

Received 6 April 2023, accepted 11 May 2023, date of publication 17 May 2023, date of current version 23 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3277254

RESEARCH ARTICLE

Mental Overload Assessment Method Considering the Effects of Performance Shaping Factors

XIN LU¹, JIANBIN GUO¹, SHENGKUI ZENG¹, AND HAIYANG CHE^{2,3}

¹School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China

²School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China

³Science and Technology on Reliability and Environmental Engineering Laboratory, Beihang University, Beijing 100191, China

Corresponding author: Haiyang Che (chehaiyang@buaa.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 72201018, and in part by the Funding Project of Science and Technology on Reliability and Environmental Engineering Laboratory under Grant 6142004210104.

ABSTRACT The assessment of mental overload is essential as mental overload has been proved to be a critical cause of many accidents. Mental overload is affected by various performance shaping factors (PSFs), and harsh PSFs will either increase the task demand or decrease the overload threshold, which is rarely considered in the current mental overload assessment method. This research proposes a VACP-based mental overload assessment model which considers the effects of PSFs. In the VACP model, mental overload can be identified when the sum of the task demand values of every unit task is greater than the given threshold. In human reliability analysis, PSFs are mainly used as weighting factors to modify basic human error probability. In our research, by virtue of the quantitative relationship between human error and mental workload, the weighting factors of PSFs are converted to modify the task demand values and threshold for VACP activities. Furthermore, Bayesian Network (BN) is used to model the influence of PSFs and to calculate the probability of mental overload. The proposed method is applied to an accident involving a helicopter crash that occurred in Maryland, and the results show that, in comparison with the VACP method, the proposed method can more effectively identify the state of mental overload and provide a more rational explanation of the process of the accident.

INDEX TERMS Mental overload, visual-auditory-cognitive-psychomotor (VACP), performance shaping factors, Bayesian network.

I. INTRODUCTION

The safe and efficient performance of given tasks in a human-machine system requires that the mental workload imposed on operators does not exceed their capacity. An excessive mental workload will overwhelm an operator's information-processing, degrade an operator's vigilance, and eventually lead to "cognitive tunneling", which describes the phenomenon in which an operator is unable to reallocate his/her attention from one task to another [1], [2]. The National Safety Council (NSC) reports 21% of all fatal accidents are attributed to excessive mental workload [3]. It is widely accepted that excessive mental workload, i.e., mental

overload, is an important safety-related factor; furthermore, the consensus is that mental overload assessments should be taken into consideration in the design stage [4].

In essence, two elements are incorporated in assessments of mental overload: the mental workload caused by task demands and the operator's mental workload capacity. The mental workload caused by task demands (also known as "mental workload level") represents "how hard the brain is working to meet task demands" [5]. Meanwhile, the operator's mental workload capacity (also known as "the operator's capacity") represents "the operator's limited mental resource in processing task demand" [6]. Mental overload is considered to have occurred when the mental workload level reaches beyond the level of operator's capacity [7]. Researchers tend to use Cognitive Load Theory (CLT) to

The associate editor coordinating the review of this manuscript and approving it for publication was Catherine Fang.

theoretically explain mental overload [8]. According to CLT, an operator's mental resources for handling task demands are heavily constrained because new information should be stored first in the working memory, and the working memory is limited both in terms of storage length and duration. In this sense, task demands can reflect the mental workload level, and the working memory can reflect the operator's capacity.

There are many researches about mental overload assessment, and those researches can be categorized as experimental methods and simulation methods.

For experimental methods, by virtue of the high-fidelity training device and virtual reality technology, mental overload can be assessed through man-in-loop experiment. Researchers tend to use physiological parameters and subjective assessment to reflect mental workload [9], [10], [11]. However, different experiments can only assess mental workload for a certain scenario, which means that with the change of application scenario, the experimental environment must be rebuilt to adapt it. In addition, the experimental method is both time-consuming and costly, therefore, the results obtained by such studies are not suitable for simple and quick analysis [12].

For simulation methods, they can be categorized as time-based and task-based. Time-based simulation methods, such as cognitive architectures like Adaptive Control of Thought-Rational (ACT-R), Queuing Network-Model Human Processor (QN-MHP) and Executive Process/Interactive Control (EPIC) [13], mainly use the ratio of total time spent for production firing and chunk exchanging in cognitive architecture to the available time as a natural index of mental workload [14]. Task-based methods, such as W/INDEX, Task Analysis/Workload (TAWL) model and Improved Performance Research Integration Tool (IMPRINT) [15], mainly use a rating scale to determine mental workload values based on different task type. Usually, the rating scale of these methods is based on Visual-Auditory-Cognitive-Psychomotor (VACP) method [16]. Compared with time-based simulation methods, task-based simulation methods are simple to implement and with higher versatility. Actually, VACP has shown its promising application in mental overload assessment [17], [18]. For example, Wang et al. used the VACP to assess the mental overload state of train driver so as to optimize task flow, and the experimental results proved that the optimized task flow balances the relationship between mental workload and reaction time well [19]. The VACP can provide a continuous prediction of mental workload level (as reflected by "task demand" in VACP), as well as a fixed value for the operator's capacity (as reflected by "threshold" in VACP), which is sufficient for performing real-time mental overload assessments. The VACP method has been applied to the study of many areas, including manufacturing [20], the operation of unmanned aviation vehicles [21], driving [22], and so on.

However, as with human error, mental overload is related to the contextual factors attending the situation in which

the performance occurs [23], which is rarely considered in application of the aforementioned methods. In fact, several researchers have studied the influence of contextual factors on mental workload level and operator's capacity. In regard to mental workload level, Tsao et al. proved that an improper work/rest rhythm and adverse physical working environment could significantly increase the mental workload level [24]. Kaptan et al. pointed out that the mental workload level, whatever it be in a normal operation process or in an abnormal operation process, decreases as the automation level increases [25]. With respect to the operator's capacity, Young et al. noted that skill is a factor that can influence an operator's capacity [26] and Kim et al. used an experimental method to show that skilled operators would have a higher capacity [27]. In actuality, all of these factors can be considered to be performance shaping factors (PSFs). Therefore, the assessment of mental overload should consider the influence of PSFs.

PSFs is often used in Human Reliability Analysis (HRA). HRA is proposed to systematically incorporate human as a part of a Probabilistic Risk Assessment (PRA) activity [28], and PSFs are used to characterize the context of human tasks, which is assumed to enhance or degrade human performance [29]. The term "PSFs" varies from different HRA methods, such as PSFs, error-producing conditions (EPCs), common performance conditions (CPCs), and performance influencing factors (PIFs). Despite their different forms, the concepts are same. For the sake of consistence, they are all called as PSFs. One of the most common applications of PSFs in HRA is to adjust Basic Human Error Probability (BHEP) [30]. With the development of HRA method, the PSFs also experience the evolution. In the first-generation HRA methods, cognition is not particularly considered among the PSFs, which, to a certain degree, cannot explain how PSFs exert influence on performance. While in the second-generation, PSFs were derived by focusing on the cognitive impacts on operators. Nowadays, the application of PSFs has beyond the scope of HRA. There are several other research topics addressing PSFs, including human behavior models, human-related event analysis method and human performance database [31].

In our methodology, PSFs were used to modify the task demands and the threshold of VACP, so choosing the appropriate PSF set is necessary. Di Mascio et al. classified the PSFs that affect mental overload into three categories: the work environment in which the controller operates, some variable personal factors and physical preconditions [32]. Azadeh et al. further extended the categories, describing them as health-related PSFs, safety-related PSFs, environment-related PSFs and ergonomics-related PSFs [33]. In our research, Azadeh's classification was applied because it covers more aspects of the relevant contextual factors. However, in design stage, a designer must consider the diversity of potential operators, and thus, individual differences are temporarily set aside. As a result, all of the PSF categories should be taken into consideration, apart from health-related PSFs.

The introduction of PSFs poses a problem: the relationship between PSFs and mental overload should be established. The existing investigation have mainly employed experiments to qualitatively test the effect of a particular PSF on mental overload. For instance, Bhavsar et al. proposed a qualitative methodology based on pupillometry (the measurement of pupil diameter) to noninvasively estimate the mental overload of control room operators in real time during process operations [34]. Casner tested the advanced cockpit system's effect on mental overload via an in-flight experiment, and the results indicate that the use of the advanced cockpit helps to relieve mental overload in some situations [35]. However, in order to quantitatively assess mental overload, solely considering qualitative relationship is not enough. Thus, establishing quantitative relationship between PSFs and mental overload is necessary.

Another problem introduced by including PSFs in the model is that differing levels of probability distributions exist for the various PSFs. To address this issue, employing a Bayesian Network (BN) seems to be a viable approach. BN has long been used to handle the probability distributions of PSFs in HRA [36], [37], [38]. Furthermore, from the perspective of modeling, BN also has many obvious advantages such as stronger theoretical roots and lower levels of computational complexity, as compared with other tools [39].

The present study, thus, aims to propose a method for assessing mental overload with consideration given to PSFs. In our research, by virtue of the quantitative relationship between human error and mental workload, the PSFs' weighting factors have been converted to modify the task demand value and threshold of VACP. Furthermore, BN is applied so as to model the influence of each PSF and to calculate the associated probability of mental overload.

The structure of this paper is as follows: The next section provides information on the basic concepts related to the proposed method. Section III describes the framework of the proposed mental overload model and steps taken in the construction of the mental overload assessment model. In Section IV, the proposed method is demonstrated through its application to an actual helicopter crash that occurred in Maryland. Section V presents our conclusions.

II. PRELIMINARIES

A. VACP METHODS

The VACP method presents mental workload as relying on four resource channels: visual, auditory, cognitive and psychomotor. The amount of demand required of each channel to perform tasks is estimated on McCracken and Aldrich's seven-point scale, as shown in TABLE 1. Mental workload is defined as the sum of every channel's task demand value, and a score of 40 is regarded as the threshold [40]. When the calculated mental workload value is greater than 40, mental overload is assumed to occur. To perform tasks safely and effectively in a dynamic complex environment, the mental

workload required of the operators should not exceed that threshold.

