Xi, YT, Yang, ZL, Fang, QG, Chen, WJ and Wang, J

A new hybrid approach to human error probability quantification-applications in maritime operations

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Xi, YT, Yang, ZL, Fang, QG, Chen, WJ and Wang, J (2017) A new hybrid approach to human error probability quantification-applications in maritime operations. Ocean Engineering, 138. pp. 45-54. ISSN 0029-8018

A new hybrid approach to human error probability quantification - applications in maritime operations

Y.T. Xi[a,b], Z.L. Yang[b*], Q.G. Fang[a], W.J. Chen[a], J. Wang[b]

[a]Merchant Marine College, Shanghai Maritime University, Shanghai 201306, P R China
[b]Liverpool Logistics Offshore and Marine (LOOM) Research Institute, School of Engineering, Technology and Maritime Operations, Liverpool John Moores University, Liverpool L3 3AF, UK

Abstract

Human Reliability Analysis (HRA) has always been an essential research issue in safety critical systems. Cognitive Reliability Error Analysis Method (CREAM), as a well-known second generation HRA method is capable of conducting both retrospective and prospective analysis, thus being widely used in many sectors. However, the needs of addressing the use of a deterministic approach to configure common performance conditions (CPCs) and the assignment of the same importance to all the CPCs in a traditional CREAM method reveal a significant research gap to be fulfilled. This paper describes a modified CREAM methodology based on an Evidential Reasoning (ER) approach and a Decision Making Trial and Evaluation Laboratory (DEMATEL) technique for making human error probability quantification in CREAM rational. An illustrative case study associated with maritime operation is presented. The proposed method is validated by sensitivity analysis and the quantitative analysis result is verified through comparing the data and the benchmarking with the real data collected from Shanghai coastal waters. Its main contribution lies in that it for the first time addresses the data incompleteness in HEP, given that the previous relevant studies mainly focus on the fuzziness in data. The findings will provide useful insights for quantitative assessment of seafarers' errors to reduce maritime risks due to human errors.

Keywords: Human reliability analysis, CREAM, HEP quantification, maritime transport, maritime risk

1. Introduction

Human error has caused many industrial accidents and disasters. Human Reliability Analysis (HRA) has therefore been an essential research issue. With the development of new technologies, the emerge of complex systems such as Nuclear Power Plant (NPP) and Very Large Crude Carrier (VLCC), makes the consequences of accidents more and more serious. Many approaches have been developed for facilitating human error quantification and human reliability analysis, such as Human Cognitive Reliability model (HCR) (Hannaman et al., 1984), Success Likelihood Index Method (SLIM) (Embrey et al., 1984), Technique for Human Error Rate Prediction (THERP) (Swain and Guttmann, 1983), Human Error Assessment and Reduction Technique (HEART) (Williams, 1988), A Technique for Human Error Analysis (ATHEANA) (Cooper et al., 1996), Cognitive Reliability and Error Analysis

* Corresponding author. E-mail: z.yang@ljmu.ac.uk (Zaili Yang).
Tel: 0044 151 231 2531; fax: 0044 151 231 2453.

Method (CREAM) (Hollnagel, 1998), SPAR-H (Gertman et al., 2005) and Bayesian Network approach (Baraldi et al., 2015). There are also some specific methods developed in the literature for Human Error Probability (HEP) quantification (Sun et al., 2012). Such approaches have been widely applied to deal with human error and human factors in various sectors including nuclear (Alvarenga et al., 2014; Jang et al., 2013), spaceflight (Calhoun et al., 2013, 2014), marine and maritime (Akyuz and Celik, 2015a, 2015b, 2016; Yang et al., 2013; Akyuz, 2016; Chen et al., 2013; Wang et al., 2011), and civil infrastructure (Nan and Sansavini, 2016), etc.

