR-H) [24–26], and others [27, 28], among which THERP has been one of the most representative and widely used method. In THERP, five dependence levels, i.e., zero dependence (ZD), low dependence (LD), moderate dependence (MD), high dependence (HD) and complete dependence (CD), are defined for evaluating the level of dependence between two tasks regarding several influential factors. However, despite its simplicity, THERP still received some criticism for its lack of traceability and repeatability [15, 29].

To overcome these limitations, several methods have been adopted to extend the THERP model, including decision tree (DT) [30, 31], evidence theory [3, 32, 33], fuzzy expert system (FES) [15, 29, 34] and evidential reasoning (ER) [10, 35]. For example, Paglioni and Groth [36] systemically analyzed fundamental concept of dependency in HRA methods, and proposed a standardized library of key terms and mathematics to provide a basis for the development of a dependency framework. Zio et al. [29] used fuzzy expert system in dependence assessment, where a fuzzy rule base is constructed based on experts knowledge, and the assessment is conducted based on the FES. Su et al. [32] adopted the evidence theory into dependence assessment by using basic probability assignments to represent the experts judgments on the dependence level. Zhang et al. [15] studied dependence assessment using the belief rule-based system, where the belief rule-based system is employed to model uncertainties in experts' knowledge and the interval belief distribution is used to model interval uncertainty in expert's judgment. Bi et al. [35] proposed a novel dependence assessment method based on the interval evidential reasoning algorithm. Liu et al. [37] employed a system dynamics approach to the modeling and analysis of the dependencies of performance shaping factors (PSFs) within the standardized plant analysis of risk–human reliability analysis method. Gao et al. [38] integrated the 2-tuple linguistic variables and DEMATEL method, and proposed a novel way to assess the dependence among human actions in HRA, where 2-tuple linguistic variables are used to model the linguistic judgments of the experts. Gao et al. [39] proposed a novel dependence assessment method based on the probabilistic linguistic term set (PLTS) to assess dependence among human actions, where the judgments on the influential factors are modeled using PLTS.

By summarizing the development of dependence assessment methods in HRA, two main issues have received the most attention. One is how to represent the experts' knowledge of the dependence level of the influential factors under uncertainty. Fuzzy numbers and belief functions are the most commonly used [29, 32, 35]. The other is how to accurately and intuitively express the dependence assessment under uncertainty, where current methods mainly employ fuzzy numbers, belief functions and D numbers. However, in previous studies, though the assessment result is mostly reliable, it may not always be straightforward, and the uncertainty in the assessment is sometimes ignored.

To address these limitations, a novel dependence assessment based on cloud model and best-worst method (BWM) is presented in this paper. As a new method for linguistic computation based on probability theory and fuzzy set theory, the cloud model can effectively model uncertainty and provide a more intuitive representation in the form of cloud [40–42]. In this paper, the cloud model is employed to represent the experts' judgments under uncertainty, and the AHP is adopted to determine the weights of different influential factors. In addition, a novel subjective weighting method based on social network trust graph is developed to determine the weights of experts' judgments, which makes the results more reliable. A case study is presented to demonstrate the effectiveness of the proposed method. By integrating the cloud model and the BWM into dependence assessment, the proposed method could effectively represent the experts' judgments and provide a more reliable and intuitive assessment result.

The rest of the paper is organized as follows. Section 2 briefly introduces the cloud model. Section 3 presents the proposed dependence assessment method. Section 4 presents two practical cases, and some discussions are provided in Section 5. Finally, Section 6 concludes the paper.

2. Preliminaries

2.1. Cloud model

As the foundation of cloud-based reasoning, computing and control, the cloud model is an uncertain transformation model for addressing qualitative concepts and quantitative descriptions. The cloud model can represent the process from qualitative concept to quantitative representation through forward cloud generator, and it can also represent the process from quantitative representation to qualitative concept through reverse cloud generator.

Let U be a qualitative domain, and C be the corresponding qualitative concept on U. Suppose x is a random number that obeys a normal distribution with $x \in U$, and the membership degree $\mu(x)$ of x for C is a random number with a stable inclination that satisfies $\mu(x) \in [0, 1]$. Then x and its distribution on U is called as cloud droplets and clouds, respectively.

In the cloud model, the uncertainty of the data x is expressed through three values: The expected value Ex, which can best reflect this qualitative concept in the argument domain space and is the location of the cloud center of gravity. The entropy En, which represents the desirable range of assessment results and the degree of cloud droplet clustering, i.e., its randomness and fuzziness. The hyper entropy He, which reflects the dispersion degree of the cloud droplets. The characteristics of the cloud model $C = (Ex, En, He)$ can be calculated as:

$$Ex = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{1}$$

$$En = \sqrt{\frac{\pi}{2}} \times \frac{1}{n}\sum_{i=1}^{n} |X_i - Ex| \tag{2}$$

$$He = \sqrt{\left|\frac{1}{n-1}\sum_{i=1}^{n} (X_i - Ex)^2 - En^2\right|} \tag{3}$$

where X_i $(i = 1, 2, \ldots, n)$ represents the ith data from the distribution, and n is the number of data in the distribution.

By using the forward cloud generator based on the characteristics obtained by the cloud model, a positive random number $x \sim N(Ex, En^2)$ can be obtained, as shown in Fig 1, and a cloud droplet is defined as $(x, \mu(x))$, where the cloud droplet membership degree $\mu(x)$ is calculated by:

$$\mu(x) = exp\left[\frac{-(x - Ex)^2}{2En^2}\right] \tag{4}$$

Let there be two clouds $C_i = (Ex_i, En_i, He_i)$ and $C_j = (Ex_j, En_j, He_j)$, there is:

$$\begin{aligned}
C_i + C_j =& (Ex_j + Ex_j, \sqrt{En_i^2 + En_j^2}, \sqrt{He_i^2 + He_j^2}) \\
C_i - C_j =& (Ex_j - Ex_j, \sqrt{En_i^2 + En_j^2}, \sqrt{He_i^2 + He_j^2}) \\
C_i \times C_j =& (Ex_i \times Ex_j, \sqrt{(En_i Ex_j)^2 + (En_j Ex_i)^2}, \sqrt{(He_i Ex_j)^2 + (He_j Ex_i)^2}) \\
\lambda C_i =& (\lambda Ex_i, \sqrt{\lambda} En_i, \sqrt{\lambda} He_i)
\end{aligned} \tag{5}$$

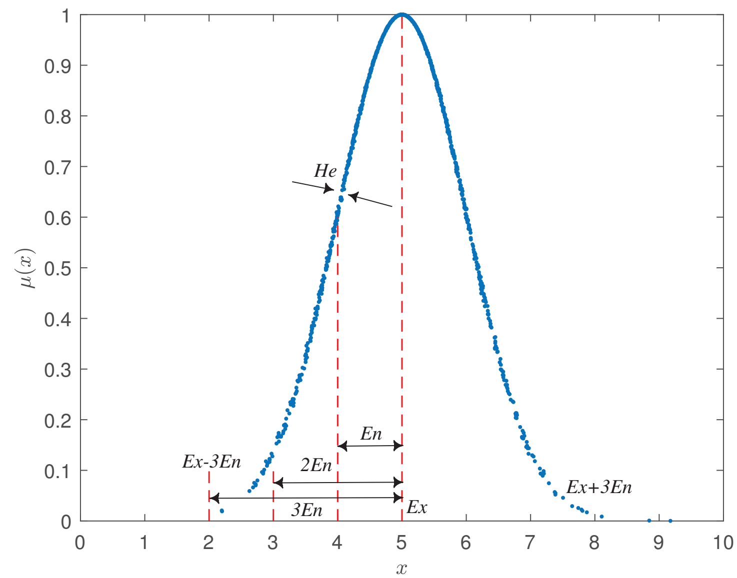

Figure 1: Illustration of the cloud model.

Let $C_i = (Ex_i, En_i, He_i)$ and $C_j = (Ex_j, En_j, He_j)$ be two clouds, then the distance between them is defined as:

$$d_{i,j} = \sqrt{|Ex_i - Ex_j|^2 + |En_i - En_j|^2 + |He_i - He_j|^2} \tag{6}$$

Let $C_i = (Ex_i, En_i, He_i)$ $(i = 1, 2, \dots, n)$ be a set of clouds in the domain U, the cloud weighted average (CWA) operator is defined as a mapping as:

$$\begin{aligned} CWA(C_1, C_2, \dots, C_n) &= \sum_{i=1}^{n} w_i C_i \\ &= \left(\sum_{i=1}^{n} w_i Ex_i, \sqrt{\sum_{i=1}^{n} w_i (En_i)^2}, \sqrt{\sum_{i=1}^{n} w_i (He_i)^2} \right) \end{aligned} \tag{7}$$

where $(w_1, w_2, \dots, w_n)$ is the weight vector with $0 \le w_i \le 1$ and $\sum_{i=1}^{n} w_i = 1$.

In the cloud model-based assessment, the $3En$ principle is often applied to analyze the assessment results as the droplets in the cloud diagram are mainly concentrated in the $[Ex - 3En, Ex + 3En]$ interval, as shown in Fig 1. It should be noted that different distribution locations of the cloud droplets represent different qualitative assessments, which are mainly divided into three part: the main part $(Ex - En, Ex + En)$, which has the highest membership degree, the secondary part $(Ex-2En, Ex-En) \cup (Ex+En, Ex+2En)$ and the minor part $(Ex - 3En, Ex - 2En) \cup (Ex + 2En, Ex + 3En)$. Cloud droplets beyond the interval are normally not used as the basis for qualitative description of the assessment.

