Research on Quantitative Evaluation Model of Human Error Probability of High-speed Railway Train Dispatchers

Yanhao Sun[*1,2], Tao Zhang[1,2], Shuxin Ding[1,2], Zhi Li[1,2], Ang Li[1]
[1] Signal & communication research institute, China Academy of Railway Sciences Co. Ltd, Beijing, China.
[2] The Center of National Railway Intelligent Transportation System Engineering and Technology, China Academy of Railway Sciences Corporation Limited, Beijing China.
sunyanhao@163.com*

ABSTRACT

In order to be able to accurately assess the Human Error Probability (HEP) of high-speed rail train dispatchers. A HEP quantitative evaluation model based on Improved Weighted Cognitive Reliability and Error Analysis (CREAM) was constructed, which corrected the shortcomings of traditional CREMA. Using bipolar 2-tuples as the Common Performance Condition (CPC) evaluation linguistics, the subjective and objective weights of CPC are calculated through the Analytic Hierarchy Process and the Criteria Importance Though Intercriteria Correlation (CRITIC), and then the combined weighting method is used to further obtain the comprehensive weight of CPC ; At the same time, the idea of group decision-making is used to reduce the subjectivity of CPC evaluation, and the weight of experts is calculated by using the adaptive dynamic weight adjustment method of grey correlation degree; on the basis of obtaining CPC and expert weights, the evaluation value of CPC is calculated. The weighted operation then obtains the Context Influence Index (CII) value; finally, the quantitative calculation of HEP is obtained by using the CII value. The model is validated by taking the high-speed train dispatcher's normal and abnormal train reception as examples. The research results show that: under normal conditions, the HEP of high-speed rail train dispatchers is 9.1586×10^{-5}, and under abnormal conditions, the HEP of high-speed rail train dispatchers is 2.1189×10^{-3}. Under abnormal circumstances, the human error probability of high-speed train dispatchers is more than 20 times that under normal circumstances. At the same time, compared with other methods, the model has certain advantages in terms of weight sensitivity, data utilization sufficiency, and model solution domain. The improved CREAM model can better reflect the impact of different CPCs on high-speed rail train dispatchers, and can provide a calculation method for the HEP quantification of high-speed rail train dispatchers.

Keywords- High-speed railway train dispatcher; Human error probability; Quantitative evaluation; Cognitive reliability and error analysis method; Bipolar 2-tuples; Common performance condition; Context influence index

1. INTRODUCTION

As the nerve center of the railway transportation system, the high-speed train dispatching system plays an important role in ensuring the safety and punctuality of trains. The high-speed rail train dispatching system is a typical "man-machine" complex system, and the train dispatcher, as a system element with independent thinking, plays an important role in coordinating and controlling the system. With the rapid advancement of science and technology, the reliability of system equipment has been greatly improved, but the human reliability of train dispatchers has become a bottleneck restricting system safety. At present, the research on human error in railway system, especially railway dispatching system, is in its infancy at home and abroad, but it has also achieved certain results. Therefore, it is of great significance to carry out quantitative research on the human reliability of high-speed railway train dispatchers.

At present, the research on human error in railway system especially in railway dispatching system is in the initial stage at home and abroad, but it has also achieved certain results. WANG et al.[1] improved the network analysis method by using triangular fuzzy numbers, and evaluated the dispatcher's reliability by combining the Human Error Assessment Reduction Technique (HEART). Xu et al.[2] combined the Markov principle with Technique for Human Error Rate Prediction to obtain the Human Error Probability (HEP) value of the train dispatcher and the dynamic reliability change law of human factors under the influence of pressure factors. Wu[3] took individual, organization, equipment and environment as the main factors affecting the human factor reliability of train dispatchers, and combined with Bayesian network to quantitatively calculate the human factor reliability of the subway dispatching system. CIANI et al.[4] used fuzzy logic to improve the train driver reliability evaluation method, and realized the quantitative output of the train driver's HEP value.

Seventh International Conference on Traffic Engineering and Transportation System (ICTETS 2023), edited by Ali Reza Ghanizadeh, Hongfei Jia, Proc. of SPIE Vol. 13064, 130642B

ZHOU et al.[5] proposed a hybrid HEART to evaluate the HEP during locomotive driving, and verified this hybrid HEART with Monte Carlo simulation. JANOTA et al.[6] constructed a set pair analysis method and Markov chain human factor evaluation model, which was used to analyze the human factors of railway transportation system operators. DINDAR et al.[7] used fuzzy Bayesian network to model and analyze the derailment factors caused by human errors. Li et al.[8] researched and analyzed the influencing factors of human errors of railway maintenance personnel through Decision-making Trial and Evaluation Laboratory and Interpretative Structural Modeling Method. Zhang[9] analyzed and explained the human errors of subway train drivers from the perspective of personality traits.