1.1. CREAM

The well-known CREAM, established by Hollnagel in 1998 (Hollnagel, 1998), has been applied to human error quantification of safety-critical systems. Both retrospective and prospective analyses can be carried out for the diagnosis and prediction of industrial accidents and events. For the prospective analysis, two steps are designed for human error quantification which are a basic method and an extended method, respectively. The basic method is used for determination of control modes and corresponding error rate intervals at a screening stage, while the extended method is employed for error quantification of cognitive functions. However, the two inherently deterministic methods arguably lack capability of dealing with the uncertainties in common performance condition (CPC) configuration (Kim et al., 2006) and different weight assignments (He et al., 2008) to the CPCs in traditional CREAM. Recently, some studies have been conducted for the improvement of CREAM and HEP estimation by means of probabilistic techniques (Kim et al., 2006; Fujita and Hollnagel, 2004), fuzzy approaches (Konstandinidou et al., 2006; Marseguerra et al., 2006, 2007; Wang et al., 2001; Yang et al., 2013; Ung et al., 2015), simplification (Sun et al., 2012; He et al., 2008) and combination (Akyuz, 2015; Ribeiro, 2016). Fujita and Hollnagel (2004) designed a new version of basic method of CREAM and Kim et al. (2006) described a probabilistic approach for determining the control modes. Konstandinidou et al. (2006), Marseguerra et al. (2006, 2007 and Wang et al. (2001) presented fuzzy sets, fuzzy rules and fuzzy-clonal selection methods for contextual and reliability evaluation. However, the fuzzy model of CREAM brings on many redundant, self-contradictory rules, which would consume computational time, and lose the truth degree of the results (Wang et al., 2001). Sun et al. (2012) and He et al. (2008) simplified the CREAM for HEP point estimation while Lee et al. (2011) designed a CREAM-based communication error analysis method (CEAM) for communication error analysis. Akyuz (2015) constructed a risk-based CREAM model for HEP quantification towards the gas inerting process on-board crude oil tankers and Ribeiro et al (2016) presented a hybrid THERP-CREAM method to analysis the human reliability of Tokai-Mura accidents. Although showing attractiveness in terms of providing solutions to some of the inherent drawbacks of CREAM, such methods have not yet well addressed the incompleteness in subjective data from experts, revealing the need for further studies in the field.

1.2. Evidential Reasoning (ER) and DEMATEL technique

Dempster-Shafer Theory (DST) was developed by Dempster in 1967 (Dempster, 1967) and later refined by Shafer in 1976 (Shafer, 1976). Subsequently, DST is known as one of the most powerful tools to deal with uncertainty problems. The evidential reasoning (ER) approach (Yang and Singh, 1994; Yang et al., 2006; Xu et al., 2006), which was initially oriented to model multiple attribute decision analysis (MADA) problems is further developed for overcoming some weaknesses in DST. It is conducted by generating basic probability assignments (BPA) through the combination of degrees of belief (DoBs) and normalised weights. It is because of the advantage of ER in modelling incompleteness that recently it has been widely applied in many domains such as safety assessment (Xi et al., 2008), environmental impact assessment (Wang et al., 2006) and software selection (Fu and Yang, 2010). In this paper, the ER approach is employed for evaluating the CPCs and combining the DoBs to generate the BPA of CPCs in the modified CREAM methodology. The DEMATEL technique was developed by the Geneva Research Centre of the Battelle Memorial Institute (Fontela and Gabus, 1976; Gabus and Fontela, 1973). It presents the cause and effect groups within a system or subsystem by applying matrices and digraphs to visualise the structure of complicated causal relationships. The advantages of DEMATEL lie in its capability of effectively modelling and quantifying the causal relationships among interdependent factors. The DEMATEL has therefore been successfully applied in many fields including business (Tseng, 2009; Wu and Lee, 2007), engineering (Seyed-Hosseini et al., 2008), education (Tzeng et al., 2007) and social studies (Tamura and Akazawa, 2005). In this paper, the technique is used to model the dependency among CPCs and further, together with the assigned weights of CPCs, to determine the values of adjusting indices. Given the strengths of ER and DEMATEL in modelling uncertainties, this paper proposes a new modified CREAM method to calculate HEP in a rational way.

Given their strengths in tackling different uncertainties exposed in traditional CREAM, ER and DEMATEL are combined to construct a new modified HEP quantification model of two stages, a general analysis stage and a quantification evaluation stage. The general analysis in Section 2 proposes an ER approach to model the incompleteness associated with CPC configuration and combination of DoBs, and to determine the corresponding control mode(s), probability intervals and the total state of context. The quantification evaluation in Section 3 describes a DEMATEL technique to simulate the interdependency of CPCs and to determine the different weights of CPCs and adjusting index values. Consequently, a rational error probability can be obtained at this stage. In Section 4, an illustrative case study in maritime operations is presented and real data has been recorded to benchmark and validate the modified method and research findings. Section 5 concludes this paper.