2.2. Transform between linguistic terms and clouds

Let $U = [x_{\min}, x_{\max}]$ be the effective domain and $S = \{s_0, s_1, \ldots, s_T\}$ be a linguistic term set, then $T + 1$ basic clouds can be generated based on the golden segmentation method as [43]:

$$C_1 = (Ex_0, En_0, He_0), C_1 = (Ex_1, En_1, He_1), \ldots, C_T = (Ex_T, En_T, He_T) \tag{8}$$

For example, for a linguistic term set with five linguistic terms, five basic clouds can be obtained as:

$$\begin{aligned}
C_0 &= \left(x_{\min}, \frac{En_1}{0.618}, \frac{He_1}{0.618}\right) \\
C_1 &= \left(Ex_2 - 0.382\frac{x_{\max} + x_{\min}}{2}, 0.382\frac{x_{\max} + x_{\min}}{6}, \frac{He_2}{0.618}\right) \\
C_2 &= \left(\frac{x_{\max} + x_{\min}}{2}, 0.618En_1, He_2\right) \\
C_3 &= \left(Ex_2 + 0.382\frac{x_{\max} + x_{\min}}{2}, 0.382\frac{x_{\max} + x_{\min}}{6}, \frac{He_2}{0.618}\right) \\
C_4 &= \left(x_{\max}, \frac{En_3}{0.618}, \frac{He_3}{0.618}\right)
\end{aligned} \tag{9}$$

where $x_{\min}$, $x_{\max}$ and He_2 are given in advance.

3. Method

In this section, a novel dependence assessment method based on the cloud model and BWM is developed to assess the dependence among human errors in HRA, which is consisted of five parts: (1) Influential factors identification; (2) Expert weight calculation; (3) Influential factor weight calculation; (4) Cloud model-based influential factor assessment; (5) CHEP calculation. The detailed process of the proposed method is shown in Fig 2.

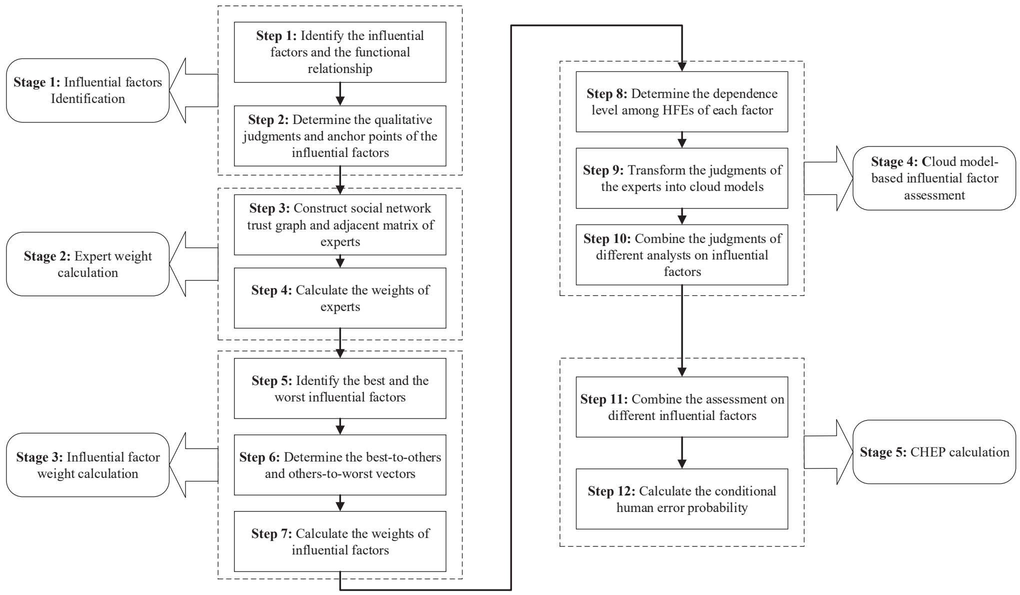

Figure 2: Procedure of the proposed method.

Step 1: Identify the influential factors and the functional relationship

The first step of the dependence assessment is to identify the factors that have influences on the dependence of human actions. For example, in previous research, five influential factors are identified in the THERP model, namely, "closeness in time", "similarity of performers", "task relatedness", "similarity of cues" and "similarity of goals" [18]. It is worth noting that in THERP model, it is suggested that the influential factors could consider closeness in time and space, functional relatedness (e.g., tasks related to the same subsystem), stress, similarity of the performers (status, training, responsibility, and many social and psychological factors), hence, similar influential factors could be determined for different situations [44, 45]. Furthermore, it is also worth noting that several systemic approaches such as grounded theory, process hazards analysis (PHA) and process flow failure modes (PFFM) analysis could be used for identifying the influential factors, and the identification of the influential factors could be connected to the specific situation, when the situation changes, the influential factors may change accordingly.

Moreover, as there are normally several influential factors for dependence assessment, it is also necessary to determine the relationship among these factors, especially for cases where the influential factors are not at the same level. For instance, the functional relationship among the five influential factors introduced above is shown in Fig 3. It can be seen that for all five influential factors, four of them are input factors that can be directly assessed through the judgments of the experts, and one factor, "task relatedness", is an intermediate factor that should be assessed based on the assessments of its sub-factors [18]. It is worth noting that the functional relationship among the influential factors is mainly determined based on the analysis of the factors and their corresponding relations, and different approaches including the grounded theory, process hazards analysis (PHA), and process flow failure modes (PFFM) analysis could be used to determine the relationship.

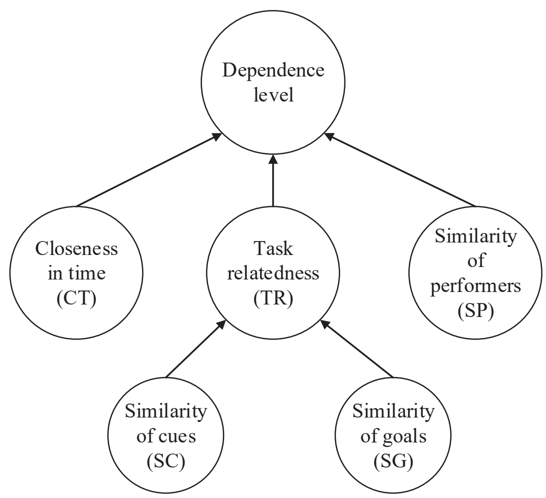

Figure 3: Example of the functional relationship among the influential factors.

Step 2: Determine the qualitative judgments and anchor points of the influential factors

The qualitative judgments and the corresponding anchor points provide guidance for the judgments of experts on the influential factors. Hence, in order to assess the dependence among human errors, once the influential factors are determined, their qualitative judgments and anchor points should be determined accordingly, where the qualitative judgments qualitatively describe the effect of the factor on the dependence level and the anchor points represent the possible situation of the factor. For example, for the factor "closeness in time", different values would result in different effects on the dependence between two actions, such as "5 min" may indicate "Complete Dependence", "30 min" may indicate "Moderate Dependence", and "8 h" may indicate "Zero Dependence".

It is worth noting that in the THERP model, five qualitative judgments are identified, namely, "Zero dependence", "Low dependence", "Medium dependence", "High dependence", and "Complete dependence", which are used for dependence assessment in many previous studies and are used in this study. As for anchor points, it should be noted that they are mainly determined based on the analysis of the influential factors and the knowledge of experts [18].

Step 3: Construct social network trust graph and adjacent matrix of experts

In the proposed method, a social network that is defined by a directional graph $G(E, L)$ is used to reflect the trust relation of different experts, which could be regarded as different importance of experts as being trusted by other experts obviously means higher importance. In the social network trust graph, edge (E_i, E_j) represents the trust from E_i to E_j, and (E_j, E_i) represents the trust from E_j to E_i. Hence, a bidirectional trust graph could be constructed to model the trust relation among the experts.

Moreover, as the trust graph only represents the trust relation not the specific trust degree, the adjacent matrix is constructed based on the trust graph to represent the trust among the experts, which is expressed as:

$$A = [a_{ij}]_{l \times l} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1l} \\ a_{21} & a_{22} & \cdots & a_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ a_{l1} & a_{l2} & \cdots & a_{ll} \end{bmatrix} \tag{10}$$

where a_{ij} represents the trust degree from E_i to E_j.

It should be noted that the number of experts may vary for different situations, and there are no strict guidelines on how many experts are required. However, in many cases, the number of experts mainly lies between 3 to 10.

Step 4: Calculate the weights of experts

In the proposed method, the weights of experts are calculated based on the in-degree centrality index, which represents the overall trust an expert could receive from other experts, and the in-degree centrality index $C(E_k)$

of expert E_k is calculated by:

$$C(E_k) = \sum_{i=1}^{n} E(a_{ik}) \tag{11}$$

where $E(a_{ik})$ is the converted crisp value of a_{ik}.

Generally, the higher the value of the in-degree centrality index, the higher trust the expert receives from other experts, and the more important the expert is. Hence, the in-degree centrality index could be regarded as the importance of the expert, and the weights of experts could be calculated as:

$$\phi_k = \frac{C(E_k)}{\sum_{k=1}^{l} E_k} \tag{12}$$

where $0 \leq \phi_k \leq 1$, $\sum_{k=1}^{l} C(E_k) = 1$.

Step 5: Identify the best and the worst influential factors

In dependence assessment, there normally are several influential factors with functional relationship, furthermore, these influential factors may have different importance. Therefore, it is necessary to determine the weights of the influential factors, which describe the degree to which the factors influence the dependence level. In the proposed method, the BWM is adopted to weigh the influential factors [46]. For the n influential factors $F_j(j = 1, 2, \ldots, n)$, each expert E_k is invited to identify the best, i.e., most important, influential factor F_B^k and the worst, i.e., least important, influential factor F_W^k based on their understanding and knowledge of the problem.

