Comprehensive Evaluation of Cross-Sea Cluster Facility Safety Emergency Drill Based on Improved Two-Tuple Linguistic Information

Qiang Jing[1], Xingwang Liu[2](✉), and Ni Zhou[1]

[1] Hong Kong-Zhuhai-Macao Bridge Authority, Zhuhai CN 519060, China
[2] ZhongJiao HuaAn Technology Co., Ltd, Beijing CN 100088, China
yingwang2023@126.com, LiuXingWang1981@126.com

Abstract. The evaluation of an emergency drill is the key to testing the quality of the emergency plan and investigating the effect of the emergency drill. In this paper, taking the evaluation of cross-sea cluster facilities as the research subject, the emergency drill content was analyzed, and a comprehensive evaluation index system of emergency drills got established from four aspects: the scientificity of emergency plans U_1, the rationality of emergency drill organization U_2, the effectiveness of drill execution U_3 and the post-drill evaluation and optimization effectiveness U_4. The analytic network process (ANP) and entropy weight method were next used to calculate the subjective weights and objective weights respectively, and the combined weights are obtained based on the minimum deviation theory. Then, the improved two-tuple linguistic information evaluation model was employed to evaluate the emergency drill. Finally, by illustrating the emergency drill of the Hong Kong-Zhuhai-Macao Bridge, the application practice and demonstration of the evaluation method were carried out. The application suggests that the proposed emergency drill evaluation index system and evaluation method can be well applied to emergency drills and realize the continuous optimization of the emergency plan.

Keywords: Safety Guarantee and Protection · Safety Emergency · Cross-sea Cluster Facility · Emergency Drill Evaluation · Two-tuple Linguistic Information

1 Introduction

With the increasing development and maturity of bridge and tunnel engineering technology, equipment, talents, and other fields in China, a number of cross-sea cluster facilities such as Hangzhou Bay Cross-sea Bridge, Hong Kong-Zhuhai-Macao Bridge, and Shanghai Yangtze River Tunnel and Bridge Project (Chongming Crossing) have been built and open for traffic, providing convenient and efficient transportation channels for economic development and cultural exchanges between cities and regions. After the completion of the cross-sea cluster facilities, tasks and focus were shifted to operational management and service, setting safety and emergency management as the top priority. To ensure that the emergency management department of cross-sea cluster facilities carries out

S. H. B. D. M. Zailani et al. (Eds.): ICMSEM 2023, 259, pp. 203–217, 2024.
https://doi.org/10.2991/978-94-6463-256-9_21

emergency response and disposal in a timely, orderly, scientific, and effective manner in the face of emergencies, it is necessary to perform emergency drills and assessments in time.

Wu et al. put forward the evaluation method of fire drill effect for ships and personnel at sea from the three aspects of drill planning, implementation, and recovery according to the new system of port state supervision [1]. Burns et al. developed the evaluation content and scoring mechanism of the air medical service department in a large emergency exercise in Connecticut, the United States, and evaluated the training effect of the department [2]. Li et al. constructed an evaluation index system of emergency drills for offshore platform accidents used the AHP method to calculate and test the index weight and designed the index scoring mechanism [3]. Luan et al. evaluated the safety emergency drill of industrial production with the AHP-FCEM method [4]. Chen established an evaluation framework for urban rail network emergency drills, and employed the ANP-DEA combination method to comprehensively evaluate the effect of drills [5]. Feng et al. built a Kirkpatrick four-level evaluation framework from four aspects: reaction, learn, behavior, and result, and evaluated the training effect of the emergency plan using the two-tuple linguistic-FAHP combination method, and verified the effectiveness of the method with the help of example cases [6]. Li et al. established an emergency drill evaluation model library and evaluation system based on the main line of emergency drill evaluation business using the AHP method and software development technology [7]. In addition, the Bayesian network [8] and BP neural network are used in the existing research as well [9]. The above research involved the evaluation method of emergency drills in multiple fields, and from different levels and angles, testing the effectiveness of a certain type of emergency plan, or evaluating the effect of an emergency drill. However, no literature is dedicated to the safety emergency drill of cross-sea cluster facilities, and further research is thus required.

In this paper, the cross-sea cluster facility emergency drill is taken as the research subject to analyze the evaluation content of the emergency drill, establish a comprehensive evaluation index system of emergency drill, and draw on the existing experience and related methods. A comprehensive evaluation method of safety emergency drill is proposed in order to test the quality of an emergency plan, evaluate the emergency response of the emergency team, and provide a scientific basis for continuous improvement of plans and emergency capacity building.