2. Modified CREAM: a general analysis

The core of CREAM is the Contextual Control Model (COCOM). COCOM focuses on a principle that human cognition is not only a response to a serious input but also a close loop process of continuous purposive adjustment for intension. Four kinds of control modes are defined according to the human cognition and the context, which are determined by nine Common Performance Conditions (CPCs). The four control modes are "*Scrambled*", "*Opportunistic*", "*Tactical*" and "*Strategic*" respectively (Fig.1) while the nine CPCs are "*Adequacy of organisation (CPC_1)*", "*Working conditions (CPC_2)*", "*Adequacy of man-machine interface and operational support (CPC_3)*", "*Availability of procedures and plans (CPC_4)*", "*Number of simultaneous goals (CPC_5)*", "*Available time (CPC_6)*", "*Time of day (CPC_7)*", "*Adequacy of training and experience (CPC_8)*" and "*Crew collaboration quality (CPC_9)*", respectively (Table 1). Each CPC may be evaluated at different levels indicating an *improved, not significant* or *reduced* effect on human performance accordingly. The control mode and its wide failure interval are determined by the couple of $(\Sigma_{Improved}, \Sigma_{Reduced})$. For example, if an event is evaluated to have 5 CPCs of *improved* effects and 4 CPCs of *reduced* effects, then its corresponding COCOM will be "*Tactical*" according to Fig. 1 and failure probability interval will be 1E-3 $<$ P $<$ 1E-1 from Table 2.

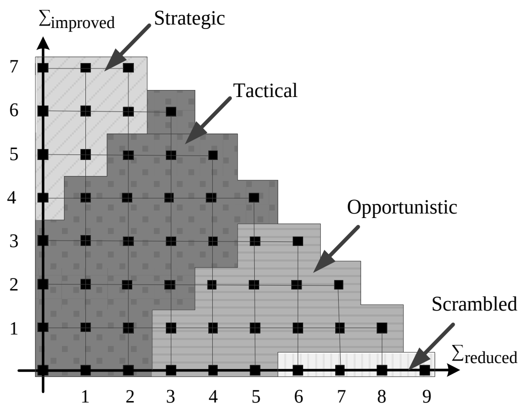

Fig.1. Relation between CPCs and control modes (Hollnagel, 1998)

Table1.
Description of CPCs and associated linguistic variables (Hollnagel, 1998)

CPC	Level	Effect
CPC_1 Adequacy of organization	L_{11} Very efficient	Improved
(AOO)	L_{12} Efficient	Not significant
	L_{13} Inefficient	Reduced
	L_{14} Deficient	Reduced
CPC_2 Working conditions	L_{21} Advantageous	Improved
(WCO)	L_{22} Compatible	Not significant
	L_{23} Incompatible	Reduced
CPC_3 Adequacy of MMI and operational support	L_{31} Supportive	Improved
(MMI)	L_{32} Adequate	Not significant
	L_{33} Tolerable	Not significant
	L_{34} Inappropriate	Reduced
CPC_4 Availability of procedures/plans	L_{41} Appropriate	Improved
(AOP)	L_{42} Acceptable	Not significant
	L_{43} Inappropriate	Reduced
CPC_5 Number of simultaneous goals	L_{51} Fewer than capacity	Not significant
(NOS)	L_{52} Matching current capacity	Not significant
	L_{53} More than capacity	Reduced
CPC_6 Available time	L_{61} Adequate	Improved
(ATI)	L_{62} Temporarily inadequate	Not significant
	L_{63} Continuously inadequate	Reduced
CPC_7 Time of day	L_{71} Day-time (0600-1800) (adjusted)	Not significant
(TOD)	L_{72} Night-time (1700-2400) (unadjusted)	Reduced
	L_{73} Night-time(0000-0700) (unadjusted)	Reduced
CPC_8 Adequacy of training and preparation	L_{81} Adequate, high experience	Improved
(AOT)	L_{82} Adequate, limited experience	Not significant
	L_{83} Inadequate	Reduced
CPC_9 Crew collaboration quality	L_{91} Very efficient	Improved
(CCQ)	L_{92} Efficient	Not significant
	L_{93} Inefficient	Not significant
	L_{94} Deficient	Reduced

2.1. Evaluating the CPCs using a DoB approach

The traditional method for evaluating the level and its effect to human reliability is easy and visible. However, it only shows a general principle (Kim et al., 2006) thus revealing some problems in its practical applications. It needs to be further improved and developed in determining the levels of CPCs and its effect to human reliability rationally. It is well known that an exact evaluation of CPCs is important but very difficult for the performance prediction. Nevertheless, it is not always easy to specify CPCs exclusively, due to the insufficiency of information and data of the context under investigation. There would be a great amount of uncertainties in the specification of CPCs. For tackling such uncertainties, a new approach which can model the DoBs distribution in the specification of CPCs is proposed. For example, an expert can give DoBs such as {0.2 *Very efficient*, 0.5 *Efficient*, 0.3 *Inefficient*, 0 *Deficient*} instead of a "crisp" judgement when evaluating CPC_1. One of the advantages of the DoB approach (Yang and Singh, 1994) is that the uncertainty associated with a CPC can be expressed in a rational way as required in practice. Second, a number of $(\Sigma_{Improved}, \Sigma_{Reduced})$ couples can be produced to indicate the possible contexts and to optimise the simulation of the associated operations. Third, the DoB approach allows the synthesis of CPCs specifications from multiple experts. It is noteworthy that the DoB approach for CPCs specification is an extension of the deterministic approach. When some CPCs are clearly specified, the DoB can be set to be 1 for the corresponding CPCs levels. The algorithm of combining the DoB structures given by multiple experts on CPCs levels is represented in Section 2.2.