Step 6: Determine the best-to-others and others-to-worst vectors

The preferences of the best influential factor to other influential factors and the preferences of other influential factors to the worst influential factor are assessed by experts using linguistic terms, and the best-to-others and others-to-worst vectors of E_k are obtained as:

$$\begin{aligned} BO_k &= \left(v_{B1}^k, v_{B2}^k, \ldots, v_{Bn}^k\right) \\ OW_k &= \left(v_{1W}^k, v_{2W}^k, \ldots, v_{nW}^k\right) \end{aligned} \tag{13}$$

where v_{Bj}^k denotes the equivalently converted crisp values of the expert's linguistic judgment of F_B^k over the jth influential factor, and v_{jW}^k denotes the equivalently converted crisp values of the expert's linguistic judgment of the jth influential factor over the worst influential factor F_W^k, $v_{BB}^k = v_{WW}^k = 1$.

It should be noted that the elements in the best-to-others and other-to-worst vectors represent the preference between the influential factors, expressed by crisp values 1–9 corresponding to linguistic terms, where higher values indicate higher importance.

Step 7: Calculate the weights of influential factors

The optimal weights of influential factors are the one where the maximum absolute distance between $|w_B^k/w_j^k - v_{Bj}^k|$ and $|w_j^k/w_W^k - v_{jW}^k|$ is minimized. Therefore, in order to calculate the weights of influential factors $(w_1^{k*}, w_2^{k*}, \ldots, w_n^{k*})$, in this step, the following optimization model is established:

$$\begin{aligned} &\min \ \max_j \left\{ \left| \frac{w_B^k}{w_j^k} - v_{Bj}^k \right|, \left| \frac{w_j^k}{w_W^j} - v_{jW}^k \right| \right\} \\ &s.t. \begin{cases} \sum_{j=1}^{n} w_j^k = 1 \\ 0 \le w_j^k \le 1 \end{cases} \end{aligned} \tag{14}$$

where ω_B^k, ω_W^k, and ω_j^k denotes the weight of the best factor, the worst factor, and the jth factor determined by E_k, respectively.

Eq. (14) can be equivalently converted to the following optimization model:

$$\begin{aligned} &\min \ \xi^k \\ &s.t. \begin{cases} \left| \frac{w_B^k}{w_j^k} - v_{Bj}^k \right| \le \xi^k \\ \left| \frac{w_j^k}{w_W^k} - v_{jW}^k \right| \le \xi^k \\ \sum_{j=1}^{n} w_j^k = 1 \\ 0 \le w_j^k \le 1 \end{cases} \end{aligned} \tag{15}$$

Hence, by solving the above optimization model, the weights of the influential factors $(w_1^{k*}, w_2^{k*}, \ldots, w_n^{k*})$ from E_k can be determined.

By combining the calculated weights of influential factors from different experts, the overall weights of influential factors can be obtained by:

$$w_i = \sum_{k=1}^{l} \phi_k w_i^{k*} \tag{16}$$

Step 8: Determine the dependence level among HFEs of each factor

In the dependence assessment, experts are asked to provide their judgments on the input factors by referring to the anchor points and qualitative judgments. By referring to the anchor points, the subjectivity of the judgments can be reduced. However, as can be seen from Step 2, since the judgments on the influential factors are often in the term of linguistic terms, the judgments provided by the experts could also be in the form of linguistic terms, and a linguistic assessment matrix can be obtained as:

$$Y = [y_{ij}]_{M \times N} \tag{17}$$

where y_{ij} denotes the linguistic judgment of the ith expert on the jth influential factor with the linguistic term set $S = \{s_0, s_1, \ldots, s_T\}$, M is the number of experts and N is the number of influential factors.

Step 9: Transform the judgments of the experts into cloud models

In dependence assessment, the judgments provided by the experts are often in the form of linguistic terms, which could then be transformed into cloud models for the subsequent calculation. According to the conversion method between linguistic terms and clouds, each linguistic judgment can be converted into a normal cloud, denoted as $C_{ij} = (Ex_{ij}, En_{ij}, He_{ij})$, and the cloud assessment matrix can be obtained as:

$$Y = [C_{ij}]_{M \times N} \tag{18}$$

where C_{ij} represents the transformed cloud model of the ith expert's judgment on the jth influential factor.

Step 10: Combine the judgments of different experts on influential factors

After determining the weights of different judgments on the same influential factor, the judgments of different experts could be combined to obtain a comprehensive assessment of the factor. Since the judgments are in the form of clouds and weights of different judgments are considered, the judgments of different experts could be combined using the CWA operator as:

$$CWA(C_{1j}, C_{2j}, \ldots, C_{Mj}) = \left(\sum_{i=1}^{M} w_i Ex_{ij}, \sqrt{\sum_{i=1}^{M} w_i (En_{ij})^2}, \sqrt{\sum_{i=1}^{M} w_i (He_{ij})^2} \right) \tag{19}$$

where w_i denotes the weight of the ith experts' judgment, $C_{ij} = (Ex_{ij}, En_{ij}, He_{ij})$ represents the judgment of the ith expert on the jth factor.

Step 11: Combine the assessment on different influential factors

After obtaining the combined assessments on the influential factors, the assessments should then be combined to obtain the assessment of the influential factors at the upper level. It can be noted that since the influential factors have certain kinds of functional relationship, only factors that are at the same level of the functional relationship will be combined, and the assessment of a factor at the upper level is obtained by combining the assessments of its sub-factors. For example, for the influential factors shown in Fig 3, the assessment of factor "Task relatedness" is obtained by combining the assessments of factors "Similarity of cues" and "Similarity of goals".

Step 12: Calculate the conditional human error probability

In order to obtain a more illustrative dependence assessment result, the CHEP value should be calculated. Suppose the combined dependence assessment result is $C = (Ex, En, He)$, by adopting the $3En$ principle, it can be noted that the CHEP value would be an interval as:

$$\begin{aligned} p^-(B|A) &= Ex - 3En \\ p^+(B|A) &= Ex + 3En \end{aligned} \tag{20}$$

Thus, the CHEP value can be obtained as $p(B|A) \in [p^-(B|A), p^+(B|A)]$. Furthermore, for illustrative purpose, a representative CHEP value could be obtained as:

$$p_{avg}(B|A) = Ex \tag{21}$$

4. Case study

In order to verify the effectiveness and validity of the proposed method, two numerical examples are studied in this section.

4.1. Case 1: Post-initiator human failure events of nuclear power plant

4.1.1. Case description

This case study refers to a set of operator actions to avoid excessive boron dilution in the reactor cooling system in case of an anticipated transient without scram (ATWS) at a nuclear boiling water reactor (BWR). It is assumed that the standby liquid control system (SLCS) has been successfully initiated by the operators to shut the reactor down. The operators are required to increase the voiding and inhibit the actuation of the automatic depressurization system (ADS) to facilitate the reactor shutdown. The operator tasks include the prevention of the ADS (Action A) and the control of the reactor vessel level (Action B) to prevent diluting boron concentration after the ADS failure. The probability of human failure in controlling the reactor vessel level is used as the output [39].

4.1.2. Implementation

Based on the proposed method presented in last section, the dependence degree between two human operations is evaluated as below.

Step 1: Identify the influential factors and the functional relationship

Based on previous studies [15, 29], five factors, namely, "closeness in time", "similarity of performers", "task relatedness", "similarity of cues" and "similarity of goals" are identified as influential factors to evaluate the dependence degree of human operations. Among these five factors, three of which directly impact the dependence degree, namely, "closeness in time", "similarity of performers", "task relatedness", "task relatedness" is further divided into two sub-factors, "similarity of cues" and "similarity of goals". The hierarchical structure of the influential factors is shown in Fig 3.

Step 2: Determine the qualitative judgments and anchor points of the influential factors

For each influential factor, different dependence levels with different anchor points and corresponding qualitative judgments are provided based on the knowledge of experts. In this case, five dependence levels are defined based on previous research, namely, "Zero dependence", "low dependence", "Medium dependence", "High dependence" and "Complete dependence". The anchor points and the corresponding qualitative judgments for each influential factor are shown in Tables 1-4. It should be noted that there are no anchor points for "task relatedness" as it is an intermediate factor.

Table 1: Anchor points and qualitative judgments for factor "closeness in time"

Anchor point	Dependence level	Qualitative judgment
8 h	ZD	Two tasks are widely separated in time
1 h	LD	The time difference between tasks is less than wide
30 min	MD	It is not relevant in the dependence assessment
20 min	HD	The tasks are in a short time window, but not close enough
5 min	CD	The two tasks are close in time

Table 2: Anchor points and qualitative judgments for factor "similarity of performers"

Anchor point	Dependence level	Qualitative judgments
TSC vs control shift room	ZD	No similarity of performers is present between tasks
Different steam	LD	A low level of performer similarity exists
Different individual with same qualification	MD	The level of performer similarity is medium
Same team	HD	High level of performer similarity is present between tasks
Same person	CD	The tasks are accomplished by the same individual. In this case, the similarity of performer is complete.

Table 3: Anchor points and qualitative judgments for factor "similarity of cues"

Anchor point	Dependence level	Qualitative judgment
Different sets of indicators for different parameters	ZD	No similarity of cues is present between tasks
Different sets of indicators for the same parameters	LD-MD	An intermediate level of cues similarity exists although not fully medium
Single indicator for the same parameter	MD-HD	The level of cues similarity is more than medium
Different sets of indicators for the same physical quantity	HD-CD	Slightly more than high level of similarity of cues is present between the tasks
Same sets of indicators for the same sets of parameters	CD	The tasks present complete similarity of cues

Table 4: Anchor points and qualitative judgments for factor "similarity of goals"

Anchor point	Dependence level	Qualitative judgment
Different functions by different systems	ZD	No similarity of cues is present between tasks
Different functions by same system	LD	A low level of goals similarity exists
Same function by different systems	HD	The level of goals similarity is high
Same function by same system	CD	A complete level of similarity of goals is present between tasks

Step 3: Construct social network trust graph and adjacent matrix of experts

In this case, three experts are invited, and a five-level linguistic term set is used for experts to represent their trust degrees, as shown in Table

5. By analyzing the trust relation, the bidirectional social network trust graph is constructed, as shown in Fig 4, where the trust relations between experts are bidirectional. Based on the trust graph, each expert provides his/her trust degree to other experts, and the adjacent matrix is established, as shown in Table 6. It should be noted that the trust degrees of the experts themselves are set to be the highest, i.e., s_5, as they have absolute trust in their assessments.