The above research results have great reference for the quantitative evaluation of human reliability of high-speed train dispatchers, but there are still some deficiencies. Fewer, and most of the methods are limited to the first generation of human error analysis methods. This type of method compares the human error mechanism to the hardware failure or failure mechanism, and uses the decomposition technology similar to the hardware failure analysis to analyze the human error, so there are many defects[10]. With the development of cognitive science, human factor reliability technology has developed rapidly. CREAM as a new generation of human error analysis method, is widely used in fields requiring high reliability[11-14]. Moreover, the CREAM method has strong versatility and operability, and is easy to be transplanted to the research on human error of high-speed railway train dispatchers. In order to enable CREAM to calculate the HEP value more accurately, this paper improves it, constructs an improved weighted CREAM quantitative evaluation model of human error probability, and uses this model to quantitatively evaluate the HEP of high-speed rail train dispatchers.

2. BASIC THEORY

2.1. CREAM

CREAM is divided into two versions, one is the basic version and the other is the extended version. But in 2012, HOLLNAGEL, the founder of CREAM, issued a disclaimer for the extended version, pointing out that the extended version has some inevitable defects. Therefore, this article only improves the basic version of CREAM, and the default version of CREAM in this article is the basic version. CREAM believes that human performance is the outcome of the purposive use of competence adjusted to specific working conditions rather than of the pre-determined sequence of response to given events. CREAM therefore defines 4 characteristic control modes according to the human cognition and action context, which are determined by 9 CPC[15]. The names and HEP intervals of the 4 control modes are shown in Table 1.

Table 1. HEP intervals under different control modes

Control modes	Probability interval
Strategic	$(0.5\times10^{-4}，0.1\times10^{-1})$
Tactical	$(0.1\times10^{-2}，0.1)$
Tactical	$(0.1\times10^{-1}，0.5)$
Scrambled	(0.1，1.0)

CREAM identifies 9 CPCs, which are shown in Table 2. Each CPC is divided into different levels, and its influence on people can be divided into 3 effects: positive "improved", neutral "insignificant" and negative "reduced".

Table 2. 9 CPCs

CPC number	CPC name
C_1	Adequacy of organization
C_2	Working conditions
C_3	Adequacy of MMI and operational support
C_4	Availability of procedures/plans
C_5	Number of simultaneous goals
C_6	Available time
C_7	Time of the day (circadian rhythm)
C_8	Adequacy of training and experience
C_9	Crew collaboration quality

By calculating the "improved" and "reduced" CPC numbers, use ($\Sigma_{improved}$, $\Sigma_{reduced}$) coordinates to determine the control mode, and get the value of HEP according to Table 1. The coordinate diagram of the control mode is shown in Figure 1.

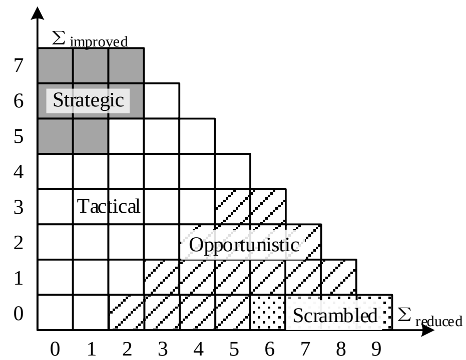

Figure 1. Determination of control modes.

2.2. Bipolar 2-tuples

For some information that is difficult to express with real numbers, it is generally represented by fuzzy sets or fuzzy numbers, but the former will lose information in the process of information aggregation, so Herrera proposed a binary semantic representation model to deal with fuzzy information[16]. On this basis, Yang extended 2-tuples to bipolar binary 2-tuples[17].

Definition 1: Let S=($s_{-g},\ldots,s_{-1},s_0,s_1,\ldots,s_g$) be a bipolar linguistic term set with odd cardinality, and β($\beta\in$[-g, g]) be the result of a symbolic aggregation operation. i = round(β), ("round" is the common rounding operation), $\alpha_i = \beta - i$, make $i\in$[-g,…,-1,0,1,…,g], $\alpha_i \in$[-0.5, 0.5), and α_i is the symbolic translation value s_i.

Definition 2: Suppose S=($s_{-g},\ldots,s_{-1},s_0,s_1,\ldots,s_g$) is a bipolar linguistic term set, then there is a function Δ that converts β into bipolar 2-tuples.

$$\begin{aligned}&\Delta[-g,g]\to S\times[-0.5,0.5)\\&\beta\to(s_i,\alpha_i)=\begin{cases}s_i & i=round(\beta)\\ \alpha_i & \alpha_i\in[-0.5,0.5)\end{cases}\end{aligned} \tag{1}$$

It can be seen that for any real number on [-g, g], there is a bipolar 2-tuples (s_i, α_i) corresponding to it. At the same time, there is an inverse function Δ^{-1} to convert the bipolar 2-tuples into the corresponding value $\beta\in$[-g, g]:

$$\begin{aligned}&\Delta^{-1}:S\times[-0.5,0.5)\to[-g,g]\\&\Delta^{-1}(s_i,\alpha_i)=i+\alpha_i=\beta\end{aligned} \tag{2}$$

Definition 3: Suppose S=($s_{-g},\ldots,s_{-1},s_0,s_1,\ldots,s_g$) is a bipolar linguistic term set, (s_i, α_i) is a bipolar 2-tuples, then there is a function φ that converts bipolar 2-tuples translate to an interval of [-1,1]..

$$\varphi(s_i,\alpha_i)\to\varphi:(\beta)=(\beta-i)(c_i-b_i)+b_i \tag{3}$$

Where, (a_i, b_i, c_i) is the triangular fuzzy value of s_i. For the convenience of calculation, Yang gave the triangular fuzzy value of s_i when g is 3[17].