Although the rules of modelling the interdependence adjustment of CPCs in the traditional basic CREAM are presented, it was argued that the influence of such adjustment deems to be insignificant at a general analysis stage (Kim et al., 2006). The detailed consideration of modelling the CPC adjustment rules is therefore investigated in the detailed quantification analysis stage in Section 3.

2.2. Combining the evaluations by experts using ER algorithm

Let L_{in} denote the *n*th level of CPC_i, then $L_i = \{L_{i1}, \dots L_{in}, \dots, L_{iN}\}$ be the frame of discernment which represents the levels of CPC_i ($1 \le i \le 9$; $n=1,\dots,N$) and it is required to be a collectively exhaustive and mutually exclusive set of hypotheses. If CPC_i is evaluated to a level L_{in} by the expert E_g ($1 \le g \le G$) with a degree of belief β_{in}^{g}, we denote this evaluation by $S(E_g) = \{(L_{in}, \beta_{in}^{g})\}$, which is a distributed evaluation and referred to as a belief structure (Yang et al., 2006), where $\beta_{in}^{g} \ge 0$ and

$$\sum_{n=1}^{N} \beta_{in}^{g} \le 1 \qquad (1)$$

If $\sum_{n=1}^{N} \beta_{in}^{g} = 1$, it is called a complete evaluation. If $\sum_{n=1}^{N} \beta_{in}^{g} < 1$, the evaluation is incomplete, the difference ($1-\sum_{n=1}^{N} \beta_{in}^{g}$) shows the degree of ignorance.

In the first step, a DoB given to an evaluation level L_{in} by E_g is transformed into a basic probability mass by multiplying the given DoB β_{in}^{g} by w_g[†] ($\sum_{g=1}^{G} w_g = 1$) which is defined as the relative importance of E_g. The liner function which transformed the DoB given to an evaluation L_{in} by E_g into a basic probability mass can be established as follows

$$m_{in}^{g} = w_g \beta_{in}^{g}, n = 1, \dots, N; g = 1, \dots, G \qquad (2)$$

If β_{in}^{g} satisfies Eq. (1), then $\sum_{n=1}^{N} m_{g}^{in} \le 1$ and accordingly (Yang et al., 2006; Xu et al., 2006; Yang et al., 2009; Yang and Wang, 2015).

$$m_{L_i}^{g} = 1 - \sum_{n=1}^{N} m_{in}^{g} = 1 - w_g \sum_{n=1}^{N} \beta_{in}^{g} \qquad (3)$$

[†] Expert knowledge is influenced by individual perspectives and goals (Ford and Serman, 1998). Therefore, complete impartiality of expert knowledge is often difficult to achieve. An important consideration is the weighting factors of experts. In this study, the weighting factors of experts are determined according to their professional position, service time and education level.

$$\bar{m}_{L_i}^{g} = 1 - w_g \tag{4}$$

$$\tilde{m}_{L_i}^{g} = w_g\left(1 - \sum_{n=1}^{N} \beta_{in}^{g}\right) \tag{5}$$

where m_{in}^{g} is the basic probability mass which measures the belief exactly assigned to L_{in} and represents how strongly the expert E_g supports L_{in}. $m_{L_i}^{g} = \bar{m}_{L_i}^{g} + \tilde{m}_{L_i}^{g}$ for all $g=1,2,\ldots,G$. The probability mass of E_g ($m_{L_i}^{g}$) unassigned to any individual L_{in} is split into two parts, one caused by the relative importance of E_g ($\bar{m}_{L_i}^{g}$), and the other due to the incompleteness of the belief degree assessment β_{in}^{g} ($\tilde{m}_{L_i}^{g}$).