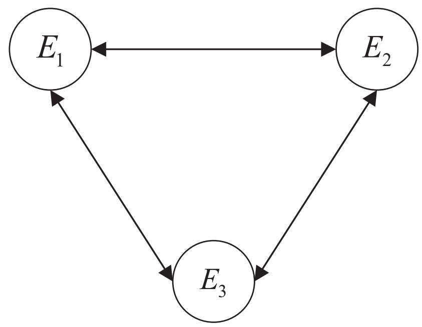

Figure 4: Social network trust graph of experts.

Table 5: Linguistic terms for expert trust degree

Linguistic term	Trust level	Crisp value
s_1	Very low	1
s_2	Low	3
s_3	Moderate	5
s_4	High	7
s_5	Very high	9

Table 6: Adjacent matrix of experts

Expert	E_1	E_2	E_3
E_1	s_5	s_3	s_4
E_2	s_1	s_5	s_4
E_3	s_2	s_3	s_5

Step 4: Calculate the weights of experts

According to the adjacent matrix, the in-degree centrality index of the experts could be calculated using Eq. (11) as $C = C(E_k) = (13, 19, 23)$. Hence, via Eq. (12), the weights of experts are obtained as $\phi = (0.2363, 0.3455, 0.4182)$.

Step 5: Identify the best and the worst influential factors

In this case, the influential factors are with hierarchical relationship, thus, the best and the worst influential factors are identified separately, corresponding to the hierarchical level. For the first level, i.e., "closeness in time" (CT), "similarity of performers" (SP), "task relatedness" (TR), each expert provides his/her judgments on the best and the worst influential factors, as listed in Table 7. Then, for the second level, i.e., "similarity of cues" (SC) and "similarity of goals" (SG), each expert identifies the best and the worst influential factors, shown in Table 7.

Table 7: Best and worst influential factors

Expert	First level		Second level	
	Best	Worst	Best	Worst
E_1	SC	SG	TR	CT
E_2	SC	SG	TR	CT
E_3	SC	SG	TR	SP

Step 6: Determine the best-to-others and others-to-worst vectors

Corresponding to the two levels, two sets of best-to-others and others-to-worst vectors are determined by the experts, as listed in Table 8.

In terms of all the influential factors with hierarchical relationship, the weights of the influential factors that are at the second level should be determined first, that is, influential factors "similarity of cues" and "similarity of goals". The best and the worst factors among these two factors, as well as the best-to-others and others-to-worst vectors are obtained from the judgments of each expert, as listed in Table 7. It should be noted that the values of the best-to-others vector and others-to-worst vector are determined based on the judgments of the experts on the relative preference among the different influential factors, which are converted from corresponding linguistic judgments, where higher values indicate higher importance.

Step 7: Calculate the weights of influential factors

Hence, the optimization model could be established to obtain the weights of the influential factors from different experts. For example, the optimiza-

Table 8: Best-to-others and others-to-worst vectors of the experts

Expert	Vector	First level			Second level	
		CT	TR	SP	SC	SG
E_1	Best-to-others	5	3	1	1	2
	Others-to-worst	1	3	5	2	1
E_2	Best-to-others	5	2	1	1	3
	Others-to-worst	1	3	5	3	1
E_3	Best-to-others	3	1	5	1	2
	Others-to-worst	3	5	1	2	1

tion model regarding E_1 for the second-level factors is constructed as:

$$\min \ \xi^1$$
$$s.t. \begin{cases} \left|\dfrac{w_1^1}{w_j^1} - v_{1j}^1\right| \leq \xi^1 \\ \left|\dfrac{w_j^1}{w_2^1} - v_{j2}^1\right| \leq \xi^1 \\ \sum\limits_{j=1}^{2} w_j^1 = 1 \\ 0 \leq w_j^1 \leq 1, j = 1, 2 \end{cases}$$

By solving the optimization model, the optimal weight from E_1 is obtained as $w_{SC}^1 = \frac{2}{3}, w_{SG}^2 = \frac{1}{3}$. Hence, via Eq. (16), the overall weights of influential factors "similarity of cues" w_{SC} and "similarity of goals" w_{SG} are obtained as $w_{SC} = 0.6955, w_{SG} = 0.3045$.

Similarly, the weights of influential factors "closeness in time", "similarity of performers", "task relatedness" are calculated by solving corresponding optimization models, and are obtained as $w_{CT} = 0.1047, w_{SP} = 0.2583, w_{TR} = 0.6370$.

Step 8: Determine the dependence level among HFEs of each factor

Three experts are invited to give their assessment on the dependence level of each influential factor based on the anchor points of the influential factors. The assessment results are shown in Table 9.

Step 9: Transform the judgments of the experts into cloud models

Table 9: Expert assessments on the dependence level of each influential factor

Expert	Influential factor	Dependence level
E_1	closeness in time	MD
	similarity of performers	HD
	similarity of cues	LD
	similarity of goals	LD
E_2	closeness in time	LD
	similarity of performers	CD
	similarity of cues	MD
	similarity of goals	LD
E_3	closeness in time	HD
	similarity of performers	MD
	similarity of cues	ZD
	similarity of goals	MD

Based on Table 9, the linguistic judgments of experts on the dependence level could be transformed into cloud models. First, five basic clouds corresponding to five qualitative judgments could be generated to model the judgments of the experts. In this case, it is set that $x_{\min} = 0$, $x_{\max} = 1$ and $He_2 = 0.02$, and the numerical characteristics of the five basic clouds are determined as:

$$
\begin{aligned}
&C_0 = (0.000, 0.103, 0.052), \quad C_1 = (0.309.0.064, 0.032),\\
&C_2 = (0.500, 0.039, 0.002), \quad C_3 = (0.691, 0.064, 0.032),\\
&C_4 = (1.000, 0.103, 0.052)
\end{aligned}
$$

Then, the linguistic judgments of the experts are transformed into cloud models to establish the cloud assessment matrix. Table 10 shows the cloud model of expert judgments on the dependence level.

Step 10: Combine the judgments of different experts on influential factors

By using Eq. (7), the combined cloud models for each influential factor could be obtained based on the cloud models and weights of different experts, shown in Table 11. The clouds of these influential factors are shown in Figs 5-8.

Step 11: Combine the assessment on different influential factors

In this case, as the influential factors have a hierarchical relationship, the

Table 10: Cloud models of expert judgments on the dependence level

Expert	Influential factor	Cloud model
E_1	closeness in time	(0.500,0.039,0.002)
	similarity of performers	(0.691,0.064,0.032)
	similarity of cues	(0.309.0.064,0.032)
	similarity of goals	(0.309.0.064,0.032)
E_2	closeness in time	(0.309.0.064,,0.032)
	similarity of performers	(1.000,0.103,0.052)
	similarity of cues	(0.500,0.039,0.002)
	similarity of goals	(0.309.0.064,0.032)
E_3	closeness in time	(0.691,0.064,0.032)
	similarity of performers	(0.500,0.039,0.002)
	similarity of cues	(0.000.0.103,0.052)
	similarity of goals	(0.500,0.039,0.002)

Table 11: Combined cloud models for each factor

Influential factor	Cloud model
closeness in time	(0.500,0.056,0.026)
similarity of performers	(0.720,0.072,0.035)
similarity of cues	(0.279,0.071,0.035)
similarity of goals	(0.368,0.057,0.027)

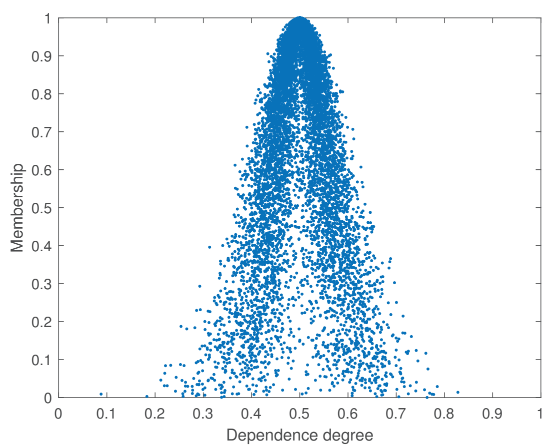

Figure 5: Cloud model of the assessment for influential factor “closeness in time”.

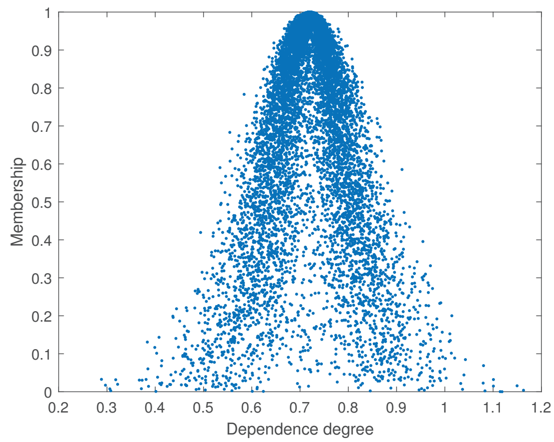

Figure 6: Cloud model of the assessment for influential factor “similarity of performers”.

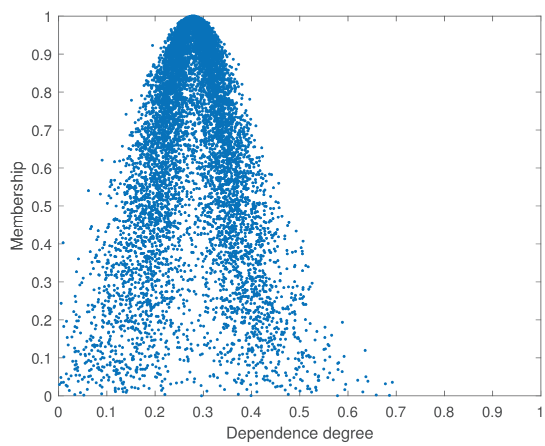

Figure 7: Cloud model of the assessment for influential factor “similarity of cues”.

combination of the assessments can be divided into two parts, corresponding to the hierarchical structure.