The VACP method has the following advantages: (1) VACP provides a more objective assessment of mental workload and can eliminate rater bias; (2) VACP uses verbal anchors to provide an objective and consistent use of scale; (3) VACP depicts mental workload as relying on resources from multiple channels; (4) VACP can provide a unique workload value for an interval time once the necessary information has been determined [21].

B. PERFORMANCE SHAPING FACTORS (PSFS)

The PSFs that influence mental overload are attributed to three categories: safety-related PSFs, environment-related PSFs and ergonomics-related PSFs. The details of the influential factors of these categories are summarized in TABLE 2.

There are many suitable PSF sets that can cover the elements mentioned in TABLE 2, such as nine PSFs included in the Cognitive Reliability and Error Analysis Method (CREAM) [41], those being "Adequacy of organization", "Working conditions", "Adequacy of MMI and operational support", "Availability of procedures/plans", "Number of simultaneous goals", "Available time", "Time of day", "Adequacy of training and preparation" and "Crew collaboration quality". CREAM's set of PSFs is widely accepted and used in HRA [42], and we adopt it here to describe the factors that influence mental overload.

Additionally, correlations exist between PSFs, and the state of a specific PSF will be influenced by the correlated PSFs. For example, researchers have found that the "Adequacy of MMI and operation support" will significantly affect task-related PSFs such as "Number of simultaneous goals" [43]. This should also be considered in the modeling process.

C. BAYESIAN NETWORK (BN)

BN, also known as Bayesian Belief Network, is a directed acyclic graph that consists of two parts: the net structure and the conditional probability table (CPT) [44]. Net structure includes nodes and causalities. Nodes stand for random variables of interest, and causalities stand for the likely causal relationships between nodes. Here if we use U to represent the collection of all nodes, then $U = \{A_1, A_2, \dots, A_n\}$. If a directed arc exists between two nodes A_i and A_j , that means A_j is the parent node of A_i , written as $Pa(A_i)$. Correspondingly, A_i is the child node of A_j . By assigning a CPT, according to the chain rule, the joint probability can be given as (1):

$$P(U) = \prod_{i=1}^n P(A_i | Pa(A_i)) \quad (1)$$

where $P(U)$ is the joint probability of the whole BN.

It should be pointed out that once the evidence e of node A emerges (some nodes are in a particular state), the BN

TABLE 1. VACP scales [21].

Value	Activities
Visual	
0	No Visual Activity
1	Visual Register/Detect (detect occurrence of image)
3	Visual Inspect/Check (discrete inspection/static condition)
4	Visual Locate/Align (selective orientation)
4.4	Visual Track/Follow (maintain orientation)
5	Visual Discriminate (detect visual difference)
5.1	Visual Read (symbol)
6	Visual Scan/Search/Monitor (continuous/series inspection, multiple conditions)
Auditory	
0	No Auditory Activity
1	Detect/Register Sound (detect occurrence of sound)
2	Orient to Sound (general orientation/attention), Simple Speech (1-2 words)
3	Interpret Semantic Content (speech, simple, 1-2 words)
4	Complex Speech (Sentence)
4.2	Orient to Sound (selective orientation/attention)
4.3	Verify Auditory Feedback (detect occurrence of anticipated sound)
6	Interpret Semantic Content (speech, complex, sentence)
6.6	Discriminate Sound Characteristics (detect auditory differences)
7	Interpret Sound Patterns (pulse rates, etc.)
Cognitive	
0	No Cognitive Activity
1	Automatic (simple association)
1.2	Alternative Selection
4.6	Evaluation/Judgment (consider single aspect)
5	Sign/Signal Recognition
5.3	Encoding/Decoding, Recall
6.8	Evaluation/Judgment (consider several aspects)
7	Estimation, Calculation, Conversion
Psychomotor	
0	No Fine Motor Activity
2.2	Discrete Actuation (button, toggle, trigger)
2.6	Continuous Adjustive (flight controls, sensor control)
4.6	Manipulative (tracking)
5.5	Discrete Adjustment (rotary, vertical thumbwheel, lever position)
6.5	Symbolic Production (writing)
7	Serial Discrete Manipulation (keyboard entries)

TABLE 2. PSFs categories and related influential factors.

PSFs categories	Related influential factors
Safety	The cooperation quality between team members, the availability of plans in operating and emergency procedures, the number of concurrent tasks, the available time to perform a task and so on
Environment	Physical environment like physical working conditions, natural environment like the time of day, organization condition like the adequacy of organization
Ergonomics	The support of control panels, computerized workstations, and so on

can update the probability; this evaluation process is often referred to as “BN updating” [45]. The whole updating process is based on (2):

$$P(A|e) = \frac{P(e|A)P(A)}{P(e)} = \frac{P(e|A)P(A)}{\sum_{i=1}^n P(e|A_i)P(A_i)} \quad (2)$$

where $P(A)$ is the prior probability of node A , $P(e)$ is the probability of evidence and can be replaced with

$\sum_{i=1}^n P(e|A_i)P(A_i)$, $P(e|A)$ is the likelihood, and $P(A|e)$ is posterior probability.

III. MENTAL OVERLOAD ASSESSMENT AND PROCESS

A. MENTAL OVERLOAD ASSESSMENT MODEL

The development of a novel BN-based mental overload assessment model with consideration given to PSFs is

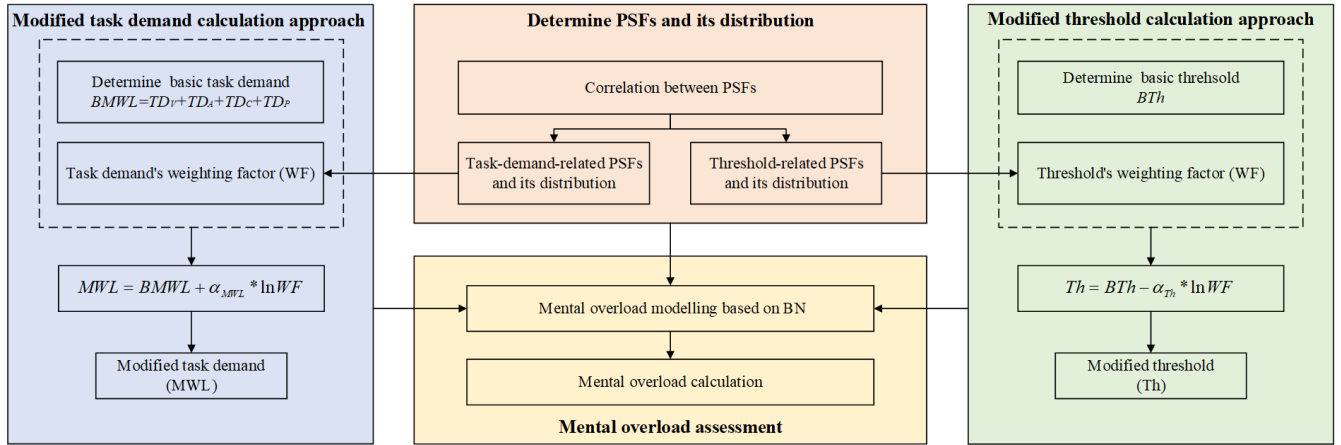


FIGURE 1. Framework of the proposed model.

presented in this section. FIGURE 1 shows the framework of the proposed model. This research introduces PSFs as weighting factors to modify the task demand and threshold of VACP. The introduction of PSFs will contribute to two improvements: First, the influences of various PSFs on task demand and threshold are established. By virtue of the quantitative relationship between human error probability (HEP) and mental workload level (MWL), the weighting factors of the PSFs are converted to modify the task demand value and threshold of VACP. Second, probability distributions exist within PSF levels; therefore, BN is used to model the influence of PSFs and to calculate the probability of mental overload. A detailed description of this process is as follows.

1) DETERMINING THE TASK-DEMAND-RELATED AND THRESHOLD-RELATED PSFs AND THEIR DISTRIBUTION

Generally speaking, PSFs affect mental overload in two ways: modification of the task demand and modification of the threshold. For example, considering the PSF “Adequacy of MMI and operational support”, if the interface layout is reasonable, then the information displayed in the interface will be easily identified, and the task demand that the task generates will be lower. In addition, when considering the PSF “Adequacy of training and preparation”, operators with good training and rich experience have higher mental overload thresholds, i.e., they can perform more tasks at the same time. This illustrates that the influence of PSFs can be categorized into two groups: those factors influencing the threshold and those factors influencing the task demand.

How should we further categorize PSFs? Generally, PSFs such as “Adequacy of MMI and operational support”, which influence the task demand, are directly related to ongoing tasks, and PSFs such as “Adequacy of training and preparation”, which influence threshold, are more related to factors such as rhythm, working conditions, and so on, which indirectly influence ongoing tasks. Based on this rule, CREAM’s PSFs can be categorized into two classes, as shown in TABLE 3.

As mentioned earlier, distributions exist between PSFs. So, the method by which to determine the distribution is another important issue. We will introduce the principles governing the distribution of PSFs in detail in the BN modelling section.

2) THE PREDICTION APPROACH TO MENTAL WORKLOAD LEVEL WITH THE INFLUENCE OF PSFs TAKEN INTO CONSIDERATION

Existing studies show that, for a pilot’s aviation activities, if HEP is less than 0.5, the relationship between human error and mental workload is very close to a simple exponential function [46], [47]. Therefore, the relationship between HEP and MWL can be presented as (3):

$$HEP = BHEP * \exp((MWL - BMWL)/\alpha_{MWL}) \quad (3)$$

where BHEP is basic human error probability, BMWL is basic mental workload, and α_{MWL} is the coefficient for MWL.