The second step is to obtain the total DoB (β_{in}) by combining the basic probability masses (m_{in}^{g}) generated above. Suppose the evaluation levels are mutually exclusive and collectively exhaustive, Dempster's rule of combination (Yang et al., 2006) can be applied to combine the basic probability masses. Suppose $m_{in}^{I(g)}$ is the combined belief degree L_{in} by aggregating all the basic probability masses from the g experts and $m_{L_i}^{I(g)}$ is the remaining belief degree unassigned to any L_{in}. Let $m_{in}^{I(1)} = m_{in}^{1}$ and $m_{L_i}^{I(1)} = m_{L_i}^{1}$. Then the overall combined belief degree in L_{in} can be generated as follows (Yang et al., 2006; Xu et al., 2006; Yang et al., 2009; Yang et al., 2014).

$$\{L_{in}\}: m_{in}^{I(g+1)} = K_{I(g+1)}\left[m_{in}^{I(g)} m_{in}^{g+1} + m_{in}^{I(g)} m_{L_i}^{g+1} + m_{L_i}^{I(g)} m_{in}^{g+1}\right] \tag{6}$$

$$m_{L_i}^{I(g)} = \bar{m}_{L_i}^{I(g)} + \tilde{m}_{L_i}^{I(g)} \quad g = 1,2,\ldots,G-1 \tag{7}$$

$$\{L_i\}: \tilde{m}_{L_i}^{I(g+1)} = K_{I(g+1)}\left[\tilde{m}_{L_i}^{I(g)} \tilde{m}_{L_i}^{g+1} + \tilde{m}_{L_i}^{I(g)} \bar{m}_{L_i}^{g+1} + \bar{m}_{L_i}^{I(g)} \tilde{m}_{L_i}^{g+1}\right] \tag{8}$$

$$\bar{m}_{L_i}^{I(g+1)} = K_{I(g+1)}\left[\bar{m}_{L_i}^{I(g)} \bar{m}_{L_i}^{g+1}\right] \tag{9}$$

$$K_{I(g+1)} = \left[1 - \sum_{n=1}^{N} \sum_{\substack{h=1 \\ h \neq n}}^{N} m_{in}^{I(g)} m_{ih}^{g+1}\right]^{-1}, g = 1,2,\ldots,G-1 \tag{10}$$

$$\{L_{in}\}: \beta_{in} = \frac{m_{in}^{I(G)}}{1 - \bar{m}_{L_i}^{I(G)}} \ (n = 1,2,\ldots,N) \tag{11}$$

$$\{L_i\}: \beta_{L_i} = \frac{\tilde{m}_{L_i}^{I(G)}}{1 - \bar{m}_{L_i}^{I(G)}} \tag{12}$$

where β_{in} shows the combined DoB assigned to L_{in} in the levels of CPC_i (L_i) and β_{L_i} indicates the normalised remaining DoB unassigned to any L_{in}.

2.3. Calculating HEP of the whole task

In terms of the DoB approach for CPCs specification and the ER algorithm, the corresponding couple ($\Sigma_{Improved}, \Sigma_{Reduced}$) can be obtained and presented by decimal figures instead of integer ones. It is therefore difficult to directly use Figure 1 and Table 2 to determine the probability intervals. For example, if a couple (1.5, 3.2) is obtained, the control mode is at the boundary between Strategic and Tactical from Fig. 1, indicating that it belongs to both Strategic and Tactical control modes. Therefore, its probability interval can be determined by the intersection of the probability intervals of Strategic and Tactical modes. The probability intervals of the three boundaries between the four control modes are obtained through deriving the intersection between two neighbouring intervals. For example, the intersection between the probability intervals of Strategic and Tactical modes is 1E-3＜P＜1E-2. Consequently, all the probability intervals describing the original 4 control modes and the three boundaries between them are presented in Table 2.

Table 2.
The control modes and probability intervals

Control mode	Probability interval
Strategic	5E-5＜P＜1E-2
Boundary between St. and Ta.	*1E-3＜P＜1E-2*
Tactical	1E-3＜P＜1E-1
Boundary between Ta. and Op.	*1E-2＜P＜1E-1*
Opportunistic	1E-2＜P＜5E-1
Boundary between Op. and Sc.	*1E-1＜P＜5E-1*
Scrambled	1E-1＜P＜1E-0

Although the above analysis can assist in narrowing the intervals, the failure probabilities obtained are still presented in a wide range interval and may not be sufficient for appropriate screening and control purposes. Further investigation at the general analysis stage is carried out based on a context variable which describes the total state of the nine CPCs. For better description of the context variable, the following assumptions are made (Sun et al., 2012; He et al., 2008; Apostolakis et al., 1988).

(1) The context variable is continuous.
(2) The HEP is also continuous.
(3) HEP exponentially increases or decreases with the context variations.