Firstly, the dependence level of factor “task relatedness” is calculated based on the assessments on factors “similarity of cues” and “similarity of goals”, which is presented in the form of cloud, shown in Fig 9, as:

$$(0.309, 0.067, 0.033)$$

Then, by combining the clouds of factors “closeness in time”, “similarity

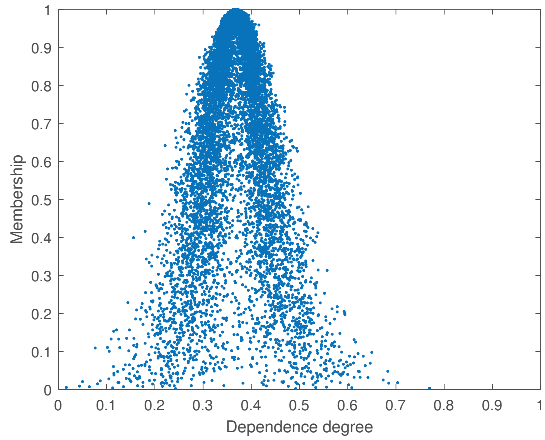

Figure 8: Cloud model of the assessment for influential factor “similarity of goals”.

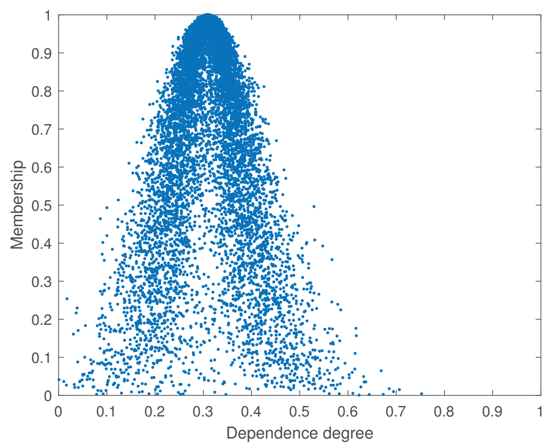

Figure 9: Cloud model of the assessment for influential factor “task relatedness”.

of performers” and “task relatedness”, the final assessment could be obtained as:

$$(0.436, 0.067, 0.033),$$

and the cloud is shown in Fig 10.

Step 12: Calculate the conditional human error probability

Following the process of the last step, the CHEP value could be calculated as $p(B|A) = [0.235, 0.637]$, with the representative CHEP value obtained as

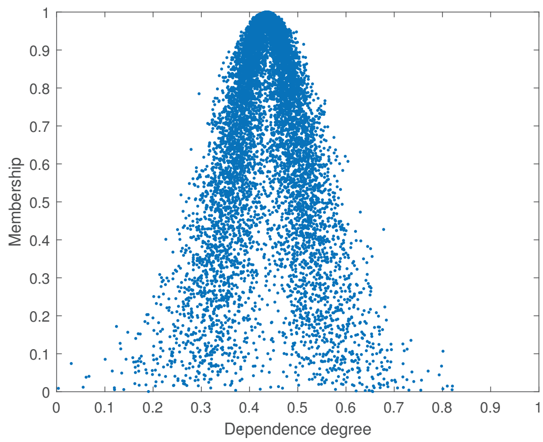

Figure 10: Cloud model of the final dependence assessment.

$p_{avg}(B|A) = 0.436.$

4.2. Case 2: Human error in blood transfusion

4.2.1. Case description

In modern medical treatment, blood transfusion is one of the most important and essential means to save lives and reduce incidence. When conducted properly, blood transfusion could help patients receive blood and reduce harm, however, when conducted improperly, it could lead to infection and even diseases. Thus, safe and reliable blood transfusion practice has been studied by minimizing the risk of human errors during transfusion operation [16]. In this case, the proposed method is applied to analyze and assess CHEP during blood transfusion operation. Five pairs of sequential tasks are considered in this case, as shown in Table 12.

Table 12: HRA of blood transfusion operation

Sequential task	Action	HFE	Cause of HFE
T_1	Preoperative assessment	Insufficient preoperative assessment of blood requirement	Incorrect assessment of potential blood loss
T_2	Request form filling	Incorrect information on request form	Incorrect application form
T_3	Transfusion preparation	Extensive time before injection	Delay in delivering blood to clinics
T_4	Transfusing blood	Inappropriate timing of transfusion	Inappropriate transfusion time
T_5	Transfusion monitoring	Reactions occur in the transfusion	Patients are not properly monitored

In this case, five experts, denoted as E_1, E_2, E_3, E_4 and E_5, are invited to provide their assessment on the dependence level. According to their different backgrounds and experiences, their assessments could vary.

4.2.2. Implementation

In this case, three influential factors are considered, namely, "time closeness" (F_1), "task relatedness" (F_2), and "performers similarity" (F_3). Time closeness represents the time relationship among human actions. Task relatedness is assessed by similarity of cues and similarity of goals, and it is used to represent the relationship among the tasks. Performer similarity represents factors such as training and status in human actions. It should be noted that all these three factors are at the same level.

For assessing the dependence level of different factors, five dependence levels are used, namely, "Zero dependence (ZD)", "Low dependence (LD)", "Medium dependence (MD)", "High dependence (HD)" and "Complete dependence (CD)". The qualitative judgments and corresponding anchor points of each influential factor are listed in Table 13.

Table 13: Anchor points for the influential factors

Time closeness	Task relatedness	Performer similarity	Qualitative judgment
Tasks are widely separated in time	Tasks are unrelated	No similarity of performers is present	ZD
Tasks are moderately farness in time	Tasks are slightly related	Tasks are accomplished by different teams	LD
Closeness in time is not relevant in the dependence assessment	Tasks are moderately related	Tasks are accomplished by different individuals with same qualification	MD
Tasks are moderate close in time	Tasks are highly related	Tasks are accomplished by the same team	HD
Tasks are strong close in time	Tasks are closely related	Tasks are accomplished by the same individual	CD

Then, the trust degrees of the experts to others and to themselves are expressed by a five-level linguistic term set, as shown in Table 5. The bidirectional social trust graph is constructed by analyzing the trust relation among experts, as shown in Fig 11.

Based on the trust graph, the experts are invited to provide their trust degrees to other experts, and the adjacent matrix is constructed, as shown in Table 14, where the trust degrees of the experts to themselves are set to be the highest.

Based on the adjacent matrix, the in-degree centrality index of the experts are calculated via Eq. (11) as $C = C(E_k) = (29, 23, 29, 37, 25)$. Thus, according to Eq. (12), the weights of experts are calculated as $\phi = (0.2028, 0.1608, 0.2028, 0.2587, 0.1748)$.

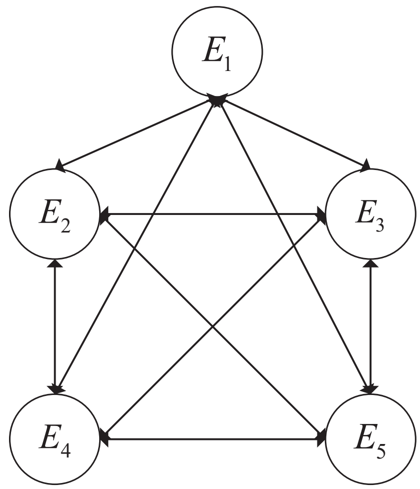

Figure 11: Social network trust graph with five experts.

Table 14: Adjacent matrix of experts

Expert	E_1	E_2	E_3	E_4	E_5
E_1	s_5	s_2	s_4	s_3	s_1
E_2	s_2	s_5	s_3	s_5	s_3
E_3	s_3	s_2	s_5	s_4	s_4
E_4	s_4	s_3	s_2	s_5	s_2
E_5	s_3	s_2	s_3	s_4	s_5

Each expert identifies the best and the worst influential factors according to their understanding and knowledge, and the identified best and worst factors are listed in Table 15.

Table 15: The best and the worst influential factors

Expert	Best influential factor	Worst influential factor
E_1	Task (F_2)	Time (F_1)
E_2	Time (F_1)	Task (F_2)
E_3	Performer (F_3)	Task (F_2)
E_4	Task (F_2)	Performer (F_3)
E_5	Time (F_1)	Performer (F_3)

By using a five-level linguistic term set as listed in Table 5, the best-to-others and others-to-worst vectors of five experts are obtained, as listed in

Table 16.

Table 16: Best-to-others and others-to-worst vectors of the experts

Expert	Vector	F_1	F_2	F_3
E_1	Best-to-others	7	1	3
	Others-to-worst	1	7	5
E_2	Best-to-others	1	5	3
	Others-to-worst	7	1	3
E_3	Best-to-others	5	7	1
	Others-to-worst	3	1	5
E_4	Best-to-others	3	1	7
	Others-to-worst	5	7	1
E_5	Best-to-others	1	5	7
	Others-to-worst	7	3	1

According to Eq. (15), five optimization models are established to obtain the weights of influential factors regarding each expert. For instance, the optimization model concerning E_1 is constructed as:

$$\min\ \xi^1$$
$$s.t.\ \begin{cases} \left|\dfrac{w_2^1}{w_j^1} - v_{2j}^1\right| \leq \xi^1 \\ \left|\dfrac{w_j^1}{w_1^1} - v_{j1}^1\right| \leq \xi^1 \\ \displaystyle\sum_{j=1}^{3} w_j^1 = 1 \\ 0 \leq w_j^1 \leq 1, j = 1, 2, 3 \end{cases}$$

By solving the above model, the optimal weights provided by E_1 are obtained as $w^{1*} = (0.0769, 0.6154, 0.3077)$. Thus, via Eq. (16), the overall weights of influential factors are calculated as $w = (0.3512, 0.3605, 0.2883)$.