Definition 4: Suppose S=($s_{-g},\ldots,s_{-1},s_0,s_1,\ldots,s_g$) is a bipolar linguistic term set, ω_i is the weight and $w_i\in$[0, 1], $\sum_{i=1}^{g} w_i = 1$., then arithmetic weighted average operator $\Phi(S)$ of bipolar 2-tuples is defined as follows.

$$\Phi(S)=\Delta\left(\sum_{i=1}^{n}\omega_i\Delta^{-1}\left(s_i,\alpha_i\right)\right) \tag{4}$$

3. IMPROVED CREAM MODEL

Although the CREAM method is efficient and simple, it has several obvious deficiencies:

① The evaluation of CPC has great ambiguity and uncertainty, and it is difficult to use real numbers to represent it; ②The weight of CPC is not considered, and different CPCs have different effects on human error in reality; ③The way to determine the control mode is discretized, and the output HEP is an interval value. Therefore, in view of the above deficiencies, it is necessary to improve CREAM accordingly. For the convenience of description, the following assumptions are made.

Assume that an expert group is composed of k（$k = 1,2,\ldots,l$）experts. The weight of expert D_k is λ_k， where $\lambda_k\in[0, 1]$, and $\sum_{k=1}^{l}\lambda_k=1$. Let C_j ($j = 1,2,\ldots,n$) be CPC and w_j be its weight, where $w_i\in[0, 1]$, and $\sum_{n=1}^{j}w_j=1$. The expert group evaluates m operation tasks A_i ($i = 1, 2,..., m$). The evaluation matrix is $\boldsymbol{X}^k=(x_{ij}^k)_{m\times n}$, where x_{ij}^k is the bipolar 2-tuples evaluation value of the k-th expert on the j-th CPC under the i-th task.

3.1. Calculation of CPC weight

When calculating the CPC weight, not only its subjective weight but also its objective weight should be considered. In order to make the calculation of CPC weight more accurate, the combined weighting method is used to calculate the weight of CPC. For the subjective weight, the most common AHP is used for calculation. Considering the length of the paper, I will not repeat it here. For the objective weight, due to the certain correlation between the CPCs, the general objective weighting method will have a certain deviation when calculating the weight. In order to make the weight calculation more accurate, CRITIC is used to obtain the objective weight of the CPC. However, there is a problem in CRITIC that the correlation between indicators has nothing to do with positive or negative. If the correlation coefficient is negative, the results obtained cannot faithfully reflect the conflict between indicators, and the smaller the correlation coefficient, the greater the conflict between indicators. Therefore, the dimensionless Gini coefficient is introduced to represent the difference of indicators to realize the effective calculation of CPC weight.

Step 1: The evaluation matrices of all experts are weighted and assembled to obtain the comprehensive evaluation matrix ***X***.

$$\boldsymbol{X}=(\lambda_k x_{ij}^k)_{m\times n}=(x_{ij})_{m\times n} \tag{5}$$

Step 2: Standardize the comprehensive evaluation matrix to obtain the standard evaluation matrix ***Y***. The elements yij in the matrix are:

$$y_{ij}=\frac{\max\limits_{j}\left(\Delta^{-1}(x_{ij})\right)-\Delta^{-1}(x_{ij})}{\max\limits_{j}\left(\Delta^{-1}(x_{ij})\right)-\min\limits_{j}\left(\Delta^{-1}(x_{ij})\right)} \tag{6}$$

Step 3: Obtain the correlation coefficient matrix ***Z*** according to the Pearson correlation coefficient. The element z_{jt} in the matrix is

$$z_{jt}=\frac{\sum_{i=1}^{m}(y_{ij}-\overline{y}_j)(y_{it}-\overline{y}_t)}{\sqrt{\sum_{i=1}^{m}(y_{ij}-\overline{y}_j)^2}\sqrt{\sum_{i=1}^{m}(y_{it}-\overline{y}_t)^2}} \tag{7}$$

Where, z_{ij} and z_{it} respectively represent the standardized values of the j-th CPC and the t-th CPC under the i-th task. And are the mean values.