2 Methodology

In this section, the specific work of the emergency drill for the cross-sea cluster facilities is first analyzed, and the emergency drill evaluation index system is established concerning the relevant specifications and standards. Then, how to determine the subjective and objective weights of the evaluation indicators, and calculate the comprehensive weights of both subjective and objective is discussed. Next, the shortcomings of the existing methods are fixed as required by real demands, and a comprehensive evaluation method for emergency drills is proposed. Finally, based on the above research, the emergency drill evaluation steps and evaluation grading criteria are given.

2.1 Construction of Emergency Drill Evaluation Index System

Cross-sea cluster facilities generally refer to combined structures of bridges, islands, and tunnels. The geographical environment and climatic conditions are relatively complex while cross-sea bridges have large spans and navigation requirements. However, the driving space of immersed tunnels is closed, making the evacuation of personnel and vehicles difficult. During the operation period, it may face accidents or incidents such as production and social safety accidents, natural disasters, and public health incidents. The safety emergency drill is performed based on the hazard sources or factors identified during the operation of cross-sea cluster facilities, and pre-set emergency scenarios. And each emergency working group simulates a series of activities such as accident risk identification, emergency response, command coordination, alert evacuation, emergency rescue, and recovery guarantee according to the emergency plan, so as to achieve the purposes of testing the plan and running-in mechanism, team training and improving emergency preparedness.

The organization and implementation of emergency drills are generally divided into four stages: drill planning, implementation, evaluation and summary, and improvement. The comprehensive evaluation of emergency drills is completed by an evaluation expert group formed by the drill organization to supervise the emergency drills from an all-round perspective. Therefore, on the basis of referring to the policy documents such as Guidelines for Emergency Drills and Hong Kong-Zhuhai-Macao Bridge Administration Emergency Comprehensive Emergency Plan, and drawing on the existing research results [3–7], in this paper, a comprehensive evaluation index system of safety emergency drills is built up, comprising 4 first-level indexes and 26 s-level indexes, as shown in Table 1.

2.2 Weight Calculation of Emergency Drill Evaluation Index

The weight calculation methods can be divided into three categories: subjective weighting, objective weighting, and comprehensive weighting. In order to take into account both subjectivity and objectivity, the network analytic network process and the entropy weight method were employed to determine the subjective weight and objective weight of each index respectively, and the minimum deviation theory was utilized to obtain the comprehensive weight in this study.

2.2.1 Subjective Weight Calculation Based on Analytic Network Process

In the evaluation index system of emergency drills of cross-sea cluster facilities, the coupling effect makes each index interrelated and interdependent. For example, the comprehensiveness of the emergency plan and the viability of emergency response measures will affect the effectiveness of drill execution such as the abilities of accident emergency response, rescue, and recovery.

Professor T.L. Saaty designed a scientific decision-making method called the Analytic Network Process (ANP) which was transformed into an internal dependency network structure based on the independent and hierarchical structure of the Analytic Hierarchy Process (AHP). In the calculation process, using the ANP method, the index and

Table 1. Cross-sea cluster facilities emergency drill evaluation index system

First-level index	Second-level index
The scientificity of emergency plan U_1	The comprehensiveness of emergency plan content U_{11}
	The clarity of organizational responsibilities U_{12}
	The depth of accident risk analysis U_{13}
	The accuracy of accident classification response U_{14}
	The effectiveness of emergency prevention U_{15}
	The operability of emergency measures U_{16}
	The rationality of information report and release U_{17}
The rationality of drill organization U_2	The level of exercise plan formulation U_{21}
	The preparations for the drill U_{22}
	The emergency material support performance U_{23}
	The enthusiasm for personnel training U_{24}
The effectiveness of drill execution U_3	The capabilities of emergency early-warning U_{31}
	The ability of emergency response U_{32}
	The ability of Emergency command decision-making U_{33}
	The ability of emergency personnel cooperation U_{34}
	The capability of on-site alert and evacuation U_{35}
	The capability of on-site emergency rescue U_{36}
	The capacity to restore traffic at site U_{37}
	The capabilities of emergency support U_{38}
	The capabilities of report and communication U_{39}
The effectiveness of evaluation and optimization after drill U_4	The degree of completion of the exercise target U_{41}
	The overall performance of the drill personnel U_{42}

(*continued*)

Table 1. (*continued*)

First-level index	Second-level index
	The evaluation and summary of the exercise U_{43}
	The drill evaluation and summary U_{44}
	The modification and improvement of plan U_{45}

index set are determined first, which is the evaluation index system of the emergency drill. Second, the control layer is constructed according to the evaluation objectives and criteria, the network layer is built considering the internal and external independence and dependence, and the hierarchical structure diagram of the emergency drill evaluation network is established. Then, the judgment matrices of the relevant index sets are compared in pairs, with weights being calculated respectively. The relevant indexes within the index set and between the index sets are compared in pairs one by one, and the relative weights of the judgment matrices are measured. The initial super-matrix is constructed in order. Finally, the weighted super-matrix is obtained using the sorting vector, and then the fixed value is obtained by limiting convergence, which is the weight of the ANP method. In order to simplify the workload of many indexes in the evaluation index system of emergency drills, *YAANP* software is used as well to assist the solution.