For each influential factor of four sequential tasks, five experts provide their assessments based on the anchor points of the influential factors, as listed in Table 17.

Based on the linguistic judgments of the experts, each assessment could be equivalently transformed into corresponding cloud models to represent

Table 17: Expert judgments on the dependence level

Expert	Factor	Sequential task				
		T_1	T_2	T_3	T_4	T_5
E_1	F_1	LD	ZD	HD	MD	MD
	F_2	MD	CD	LD	ZD	LD
	F_3	ZD	LD	MD	HD	LD
E_2	F_1	ZD	HD	LD	MD	HD
	F_2	MD	LD	ZD	LD	MD
	F_3	CD	HD	MD	HD	HD
E_3	F_1	MD	ZD	HD	CD	LD
	F_2	HD	ZD	MD	LD	HD
	F_3	LD	MD	HD	MD	CD
E_4	F_1	HD	LD	MD	HD	MD
	F_2	CD	HD	CD	HD	LD
	F_3	MD	HD	CD	ZD	LD
E_5	F_1	LD	HD	MD	MD	MD
	F_2	HD	CD	LD	HD	ZD
	F_2	ZD	HD	MD	LD	ZD

the uncertainty and fuzziness of the assessment. In this case, the five basic clouds are determined as:

$$\begin{aligned} &C_0 = (0.000, 0.103, 0.052), \quad C_1 = (0.309, 0.064, 0.032), \\ &C_2 = (0.500, 0.039, 0.002), \quad C_3 = (0.691, 0.064, 0.032), \\ &C_4 = (1.000, 0.103, 0.052) \end{aligned}$$

Hence, the linguistic assessments of experts are transformed into cloud models to establish the cloud assessment matrix, as shown in Table 18.

For each influential factor, the judgments of different experts could be combined to obtain the combined cloud models by using Eq. (11), and the combined clouds of the influential factors are listed in Table 19.

Since all three influential factors are within the same level, the cloud models of these influential factors could be combined directly to obtain the overall assessment on the dependence level, and the results are listed in Table 20.

Based on the overall cloud models, the CHEP of different sequential tasks is computed, as shown in Table 20. As can be seen from Table 20, the third

Table 18: Transformed cloud models of expert judgments

Expert	Factor	Sequential task				
		T_1	T_2	T_3	T_4	T_5
E_1	F_1	(0.309,0.064,0.032)	(0.000,0.103,0.052)	(0.691,0.064,0.032)	(0.500,0.039,0.002)	(0.500,0.039,0.002)
	F_2	(0.500,0.039,0.002)	(1.000,0.103,0.052)	(0.309,0.064,0.032)	(0.000,0.103,0.052)	(0.309,0.064,0.032)
	F_3	(0.000,0.103,0.052)	(0.309,0.064,0.032)	(0.500,0.039,0.002)	(0.691,0.064,0.032)	(0.309,0.064,0.032)
E_2	F_1	(0.000,0.103,0.052)	(0.691,0.064,0.032)	(0.309,0.064,0.032)	(0.500,0.039,0.002)	(0.691,0.064,0.032)
	F_2	(0.500,0.039,0.002)	(0.309,0.064,0.032)	(0.000,0.103,0.052)	(0.309,0.064,0.032)	(0.500,0.039,0.002)
	F_3	(1.000,0.103,0.052)	(0.691,0.064,0.032)	(0.500,0.039,0.002)	(0.691,0.064,0.032)	(0.691,0.064,0.032)
E_3	F_1	(0.500,0.039,0.002)	(0.000,0.103,0.052)	(0.691,0.064,0.032)	(1.000,0.103,0.052)	(0.309.0.064,0.032)
	F_2	(0.691,0.064,0.032)	(0.000,0.103,0.052)	(0.500,0.039,0.002)	(0.309,0.064,0.032)	(0.691,0.064,0.032)
	F_3	(0.309,0.064,0.032)	(0.500,0.039,0.002)	(0.691,0.064,0.032)	(0.500,0.039,0.002)	(1.000,0.103,0.052)
E_4	F_1	(0.691,0.064,0.032)	(0.309,0.064,0.032)	(0.500,0.039,0.002)	(0.691,0.064,0.032)	(0.500,0.039,0.002)
	F_2	(1.000,0.103,0.052)	(0.691,0.064,0.032)	(1.000,0.103,0.052)	(0.691,0.064,0.032)	(0.309,0.064,0.032)
	F_3	(0.500,0.039,0.002)	(0.691,0.064,0.032)	(1.000,0.103,0.052)	(0.000,0.103,0.052)	(0.309,0.064,0.032)
E_5	F_1	(0.309,0.064,0.032)	(0.691,0.064,0.032)	(0.500,0.039,0.002)	(0.500,0.039,0.002)	(0.500,0.039,0.002)
	F_2	(0.691,0.064,0.032)	(1.000,0.103,0.052)	(0.309,0.064,0.032)	(0.691,0.064,0.032)	(0.000,0.103,0.052)
	F_2	(0.000,0.103,0.052)	(0.691,0.064,0.032)	(0.500,0.039,0.002)	(0.309,0.064,0.032)	(0.000,0.103,0.052)

Table 19: Combined cloud models of the influential factors

Sequential tasks	Influential factors		
	F_1	F_2	F_3
T_1	(0.3698,0.0680,0.0330)	(0.7014,0.0696,0.0330)	(0.3528,0.0833,0.0408)
T_2	(0.3118,0.0821,0.0413)	(0.6060,0.0887,0.0447)	(0.5747,0.0598,0.0286)
T_3	(0.5467,0.0546,0.0241)	(0.4768,0.0794,0.0390)	(0.6680,0.0663,0.0302)
T_4	(0.6508,0.0635,0.0286)	(0.4119,0.0736,0.0369)	(0.4067,0.0725,0.0354)
T_5	(0.4919,0.0496,0.0194)	(0.3631,0.0694,0.0340)	(0.4565,0.0810,0.0407)

Table 20: Overall assessment of different sequential tasks

Sequential tasks	Overall assessment	CHEP
T_1	(0.4939,0.0733,0.0354)	[0.2740,0.7138]
T_2	(0.4937,0.0790,0.0394)	[0.2567,0.7307]
T_3	(0.5565,0.0677,0.0319)	[0.3534,0.7596]
T_4	(0.4943,0.0699,0.0338)	[0.2846,0.7040]
T_5	(0.4353,0.0670,0.0320)	[0.2343,0.6363]

pair of sequential tasks has the highest CHEP, and could be regarded as the one with the highest dependence level.

5. Evaluation and validation

In order to validate the effectiveness of the proposed method, Case 2 is further analyzed and discussed in this section.

5.1. *Sensitivity analysis*

In order to better analyze the performance and behavior of the proposed method, sensitivity analysis is conducted in this section. In this section, two sets of sensitivity analyses are conducted, one with the different setting parameters, and one with different influential factor weights.

Firstly, consider the setting parameters used in the cloud model, a sensitivity analysis with 9 different sets of setting parameters is conducted, as listed in Table 21. The results of different scenarios are shown in Table 22.

Table 21: Parameter settings of cloud model for sensitivity analysis

Scenario	$x_{\min}$	$x_{\max}$	He_2
Scenario 1	0	1	0.02
Scenario 2	0.05	0.95	0.02
Scenario 3	0.1	0.9	0.02
Scenario 4	0	1	0.01
Scenario 5	0.05	0.95	0.01
Scenario 6	0.1	0.9	0.01
Scenario 7	0	1	0.03
Scenario 8	0.05	0.95	0.03
Scenario 9	0.1	0.9	0.03

Table 22: Assessment results with different parameter settings

Scenario	T_1		T_2		T_3		T_4		T_5	
	Overall assessment	CHEP	Overall assessment	CHEP	Overall assessment	CHEP	Overall assessment	CHEP	Overall assessment	CHEP
Scenario 1	(0.494,0.073,0.035)	[0.274,0.714]	(0.494,0.079,0.039)	[0.257,0.731]	(0.557,0.068,0.032)	[0.353,0.760]	(0.494,0.070,0.034)	[0.285,0.704]	(0.435,0.067,0.032)	[0.235,0.636]
Scenario 2	(0.495,0.073,0.037)	[0.276,0.715]	(0.498,0.079,0.040)	[0.261,0.734]	(0.551,0.068,0.034)	[0.348,0.754]	(0.498,0.070,0.035)	[0.289,0.707]	(0.438,0.067,0.034)	[0.237,0.639]
Scenario 3	(0.497,0.073,0.037)	[0.277,0.716]	(0.502,0.079,0.040)	[0.265,0.738]	(0.545,0.068,0.034)	[0.342,0.749]	(0.502,0.070,0.035)	[0.293,0.711]	(0.441,0.067,0.034)	[0.240,0.641]
Scenario 4	(0.494,0.073,0.019)	[0.274,0.714]	(0.494,0.079,0.020)	[0.257,0.730]	(0.556,0.068,0.017)	[0.353,0.760]	(0.494,0.070,0.018)	[0.285,0.704]	(0.435,0.067,0.017)	[0.235,0.636]
Scenario 5	(0.495,0.073,0.019)	[0.276,0.715]	(0.498,0.079,0.020)	[0.261,0.734]	(0.551,0.068,0.017)	[0.348,0.754]	(0.498,0.070,0.018)	[0.289,0.707]	(0.438,0.067,0.017)	[0.237,0.639]
Scenario 6	(0.497,0.073,0.019)	[0.277,0.716]	(0.502,0.079,0.020)	[0.265,0.738]	(0.545,0.068,0.017)	[0.342,0.749]	(0.502,0.070,0.018)	[0.293,0.711]	(0.441,0.067,0.017)	[0.240,0.641]
Scenario 7	(0.494,0.073,0.056)	[0.274,0.714]	(0.494,0.079,0.060)	[0.257,0.730]	(0.556,0.068,0.052)	[0.353,0.760]	(0.494,0.070,0.053)	[0.285,0.704]	(0.435,0.067,0.051)	[0.235,0.636]
Scenario 8	(0.495,0.073,0.056)	[0.276,0.715]	(0.498,0.079,0.060)	[0.261,0.734]	(0.551,0.068,0.052)	[0.348,0.754]	(0.498,0.070,0.053)	[0.289,0.707]	(0.438,0.067,0.051)	[0.237,0.639]
Scenario 9	(0.497,0.073,0.056)	[0.277,0.716]	(0.502,0.079,0.060)	[0.265,0.738]	(0.545,0.068,0.052)	[0.342,0.749]	(0.502,0.070,0.053)	[0.293,0.711]	(0.441,0.067,0.051)	[0.240,0.641]