Step 4: Obtain the Gini coefficient. Calculated as follows:

$$\xi_j = \frac{\sum_{i=1}^{m}\sum_{t=1}^{n}\left|z_{ij} - z_{it}\right|}{2n\sum_{j=1}^{n} z_{ij}} \tag{8}$$

Where, $\xi_j \in [0,1]$, from 0 to 1 means that the contrast strength of the index is the largest, and 0 means that the contrast strength of the index is the smallest. Finally, the combined weighting method is used to calculate the combined weight of CPC, and the objective function is constructed by referring to the principle of minimum identification information[18].

$$\begin{cases} \min J(w) = \sum_{j=1}^{n}\left(w_j \left(\ln \frac{w_j}{w_{oj}} + \ln \frac{w_j}{w_{sj}} \right)\right) \\ s.t. \qquad \sum_{j=1}^{n} w_j = 1, w_j \geq 0 \qquad j = 1, 2, ..., n \end{cases} \tag{9}$$

Where, w_{sj} is the subjective weight, and w_{oj} is the objective weight. Solving the objective function, the comprehensive weight of CPC is obtained as:

$$w_j = \frac{\sqrt{w_{sj} w_{oj}}}{\sum_{j=1}^{n}\sqrt{w_{sj} w_{oj}}} \tag{10}$$

3.2. Calculation of expert weight

Existing research on CREAM basically uses the evaluation of a single expert. Such evaluation information is highly subjective, which reduces the credibility of the calculation results. In order to further reduce the subjectivity of evaluation information, the method of group decision-making is used to evaluate the information of CPC. Group decision-making reflects a process in which experts collectively negotiate and finally reach a consensus. It reflects the compromise of expert opinions, and the final result should tend to be consistent. Therefore, the weight of experts is generally dynamic. By assigning smaller weights to experts with large differences in evaluation information for weight adaptive adjustment, based on the idea of gray relational degree, an adaptive and dynamic weight adjustment method based on gray relational degree is proposed.

Step1: Determine the subjective weight of experts. According to the expert's experience and knowledge, the expert's subjective weight λ_{sk} is given as the initial weight.

Step2: Determine the reference sequence and comparison sequence. The sequence after the weighted aggregation of the CPC evaluation value of the operation task by a single expert Dk is used as the comparison sequence $\overline{\boldsymbol{X}}^k = \left(\overline{x}_i^k\right)_{1\times m}$, and the sequence $\overline{\boldsymbol{X}} = \left(\overline{x}_i\right)_{1\times m}$ after the comparison sequence is assembled according to the expert's initial weight is used as the reference sequence. in:

$$\begin{cases} \overline{x}_i^k = \sum_{j=1}^{n} w_j x_{ij}^k \\ \overline{x}_i = \sum_{k=1}^{l} \lambda_{sk} \overline{x}_i^k \end{cases} \tag{11}$$

Step3: Calculate the correlation coefficient between the comparison sequence and the reference sequence.

$$\eta_{ki} = \frac{\min_k \min_i \left|\Delta^{-1}(\overline{x}_i) - \Delta^{-1}(\overline{x}_i^k)\right| + \mu \max_k \max_i \left|\Delta^{-1}(\overline{x}_i) - \Delta^{-1}(\overline{x}_i^k)\right|}{\left|\Delta^{-1}(\overline{x}_i) - \Delta^{-1}(\overline{x}_i^k)\right| + \mu \max_k \max_i \left|\Delta^{-1}(\overline{x}_i) - \Delta^{-1}(\overline{x}_i^k)\right|} \tag{12}$$

Among them, μ is the resolution coefficient, $\mu\in$ [0,1], without loss of generality, the value of μ is 0.5[19].

Step4: Calculate the correlation between the reference sequence and the comparison sequence.

$$R_{0k} = \frac{1}{m}\sum_{i=1}^{m}\eta_{ki} \tag{13}$$

Step5: The weight is dynamically adjusted. In order to prevent too much pursuit of the consensus of opinions and ignore the influence of experts on the results, the weights of experts are revised, and the subjective weights are adjusted as the initial weights. Initially there are:

$$\lambda_k = \frac{\lambda_{sk} R_{0k}}{\sum_{k=1}^{l}\lambda_{sk} R_{0k}} \tag{14}$$

When adjusting again, replace λ_{sk} with λ_k as the new weight, and iterate in turn.

Step6: Determine the final weight. Bring the iterative weight into Equation (11) to get a new aggregation value, and define the difference between and as:

$$d(\lambda)=\sqrt{\sum_{i=1}^{m}\left(\Delta^{-1}(\overline{x}_i') - \Delta^{-1}(\overline{x}_i)\right)^2} \tag{15}$$

Under the condition of not affecting the accuracy, in order to simplify the calculation, when $d(\lambda)\leq 0.00005$, it is considered that the evaluation results tend to be stable and reach a certain consensus. The weight adjustment is over, and the λ_k at this time is the weight of the final expert

3.3. Calculation of HEP

There are 3 proven hypotheses for the discrete problems in the CREAM method[20],[21].

① Both control mode and HEP are continuous;

② HEP changes exponentially with the change of CPC;

③ When (Σ_{improved}, Σ_{reduced})=(7, 0), it means that the reliability of completing the task is the highest, and HEP has achieved the minimum value at this time. On the contrary, if (Σ_{improved}, Σ_{reduced})=(0, 9), it means completion The reliability of the task is the lowest, and HEP reaches the maximum value at this time. If (Σ_{improved}, Σ_{reduced})=(0, 0), it means that the comprehensive impact of CPC on people is 0, that is, there is no impact on the environment, and the failure probability at this time is the basic failure probability.