In HRA, a PSF is used to provide weighting factor (denoted as WF) to modify BHEP, so HEP yields the following expression:

$$HEP = BHEP * WF \quad (4)$$

Based on (3) and (4), the influence of PSFs on mental workload is derived as (5):

$$MWL = BMWL + \alpha_{MWL} * \ln(WF) \quad (5)$$

BMWL can be regarded as the task demand value of different channels estimated by VACP, and MWL can be regarded as the modified task demand. If TD_V , TD_A , TD_C , and TD_P represent the task demand value of the visual, auditory, cognitive and psychomotor channels in VACP, TD_V^M , TD_A^M , TD_C^M , and TD_P^M represent the modified task demand values of the visual, auditory, cognitive and psychomotor channels, then

we can obtain:

$$TD_V^M = TD_V + \alpha_{MWL} * \ln WF_V^{TD} \quad (6)$$

$$TD_A^M = TD_A + \alpha_{MWL} * \ln WF_A^{TD} \quad (7)$$

$$TD_C^M = TD_C + \alpha_{MWL} * \ln WF_C^{TD} \quad (8)$$

$$TD_P^M = TD_P + \alpha_{MWL} * \ln WF_P^{TD} \quad (9)$$

where WF_V^{TD} , WF_A^{TD} , WF_C^{TD} , and WF_P^{TD} stand for PSFs' WF of the visual, auditory, cognitive and psychomotor task demand.

Therefore, the total modified task demand MWL is calculated using (10):

$$MWL = \sum_n^{V,A,C,P} TD_n^M = \sum_n^{V,A,C,P} TD_n + \sum_n^{V,A,C,P} \alpha_{MWL} * \ln WF_n^{TD} \quad (10)$$

where n stands for different channels in VACP, and TD_n^M , TD_n , and WF_n^{TD} represent modified task demand, original task demand, and the PSFs' WF of different channels respectively.

As for the calculation of WF, we can refer to CREAM. In CREAM, each level of PSF has a multiplier θ . Once the level of each PSF is determined, WF can be obtained by multiplying these multipliers [42]. In our research, similarly, WF in (6) – (9) is calculated by multiplying the multipliers of these task-demand-related PSFs, denoted as (11):

$$WF_n^{TD} = \prod_{i=1}^4 \theta_i^n \quad (11)$$

where θ_i^n represents the i th task-demand-related PSF's multiplier of channel n . The multiplier value is shown in TABLE 4. It should be noted that in CREAM, the visual and auditory channels are not distinguished, but are rather unified into one perceptive channel. Thus, it is assumed that visual and auditory channels share the same multiplier.

3) THE ADJUSTMENT APPROACH OF THRESHOLD CONSIDERING THE INFLUENCE OF PSFS

Harsh PSFs will increase the task demand and lower the mental overload threshold. So, the equation is similar in character, and it can be expressed as (12):

$$Th = BTh - \alpha_{Th} * \ln WF \quad (12)$$

where Th is the modified threshold, BTh is the basic threshold, WF is the weighting factor provided by threshold-related PSFs, and α_{Th} is the coefficient for threshold.

In combination with VACP, the total modified threshold Th is shown as (13):

$$Th = \sum_n^{V,A,C,P} BTh_n - \sum_n^{V,A,C,P} \alpha_{Th} * \ln WF_n^{Th} \quad (13)$$

where n stands for the different channels of VACP, BTh_n is channel n 's basic threshold, and WF_n^{Th} is threshold's WF of channel n . In VACP, $\sum_n^{V,A,C,P} BTh_n$ is often regarded as 40. WF_n^{Th} is calculated in a manner similar to (11), with the exception that the task-demand-related PSFs are replaced with threshold-related factors, as shown in (14). TABLE 5 represents the multipliers of the threshold-related PSFs.

$$WF_n^{Th} = \prod_{i=1}^5 \theta_i^n \quad (14)$$

Based on (10) and (13), mental overload can be judged using (15):

$$Th < MWL \quad (15)$$

4) MENTAL OVERLOAD ASSESSMENT MODEL BASED ON BAYESIAN NETWORK

a: THE STRUCTURE OF THE BN-BASED MENTAL OVERLOAD ASSESSMENT MODEL

In this section, we describe the construction of our BN-based mental overload assessment model. Based on the discussions in Section III-A2 and III-A3, (11) and (14) are the starting points for all the calculation processes as they provide the means by which to obtain weighting factors. Furthermore, the correlation among PSFs will change the distribution of some PSFs, so the PSF nodes should consist of the primary PSFs nodes and the PSF nodes adjusted in consideration of the correlations. For the sake of presentation alone, the primary PSF nodes are denoted as PSFs, and the adjusted PSF nodes are denoted as adjusted PSFs. The next step in the mental overload assessment process is based on (10) and (13). The causality described in (10) and (13) is clear, it derives from weighting factor nodes, basic task demand nodes, threshold nodes to modified task demand and threshold nodes. The last step in the assessment process is (15), namely to make the comparison between modified task demand and modified threshold so as to determine the mental overload state. We will introduce the mental overload probability node to display the mental overload assessment result.

In conclusion, there are six classes of nodes and four classes of CPT that are used to build this BN-based mental overload assessment model. The six classes of nodes are:

- PSFs;
- Adjusted PSFs;
- Weighting factors;
- Basic task demands and threshold;
- Modified task demands and threshold;
- Mental overload probability.

Four classes of CPT are:

- Those between PSFs and adjusted PSFs;
- Those between PSFs, adjusted PSFs and weighting factor;
- Those between basic task demand, basic threshold, weighting factor and modified task demand, modified threshold;

TABLE 3. CREAM’s PSFs that related to task demand and threshold.

Task-demand-related PSFs	Threshold-related PSFs
Adequacy of MMI and operational support	Adequacy of organization
Availability of procedures/plans	Working conditions
Number of simultaneous goals	Time of day
Available time	Adequacy of training and preparation
	Crew collaboration quality

TABLE 4. Multipliers of task-demand-related PSFs [41].

PSF	PSF level	Multiplier		
		Visual/Auditory	Cognitive	Psychomotor
Adequacy of MMI and operational support	Supportive	0.50	1.00	0.50
	Adequate	1.00	1.00	1.00
	Tolerable	1.00	1.00	1.00
	Inappropriate	5.00	1.00	5.00
Availability of procedures/plans	Appropriate	0.80	0.50	0.80
	Acceptable	1.00	1.00	1.00
	Inappropriate	2.00	5.00	2.00
Number of simultaneous goals	Fewer than capacity	1.00	1.00	1.00
	Matching current capacity	1.00	1.00	1.00
	More than capacity	2.00	5.00	2.00
	Adequate	0.50	0.50	0.50
Available time	Temporarily inadequate	1.00	1.00	1.00
	Continuously inadequate	5.00	5.00	5.00

TABLE 5. Multipliers of threshold-related PSFs [41].

PSF	PSF level	Multiplier		
		Visual/Auditory	Cognitive	Psychomotor
Adequacy of organization	Very efficient	1.00	0.80	0.80
	Efficient	1.00	1.00	1.00
	Inefficient	1.00	1.20	1.20
	Deficient	1.00	2.00	2.00
Working Conditions	Advantageous	0.80	1.00	0.80
	Compatible	1.00	1.00	1.00
	Incompatible	2.00	1.00	2.00
Time of day	Day-time	1.00	1.00	1.00
	Night-time	1.20	1.20	1.20
Adequacy of training and preparation	Adequate, high experience	0.80	0.50	0.80
	Adequate, low experience	1.00	1.00	1.00
	Inadequate	2.00	5.00	2.00
Crew collaboration quality	Very efficient	0.50	0.50	0.50
	Efficient	1.00	1.00	1.00
	Inefficient	1.00	1.00	1.00
	Deficient	2.00	2.00	5.00

- Those between modified task demand, modified threshold and mental overload probability.

The BN-based mental overload assessment model is shown in FIGURE 2. The BN for this work was constructed using the commercially available software GeNIe 2.4 [48].

b: THE QUANTIFICATION OF ROOT NODES

i) TASK DEMAND AND THRESHOLD QUANTIFICATION

VACP uses verbal anchors to assess task demand levels to aid in increasing consistency and reducing inter-rater

variability [21]. Verbal anchors provide key words that are used to classify the unit tasks, and users are encouraged to assess unit tasks by combining task background, task type, and other task-related information rather than performing the assessment mechanically by simply matching keywords. For example, when a novice performs the same task as an experienced operator, such as driving on the same road, the cognitive task demands are significantly different. According to TABLE 1, the novice may have a higher cognitive task demand, perhaps rated at 6.8 out of 7.0, which has the verbal anchor “evaluation/judgment consider several actions’, while

for the experienced operator, the cognitive task demand is rated at 1.0 out of 7.0, which has the verbal anchor “automatic simple association”. As for the threshold, it is considered to be 40.

ii) DISTRIBUTION QUANTIFICATION OF PSFS

For PSFs nodes, with the support of data, the prior distribution of PSFs can be determined. However, the data are usually sparse, so we can assume that if the PSFs cannot be determined, then, with the exclusion of the impossible levels, the remaining levels are considered as having even distribution. For example, in the area of aviation, all pilots must be fully trained before obtaining permission to fly an aircraft, so for the PSF “Adequacy of training and preparation”, the level “inadequate” is impossible. Then, the remaining levels, “Adequate, high experience” and “Adequate, low experience” are under an even distribution.

c: THE NODES AND CPT OF THE BN-BASED MENTAL OVERLOAD ASSESSMENT MODEL

i) IDENTIFICATION OF THE NODES OF THE TARGET MODEL

As discussed in previous section, we introduced CREAM’s PSFs to describe the contextual conditions of the assessment, so the nine PSFs are identified as root nodes of BN-based model, as shown in FIGURE 2. Four PSFs, namely “working conditions”, “number of simultaneous goals”, “available time” and “crew collaboration quality”, are prescribed in CREAM such that they will be influenced by other PSFs, so we added another four nodes to reflect that. Next, the influence of each PSFs is represented by weighting factors, according to (10) and (13); these weighting factors are classified into four groups: the threshold’s weighting factor, the visual and auditory weighting factor, the cognitive weighting factor, and the psychomotor weighting factor. Accordingly, there are four nodes used to calculate the weighting factors. As for the basic task demand and threshold, in accordance with the weighting factor nodes, there are also four nodes. Finally, two nodes are used to represent the modified task demand and threshold, and one node is used to display the mental overload result.

ii) ALLOCATION OF THE CPT TO THE NODES

• Between PSFs and adjusted PSFs

Theoretically, if the distributions of some PSFs are known, the distributions of their correlated PSFs will be adjusted to adapt them. For example, “Crew collaboration quality” has a correlation with “Adequacy of organization” and “Adequacy of training and preparation”, when the states of “Adequacy of organization” and “Adequacy of training and preparation” are at their highest level, “Crew collaboration quality” will also have a higher probability at its best level rather than continuing to with its prior values. In CREAM, the correlation can be summarized as shown in TABLE 6. When the levels of the PSFs in the left column can be determined, the

distributions will not be adjusted; if not, then the distribution will be adjusted according to the states of the PSFs in the right column.