From the results in Table 22, it can be found that when the parameter settings changes, the assessment results change accordingly However, it is worth noting that the ranking of the sequential tasks generally remains stable, as T_3 always has the highest CHEP and T_5 has the lowest. Hence, the sensitivity results show that the proposed method could provide reliable and consistent dependence assessment results. Moreover, it is worth noting that when the value of He_2 changes, the CHEP remains unchanged, and that is because He_2 mainly affects the dispersion degree of the cloud droplets, which is not considered when calculating CHEP.

Secondly, in the proposed method, several influential factors are considered for dependence assessment, and the weights assigned to these factors are determined using the BWM. Since the weights of the influential factors can significantly influence the evaluation results, a sensitivity analysis is performed using 10 sets of randomly generated factor weights to assess their effects.

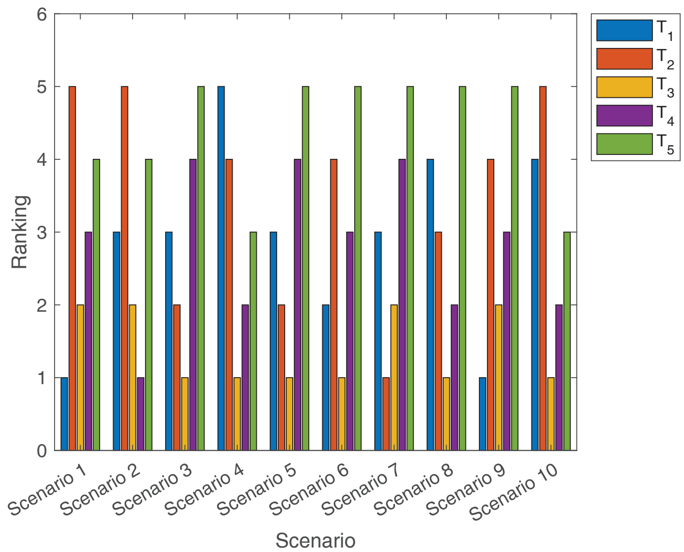

Figure 12: Dependence assessment results of different influential factor weights.

Fig. 12 shows the rankings of the sequential tasks with different factors weights based on the CHEP value. From the results, it can be observed that altering the factor weight could affect the dependence assessment results and the ranking of the sequential tasks. Though T_3 is assessed to be the top 2 tasks, the results of other tasks vary, highlighting the influence and importance of the factor weights on the dependence assessment results. Therefore, the sensitivity analysis result further emphasizes the significance of carefully considering and determining the weights of the influential factors using the BWM in this study, as they can have a notable impact on overall dependence assessment results.

5.2. Comparison analysis

In order to validate the effectiveness and feasibility of the proposed method, a comparison analysis with several existing dependence assessment methods is conducted. The results of the proposed method are compared with the results of the evidence theory method [32], the probabilistic linguistic term method [39], the interval evidential reasoning method [35], and the improved evidential method [47]. It should be noted that for these methods, the judgments of the experts are directly aggregated, and the human error probability of the five dependence levels are listed in Table 23. Based on the results of these methods, the ranking of the five pairs of sequential tasks is obtained, as shown in Fig 13.

Table 23: Human error probability of the dependence levels for comparative methods

Dependence level	$p(B\|A)$
ZD	0
LD	0.0595
MD	0.1514
HD	0.5050
CD	1

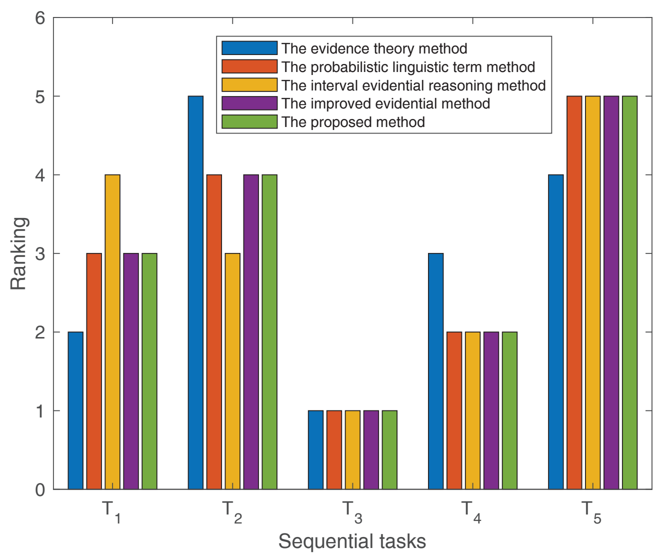

Figure 13: Dependence assessment results of different methods.

From Fig 13, it can be found that T_3 has the highest dependence level among all five pairs of sequential tasks according to all five dependence assessment methods. It is also worth noting that by using the proposed method, the probabilistic linguistic term method, the interval evidential method and the improved evidential method, T_3 and T_4 are assessed to be the top two highly dependent sequential tasks. Moreover, the sequential tasks with the lowest dependence level are the same for the probabilistic linguistic term method, the interval evidential reasoning method, the improved evidential method and the proposed method, i.e., T_5, which is in line with the experts' analysis that the HFE of "reactions occur in the transfusion" is not highly rely on the failure of its preceding task "patients are not properly monitored". From these results, the effectiveness and feasibility of the proposed method are further proved.

On the other hand, it is worth noting that there are some differences between the results obtained using the proposed method and those determined through the evidence theory method and the interval evidential reasoning method. The main reasons for that can be summarized in the following: (1) The compared method adopted evidence theory and interval evidential reasoning to represent the assessments of experts. However, these two methods have some limitations in handling the uncertainty in experts' assessments. (2) These methods used AHP to determine the weight of influential factors. However, the AHP has limitations in dealing with cases with many influential factors as it would require a significant amount of pairwise comparisons. (3) These methods ignored the complex trust relation among different experts, which could lead to inaccuracies when determining the weights of experts.

5.3. Discussion

The results of the proposed method can be presented in two parts. Firstly, it is the final interval CHEP value and the corresponding representative value, which could be used for quantitative analysis. Secondly, the final results are also in the form of cloud model, as shown in Fig 10, which allows a more intuitive and representative understanding of the assessment result. The uncertainty in the experts judgments lies in the interval value and cloud model, which enables the proposed method to provide reliable results under uncertainty. Since the credibility of different experts is taken into consideration during the assessment, the proposed method focuses more on the assessments that have more agreements, thus reducing the uncertainty of the final result to some extent.

The result can be partially explained by examining the assessments and weights of the influential factors. The weights of the influential factors suggest that the factor "task relatedness" is more important than other factors, in other words, the assessment of "task relatedness" would have more influence on the final result. As shown above, the final result relies more on "task relatedness" than the other two factors, in line with intuition. In a word, the proposed method provides a reliable and intuitive way for human dependence assessment in HRA.

6. Conclusion

Dependence assessment is one of the most important issues in human reliability, and how to effectively and accurately assess dependence level of human errors under uncertainty remains challenging. To this end, this paper proposes a new dependence assessment approach integrating cloud model and best-worst method, where the cloud model is used to model the uncertainty of linguistic judgments of experts and the best-worst method is employed to determine the weights of different influential factors. First, the linguistic judgments of the experts on the dependence level of the influential factors are transferred into cloud models. Then, the judgments of different experts are combined, where the weight of the experts' judgment is determined using a subjective weighting method. Next, the assessments of different influential factors are combined based on the weights determined by using the best-worst method. Finally, the cloud for the final assessment is calculated, and the conditional human error probability is calculated. By using the cloud model to represent and aggregate experts' judgments under uncertainty, the proposed method could enhance the reliability and accuracy of the dependence assessment results. Moreover, the adoption of the best-worst method ensures that the influential factors are accurately weighed, further enhancing the reliability of the results. In order to demonstrate the effectiveness of the proposed method two cases are examined, and the results show that the proposed method could provide reliable and intuitive dependence assessment results under uncertainty, which is further validated through sensitivity analysis and comparison analysis. It can be concluded that the proposed method provides a novel and effective mechanism for dependence assessment in human reliability analysis under uncertainty, which could be used in the future.

Future research could be carried out in the following directions. First, the influential factors used in this study are relatively simple, and more factors could be considered for other problems. Second, the proposed method is restricted to a small group of experts, the application of the proposed method to large-group problems may be needed. Third, some dependence assessment problems may have dynamic features, hence, the proposed method could be extended to dynamic situations. Fourth, the proposed method mainly focuses on the assessment of human dependence levels, future studies could integrate machine learning models such as Bayesian network to obtain human error probability to provide more reliable and reasonable results. Moreover, it is worth noting that in this study, the influential factors and their functional relationships are determined based on previous studies intuitively, and systemic approaches such as grounded theory could be applied to enhance its reliability.

Acknowledgments

This work was supported by the Shandong Provincial Natural Science Foundation under Grant No. ZR2023QF148.