2.2.2 Objective Weight Calculation Based on Entropy Weight Method

The entropy weight method intends to determine the objective weight by measuring the degree of variation of the index. In the evaluation process, a smaller information entropy of a certain index causes a greater degree of variation, more effective information provided, greater prominence in the evaluation process, and greater weight. Assuming that n indexes are given, then there are m sample values and the original data matrix is established, which is recorded as $X = (x_{ij})_{n \times m}$.

The indexes are divided into positive and negative through dimensionless processing:

$$y_{ij} = \begin{cases} \frac{x_{ij} - \min(X_i)}{\max(X_i) - \min(X_i)}, \text{ positive indexs} \\ \frac{\max(X_i) - x_{ij}}{\max(X_i) - \min(X_i)}, \text{ negativeindexs} \end{cases} \tag{1}$$

where X_i is the index i sample set, $X_i = \{x_{i1}, x_{i2}, \ldots, x_{im}\}$.

If $y_{ij} = 0$, then the subsequent calculation of information entropy is meaningless. So standardization are performed as follows:

$$p_{ij} = \frac{y_{ij} + 10^{-4}}{\sum_{i=1}^{n} \left(y_{ij} + 10^{-4}\right)} \tag{2}$$

The information entropy of index i ($e_i \in [0, 1]$) is calculated as:

$$e_i = -\frac{1}{\ln m}\sum_{j=1}^{m} p_{ij} \ln p_{ij} \tag{3}$$

The weight of index information entropy is calculated as:

$$w_i = \frac{1 - e_i}{\sum_{i=1}^{n} (1 - e_i)} \tag{4}$$

2.2.3 Comprehensive Weight Calculation Based on the Minimum Deviation

Considering the weights of subjective and objective weighting methods, linear weighting is the more commonly-used combination method, as shown below:

$$w_i = \alpha w_{i1} + (1 - \alpha) w_{i2} \tag{5}$$

where w_{i1} is the weight of the ANP method; w_{i2} is the weight of the entropy weight method; w_i is the combination weight; $\alpha(\alpha \in [0. 1])$ is the subjective weight preference coefficient with a significant subjectivity, generally determined by expert experience.

In order to consider both subjectivity and objectivity in decision-making, the minimum deviation theory is introduced. The core idea is to minimize the deviation between the comprehensive weight and the subjective and objective weights. Accordingly, a comprehensive weight calculation model based on minimum deviation is established [10]:

$$\begin{cases} \min f = \sum_{i=1}^{n} \left[(w_i - w_{i1})^2 + (w_i - w_{i2})^2\right] \\ w_i \geq 0, \sum_{i=1}^{n} w_i = 1 \end{cases} \tag{6}$$

2.3 Comprehensive Evaluation of Emergency Drills

The evaluation indexes are complex, diverse, and difficult to quantify. Scholars often use fuzzy language to report feedback on the evaluation results. In this study, the improved two-tuple linguistic method has been employed for constructing a comprehensive evaluation model.

2.3.1 Two-Tuple Linguistic

A new fuzzy language expression method proposed by Professor F. Herrera [11] from the University of Granada, Spain in 2000. Its core idea is to represent expert evaluation through two-tuple linguistic variables (s_i, α_i), so as to avoid distortion and loss of evaluation information in aggregation operation and make evaluation information operation analysis more reliable and accurate. Assuming that the natural language evaluation term set is S = {si|i=0,1,…,m}, then m + 1 is called granularity. For example, S is a set of seven language evaluation terms, that is to say, S = {s_0, s_1, s_2, s_3, s_4, s_5, s_6}. In two-tuple linguistic variables, s_i is called a linguistic term, and $s_i \in S$; α_i is called a symbol transfer value, representing the deviation between the expert evaluation information and the closest linguistic term. In the two-tuple linguistic analysis, if the linguistic evaluation scale is improperly selected for the operation, then it is easy to produce improper evaluation information conversion and aggregation distortion. Upon referring to the existing research results [12], the improved two-tuple linguistic method was finally used to evaluate the emergency drill.