According to the above assumptions, Sun et al.[21] used CII to calculate HEP. And define CII as:

$$\text{CII} = \frac{\Sigma_{\text{improved}}}{7} - \frac{\Sigma_{\text{reduced}}}{9} \tag{16}$$

It can be seen that the value range of CII is within [-1,1], which coincides with the value range of bipolar 2-tuples.

Establish the functional relationship between HEP and CII, the relationship is as follows:

$$\text{HEP} = \text{HEP}_0 \exp(a \times \text{CII}) \tag{17}$$

Where, HEP0 is the basic failure probability, and a is the coefficient.

According to assumption ③, it can be known that:

$$\begin{cases} \text{HEP}_{\max} = \text{HEP}_0 \exp(-a) \\ \text{HEP}_{\min} = \text{HEP}_0 \exp(a) \end{cases} \tag{18}$$

HEP_{max}=1, HEP_{min}=0.00005 as shown in Table 1, the above equation is solved, and then the calculation formula of human error probability HEP is obtained:

$$\mathrm{HEP} = 7.07 \times 10^{-3} \exp(-4.9517 \times \mathrm{CII}) \tag{19}$$

Equation (16) still does not consider the weight of CPC when calculating CII, but averages the evaluation of CPC. Therefore, referring to the ideas of SUN[23], CII is defined here as:

$$\mathrm{CII}_i = \sum_{k=1}^{l}\sum_{j=1}^{n} \varphi(x_{ij}^k) \tag{20}$$

Where, x_{ij}^k is the bipolar binary 2-tuples evaluation value of the *k*-th expert on the *j*-th CPC under the *i*-th task. $\varphi(x_{ij}^k)$ is the conversion function of bipolar binary semantics x_{ij}^k on [-1,1], and CII_i represents the CII value of the *i*-th task.

4. CASE STUDY

4.1. Calculation of HEP

The operation of reception and departure trains at the station refers to the whole process that the station handles the whole process of connecting trains from the section, sending out to the section and passing through the operation according to the traffic block mode and technical equipment conditions, in accordance with the relevant procedures. It is an indispensable and important link in the operation of high-speed railways. In this paper, the operation of the high-speed train dispatcher under normal conditions is taken as task 1, and the operation of reception trains under abnormal conditions (The turnout that receives the train route loses the indication) is taken as task 2. The work flow of the two is shown in Table 3.

Table 3. Receive the train into the station operation process

Task1(normal conditions)	Task2(abnormal conditions)
Receive reception train notice	It is found that the turnout of the receiving train route has lost its indication, and the follow-up train leading to the faulty turnout is notified to stop outside the signal machine.
ready to receive train	Stop using the faulty turnout, and lock the position required for opening the turnout on the access route separately
Confirm to receive the train route(Route free)	Notify the deputy director on duty, and notify the track and signal personnel to carry out the up-channel inspection
open signal	The line where the faulty turnout is located is blocked, and the speed limit of the adjacent line is 80km/h. The dispatch order notifies the emergency personnel on duty of the vehicle, and agrees that the track and signal personnel will go to the road to check and deal with the fault.
train approaching	Inform the train emergency crew on duty to manually prepare for the route
receiving train	After confirming that the route is ready, issue a dispatch command to switch to the isolation mode and allow parking at the station
Train arrival (report arrival time)	Issue a dispatch command to cancel the blockade of the main line and limit the speed of adjacent lines to 80km/h

In order to better evaluate the human factor reliability of high-speed rail train dispatchers. A railway bureau invited 4 experts D_1, D_2, D_3 and D_4 from the safety supervision and traffic departments to form an expert group to evaluate the operation of the train dispatcher. It is known that the subjective weights of the four experts are 0.4, 0.3, 0.2 and 0.1 respectively. The expert group adopted the following bipolar language term set S={s_{-3}: very poor, s_{-2}: poor, s_{-1}: a little poor, s_0: medium, s_1: a little good, s_2:good, s_3: very good }. The evaluation information of Task 1 and Task 2 are shown in Table 4 and Table 5, respectively.

Table 4. CPC evaluation information of task 1

CPC	D_1	D_2	D_3	D_4
C_1	$(s_2,0)$	$(s_3,0)$	$(s_2,0.2)$	$(s_3,0)$
C_2	$(s_3,0)$	$(s_3,0)$	$(s_2,0)$	$(s_3,0)$
C_3	$(s_3,-0.3)$	$(s_3,0)$	$(s_3,0)$	$(s_2,0)$
C_4	$(s_2,0)$	$(s_2,0)$	$(s_2,0.3)$	$(s_2,0.1)$
C_5	$(s_3,0)$	$(s_3,0)$	$(s_3,0)$	$(s_2,0.2)$
C_6	$(s_3,0)$	$(s_3,0)$	$(s_3,0)$	$(s_3,0)$
C_7	$(s_2,0)$	$(s_2,0)$	$(s_2,-0.2)$	$(s_2,-0.1)$
C_8	$(s_3,0)$	$(s_3,0)$	$(s_3,0)$	$(s_3,0)$
C_9	$(s_3,0)$	$(s_3,-0.3)$	$(s_2,0)$	$(s_2,0)$