Subsequently, we must determine the adjustment rule for the adjusted PSFs nodes. In CREAM, every level of a PSF has an effect on performance, namely “Improved”, “Not significant”, and “Reduced”, as shown in TABLE 7. If the majority of the effects of the PSFs in the right column of TABLE 6 are the same, the state of the PSFs that have the same effect in the left column will have a higher probability. TABLE 8 shows the adjustment rules and the adjusted distributions. Let us take “Crew collaboration quality” as an example. If all of the correlated PSFs’ effects are “Improved”, then the state of “Very efficient” which has the same effect, will have a higher probability (3/4). Accordingly, the distribution of “Crew collaboration quality” will be adjusted from an even distribution to (3/4, 1/4, 0, 0). Based on the above rules, the CPT can be obtained. The CPT for the node “Crew collaboration quality” is shown in TABLE 9.

• Between PSFs, adjusted PSFs and weighting factor

Unlike the adjusted PSF nodes, the weighting factor nodes are complicated by the problem that the scale of CPT is much bigger. Let us take the node “Visual and auditory weighting factor” as an example. The scale of the CPT is 108×54 , which is difficult to calculate further. Fortunately, in GeNIe, we can set the node’s type as “equation” to solve this problem. Still using the node “Visual and auditory weighting factor” as the example, the steps for setting the node are shown in TABLE 10. First, the function “choose” is used to allocate the multiplier from TABLE 4; then (11) can be directly input into the nodes, and the node can be set to perform further calculations, which are much easier to perform.

• Between weighting factors, basic task demands, basic threshold and modified task demands, modified threshold

As with the earlier calculation, the type of node “Modified task demand” and “Modified threshold” are set as “equation”, and then the settings from TABLE 11 are input into the nodes. In these equations, parameter α_{MWL} and α_{Th} is around 5 based on the regression analysis of the data [46].

• Between modified task demands, modified threshold and mental overload probability

In the mental overload probability node, conditional functions are used to calculate mental overload probability. The function expression is shown in TABLE 12, and it takes on a value of 1 when the operator is in a state of mental overload; otherwise, it remains at a value of 0 otherwise. By calculating this expression, mental overload probability can be obtained.

The complete BN-based mental overload assessment model is shown in FIGURE 3.

B. MENTAL OVERLOAD ASSESSMENT PROCESS

The mental overload assessment process consists of four steps. First, task analysis needs to be performed to gain knowledge and insight of the target scenario. Next, the task

TABLE 6. Correlation between PSFs in cream [41].

PSF	Depends on the following PSFs				
Working conditions	Adequacy of organization	Adequacy of MMI and operational support	Available time	Time of day	Adequacy of training and experience
Number of simultaneous goals	Working conditions	Adequacy of MMI and operational support	Availability of procedures/plans		
Available time	Working conditions	Adequacy of MMI and operational support	Availability of procedures/plans	Number of simultaneous goals	Time of day
Crew collaboration quality	Adequacy of organization	Adequacy of training and experience			

TABLE 7. PSFs and its expected effects.

PSFs	PSF level	Expected effect
Adequacy of organization	Very efficient	Improved
	Efficient	Not significant
	Inefficient	Reduced
	Deficient	Reduced
Working conditions	Advantageous	Improved
	Compatible	Not significant
	Incompatible	Reduced
Adequacy of MMI and operational support	Supportive	Improved
	Adequate	Not significant
	Tolerable	Not significant
Availability of procedures/plans	Inappropriate	Reduced
	Appropriate	Improved
	Acceptable	Not significant
Number of simultaneous goals	Inappropriate	Reduced
	Fewer than capacity	Not significant
	Matching current capacity	Not significant
Available time	More than capacity	Reduced
	Adequate	Improved
	Temporarily inadequate	Not significant
Time of day	Continuously inadequate	Reduced
	Day-time	Not significant
	Night-time	Reduced
Adequacy of training and preparation	Adequate, high experience	Improved
	Adequate, low experience	Not significant
	Inadequate	Reduced
Crew collaboration quality	Very efficient	Improved
	Efficient	Not significant
	Inefficient	Not significant
	Deficient	Reduced

TABLE 8. The adjustment of PSFs' distribution.

PSF	Adjust when the follow condition is met	Adjusted distribution
Working conditions	4/5 of related PSFs' effects are improved	(2/3,1/3,0)
	4/5 of related PSFs' effects are reduced	(0,1/3,2/3)
Number of simultaneous goals	2/3 of related PSFs' effects are improved	(2/3,1/3,0)
	2/3 of related PSFs' effects are reduced	(0,1/3,2/3)
Available time	4/5 of related PSFs' effects are improved	(2/3,1/3,0)
	4/5 of related PSFs' effects are reduced	(0,1/3,2/3)
Crew collaboration quality	all the related PSFs' effects are improved	(3/4,1/4,0,0)
	all the related PSFs' effects are improved	(0,0,1/4,3/4)

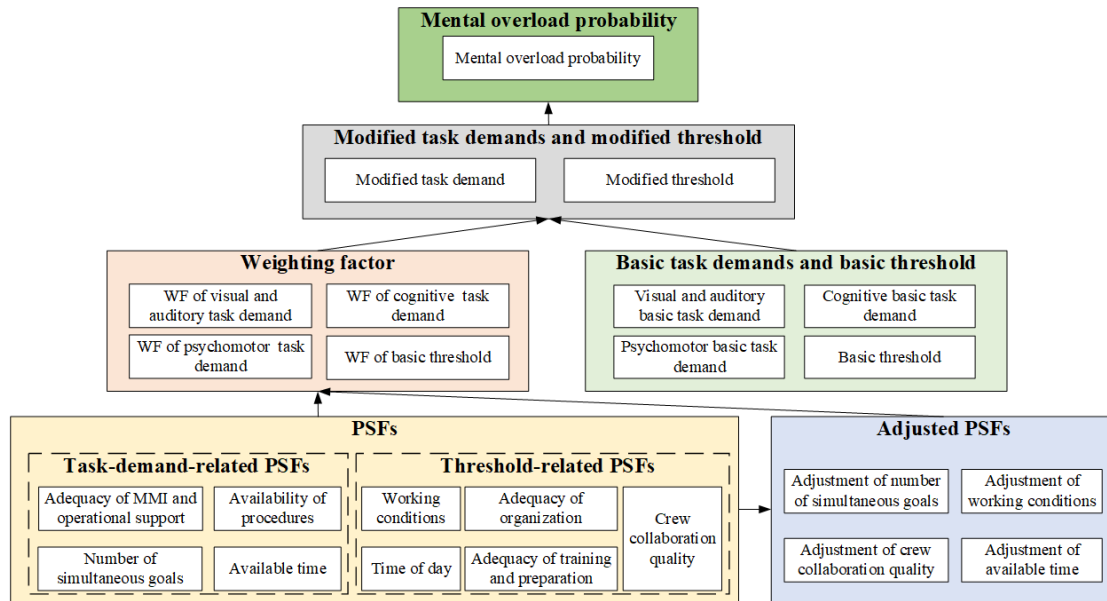


FIGURE 2. The structure of the BN-based mental overload assessment model.

TABLE 9. The CPT Of node “adjustment of crew collaboration quality” (I =improved, NS = not significant, R = reduced).

		I			NS			R		
		I	NS	R	I	NS	R	I	NS	R
Adequacy of organization										
Adequacy of training and preparation										
Adjustment of crew collaboration quality	Very efficient	0.75	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0
	Efficient	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0
	Inefficient	0	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
	Deficient	0	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.75

demands imposed by the task are determined so as to be used as inputs. Then, the prior distribution of PSFs is determined so as to assign the CPT to BN. Finally, the mental overload probability can be obtained through the mental overload assessment of the scenario. If the predicted mental overload probability cannot meet the needs of the designers, then the design requires iteration until it satisfies the design index. The following section elaborates upon these steps.

1) TASK ANALYSIS OF THE TARGET SCENARIO

To perform the task analysis, the target scenario must be identified. Several rules must be obeyed. First, the scenario must be an interactive scenario, meaning that the operator must monitor the parameters, judge the state and react to it if needed. If the scenario only requires that the pilot monitor or judge or respond, the mental workload level will usually remain in a normal state, and the designers are not encouraged to pay more attention to it. Second, the operator should have several unit tasks to address in the scenario. A single unit task is not enough to cause the mental workload level to be too high; thus, a multi-task state requires more concern from the designers. Scenarios can be filtered by applying these two rules, and preliminary scenario screening can identify whether the scenario needs to be analyzed.

After the identification of the target scenario, task analysis must be carried out. The purpose of the task analysis is to divide the tasks into unit tasks. Here, to better perform task analysis, hierarchical task analysis (HTA) is used. HTA is a goal-oriented task model used to analyze the interactive process of operators in complex socio-technical systems. Ideally, the operating manual is the best place to obtain the information required to perform the HTA. The opinions of experts and interviews with operators can be used to supplement the HTA's details.

2) IDENTIFICATION OF THE TASK DEMAND VALUES

Sections III-A4.b, which treats the topic of task demand values, describes a more fully selection criteria and precaution required for this step of process. In this section, we will present the criteria and precautions required to identify the task demand values of the operators under a specific scenario. The selection of the task demand value determines whether the assessment result will be correct, so it is of vital importance.

3) ALLOCATION OF THE PRIOR DISTRIBUTION TO THE PSFS

Relevant statistical reports are the best way to obtain the prior distributions of the PSFs. In addition, expert evaluation is

TABLE 10. The setting of node “visual and auditory weighting factor”.