The above assumptions have been used in many HRA methods and the results are validated to be acceptable in practice (Sue et al., 2012). According to the assumptions above, the functional relation between the context variable (denoted as x) and HEP can be constructed as follows (He et al., 2008):

$$HEP = f(x) = \rho exp(\varphi x) \tag{13}$$

where $x = \Sigma_{Improved} - \Sigma_{Reduced}$, ρ and φ are two constants to be determined.

Obviously, the maximum and minimum values of x are 7 and -9 (from Table 1) respectively as presented in Eq. (14).

$$HEP_{max}=\rho exp\ (-9\varphi),$$

$$HEP_{min}=\rho exp\ (7\varphi) \tag{14}$$

It can be found that HEP_{min}=0.00005 and HEP_{max}=1 from the failure probabilities in Table 2 (Hollnagel, 1998). Then, the values of ρ and φ could be calculated as 3.81E-3 and -0.619. When x=0, the context is in a neutral state. At this moment, the HEP indicates a value without any positive or negative effects received from the operational context where HEP_0 = ρ= 3.81E-3.

Thus for a certain value of x, Eq. (15) could be utilized for a general analysis:

$$HEP=HEP_0 \cdot exp\ (-0.619x) = 3.81\times10^{-3}\ exp\ (-0.619x) \quad (15)$$

Eq. (15) reflects the functional relation between the overall state of context and HEP. It can be used to determine the extent to which the context of the whole task is of positive, neutral or negative effect on human performance in the investigation of an event.

3. Modified CREAM: an extended quantification

3.1. Identifying the weights (v_i) of CPCs through DEMATEL

In the traditional CREAM, each CPC plays the same role on the influence to the operational context and HEP, though it is not a real case in practice. For example, it is often the case that the negative effect received from "night time" of CPC_7 *TOD* may be less than the one from "continuously inadequate" (available time) associated with CPC_6 *ATI*. Therefore, appropriate weights should be assigned to the CPCs to reflect their contribution towards HEP (Marseguerra et al., 2007).

The DEMATEL method is used to calculate the weights of CPCs in this section. The step of DEMATEL can be summarized as follows (Wu, 2008; Fontela and Gabus, 1976; Gabus and Fonela, 1973).

Firstly, each expert is asked to evaluate the direct influence between any two CPCs by using scores ranging from 0, 1, 2, and 3 which represent "no influence", "low influence", "medium influence" and "high influence" respectively in terms of synopsis of influence relationship among the CPCs in the traditional CREAM (Table 3). The notation of a_{ij} indicates the degree to which the expert believes CPC_i affects CPC_j. For $i = j$, the diagonal elements are set to zero. For each expert, a 9×9 (there are 9 CPCs) matrix can be established as $A^g = [a_{ij}^g]$, where g is the number of experts with $1 \leq g \leq G$, and G is the total number of experts. Thus, $A^1, A^2, \ldots, A^G$ are the metrics from G experts. The average matrix $M = [m_{ij}]$ can be obtained as follows:

$$M = [m_{ij}] = [\sum_{g=1}^{G} a_{ij}^g w_g] \quad (16)$$

where w_g is the relative importance of E_g and $\sum_{g=1}^{G} w_g = 1$.

Table 3.
The synopsis of influence relationship and adjusting rules of CPCs (Hollnagel, 1998)

Principal CPC (Influenced CPC)	Component CPCs (Influencing CPCs)
Working Condition (WCO)	*Adequacy of Organisation (AOO), Adequacy of Man Machine Interface and Shipboard Operational Support (MMI), Available Time (ATI), Time of day/ Circadian Rhythm (TOD), Adequacy of Training and Preparation (AOT)*
Number of Simultaneous Goals (NOS)	*WCO, MMI and Availability of Plans/Procedures (AOP)*
ATI	*WCO, MMI, AOP, NOS and TOD*
Crew Collaboration Quality (CCQ)	*AOO and AOT*
MMI	*AOO*
AOP	*AOO*
AOT	*AOO*

Secondly, the normalized direct relation matrix X can be obtained by Eq. (17) (Wu, 2008; Shieh et al., 2010)

$$X = bM, \tag{17}$$

where $b = \frac{1}{\max\limits_{1 \le i \le 9} \sum_{j=1}^{9} m_{ij}}$. Each element in matrix falls between zero and one.

Thirdly, the total matrix can be obtained by Eq. (18)

$$T = X(I - X)^{-1} \tag{18}$$

where I is the identity matrix. R and C are defined to represent the sum of rows and sum of columns of T, respectively. R_i, which is the sum of the ith row in matrix T, summarizes both direct and indirect effects given by CPC_i to other CPCs. C_i, which is the sum of ith column in matrix T, shows both direct and indirect effects received by CPC_i from the other CPCs. The sum (R_i+C_i) indicates the total effects given and received by CPC_i. That is, the value of (R_i+C_i) stands for the degree of importance that CPC_i plays in the entire system. On the contrary, the value of (R_i-C_i) shows the net effect that CPC_i contributes to the system. Specifically, if (R_i-C_i) is positive, CPC_i is a cause factor, while CPC_i is a result factor if (R_i-C_i) is negative.