2.3.2 Comprehensive Evaluation Model of Emergency Drills

The definition and calculation formula of improved binary semantics are as follows:

Definition 1, let $s_i \in S$ be a linguistic term, then its corresponding binary semantic form is obtained by function θ:

$$\theta(s_i) \in (s_i, 0), s_i \in S \tag{7}$$

Definition 2, let $s_i \in S$ be a linguistic term, then the real number $\beta \in (1 - a^{m/2}, a^{m/2} - 1)$ is a binary semantic variable (s_i, α_i). The result obtained by some aggregation operation is called the ensemble operation value, where a is called the rank ratio parameter, a = 1.4 [12]. The two-tuple linguistic form corresponding to the value of β can be obtained by function Δ:

$$\Delta : \left[1 - a^{m/2}, a^{m/2} - 1\right] \to S \times [-0.5, 0.5) \tag{8}$$

$$\Delta(\beta) = (s_i, \alpha_i) = \begin{cases} s_i, i = \begin{cases} m/2 + \text{round}\big(\log_a(\beta + 1)\big), \beta \geq 0 \\ m/2 - \text{round}\big(\log_a(|\beta| + 1)\big), \beta < 0 \end{cases} \\ \alpha_i = \begin{cases} \log_a(\beta + 1) + m/2 - i, \beta \geq 0 \\ -\log_a(1 - \beta) + m/2 - i, \beta < 0 \\ \alpha_i \in [-0.5, 0.5] \end{cases} \end{cases} \tag{9}$$

Definition 3, let (s_i, α_i) be a two-tuple linguistic variable, $s_i \in S$, $\alpha_i \in [-0.5, 0.5)$, then it is transformed into the corresponding value $\beta \in (1 - a^{m/2}, a^{m/2} - 1)$ by the inverse function Δ^{-1}:

$$\Delta^{-1} : S \times [-0.5, 0.5) \to \left[1 - a^{m/2}, a^{m/2} - 1\right] \tag{10}$$

$$\Delta^{-1}(s_i, \alpha_i) = \beta = \begin{cases} a^{(i+\alpha_i-m/2)} - 1, i + \alpha_i \geq m/2 \\ -a^{(m/2-i-\alpha_i)} + 1, i + \alpha_i < m/2 \end{cases} \tag{11}$$

Definition 4, Let $\{(s_i, \alpha_i)|i = 1,2,\ldots,n\}$ be a set of two-tuple linguistic variables, then the arithmetic mean $\overline{\beta}$ is:

$$\overline{\beta} = \frac{\sum_{i=0}^{n} \Delta^{-1}(s_i, \alpha_i)}{n} \tag{12}$$

Definition 5, Let $\{(s_i, \alpha_i)|i = 1,2,\ldots,n\}$ be a set of two-tuple linguistic variables and $\{(w_i, \sigma_i)|i = 1,2,\ldots,n\}$ be the corresponding two-tuple linguistic weight vector, then the weighted average $\tilde{\beta}$ is:

$$\tilde{\beta} = \frac{\sum_{i=0}^{n} \Delta^{-1}(w_i, \sigma_i) \times \Delta^{-1}(s_i, \alpha_i)}{\sum_{i=0}^{n} \Delta^{-1}(w_i, \sigma_i)} \tag{13}$$

2.3.3 Emergency Drills Evaluation Steps and Evaluation Grade Numerical Interval

Step 1: With Formula (7), the evaluation information of the evaluation group $\{e_j \mid j = 1,2,\ldots, k\}$ for each second-level index is transformed into a two-tuple linguistic form $(s_i, 0)$, and a two-tuple linguistic evaluation matrix is constructed.
Step 2: With Formula (11), the two-tuple linguistic evaluation matrix is transformed into the corresponding integrated operation value β.
Step 3: With Formula (12) and each secondary index $\beta(\beta \in \{\beta_1, \beta_2,\ldots, \beta_k\})$, the average evaluation value $\overline{\beta}$ of each second-level index is obtained.
Step 4: With Formula (13) and the comprehensive weight of the second-level index, the weighted average evaluation value $\tilde{\beta}$ of the primary index is obtained.
Step 5: With Formula (9), the weighted average evaluation value of the first-level index is transformed into the corresponding two-dimensional semantic form (s_i, α_i).
Step 6: Using the weighted average evaluation value of the first-level index and the corresponding comprehensive weight, the evaluation results of the emergency drill are obtained by Formula (13, 9). The evaluation level is divided into 7 levels, which are recorded as the evaluation information set $S = \{s_0\text{: } N, s_1\text{: } VL, s_2\text{: } L, s_3\text{: } M, s_4\text{: } H, s_5\text{: } VH, s_6\text{: } P\}$.