Steps	Function expression
1. Allocate multipliers to the PSFs node	TD-PSF1=Choose(Adequacy of MMI and operational support,0.5,1,1,5) TD-PSF2=Choose(Availability of procedures,0.8,1,2) TD-PSF3=Choose(Adjustment of number of simultaneous goals,1,1,2) TD-PSF4=Choose(Adjustment of available time,0.5,1,5)
2. Calculate weighting factors	Visual and auditory weighting factor=TD-PSF1*TD-PSF2*TD-PSF3*TD-PSF4

TABLE 11. The setting of nodes “modified task demand” and “modified threshold”.

Node	Setting of function
Modified task demand	Modified task demand=5*(Ln(visual and auditory weighting factor)+Ln(cognitive weighting factor)+Ln(psychomotor weighting factor))+visual and auditory task demand+cognitive task demand+psychomotor value
Modified threshold	Modified Threshold=Basic task demand-5*(Ln(visual and auditory weighting factor)+Ln(cognitive weighting factor)+Ln(psychomotor weighting factor))

TABLE 12. The setting of node “mental overload probability”.

Node	Setting of function
Mental overload probability	Mental overload probability=If(Modified task demand>Modified Threshold,1,0)

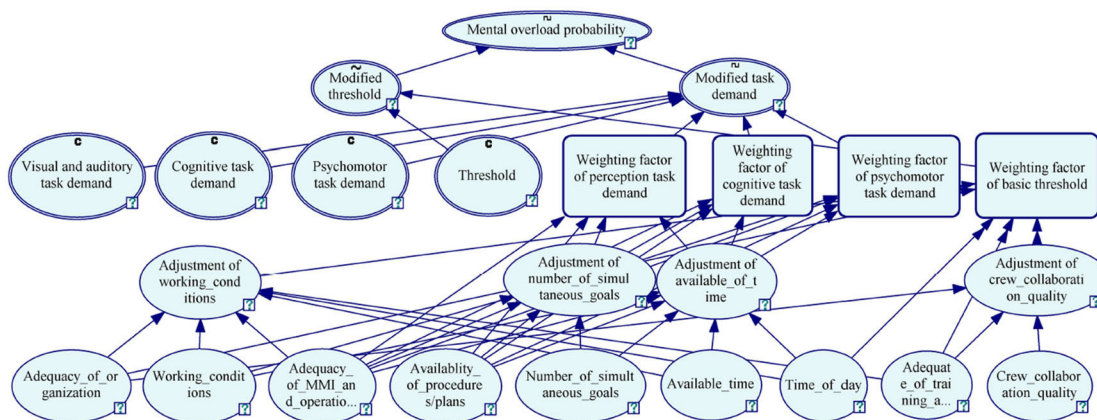


FIGURE 3. The BN-based mental overload assessment model.

also a way determine prior distributions. Readers can refer to Section III-A4.b and III-A4.c to obtain more information.

4) PERFORM MENTAL OVERLOAD ASSESSMENT OF THE SCENARIO BASED ON BN

Based on the mental overload assessment model, the mental overload assessment of the scenario can be performed. The model can output some results to assist the further analysis, including the mental overload probability, the distribution of the mental workload level, the distribution of the threshold level, and the influence of PSFs on mental overload probability. The following section demonstrates this process in detail in relation to an accident involving a helicopter crash.

IV. CASE STUDY

A. APPLICATION SCENARIO

In order to demonstrate the capacity of the proposed method, we have applied it to a helicopter accident that occurred in

Maryland. On January 10, 2005, at approximately 23:11, a helicopter crashed into the Potomac River during a low-altitude cruising flight near Oxon Hill, Maryland [49]. The pilot died, and other crew members sustained severe injuries. According to the accident report, after taking off from the Washington Hospital Center Helipad, the helicopter was en route to the Stafford Regional Airport. Due to its proximity to the target airport, the pilot chose to cruise at a low altitude. During the flight, the helicopter needed to cross over Woodrow Wilson Bridge. Hence, when helicopter approached the bridge, the pilot first climbed to a higher altitude and crossed over the bridge; then the helicopter descended so as to continue cruising at a low altitude. However, when performing these tasks, the air traffic controller informed the pilot that an Airbus craft was approaching. To avoid a possible collision, the pilot searched for Airbus visually and tried to maintain visual separation from the Airbus. Subsequently, because the pilot directed most of his

attention outside of the cockpit, ignoring the flight instrument display, the helicopter crashed into the Potomac River during the descent stage. The flight mission profile is shown in FIGURE 4.

As for the cause of the accident, the root reason may be that pilot has too many concurrent tasks to handle during the descent stage, which may have led to mental overload. As a result, the pilot abandoned the task of visually checking the flight instrumentation, believing that, once the altitude become too low, the ground proximity warning system (GPWS) would sound an alert. However, a failure occurred in the GPWS, such that it did not function to alert the pilot in time. Thus, pilot did not have a correct situational awareness of the flight parameters prior to the helicopter hitting the water.

B. MENTAL OVERLOAD ASSESSMENT OF THE SCENARIO

1) TASK ANALYSIS

To assess the mental overload state of the pilot in this scenario, we first need to conduct a HTA to divide the task into unit tasks. As Che et al. conducted a detailed analysis of this case, we will refer to their HTA results, as shown in FIGURE 5 [50]. Obviously, there were four tasks in this scenario: to climb, to cross over the bridge, to descend, and to search for and avoid Airbus. For the task “climb”, the pilot needs to contact the air traffic controller to report the climbing position. Once the permission has been obtained, the pilot starts to adjust the throttle and pedals to change the altitude, course, speed, and rate of climbing until reaching the predetermined position. For the task “cross over the bridge”, the pilot needs to carefully manipulate the helicopter and check the instruments to avoid approaching the bridge too closely. For task “descend”, the pilot needs to adjust the collective pitch, throttle and pedals to return the helicopter to its pre-climb altitude. For the task “search for and avoid the Airbus”, the pilot needs to contact the air traffic controller to obtain information about the approaching Airbus and then to search for the Airbus visually, maintaining visual separation from the Airbus. The last task is a temporary task, and the pilot is required to perform it while simultaneously performing the first three tasks.

For the sake of simplicity, the following two assumptions have been made:

- Only the pilot and air traffic controller are considered in the process;
- The accident occurred while the pilot was handling the third and fourth tasks simultaneously, so we mainly assess the mental overload state of the pilot during this period.

Based on the HTA results and assumptions, we conducted a further analysis.

2) IDENTIFICATION OF THE TASK DEMAND VALUES

To perform the VACP assessment, we need to determine the task demands of different unit tasks. As shown in FIGURE 5, among the unit tasks, 3.1.5 and 4.2 are visual tasks, task

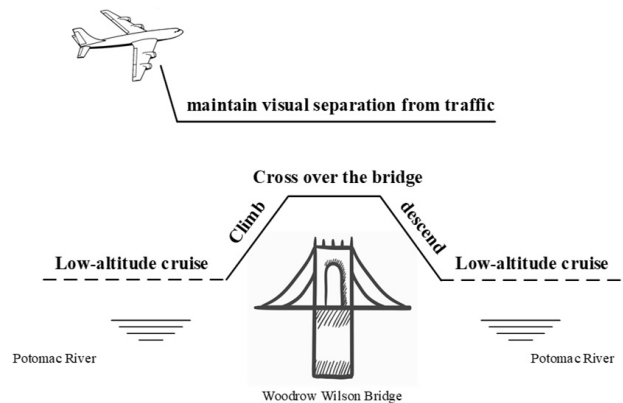


FIGURE 4. Flight mission profile of the helicopter [50].

4.1 is an auditory task, task 3.2.1 is a cognitive task, and the remaining unit tasks are psychomotor tasks. The results are summarized as TABLE 13.

For the visual tasks, the task demand value of unit task 3.1.5 is 3.0 out of 7.0, with the verbal anchor “Visual Inspect/Check” as the pilot needed to check the flight instrumentation; the task demand of unit task 4.2 is 6.0 out of 7.0, with the verbal anchor “Visual Scan/Search/Monitor” as the pilot needed to continuously search for the approaching Airbus. For the auditory task, the task demand of unit task 4.1 is 4.0 out of 7.0, with the verbal anchor “Complex speech” as the pilot and air traffic controller needed to communicate with each other in sentences. For the cognitive task, the task demand of unit task 3.2.1 is 4.6 out of 7.0, with the verbal anchor “Evaluation/Judgement (consider several aspects)” as the pilot needed to judge the appropriate altitude. Those psychomotor tasks have been treated as a whole and share a common task demand value, which is 2.6 out of 7.0, with the verbal anchor “Continuous Adjustive (flight controls, sensor control)” because flight control is a continuous procedure that consists of the adjustment of the cyclic stick, the collective pitch, the anti-torque pedal, and the throttle. If a change occurs solely in the cyclic stick, the total thrust and lift ratio will not change, thus requiring adjustment of the throttle and collective pitch accordingly. As these manipulations are always linked, they are seen as a whole task.