Lastly, the weights of CPCs can be obtained. The value of (R_i+C_i) stands for the degree of importance of CPC_i. The normalized weights of CPCs can therefore be calculated by

$$v_i = \frac{R_i + C_i}{\sum_{i=1}^{9} (R_i + C_{i)}} \times 9 \tag{19}$$

where v_i is the weight of CPC_i.

3.2. Obtaining the adjusting index (δ_{in}) of CPC levels

The value of adjusting index is based on the weights of CPCs and fine-tuned in proportion. The adjusting indices of the CPC_i's levels, which have *improved*, *not significant* and *reduced* effects, are assigned to $+v_i$, zero and $-v_i$ respectively. This is a general principle to convert weights to adjusting indices. However, if there are two levels being assigned the same effect, then the corresponding v_i value should be fine-tuned in proportion. For example, for "*crew collaboration quality*", the four levels "*very efficient*", "*efficient*", "*inefficient*" and "*deficient*" have "*improved*", "*not significant*", "*not significant*" and "*reduced*" effects on performance reliability, respectively. According to the above general principle, the adjusting indices of both "*efficient*" and "*inefficient*" are zero, which does not well reflect the reality in measuring human reliability. Given the fact that the effect of "*inefficient*" on human reliability should be slightly "*reduced*" rather than "*not significant*", its adjusting index δ_{in} of the nth level of CPC_i will be determined by proportion using a linear distribution.

3.3. Calculating the CFP of sub-tasks

In the general analysis, the context variable is defined as the overall state of context whose value is $\Sigma_{Improved}$ -$\Sigma_{Reduced}$. When $\Sigma_{Improved}$ - $\Sigma_{Reduced}$=0, the context would be the nominal state and HEP would be HEP_0. However such a simple analysis may not be well suited to modelling the scenarios in the extended quantification, because a CPC has its specific effect role on human reliability. Therefore, the value of the functional relation x should be recalculated in a specific task context.

A new calculation approach, which can disclose the specific quantitative influence of a CPC on human reliability, is developed and described as follows:

$$x_i = \sum_{n=1}^{N} \delta_{in}\beta_{in} \tag{20}$$

where, x_i is the specific influence of CPC_i on human reliability; δ_{in} is the adjusting index of nth level of CPC_i; β_{in} is the combined DoB of nth level of CPC_i as mentioned in Section 2.2.2. $N=3$ when i belongs to the set of (2, 4, 5, 6, 7 or 8) while $N=4$ when i is equal to 1, 3 or 9.

Therefore, the overall quantitative influence of context can be indicated as follows:

$$x = \sum_{i=1}^{9} x_i \tag{21}$$

Consequently, the quantitative evaluation of Cognitive Failure Probability (CFP) of each sub-task is

$$CFP=CFP_0 \cdot exp(-0.619x) \tag{22}$$

where $exp(-0.619x)$ keeps consistence with its counterpart in Eq. (15) while CFP_0 varies with respect to the cognitive function and generic failure type of sub-tasks. According to the extended CREAM method, CFP_0 representing the basic values of generic cognitive function failure types will change as listed in Table 4 (Hollnagel, 1998). In the general analysis stage, HEP_0=3.81E-3 is a holistic consideration of the whole task, while the cognitive function failure probability estimation should be carried out with respect to the adjusted CFP_0.

Table 4
Basic values for generic cognitive failure types

Cognitive function	Generic failure type	Basic value (CFP_0)
Observation	O1 Wrong object observed	1.0E-3
	O2 Wrong identification	7.0E-2
	O3 Observation not made	7.0E-2
Interpretation	I1 Faulty diagnosis	2.0E-1
	I2 Decision error	1.0E-2
	I3 Delayed interpretation	1.0E-2
Planning	P1 Priority error	1.0E-2
	P2 Inadequate plan	1.0E-2
Execution	E1 Action of wrong type	3.0E-3
	E2 Action at wrong time	3.0E-3
	E3 Action on wrong object	5.0E-4
	E4 Action out of sequence	3.0E-3
	E5 Missed action	3.0E-2

Source: Adopted from Hollnagel, (1998).