A two-tuple linguistic variable (s_i, α_i) is taken to calculate the corresponding numerical interval, where i can take the value of 0,1,2,3,4,5,6, granularity m is 6, and the symbol transfer value $\alpha_i \in [-0.5, 0.5)$. Where the evaluation grade is N, the two-tuple linguistic variable is (s_0, α_i). Where $\alpha_i = -0.5$, the corresponding value $\beta = -1.46/2 - (-0.5) + 1 = -2.2467$ is obtained by Formula (11). Similarly, where $\alpha_i = 0.5$, the corresponding value $\beta = -1.46/2 - 0.5 + 1 = -1.3191$. Therefore, the numerical interval corresponding to the N level is $[-2.2467, -1.3191)$. Similarly, the numerical intervals corresponding to other levels can be obtained, as shown in Fig. 1. For example, where the value of an evaluation result is 1.2424, the corresponding two-tuple linguistic form is $(s_5, 0.4)$, which suggests that the evaluation result is classified as the H grade and is in the upper half of the interval.

3 Case Analysis and Discussion

Taking a major traffic accident emergency drill held by the Hong Kong-Zhuhai-Macao Bridge Administration as an example, an empirical study on the evaluation method of the drill was performed.

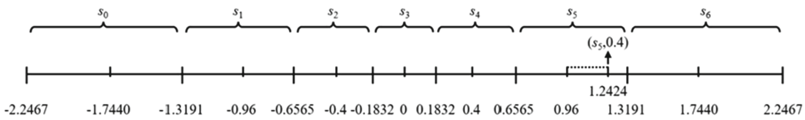

Fig. 1. Evaluation grade and numerical interval of an emergency drill

The Hong Kong-Zhuhai-Macao Bridge crosses the Lingdingyang sea area of the Pearl River Estuary and connects the Hong Kong Special Administrative Region, Zhuhai City of Guangdong Province, and the Macao Special Administrative Region. Traffic accidents, maintenance and production accidents, traffic congestion, fires, meteorological disasters, and other emergencies or events may occur after the completion and opening of the bridge. Therefore, the bridge administration regularly carries out typical emergency drills and training activities, improves the emergency system and mechanism of the unit, and enhances the emergency response capacity of each emergency working group, ensuring clarified emergency response responsibilities, timely response, and effective action.

3.1 Evaluation Results of Emergency Drills

The evaluation group (e_1, e_2, e_3, e_4, e_5) was composed of 5 experts. Referring to the actual situation of emergency drills, the group evaluated the second-level indexes of drill evaluation. The evaluation term set is $S = \{s_0: N, s_1: VL, s_2: L, s_3: M, s_4: H, s_5: VH, s_6: P\}$. According to the two-two index judgment matrix provided by experts, the subjective weight got obtained using *YAANP*. Based on the evaluation results of the second-level indexes, the objective weight was calculated by Formula (1–4). Using *LINGO* software programming, the minimum deviation comprehensive weight was obtained. The evaluation information and weight calculation results of the second-level indexes of the evaluation group are shown in Table 2.

According to the evaluation steps of the emergency drill in Sect. 2.3.3, the specific evaluation process is introduced by illustrating U_1:

(1) Convert the evaluation results of second-level indexes into two-tuple linguistic forms correspondingly. For example, the evaluation result of the comprehensiveness of the emergency plan content U_{11} is *VH*, *H*, *VH*, *VH*, *P*, which is respectively converted into a binary semantic form by Formula (7): (s_5,0), (s_4,0), (s_5,0), (s_5,0), (s_6,0).

(2) The corresponding values β are calculated by Formula (11), such as (s_5,0), $\beta = \Delta^{-1}(s_5, 0) = 1.4^{(5-3)} - 1 = 0.96$. Similarly, the evaluation values of other experts are: 0.96, 0.4, 0.96, 0.96, 1.744.

(3) The average evaluation value $\overline{\beta}$ of the index U_{11} is obtained by Formula (12). The average evaluation values of U_{12}, U_{13}, U_{14}, U_{15}, $U_{16,}$ and U_{17} are obtained respectively: −0.2336, 0.848, 0.432, 1.2736, −0.0768, 0.8128.

(4) The average evaluation value $\overline{\beta}$ of 7 s-level indexes of U_1 index and the corresponding comprehensive weight are taken into Formula (13), and the weighted average evaluation value $\tilde{\beta}$ is 0.5833.