References

[1] X. Y. Zhou, X. Y. Deng, Y. Deng, S. Mahadevan, Dependence assessment in human reliability analysis based on d numbers and ahp, Nuclear Engineering and Design 313 (2017) 243–252. doi:10.1016/j.nucengdes.2016.12.001.

[2] K. M. Groth, A. Mosleh, A data-informed pif hierarchy for model-based human reliability analysis, Reliability Engineering & System Safety 108 (2012) 154–174. doi:10.1016/j.ress.2012.08.006.

[3] S. I. Sezer, E. Akyuz, O. Arslan, An extended heart dempster–shafer evidence theory approach to assess human reliability for the gas freeing process on chemical tankers, Reliability Engineering & System Safety 220 (2022) 108275.

[4] D. Wang, Y. Wei, J. Zhan, L. Xu, Q. Lin, Human reliability assessment of home-based rehabilitation, IEEE Transactions on Reliability 70 (2020) 1310–1320.

[5] M. Catelani, L. Ciani, G. Guidi, G. Patrizi, An enhanced sherpa (e-sherpa) method for human reliability analysis in railway engineering, Reliability Engineering & System Safety 215 (2021) 107866.

[6] H.-C. Liu, J.-H. Wang, L. Zhang, Q.-Z. Zhang, New success likelihood index model for large group human reliability analysis considering non-cooperative behaviors and social network, Reliability Engineering & System Safety 228 (2022) 108817.

[7] D. T. Abreu, M. C. Maturana, E. L. Droguett, M. R. Martins, Human reliability analysis of conventional maritime pilotage operations supported by a prospective model, Reliability Engineering & System Safety 228 (2022) 108763.

[8] S. I. Ahn, R. E. Kurt, Application of a cream based framework to assess human reliability in emergency response to engine room fires on ships, Ocean Engineering 216 (2020) 108078.

[9] Q.-L. Lin, D.-J. Wang, W.-G. Lin, H.-C. Liu, Human reliability assessment for medical devices based on failure mode and effects analysis and fuzzy linguistic theory, Safety science 62 (2014) 248–256.

[10] X. Deng, W. Jiang, Dependence assessment in human reliability analysis using an evidential network approach extended by belief rules and uncertainty measures, Annals of Nuclear Energy 117 (2018) 183–193.

[11] N. Ade, S. C. Peres, A review of human reliability assessment methods for proposed application in quantitative risk analysis of offshore industries, International Journal of Industrial Ergonomics 87 (2022) 103238.

[12] X. Gao, X. Su, H. Qian, X. Pan, Dependence assessment in human reliability analysis under uncertain and dynamic situations, Nuclear Engineering and Technology (2021).

[13] L. Wang, Y. Wang, Y. Chen, X. Pan, W. Zhang, Performance shaping factors dependence assessment through moderating and mediating effect analysis, Reliability Engineering & System Safety 202 (2020) 107034.

[14] H.-C. Liu, Z. Li, J.-Q. Zhang, X.-Y. You, A large group decision making approach for dependence assessment in human reliability analysis, Reliability Engineering & System Safety 176 (2018) 135–144.

[15] A. Zhang, F. Gao, M. Yang, W. Bi, Belief rule-based dependence assessment method under interval uncertainty, Quality and Reliability Engineering International 36 (2020) 2459–2477.

[16] L. Zhang, Y.-J. Zhu, L.-X. Hou, H.-C. Liu, New method for dependence assessment in human reliability analysis based on linguistic hesitant fuzzy information, Nuclear Engineering and Technology 53 (2021) 3675–3684.

[17] L.-X. Hou, R. Liu, H.-C. Liu, S. Jiang, Two decades on human reliability analysis: a bibliometric analysis and literature review, Annals of Nuclear Energy 151 (2021) 107969.

[18] L. Podofillini, V. Dang, E. Zio, P. Baraldi, M. Librizzi, Using expert models in human reliability analysis—a dependence assessment method based on fuzzy logic, Risk Analysis: An International Journal 30 (2010) 1277–1297.

[19] E. Hollnagel, Cognitive reliability and error analysis method (CREAM), Elsevier, 1998.

[20] B. Wu, X. Yan, Y. Wang, C. G. Soares, An evidential reasoning-based cream to human reliability analysis in maritime accident process, Risk Analysis 37 (2017) 1936–1957.

[21] Q. Zhou, Y. D. Wong, H. S. Loh, K. F. Yuen, A fuzzy and bayesian network cream model for human reliability analysis – the case of tanker shipping, Safety Science 105 (2018) 149–157. doi:10.1016/j.ssci.2018.02.011.

[22] B. Gore, J. Dukelow Jr, T. Mitts, W. Nicholson, A limited assessment of the ASEP Human reliability analysis procedure using simulator examination results, Technical Report, Nuclear Regulatory Commission, 1995.

[23] W. D. Jung, W. C. Yoon, J. Kim, Structured information analysis for human reliability analysis of emergency tasks in nuclear power plants, Reliability Engineering & System Safety 71 (2001) 21–32.

[24] J. Park, A. M. Arigi, J. Kim, Treatment of human and organizational factors for multi-unit hra: Application of spar-h method, Annals of Nuclear Energy 132 (2019) 656–678.

[25] K. M. Groth, L. P. Swiler, Bridging the gap between hra research and hra practice: A bayesian network version of spar-h, Reliability Engineering & System Safety 115 (2013) 33–42.

[26] P. Liu, Z. Z. Li, Human error data collection and comparison with predictions by spar-h, Risk Analysis 34 (2014) 1706–1719.

[27] J. Forester, D. Bley, S. Cooper, E. Lois, N. Siu, A. Kolaczkowski, J. Wreathall, Expert elicitation approach for performing atheana quantification, Reliability Engineering & System Safety 83 (2004) 207–220.

[28] J. Pinto, P. F. E. Melo, P. Saldanha, A dfm/fuzzy/atheana human failure analysis of a digital control system for a pressurizer, Nuclear Technology 188 (2014) 20–33.

[29] E. Zio, P. Baraldi, M. Librizzi, L. Podofillini, V. N. Dang, A fuzzy set-based approach for modeling dependence among human errors, Fuzzy Sets and Systems 160 (2009) 1947–1964. doi:10.1016/j.fss.2009.01.016.

[30] H. Liao, K. Groth, S. Stevens-Adams, Challenges in leveraging existing human performance data for quantifying the idheas hra method, Reliability Engineering & System Safety 144 (2015) 159–169.

[31] D.-H. Ham, J. Park, Use of a big data analysis technique for extracting hra data from event investigation reports based on the safety-ii concept, Reliability Engineering & System Safety 194 (2020) 106232.

[32] X. Su, S. Mahadevan, P. Xu, Y. Deng, Dependence assessment in human reliability analysis using evidence theory and ahp, Risk Analysis 35 (2015) 1296–316.

[33] M. Musharraf, J. Hassan, F. Khan, B. Veitch, S. MacKinnon, S. Imtiaz, Human reliability assessment during offshore emergency conditions, Safety Science 59 (2013) 19–27.

[34] P. Baraldi, L. Podofillini, L. Mkrtchyan, E. Zio, V. N. Dang, Comparing the treatment of uncertainty in bayesian networks and fuzzy expert systems used for a human reliability analysis application, Reliability Engineering & System Safety 138 (2015) 176–193.

[35] W. Bi, F. Gao, A. Zhang, M. Yang, Dependence assessment in human reliability analysis based on the interval evidential reasoning algorithm under interval uncertainty, IEEE Access 8 (2020) 222187–222198. doi:10.1109/access.2020.3043848.

[36] V. P. Paglioni, K. M. Groth, Dependency definitions for quantitative human reliability analysis, Reliability Engineering & System Safety 220 (2022) 108274.

[37] J. Liu, Y. Zou, W. Wang, L. Zhang, X. Liu, Q. Ding, Z. Qin, M. Čepin, Analysis of dependencies among performance shaping factors in human reliability analysis based on a system dynamics approach, Reliability Engineering & System Safety 215 (2021) 107890.

[38] F. Gao, W. Liu, X. Mu, W. Bi, A. Zhang, Dependence assessment in human reliability analysis using the 2-tuple linguistic information and dematel method, Process Safety and Environmental Protection 173 (2023) 191–201.

[39] F. Gao, S. Bao, W. Bi, Assessing dependence in human reliability analysis using probabilistic linguistic term sets, Annals of Nuclear Energy 175 (2022) 109261.

[40] D. Li, C. Liu, W. Gan, A new cognitive model: Cloud model, International journal of intelligent systems 24 (2009) 357–375.

[41] H. C. Liu, L. E. Wang, Z. W. Li, Y. P. Hu, Improving risk evaluation in fmea with cloud model and hierarchical topsis method, IEEE Transactions on Fuzzy Systems 27 (2019) 84–95.

[42] J. Li, K. Xu, A combined fuzzy dematel and cloud model approach for risk assessment in process industries to improve system reliability, Quality and Reliability Engineering International 37 (2021) 2110–2133.

[43] C.-B. Li, Z.-Q. Qi, X. Feng, A multi-risks group evaluation method for the informatization project under linguistic environment, Journal of Intelligent & Fuzzy Systems 26 (2014) 1581–1592.

[44] A. D. Swain, H. E. Guttmann, Handbook of human-reliability analysis with emphasis on nuclear power plant applications. Final report, Technical Report, Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 1983.

[45] D. Gertman, H. Blackman, J. Marble, J. Byers, C. Smith, et al., The spar-h human reliability analysis method, US Nuclear Regulatory Commission 230 (2005) 35.

[46] J. Rezaei, Best-worst multi-criteria decision-making method, Omega 53 (2015) 49–57.

[47] L. Chen, X. Zhou, F. Xiao, Y. Deng, S. Mahadevan, Evidential analytic hierarchy process dependence assessment methodology in human reliability analysis, Nuclear Engineering and Technology 49 (2017) 123–133.