3) DETERMINATION OF THE PRIOR DISTRIBUTION OF THE PSFS

According to the accident report, the communication between the pilot and the air traffic controller was timely and accurate; both sides strictly comply with the requirements in the flight manual (Adequacy of organization: Very Efficient). It was night-time during the flight (Time of day: Night-time), and there was no outside visual reference, meaning that once the pilot crossed the bridge, he was flying into a black void (Working Conditions: Incompatible). As shown in TABLE 14, when the pilot was required to handle several tasks simultaneously, he actually abandoned the visual check

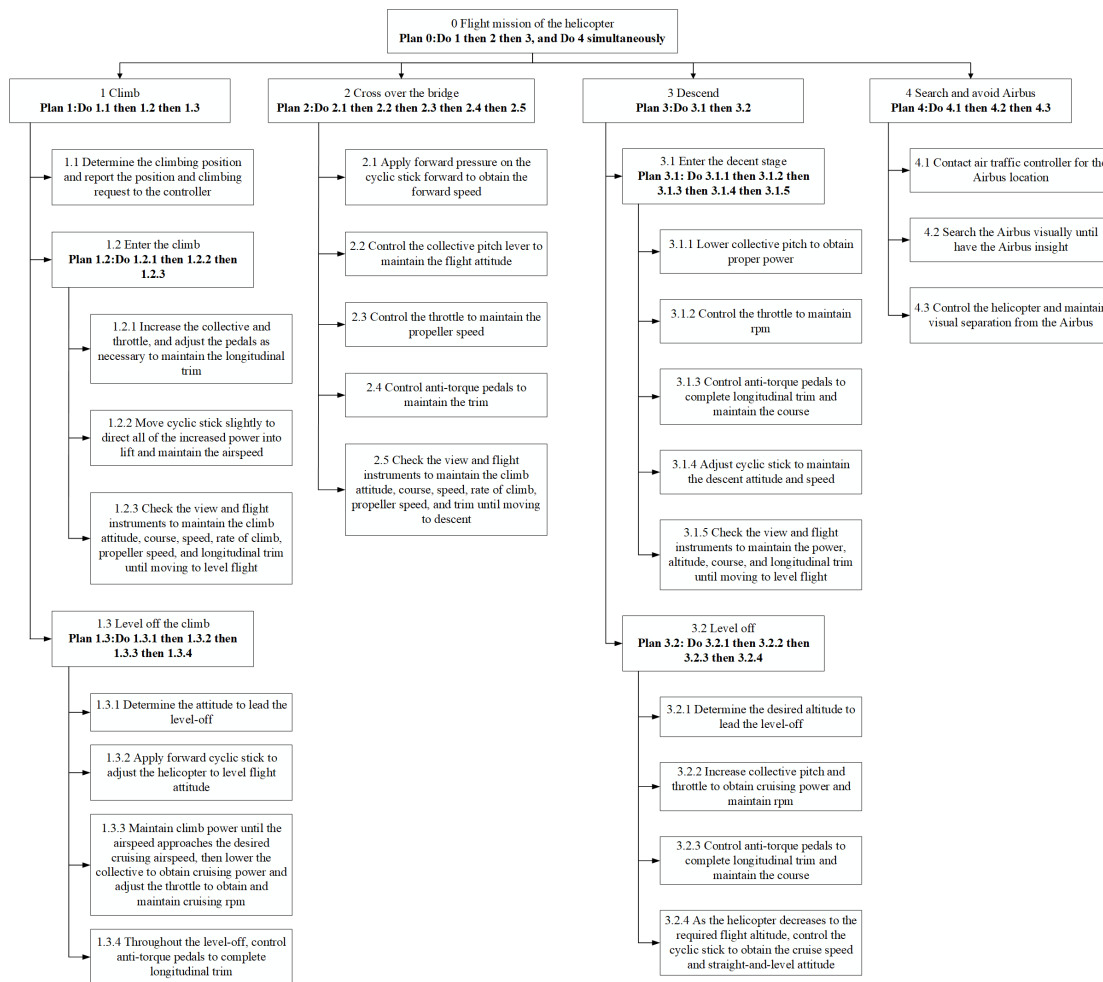


FIGURE 5. HTA result of the scenario.

TABLE 13. Task demands of unit tasks in the scenario.

Unit task	Task type	Activity	Value
3.1.5	Visual	Visual Inspect (discrete inspection /static condition)	3.0
4.2		Visual Scan/Search/Monitor (continuous/series inspection, multiple conditions)	6.0
4.1	Auditory	Complex speech (sentence)	4.0
3.2.1	Cognitive	Evaluation/Judgement (consider several aspects)	4.6
3.1.1, 3.1.2, 3.1.3, 3.1.4, 3.2.2, 3.2.3, 3.2.4, 4.3	Psychomotor	Continuous adjustive (flight control)	2.6

of the flight instruments, which proves that the number of concurrent tasks was beyond his capacity (Number of simultaneous goals: More than capacity).

4) BN MODELING AND CALCULATION

The BN structure is similar to that shown in FIGURE 3, only the adjusted nodes differs because “Working conditions” and “Number of simultaneous goals” can be determined according to the report. Thus, there is no need to adjust their distributions. Conversely, since “Available time” and “Crew

collaboration capacity” were subject to the assumed even distribution, their prior distribution must be adjusted.

The results show that the mental overload probability in this accident scenario is 68%.

C. ANALYSIS OF THE RESULTS

1) THE DISTRIBUTION OF THE MENTAL WORKLOAD LEVEL

As the assessment of VACP is a point estimation, and the assessment of the proposed method is a probabilistic estimation, we must determine where the assessment value of

TABLE 14. Distribution of PSFs.

PSFs	Prior distribution
Adequacy of organization	Very efficient: 100%
Working Conditions	Incompatible: 100%
Adequacy of MMI and operational support	Even distribution
Availability of procedures/plans	Even distribution
Number of simultaneous goals	More than capacity: 100%
Available time	Even distribution
Time of day	Night time: 100%
Adequacy of training and preparation	Even distribution between high and low experience
Crew collaboration quality	Even distribution

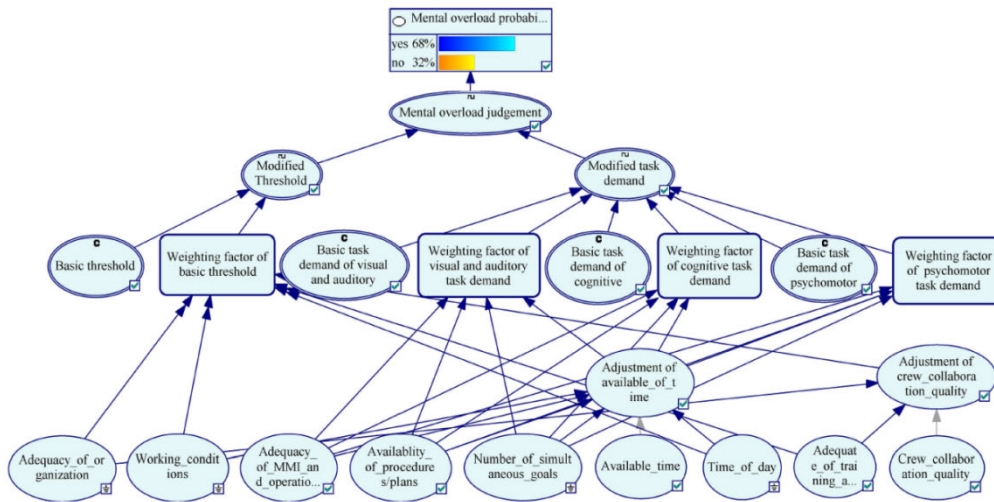


FIGURE 6. Assessment result of the proposed method.

TABLE 15. The comparison between VACP and the proposed method.

Method	Assessment value of mental workload	The probability of mental overload
VACP	20.2	0
Our proposed method	49.94	68%

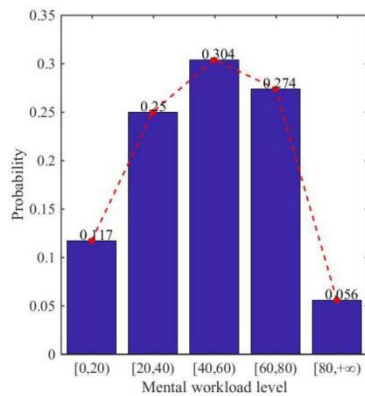


FIGURE 7. The distribution of mental workload level.

VACP falls within the probability distribution of the mental workload level. The distribution of the mental workload level is shown in FIGURE 7. We can analyze this result from two

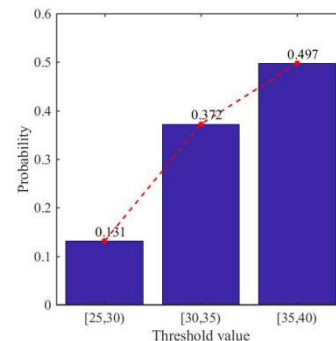


FIGURE 8. The distribution of threshold.

perspectives: (1) According to the calculation, $p(MWL < 20.2) = 0.1089$, the assessment value of VACP falls within the margins of the mental workload level distribution; (2) the mean value of the mental workload distribution is about 49.94, this is almost double the value assessed by VACP.

TABLE 16. Task demands of unit tasks in the scenario.

PSFs	PSFs Level	Probability
Adequacy of organization	Very efficient	0.68
	Efficient	0.73
	Inefficient	0.76
	Deficient	0.83
Working Conditions	Advantageous	0.44
	Compatible	0.49
	Incompatible	0.68
Number of simultaneous goals	Fewer than capacity	0.36
	Match capacity	0.37
	More than capacity	0.68
Time of day	Day-time	0.53
	Night-time	0.68

This actually shows the effect of PSFs: the degraded PSFs shift the mean value of the mental workload and change the distribution.

2) THE DISTRIBUTION OF THE THRESHOLD

As for the distribution of the threshold, reader can refer to FIGURE 8. In the described scenario, the probability that threshold is lower than 35 is almost 50%, meaning that the pilot's threshold witnessed a great drop under the impact of the degraded PSFs. For normal tasks, the threshold may not have an influence as the task demand is usually low. However, in emergency situations, the threshold can be influential because the addition of instantaneous tasks will overwhelm the threshold. The decrease in the threshold is one of the main causes of the mental overload.

3) COMPARISON BETWEEN VACP AND THE PROPOSED METHOD

VACP's assessment result is 20.2, which means that in this scenario, the pilot's mental workload was lower than the threshold (40). In contrast, our proposed method indicates with a relatively high probability (68%) that the pilot was in a stage of mental overload, as shown in TABLE 15. With the support of VACP's result, the designer may draw the wrong inference that the scenario to be assessed is safe enough. This incorrect inference would be caused by the absence of the influence of PSFs influence on the task demand and threshold. However, with the support of our proposed method's results, the designer's attention would be drawn to the discovery of the factors that influence mental workload, and the designer would therefore advance targeted suggestions.