Table 5. CPC evaluation information of task 2

CPC	D_1	D_2	D_3	D_4
C_1	$(s_2,0)$	$(s_2,0)$	$(s_2,0)$	$(s_2,0)$
C_2	$(s_3,0)$	$(s_2,0)$	$(s_2,0)$	$(s_3,0)$
C_3	$(s_3,0)$	$(s_3,0)$	$(s_1,0)$	$(s_1,0)$
C_4	$(s_{-2},0)$	$(s_{-2},-0.2)$	$(s_{-2},0)$	$(s_{-2},0)$
C_5	$(s_{-3},0)$	$(s_{-3},0)$	$(s_{-3},0)$	$(s_{-3},0.2)$
C_6	$(s_{-2},0)$	$(s_{-2},0)$	$(s_{-3},0)$	$(s_{-3},0)$
C_7	$(s_2,0)$	$(s_2,0)$	$(s_2,-0.2)$	$(s_3,0)$
C_8	$(s_2,0)$	$(s_3,0)$	$(s_3,0)$	$(s_3,0)$
C_9	$(s_3,-0.2)$	$(s_2,0)$	$(s_2,0.4)$	$(s_2,0)$

First, calculate the weight of CPC, and calculate the subjective weight, objective weight and combined weight of CPC according to Equation (5)- Equation (10). See Table 6.

Table 6. The weights of CPC

Weight category	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9
Subjective weight	0.0896	0.1027	0.1281	0.065	0.1468	0.1331	0.1092	0.1592	0.0663
Objective weight	0.1361	0.0792	0.0722	0.0586	0.1142	0.1237	0.1095	0.1493	0.1572
Combined weight	0.1125	0.0918	0.0979	0.0629	0.1319	0.1307	0.1114	0.1570	0.1040

Secondly, calculate the expert weights, and dynamically adjust the expert weights according to Equation (11)~ Equation (15) until the difference degree condition is met, and the weights of the four experts are obtained, as shown in Table 7.

Table 7. Dynamic adjustment of expert weights

Number of adjustments	0	1	3	5	6
D_1	0.4	0.3354	0.3263	0.3198	0.3193
D_2	0.3	0.2756	0.2721	0.2733	0.2634
D_3	0.2	0.1996	0.2106	0.2111	0.2204
D_4	0.1	0.1894	0.1910	0.1958	0.1969
Difference degree	0.00286	0.00143	0.00097	0.00068	0.00042

Finally, use Equation (20) to weight the evaluation value of CPC to obtain the CII values of the two tasks, and calculate the HEP value according to Equation (19). The calculation results are shown in Table 8.

Table 8. HEP value

Task	CII	HEP	Reliability
Task1	0.8766	9.1586×10^{-5}	0.999909
Task2	0.2422	2.1189×10^{-3}	0.997882

Comparing the HEP values of the human error probability between the two, it can be seen that the error probability of the high-speed train dispatcher's pick-up operation under extraordinary circumstances is 23 times higher than that under normal conditions. The on-site survey data found that the frequency of high-speed train dispatchers' mistakes under abnormal conditions is about 20 times that under normal conditions, which also verifies the effectiveness of the model from the side.

4.2. Weight Sensitivity Analysis

(1) CPC weight sensitivity analysis

Given the evaluation values of three groups of the same CPC, see Table 9. Compared with Case0, exchange the weights of C_4 and C_8 in Case1, and exchange the weights of C_4 and C_8 in Case2

Table 9. 3 sets of CPCs weights and HEP values

CPC	Evaluation value	Weight		
		Case0	Case1	Case2
C_1	$(s_0,0)$	0.1125	0.1125	0.1125
C_2	$(s_0,0)$	0.0918	0.0918	0.0918
C_3	$(s_0,0)$	0.0979	0.0979	0.0979
C_4	$(s_1,0)$	0.0629	0.1570	0.0629
C_5	$(s_1,0)$	0.1319	0.1319	0.1570
C_6	$(s_2,0)$	0.1307	0.1307	0.1307
C_7	$(s_2,0)$	0.1114	0.1114	0.1114
C_8	$(s_3,0)$	0.1570	0.0629	0.1319
C_9	$(s_2,0)$	0.1040	0.1040	0.1040
HEP		7.473×10^{-4}	1.019×10^{-3}	8.119×10^{-4}

It can be seen that after exchanging the weights of C_4 and C_8, C_5 and C_8, the HEP values of both Case1 and Case2 increase, but the change range of the former is significantly larger than that of the latter. Although the evaluation values of C_4 and C_5 are the same, the value of C_4 is higher than that of C_5 the greater the weight change, the more pronounced the impact on the HEP. Therefore, when calculating the HEP value, the weight of CPC must be considered.