Table 2. Cross-sea cluster facilities emergency drill evaluation index system

First-level index	Second-level index	Evaluation					Weight
		e_1	e_2	e_3	e_4	e_5	
The scientificity of emergency plan U_1	The comprehensiveness of plan content U_{11}	*VH*	*H*	*VH*	*VH*	*P*	0.0171
	The clarity of organizational responsibilities U_{12}	*N*	*VH*	*N*	*VH*	*H*	0.0290
	The depth of accident risk analysis U_{13}	*H*	*VH*	*VH*	*VH*	*VH*	0.0150
	The accuracy of accident classification response U_{14}	*VH*	*VH*	*H*	*H*	*M*	0.0193
	The effectiveness of emergency prevention U_{15}	*VH*	*P*	*P*	*VH*	*VH*	0.0403
	The operability of emergency measures U_{16}	*M*	*VH*	*H*	*M*	*N*	0.0255
	The rationality of information report and release U_{17}	*VH*	*H*	*VH*	*P*	*M*	0.0195
The rationality of drill organization U_2	The level of exercise plan formulation U_{21}	*VH*	*P*	*M*	*M*	*VH*	0.0354
	The preparations for the drill U_{22}	*VH*	*VH*	*P*	*VH*	*VH*	0.0786
	The emergency material support performance U_{23}	*H*	*VH*	*H*	*H*	*N*	0.0302
	The enthusiasm for personnel training U_{24}	*H*	*VH*	*VH*	*H*	*VH*	0.0576
The effectiveness of drill execution U_3	The capabilities of emergency early-warning U_{31}	*VH*	*P*	*VH*	*VH*	*VH*	0.0676
	The ability of emergency response U_{32}	*H*	*HC*	*H*	*VH*	*VH*	0.0388

(*continued*)

Table 2. (*continued*)

First-level index	Second-level index	Evaluation					Weight
		e_1	e_2	e_3	e_4	e_5	
	The ability of Emergency decision-making U_{33}	*VH*	*VH*	*VH*	*H*	*H*	0.0621
	The ability of emergency personnel cooperation U_{34}	*VH*	*VH*	*P*	*VH*	*N*	0.0278
	The capability of on-site alert and evacuation U_{35}	*M*	*M*	*H*	*H*	*N*	0.0300
	The capability of on-site emergency rescue U_{36}	*M*	*P*	*P*	*VH*	*VH*	0.0388
	The capacity to restore traffic at site U_{37}	*VH*	*H*	*VH*	*H*	*M*	0.0313
	The capabilities of emergency support U_{38}	*M*	*VH*	*P*	*VH*	*M*	0.0404
	The capabilities of report and communication U_{39}	*M*	*M*	*VH*	*H*	*VH*	0.0363
The effectiveness of evaluation and optimization after drill U_4	The degree of completion of the exercise target U_{41}	*VH*	*VH*	*P*	*VH*	*VH*	0.0622
	The overall performance of the drill personnel U_{42}	*H*	*H*	*H*	*M*	*M*	0.0501
	The evaluation and summary of the exercise U_{43}	*VH*	*P*	*VH*	*VH*	*VH*	0.0590
	The drill evaluation and summary U_{44}	*P*	*VH*	*H*	*P*	*VH*	0.0462
	The modification and improvement of plan U_{45}	*VH*	*H*	*P*	*VH*	*H*	0.0420

Note The values in the brackets represent the comprehensive weight of the index

(5) The weighted average evaluation value $\tilde{\beta}$ is converted into a binary semantic form: (s_4, 0.3657) using Formula (9). Similarly, the evaluation values and evaluation intervals of other secondary and primary indicators are obtained, as listed in Table 3.