4) THE MOST INFLUENTIAL PSFS

To assess the influence of PSFs, "target sensitivity analysis" is applied. By setting the evidence of a specific node, a target sensitivity analysis allows the user to investigate the changes in the probability of mental overload, and the results are shown in TABLE 16. Among all of the PSFs determined in the case, "Number of simultaneous goals" has been identified as being the most influential. Thus, the most possible accident chain is: the instant task of finding nearby Airbus → limited outside reference and lack of altimeter altitude → visual

search for Airbus → too many unit tasks for pilot to perform → mental overload → ignorance of visual scan of instruments → failure occurred in GPWS → accident.

5) SUGGESTION

The accident report describes the probable cause of the accident: "The probable cause of this accident was the pilot's failure to identify and arrest the helicopter's descent, which resulted in controlled flight into terrain. Contributing to the accident were the dark night conditions, limited outside visual references, and the lack of an operable radar altimeter in the helicopter". We would like to point out that the existence of too many simultaneous goals is the most influential factor, which is in accordance with the deduction in the report that "the pilot was unable to perform complex tasks in the helicopter or fly a compete mission involving several tasks in a series" which validates the effectiveness of our proposed method.

As for suggestions, there are actual two threads to follow: improving the design and optimization of the task flow to reduce the task demand; and improving PSF-related elements so as to reduce their impact on the scenario. In this scenario, visually searching for the approaching Airbus was not the ideal way, so we recommend both the use of radar to detect other aircraft and the integration of this function with the flight instrumentation. In this way, at every fixation of the flight instrument, the pilot will be able to obtain enough information about the helicopter and the surrounding situation. This suggestion is also consistent with the accident report: "If a functioning radar altimeter had been available to and used by the pilot, it could have provided a constant altitude cue, in the form of a digital readout of feet above the terrain, to enhance his awareness of the helicopter's height above the water" [49].

V. CONCLUSION

Mental workload is one of the most concerning factors in the area of safety, and the assessment of mental overload is essential to the design stage. The present study proposes a mental overload assessment model that considers the effects of PSFs. The proposed model has the following characteristics: First, the proposed method considers the influence of PSFs on mental overload, which extends the description of the traditional mental overload assessment method. Second, the relation between mental workload level and human error probability was established by converting the PSFs' weighting factors. Finally, BN was used to model the influence of PSFs and calculate the probability of mental overload. The performance of the proposed method was demonstrated through a case study of an accident involving a helicopter crash. The results show that the proposed method can efficiently determine the mental overload state of the pilot better than the traditional VACP method can, and it provides a more rational explanation of the accident's evolution.

REFERENCES

- [1] S. Pütz, V. Rick, A. Mertens, and V. Nitsch, "Using IoT devices for sensor-based monitoring of employees' mental workload: Investigating managers' expectations and concerns," *Appl. Ergonom.*, vol. 102, Jul. 2022, Art. no. 103739, doi: [10.1016/j.apergo.2022.103739](https://doi.org/10.1016/j.apergo.2022.103739).
- [2] I. M. Shuggi, H. Oh, P. A. Shewokis, and R. J. Gentili, "Mental workload and motor performance dynamics during practice of reaching movements under various levels of task difficulty," *Neuroscience*, vol. 360, pp. 166–179, Sep. 2017, doi: [10.1016/j.neuroscience.2017.07.048](https://doi.org/10.1016/j.neuroscience.2017.07.048).
- [3] A. Tjolleng, K. Jung, W. Hong, W. Lee, B. Lee, H. You, J. Son, and S. Park, "Classification of a driver's cognitive workload levels using artificial neural network on ECG signals," *Appl. Ergonom.*, vol. 59, pp. 326–332, Mar. 2017, doi: [10.1016/j.apergo.2016.09.013](https://doi.org/10.1016/j.apergo.2016.09.013).
- [4] M. Naderpour, J. Lu, and G. Zhang, "A safety-critical decision support system evaluation using situation awareness and workload measures," *Rel. Eng. Syst. Saf.*, vol. 150, pp. 147–159, Jun. 2016, doi: [10.1016/j.res.2016.01.024](https://doi.org/10.1016/j.res.2016.01.024).
- [5] C. Marchand, J. B. De Graaf, and N. Jarrassé, "Measuring mental workload in assistive wearable devices: A review," *J. Neuroeng. Rehabil.*, vol. 18, no. 1, pp. 1–15, Nov. 2021, doi: [10.1186/s12984-021-00953-w](https://doi.org/10.1186/s12984-021-00953-w).
- [6] F. Paas and J. J. G. van Merriënboer, "Cognitive-load theory: Methods to manage working memory load in the learning of complex tasks," *Current Directions Psychol. Sci.*, vol. 29, no. 4, pp. 394–398, Aug. 2020, doi: [10.1177/0963721420922183](https://doi.org/10.1177/0963721420922183).
- [7] G. Durantin, J.-F. Gagnon, S. Tremblay, and F. Dehais, "Using near infrared spectroscopy and heart rate variability to detect mental overload," *Behav. Brain Res.*, vol. 259, pp. 16–23, Feb. 2014, doi: [10.1016/j.bbr.2013.10.042](https://doi.org/10.1016/j.bbr.2013.10.042).
- [8] J. Sweller, J. J. G. van Merriënboer, and F. Paas, "Cognitive architecture and instructional design: 20 years later," *Educ. Psychol. Rev.*, vol. 31, no. 2, pp. 261–292, Jun. 2019.
- [9] K. Kia, J. Hwang, I.-S. Kim, H. Ishak, and J. H. Kim, "The effects of target size and error rate on the cognitive demand and stress during augmented reality interactions," *Appl. Ergonom.*, vol. 97, Nov. 2021, Art. no. 103502, doi: [10.1016/j.apergo.2021.103502](https://doi.org/10.1016/j.apergo.2021.103502).
- [10] R. L. Charles and J. Nixon, "Measuring mental workload using physiological measures: A systematic review," *Appl. Ergonom.*, vol. 74, pp. 221–232, Jan. 2019.
- [11] D. Lee, J. Jeong, K. Kim, B. Yu, and S. Lee, "Continuous EEG decoding of pilots' mental states using multiple feature block-based convolutional neural network," *IEEE Access*, vol. 8, pp. 121929–121941, 2020.
- [12] P. Wang, W. Fang, and B. Guo, "A measure of mental workload during multitasking: Using performance-based timed Petri nets," *Int. J. Ind. Ergonom.*, vol. 75, Jan. 2020, Art. no. 102877, doi: [10.1016/j.ergon.2019.102877](https://doi.org/10.1016/j.ergon.2019.102877).
- [13] S. Park, S. Jeong, and R. Myung, "Modeling of multiple sources of workload and time pressure effect with ACT-R," *Int. J. Ind. Ergonom.*, vol. 63, pp. 37–48, Jan. 2018, doi: [10.1016/j.ergon.2017.07.003](https://doi.org/10.1016/j.ergon.2017.07.003).
- [14] C. Wu and Y. Liu, "Queuing network modeling of driver workload and performance," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 3, pp. 528–537, Sep. 2007.
- [15] S. Jo, R. Myung, and D. Yoon, "Quantitative prediction of mental workload with the ACT-R cognitive architecture," *Int. J. Ind. Ergonom.*, vol. 42, no. 4, pp. 359–370, Jul. 2012, doi: [10.1016/j.ergon.2012.03.004](https://doi.org/10.1016/j.ergon.2012.03.004).
- [16] J. H. McCracken and T. B. T. B. Aldrich, "Analyses of selected LHX mission functions: Implications for operator workload and system automation goals," ANACAPA Sci., Santa Barbara, CA, USA, Tech. Rep. AD-A232 330, 1984.
- [17] C. D. Wickens, W. S. Helton, J. G. Hollands, and S. Banbury, *Engineering Psychology and Human Performance*. Evanston, IL, USA: Routledge, 2021.
- [18] Y. Zhang, H. Zheng, Y. Duan, L. Meng, and L. Zhang, "An integrated approach to subjective measuring commercial aviation pilot workload," in *Proc. ICIEA Conf.*, Jun. 2015, pp. 1093–1098.
- [19] P. Wang, W. Fang, and B. Guo, "Mental workload evaluation and its application in train driving multitasking scheduling: A timed Petri net-based model," *Cogn., Technol. Work*, vol. 23, no. 2, pp. 299–313, May 2021, doi: [10.1007/s10111-019-00608-w](https://doi.org/10.1007/s10111-019-00608-w).
- [20] S. C. Bommer and M. Fendley, "A theoretical framework for evaluating mental workload resources in human systems design for manufacturing operations," *Int. J. Ind. Ergonom.*, vol. 63, pp. 7–17, Jan. 2018, doi: [10.1016/j.ergon.2016.10.007](https://doi.org/10.1016/j.ergon.2016.10.007).
- [21] C. F. Rusnock and B. J. Borghetti, "Workload profiles: A continuous measure of mental workload," *Int. J. Ind. Ergonom.*, vol. 63, pp. 49–64, Jan. 2018, doi: [10.1016/j.ergon.2016.09.003](https://doi.org/10.1016/j.ergon.2016.09.003).
- [22] C. Kandemir, H. A. H. Handley, and D. Thompson, "A workload model to evaluate distracters and driver's aids," *Int. J. Ind. Ergonom.*, vol. 63, pp. 18–36, Jan. 2018, doi: [10.1016/j.ergon.2016.09.004](https://doi.org/10.1016/j.ergon.2016.09.004).
- [23] M. S. Young, K. A. Brookhuis, C. D. Wickens, and P. A. Hancock, "State of science: Mental workload in ergonomics," *Ergonomics*, vol. 58, no. 1, pp. 1–17, Jan. 2015, doi: [10.1080/00140139.2014.956151](https://doi.org/10.1080/00140139.2014.956151).
- [24] L. Tsao, J. Chang, and L. Ma, "Fatigue of Chinese railway employees and its influential factors: Structural equation modelling," *Appl. Ergonom.*, vol. 62, pp. 131–141, Jul. 2017, doi: [10.1016/j.apergo.2017.02.021](https://doi.org/10.1016/j.apergo.2017.02.021).
- [25] M. Kaptan, Ö. Uğurlu, and J. Wang, "The effect of nonconformities encountered in the use of technology on the occurrence of collision, contact and grounding accidents," *Rel. Eng. Syst. Saf.*, vol. 215, Nov. 2021, Art. no. 107886, doi: [10.1016/j.res.2021.107886](https://doi.org/10.1016/j.res.2021.107886).
- [26] M. S. Young and N. A. Stanton, "What's skill got to do with it? Vehicle automation and driver mental workload," *Ergonomics*, vol. 50, no. 8, pp. 1324–1339, Aug. 2007, doi: [10.1080/00140130701318855](https://doi.org/10.1080/00140130701318855).
- [27] J. T. Kim, S. K. Shin, J. H. Kim, and P. H. Seong, "An experimental approach to estimate operator's information processing capacity for diagnosing tasks in NPPs," *Ann. Nucl. Energy*, vol. 59, pp. 100–110, Sep. 2013, doi: [10.1016/j.anucene.2013.03.023](https://doi.org/10.1016/j.anucene.2013.03.023).
- [28] S. Rangra, M. Sallak, W. Schön, and F. Vanderhaegen, "A graphical model based on performance shaping factors for assessing human reliability," *IEEE Trans. Rel.*, vol. 66, no. 4, pp. 1120–1143, Dec. 2017, doi: [10.1109/TR.2017.2755543](https://doi.org/10.1109/TR.2017.2755543).
- [29] M. Porthin, M. Liinasuo, and T. Kling, "Effects of digitalization of nuclear power plant control rooms on human reliability analysis—A review," *Rel. Eng. Syst. Saf.*, vol. 194, Feb. 2020, Art. no. 106415, doi: [10.1016/j.res.2019.03.022](https://doi.org/10.1016/j.res.2019.03.022).
- [30] S. W. Lee, A. R. Kim, J. S. Ha, and P. H. Seong, "Development of a qualitative evaluation framework for performance shaping factors (PSFs) in advanced MCR HRA," *Ann. Nucl. Energy*, vol. 38, no. 8, pp. 1751–1759, Aug. 2011, doi: [10.1016/j.anucene.2011.04.006](https://doi.org/10.1016/j.anucene.2011.04.006).
- [31] Z. Yin, Z. Liu, and Z. Li, "Identifying and clustering performance shaping factors for nuclear power plant commissioning tasks," *Hum. Factors Ergonom. Manuf. Service Ind.*, vol. 31, no. 1, pp. 42–65, Jan. 2021, doi: [10.1002/hfm.20868](https://doi.org/10.1002/hfm.20868).
- [32] P. Di Mascio, R. Carrara, L. Frascoco, E. Luciano, A. Ponziani, and L. Moretti, "How the tower air traffic controller workload influences the capacity in a complex three-runway airport," *Int. J. Environ. Res. Public Health*, vol. 18, no. 6, p. 2807, Mar. 2021, doi: [10.3390/ijerph18062807](https://doi.org/10.3390/ijerph18062807).
- [33] A. Azadeh, M. Rouzbahman, M. Saberi, F. Valianpour, and A. Keramati, "Improved prediction of mental workload versus HSE and ergonomics factors by an adaptive intelligent algorithm," *Saf. Sci.*, vol. 58, pp. 59–75, Oct. 2013, doi: [10.1016/j.ssci.2013.03.004](https://doi.org/10.1016/j.ssci.2013.03.004).
- [34] P. Bhavsar, B. Srinivasan, and R. Srinivasan, "Pupillometry based real-time monitoring of operator's cognitive workload to prevent human error during abnormal situations," *Ind. Eng. Chem. Res.*, vol. 55, no. 12, pp. 3372–3382, Mar. 2016, doi: [10.1021/acs.iecr.5b03685](https://doi.org/10.1021/acs.iecr.5b03685).
- [35] S. M. Casner, "Perceived vs. measured effects of advanced cockpit systems on pilot workload and error: Are pilots' beliefs misaligned with reality?" *Appl. Ergonom.*, vol. 40, no. 3, pp. 448–456, May 2009, doi: [10.1016/j.apergo.2008.10.002](https://doi.org/10.1016/j.apergo.2008.10.002).
- [36] M. Musharraf, J. Hassan, F. Khan, B. Veitch, S. MacKinnon, and S. Imtiaz, "Human reliability assessment during offshore emergency conditions," *Saf. Sci.*, vol. 59, pp. 19–27, Nov. 2013.
- [37] J. Chen, D. Zhou, C. Lyu, and X. Zhu, "A method of human reliability analysis and quantification for space missions based on a Bayesian network and the cognitive reliability and error analysis method," *Qual. Rel. Eng. Int.*, vol. 34, no. 5, pp. 912–927, Jul. 2018.
- [38] S. Abrishami, N. Khakzad, and S. M. Hosseini, "A data-based comparison of BN-HRA models in assessing human error probability: An offshore evacuation case study," *Rel. Eng. Syst. Saf.*, vol. 202, Oct. 2020, Art. no. 107043, doi: [10.1016/j.res.2020.107043](https://doi.org/10.1016/j.res.2020.107043).
- [39] M. J. Akhtar and I. B. Utne, "Human fatigue's effect on the risk of maritime groundings—A Bayesian network modeling approach," *Saf. Sci.*, vol. 62, pp. 427–440, Feb. 2014, doi: [10.1016/j.ssci.2013.10.002](https://doi.org/10.1016/j.ssci.2013.10.002).
- [40] R. A. Pomranky and J. Q. Wojciechowski, "Determination of mental workload during operation of multiple unmanned systems," Army Res. Lab., Aberdeen Proving Ground, MD, USA, Tech. Rep. ARL-TR-4309, 2007.