(2) Expert weight analysis

Taking Task 1 as an example, Case0 is the final weight of the four experts, and Case1 and Case2 are the expert weights for hypothetical comparison. The comparison results are shown in Table 10.

Table 10. 3 sets of experts weights and HEP values

Case	D_1	D_2	D_3	D_4	$d(\lambda)$	HEP
Case0	0.3193	0.2634	0.2204	0.1969	0.00042	9.1586×10^{-5}
Case1	0.25	0.25	0.25	0.25	0.00121	8.5962×10^{-5}
Case2	0.1	0.2	0.3	0.4	0.00538	7.0477×10^{-4}

The expert difference threshold $d(\lambda)$ is 0.0005. It can be seen that the greater the difference between experts, the greater the impact on the HEP value. The lower the degree of expert consensus, the more serious the impact on the weight of experts, and the greater the gap with the HEP value of Case0, mainly because whether the experts reached a consensus was not considered in the process of information gathering, and there is a large difference in the evaluation information of the expert group Sometimes, there will be discrepancies between the assembled comprehensive evaluation information and the actual situation. Therefore, it is reasonable and feasible to adaptively adjust the weight of experts to reach a consensus.

4.3. Sufficiency of Data Utilization

The full use of observation data is an important way to ensure the accuracy of HEP, and the loss of data mainly occurs in the process of transforming effect value of membership degree of CPC observation value or discrimination of control mode. Take literature [22] as an example. Its membership function image is shown in Figure 2 (some literature, such as [20] uses Gaussian function, the situation is similar to it), and it can be seen in Figure 2 that when the factor score is from 0 to 10, 40 to 60, 90 to 100 When it changes, its influence effect does not change, and the control mode does not change. This process leads to information loss, so that the HEP value does not change.

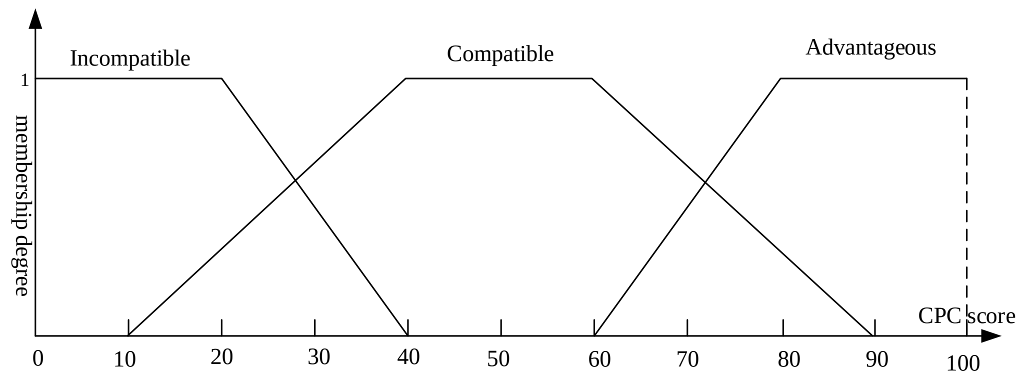

Figure 2. CPC membership function image of literature [22]

However, this paper uses bipolar 2-tuples to describe observations. Each linguistic granularity corresponds to a different linguistic value. These linguistic values are weighted and assembled to make the output of HEP continuous. Furthermore, bipolar 2-tuples can represent The degree of preference of experts, for example, assuming that the evaluation of C_4 by the expert group is (s_2,0.1), it means that the experts evaluate the factor as "a little better than good", and 0.1 means that the degree of deviation is even better. At the same time, compared with individual decision-making, group decision-making can use observation data more effectively, reduce data loss caused by individual single evaluation, improve prediction accuracy, and avoid misjudgment.

4.4. Model Solving Domain Analysis

In this paper, the model output interval [0.5×10^{-4},1] completely covers the entire solution space. However, the HEP output interval of literature [22] is [2.238×10^{-4},1], and the left side of the value range is too small, resulting in a decrease in the prediction accuracy. The HEP output interval of literature [20] is [6.28×10^{-5},6.28], resulting in the overflow of the HEP calculation value on the right, so when the HEP value is greater than 1, only the method of forced normalization can be used. The CPC factor evaluations are all "not significant" or "moderate", that is, when they are neutral, the calculation result of this model is $7.07\times10^{-3,}$ which is consistent with the HEP value obtained in literature [23] of 7.04×10^{-3} and literature [24] The weighting method used to get the HEP value 7.08×10-3 is very close.

5. CONCLUSION

1) Bipolar 2-tuples may be a good representation of the ambiguity and uncertainty of CPC, which can realize the transformation of CPC from qualitative information to quantitative representation.

2) Whether it is the weight of CPC or the weight of experts will have an important impact on the quantitative calculation of HEP, so the calculation of weight must be fully considered.

3) The improved CREAM model is superior to the basic version of CREAM in terms of weight sensitivity, data utilization sufficiency and model solution domain.