Table 3. Emergency drill evaluation index system of cross-sea cluster facilities

First-level index	$\widetilde{\beta}$	(s_k, σ_k)	Second-level index	$\widetilde{\beta}$	(s_k, σ_k)
U_1	0.5833	$(s_4, 0.3657)$	U_{11}	1.0048	$(s_5, 0.0672)$
			U_{12}	-0.2336	$(s_2, 0.3761)$
			U_{13}	0.8480	$(s_5, -0.1749)$
			U_{14}	0.4320	$(s_4, 0.0672)$
			U_{15}	1.2736	$(s_5, 0.4411)$
			U_{16}	-0.0768	$(s_3, -0.2199)$
			U_{17}	0.8128	$(s_5, -0.2320)$
U_2	0.7861	$(s_5, -0.2762)$	U_{21}	0.7328	$(s_5, -0.3662)$
			U_{22}	1.1168	$(s_5, 0.2287)$
			U_{23}	0.0832	$(s_3, 0.2375)$
			U_{24}	0.7360	$(s_5, -0.3607)$
U_3	0.6555	$(s_4, 0.4981)$	U_{31}	1.1168	$(s_5, 0.2287)$
			U_{32}	0.1952	$(s_4, -0.4700)$
			U_{33}	0.7360	$(s_5, -0.3607)$
			U_{34}	0.5760	$(s_4, 0.3519)$
			U_{35}	-0.1888	$(s_2, 0.4860)$
			U_{36}	1.0816	$(s_5, 0.1789)$
			U_{37}	0.5440	$(s_4, 0.2910)$
			U_{38}	0.7328	$(s_5, -0.3662)$
			U_{39}	0.4640	$(s_4, 0.1328)$
U_4	0.9193	$(s_5, -0.0624)$	U_{41}	1.1168	$(s_5, 0.2287)$
			U_{42}	0.2400	$(s_4, -0.3607)$
			U_{43}	1.1168	$(s_5, 0.2287)$
			U_{44}	1.1616	$(s_5, 0.2910)$
			U_{45}	0.8928	$(s_5, -0.1037)$

(6) The weighted average evaluation value of the first-level index and its weight are taken into Formulas (13, 9), and the comprehensive evaluation value of this drill is 0.7383, the two-tuple linguistic form is $(s_5, -0.3568)$, and the evaluation grade is '*VH*', respectively.

3.2 Analysis of Emergency Drill Evaluation Results

It can be seen that this emergency drill received a '*VH*' rating, illustrated as follows:

(1) The scientificity of emergency plan U_1 is 0.5833, the evaluation interval is (s_4, 0.3657), and the corresponding evaluation grade is '*VH*'. U_{12} and U_{16} are (s_2, 0.3761) and (s_3, −0.2199) respectively. Therefore, problems including unclear organization and responsibilities and weak operability of emergency responses are identified in the plan. The emergency plan is revised based on these findings.
(2) The rationality of drill organization U_2 is 0.7861, the evaluation interval is (s_5, −0.2762), and the corresponding evaluation grade is '*VH*'. However, U_{23} is (s_3, 0.2375) at the '*M*' level, indicating insufficient investment in emergency supplies in the drill.
(3) The evaluation value of the effectiveness of drill execution U_3 is 0.6555, the evaluation interval is (s_4, 0.4981), and the corresponding evaluation grade is '*H*', 0.0019 from the lower end of '*VH*'. As a result, the evaluation grade is close to '*VH*'.
(4) The evaluation value of U_4 at 0.9193 ranks the highest in the first-level indicators. The evaluation result is (s_5, −0.0624), and the corresponding evaluation grade is '*VH*'. The evaluation level of all second-level indicators is rated as the level of *VH* or *H*, indicating completed emergency drills and training, and improved emergency capability and plans.

3.3 Comparison of Emergency Drill Assessment Methods

The analytic hierarchy process (AHP), the fuzzy comprehensive evaluation method (FCE) [4], the traditional two-tuple linguistic method [6], and the improved two-tuple linguistic method were all used to evaluate the emergency drill. The evaluation results are as follows.

Figure 2 shows a good consistency of the evaluation values obtained by the four evaluation methods, verifying the effectiveness of the proposed method. The evaluation results obtained by the traditional two-tuple linguistic method and the improved two-tuple linguistic method are the closest. When the AHP and the FCE method are used to extract and aggregate the evaluation information, approximate discretization, and membership fuzzification are involved as well, which can easily lead to distortion or loss of the evaluation information.

The two-tuple linguistic method represents the semantic information in the form of two-tuple (s_i, α_i), with a continuous range. In the evaluation term set S = {s_0: *N*, s_1: *VL*,

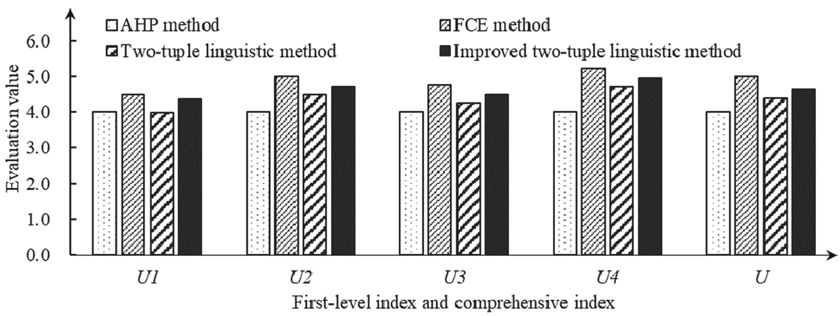

Fig. 2. Comparison of first-level indexes and comprehensive index evaluation values