- [41] E. Hollnagel, *Cognitive Reliability and Error Analysis Method (CREAM)*. New York, NY, USA: Elsevier, 1998.
- [42] D. Chen, Y. Fan, W. Li, Y. Wang, and S. Zhang, "Human reliability prediction in deep-sea sampling process of the manned submersible," *Saf. Sci.*, vol. 112, pp. 1–8, Feb. 2019, doi: [10.1016/j.ssci.2018.10.001](https://doi.org/10.1016/j.ssci.2018.10.001).
- [43] L. Wang, Y. Wang, Y. Chen, X. Pan, and W. Zhang, "Performance shaping factors dependence assessment through moderating and mediating effect analysis," *Rel. Eng. Syst. Saf.*, vol. 202, Oct. 2020, Art. no. 107034, doi: [10.1016/j.ress.2020.107034](https://doi.org/10.1016/j.ress.2020.107034).
- [44] N. Xia, P. X. W. Zou, X. Liu, X. Wang, and R. Zhu, "A hybrid BN-HFACS model for predicting safety performance in construction projects," *Saf. Sci.*, vol. 101, pp. 332–343, Jan. 2018, doi: [10.1016/j.ssci.2017.09.025](https://doi.org/10.1016/j.ssci.2017.09.025).
- [45] J. Wu, Y. Bai, H. Zhao, X. Hu, and V. Cozzani, "A quantitative LNG risk assessment model based on integrated Bayesian-catastrophe-EPE method," *Saf. Sci.*, vol. 137, May 2021, Art. no. 105184, doi: [10.1016/j.ssci.2021.105184](https://doi.org/10.1016/j.ssci.2021.105184).
- [46] K. J. Burns, C. A. Bonaceto, S. L. Estes, and J. R. Helleburg, "Quantifying the safety benefits of a digital copilot in general aviation," *J. Aerosp. Inf. Syst.*, vol. 16, no. 8, pp. 296–306, Aug. 2019, doi: [10.2514/1.1010704](https://doi.org/10.2514/1.1010704).
- [47] H. Che, S. Zeng, K. Li, and J. Guo, "Reliability analysis of load-sharing man-machine systems subject to machine degradation, human errors, and random shocks," *Rel. Eng. Syst. Saf.*, vol. 226, Oct. 2022, Art. no. 108679, doi: [10.1016/j.ress.2022.108679](https://doi.org/10.1016/j.ress.2022.108679).
- [48] *GeNIe 2.4*, BayesFusion, Pittsburgh, PA, USA, 2019.
- [49] M. V. Rosenker, R. L. Sumwalt, D. A. P. Hersman, K. O. Higgins, and S. R. Chealander, "Accident report brief," Nat. Transp. Saf. Board, Washington, DC, USA, Tech. Rep. NYC05MA039, 2007.
- [50] H. Che, S. Zeng, Q. You, Y. Song, and J. Guo, "A fault tree-based approach for aviation risk analysis considering mental workload overload," *Eksplatacja Niezawodność, Maintenance Rel.*, vol. 23, no. 4, pp. 646–658, Dec. 2021, doi: [10.17531/ein.2021.4.7](https://doi.org/10.17531/ein.2021.4.7).



XIN LU received the B.S. degree from the School of Aeronautic Science and Engineering, Beihang University, in 2015, and the M.S. degree from the School of Reliability and Systems Engineering, Beihang University, in 2018, where he is currently pursuing the Ph.D. degree. His research interests include reliability modeling, man-machine system safety analysis, and human reliability.



JIANBIN GUO received the Ph.D. degree in systems engineering from Beihang University, Beijing, China, in 2009. He is currently a Senior Lecturer with the School of Reliability and Systems Engineering, Beihang University. His research interests include reliability-based multidisciplinary design optimization, man-machine system safety analysis, and human reliability.



SHENKUI ZENG received the Ph.D. degree in systems engineering from Beihang University, Beijing, China, in 2009. He is currently a Professor and the Assistant Dean with the School of Reliability and Systems Engineering, Beihang University. His research interests include reliability modeling and optimization, man-machine system safety analysis, and human reliability.



HAIYANG CHE received the B.S. degree in industrial engineering from Jilin University, in 2014, and the Ph.D. degree from the School of Reliability and Systems Engineering, Beihang University, in 2020. He is currently a Postdoctoral Fellow with the School of Automation Science and Electrical Engineering, Beihang University. His research interests include reliability modeling, man-machine system safety analysis, and human reliability.

...