4) Due to the lack of reliable historical data, CREAM research relies on the domain knowledge and experience of analysts and experts in related fields. Therefore, how to reduce the subjectivity of experts is still the key direction of this year's research.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China under Grant (U1834211) and Research Fund of China Academy of Railway Sciences corporation limited (2021YJ097).

REFERENCES

[1] Wang Weizhong, Liu Xinwang, Qin Yong. A modified HEART method with FANP for human error assessment in high-speed railway dispatching tasks[J]. International Journal of Industrial Ergonomics, 2018, 67: 242-258.
[2] XU Peijuan, PENG Qiyuan, WEN Chao, et al. Human reliability analysis on high-speed train dispatcher based on THERP and markov theories [J]. Journal of Transportation Systems Engineering and Information Technology, 2014, 14(6): 133-140.
[3] WU Dan. Human reliability analysis of subway train dispatching system based on the bayesian networks [D]. Chengdu: Southwest Jiao tong University, 2018.
[4] CIANI L, GUIDI G, PATRIZI G, et al. Improving human reliability analysis for railway systems using fuzzy logic [J]. IEEE Access, 2021, 9: 128648-128662.
[5] Zhou Jianlan, Lei Yi, Chen Yang. A hybrid HEART method to estimate human error probabilities in locomotive driving process [J]. Reliability Engineering & System Safety, 2019, 188: 80-89.
[6] JANOTA A, PIRNÍK R, ŽDÁNSKY J, et al. Human factor analysis of the railway traffic operators [J]. Machines, 2022, 10(9): 820.
[7] DINDAR S, KAEWUNRUEN S, AN M. Bayesian network-based human error reliability assessment of derailments[J]. Reliability Engineering & System Safety, 2020, 197: 106825.
[8] LI Xiang, LI Xiao, WANG Song, et al. Study on factors leading to human errors in railway maintenance [J]. China Safety Science Journal,2022,32(06):23-30.
[9] Zhang Hanbo. Study on human error of metro drivers based on personality traits [D]. Kunming: Kunming University of Science and Technology,2022.
[10] LI Pengcheng , CHEN Guohua , ZHANG Li, et al. Research review and development trends of human reliability analysis techniques[J]. Atomic Energy Science & Technology, 2011, 45(3):329-340.
[11] AHN S I, KURT R E. Application of a CREAM based framework to assess human reliability in emergency response to engine room fires on ships [J]. Ocean Engineering, 2020, 216: 108078.
[12] Chen Denkai, Fan Y,u Ye Cong, et al. Human reliability analysis for manned submersible diving process based on CREAM and Bayesian network[J]. Quality and Reliability Engineering International, 2019, 35(7): 2261-2277.
[13] Wang Lijing, Wang Yanlong, Chen Yingchun, et al. Methodology for assessing dependencies between factors influencing airline pilot performance reliability: A case of taxiing tasks[J]. Journal of Air Transport Management, 2020, 89: 101877.
[14] Yao Kai, Yan Shengyuan, TRAN C C. A fuzzy CREAM method for human reliability analysis in digital main control room of nuclear power plants [J]. Nuclear Technology, 2022, 208(4): 761-774.
[15] HOLLNAGEL E. Cognitive reliability and error analysis method (CREAM)[M]. Elsevier, 1998.
[16] HERRERA F, MARTINEZ L. The 2-tuple linguistic computational model: Advantages of its linguistic description, accuracy and consistency [J]. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2001, 9(supp01): 33-48.
[17] YANG Yingzhi. Study on multiple attribute decision making under linguistic environment by bipolar 2-tuple fuzzy linguistic representation model [D] Nanning: Guangxi University,2014.
[18] ZHU Xuelong. Fundamentals of applied information theory [M]. Beijing: Tsinghua University publishing house,2001.
[19] ZHANG Junfeng, SHI Yaoyao, LIN Xiaojun, et al. Parameters optimization in belt polishing process of blade based on grey relational analysis[J]. Computer Integrated Manufacturing Systems, 2017(04):129-137.

[20] He Xuhong , Wang Yao , Shen Zuipei , et al. A simplified CREAM prospective quantification process and its application[J]. Reliability Engineering and System Safety, 2008, 93(2):298-306.
[21] Sun Zhiqing , Li Zhengyi , Gong Erling , et al. Estimating Human Error Probability using a modified CREAM[J]. Reliability Engineering & System Safety, 2012, 100:28-32.
[22] YANG Z L, BONSALL S, WALL A, et al. A modified CREAM to human reliability quantification in marine engineering[J]. Ocean engineering, 2013, 58: 293-303.
[23] AKYUZ E , CELIK M . Application of CREAM human reliability model to cargo loading process of LPG tankers[J]. Journal of Loss Prevention in the Process Industries, 2015, 34:39-48.
[24] WANG Ning, DU Xiuli, ZHANG Mingju, et al. A Weighted Fuzzy CREAM Model for Human Reliability Analysis in Shield Tunneling[J]. Journal of Tianjin University(Science and Technology), 2019 ,52(02):92-102.