s_2: L, s_3: M, s_4: H, s_5: VH, s_6: P}, the semantic distance between each comment is not nonlinear and isometric. However, the ($0 \sim n$) linear scale is still used in the traditional two-tuple linguistic method to classify the nonlinear semantic information linearly and isometrically, which may cause problems such as improper conversion and distortion of semantic information, and ultimately affect the accuracy and reliability of the evaluation results. According to the relationship between the subjective sensation and objective stimulus of humans, the improved binary semantic method combines the advantages of the exponential scale and ($-n \sim n$) scale to design a new composite evaluation scale ($1 - a^{m/2}$, $a^{m/2} - 1$). The scale can accurately express the meaning of semantic information, thus improving the accuracy and reliability of the evaluation model.

4 Conclusions

In this paper, the evaluation content of emergency drills for cross-sea cluster facilities is summarized, and a comprehensive evaluation index system for the emergency drill is set up in four aspects: the scientificity of the emergency plan, rationality of emergency drill organization, the effectiveness of drill execution and the effectiveness of evaluation and optimization after drills are completed. It demonstrates certain feasibility and practicability and provides a theoretical basis for similar emergency drill evaluation work.

The minimum-deviation comprehensive weighting method and the improved binary semantic method have been both used to evaluate the safety emergency drill of cross-sea cluster facilities, and a drill evaluation model has been established, providing a reference for the evaluation and summary of emergency drills in related fields.

Taking the emergency drill of cross-sea cluster facilities in face of a traffic accident as an example, the final evaluation value is 0.7383, (s_5, −0.3568), and the evaluation grade is '*VH*', which is in line with the actual drill results.

Acknowledgments. The authors would like to acknowledge fruitful discussions and assistance in emergency drill evaluation from Professor Cheng Guozhu, Dr. Chen Yongsheng, and Master Li Jinyu. This work was supported by a national key R&D program (the Integrated Application of Intelligent Operation and Maintenance Technology for Hong Kong-Zhuhai-Macao Bridge, 2019YFB1600700).

References

1. Wu J, Jin Y. Effectiveness Evaluation on Fire Drills for Emergency and PSC Inspections on Board[J]. TransNav: the International Journal on Marine Navigation and Safety of Sea Transportation, 2014, 8(2):229–236.
2. Burns K J, Robinson K, Lowe E G. Evaluation of responses of an air medical helicopter program during a comprehensive emergency response drill[J]. Air Medical Journal, 2007, 26(3):139–143.

3. Li X S, Zang Q L, Wu L, et al. Research on assessment method of emergency drill on offshore platform accidents[J]. Journal of Safety Science and Technology, 2019, 26(3): 139–143. (in Chinese)
4. Luan G H, Pei Y Q, Wu D P, et al. The AHP–fuzzy method for an integrated emergency drills assessment in industrial production[J]. Mathematics in Practice and Theory, 2014(10): 98–103. (in Chinese)
5. Chen L. Combination evaluation model of emergency exercise in urban rail transit network based on ANP & DEA[J]. Journal of Safety Science and Technology, 2014, 10(1): 59–63. (in Chinese)
6. Feng Y, Zhao J P. Study on evaluation of emergency plan training effectiveness based on Kirkpatrick model and 2-tuple linguistic information[J], Journal of Safety Science and Technology, 2016,26(02):168–174. (in Chinese)
7. Li Q, Dai D J. Research on assessment method, technology and system of emergency exercise[J]. Journal of Safety Science and Technology, 2016, 12(7): 49–54. (in Chinese)
8. Zhang X, Mei Z Y. Performance Evaluation of Emergency Response Drills for Ship Oil Spill Based on Bayesian Network[J]. Ship & Ocean Engineering, 2018, 047(002): 104–108. (in Chinese)
9. Ren L L, Guo S, Sun P J, et al. Evaluation of High-Speed Railway Emergency Drills based on BP Neural Network[J]. Railway Transport and Economy, 42(12): 83–88. (in Chinese)
10. Liu D H, Yu Q, Ma X N, et al. Emergency capability evaluation model based on minimum deviation combination weight [J]. Journal of Safety Science and Technology, 2014, 022(011): 79–86. (in Chinese)
11. F Herrera, Martinez L. Martínez, L.: A 2-Tuple Fuzzy Linguistic Representation Model for Computing with Words. IEEE Transactions on Fuzzy Systems, 2001, 8(6):746–752.
12. Bao G Y, Lian X L, He M, et al. Improved 2-tuple linguistic model based on new linguistic assessment scale [J]. Control and Decision, 2010(5): 780–784. (in Chinese)