4. Case study

The statistics of marine accidents show that ship collision is the upmost hazard and could cause serious damage to the properties, environment and human lives. Previous studies have shown that human error is the main contributor of ship collision accidents (P&I Club, 1992; Pennie et al., 2007; Graziano et al., 2016) and attempts to reduce the marine accidents caused by human errors. In this section, a case investigating the collision avoidance of a particular scenario in Shanghai coastal waters is presented (See Table 6) to demonstrate the use of the modified methodology. Keeping sharp lookout, judgment of the risk of collision, and taking anti-collision actions are key sub-tasks in collision avoidance for the involved officer on watch (OOW). Once errors or failures happened in these sub-tasks, a collision accident could occur. The question as to how the probability of the errors is obtained in a quantitative form needs to be answered.

4.1. General analysis

4.1.1. Evaluation of CPCs by experts

First of all, two experts were invited to analyse the investigated case through a collaborative research project between Shanghai Maritime University and Shanghai Pilot Station. Therefore, they both well knew the above defined investigation context and were selected to evaluate the CPCs and provide their DoBs with respect to each CPC level. Their professional positions, service time and education backgrounds are described in Table 5. Table 6 shows the evaluations of each CPC given by the two selected experts.

Table 5
Description of the selected experts

The expert	Description of the expert selected
E_1	A professor with a PhD degree who has been engaged in marine risk research for over 20 years and has a good understanding of organization of shipping and marine operations. He has also served as a chief mate on board ship continuously for 1 year during his professional career.
E_2	A captain with a bachelor degree who has worked on board ship for 18 years (2 years of being a captain) and has a good safety record.

Table 6
The states of 9 CPCs and the evaluation from two experts

CPC	Description of situations	Level	β_{in}^1	β_{in}^2
CPC₁ *AOO*	The seafarers are organized as a framework of management-operational-support level. There is a validated Safety Management System and clear responsibilities to the crew members. The shipping company has a complete organizational structure and is able to provide the powerful shore-based support for the ships.	L_{11}	0.6	0.5
		L_{12}	0.4	0.3
		L_{13}	0	0.2
		L_{14}	0	0
CPC₂ *WCO*	The bridge is small. The temperature is 30℃ due to that the air conditioner is temporarily not working. Wind is SE 10-12m/s. The vessel is slightly pitching and rolling. Environmental noise caused by VHF communications is serious.	L_{21}	0	0
		L_{22}	0.4	0.7
		L_{23}	0.6	0.3
CPC₃ *MMI*	The interface of the control panel including the monitors, knobs and buttons of the navigational aids (Radar, ARPA, AIS, GPS, engine telegraph, automatic helm, and so on) is relatively friendly.	L_{31}	0	0
		L_{32}	0.8	0.6
		L_{33}	0.2	0.3
		L_{34}	0	0.1
CPC₄ *AOP*	The checklists and procedures for navigation in heavy weather, restricted visibility, narrow channels and coast waters in SMS are well-defined, however the planning of sailing in harbour areas is slightly insufficient, particularly the awareness and preparation for heavy traffic.	L_{41}	0.8	0.7
		L_{42}	0.2	0.2
		L_{43}	0	0.1
CPC₅ *NOS*	There are multiple simultaneous goals due to the heavy traffic density. The duty officer sometimes must contact with other vessels and take an action to avoid collision with fishing vessels at the same time.	L_{51}	0	0
		L_{52}	1	1
		L_{53}	0	0
CPC₆ *ATI*	Ship speed is reduced according to the speed limitation of traffic regulation, so the captain or officer has enough time to change the speed or alter the course for anti-avoidance. But sometimes the operator is under a little high time pressure.	L_{61}	0.6	0.7
		L_{62}	0.4	0.3
		L_{63}	0	0
CPC₇ *TOD*	The time of the operation is from 1600 to 1900. The officer feel a little tired due to the preparation for sea at last midnight after discharging and loading.	L_{71}	0.4	0.6
		L_{72}	0.6	0.4
		L_{73}	0	0
CPC₈ *AOT*	The crew members with competence certifications have been trained at a maritime university and on-board. They familiar with the responsibilities. It is the 2nd time for the	L_{81}	0	0.2
		L_{82}	0.8	0.8

	officer on watch (C/O) navigating in Shanghai waterways and coaster waters.	L_{83}	0.2	0
CPC_9 *CCQ*	The captain, duty officer and helmsman have worked for the same shipping company for several years. They have been on board this ship and worked together for 2 months.	L_{91}	0.1	0
		L_{92}	0.8	0.6
		L_{93}	0.1	0.4
		L_{94}	0	0

4.1.2. Combination of the evaluations

The weights of the two experts are assigned as (agreed among them) $w_{1}= 0.6$ and $w_{2}=0.4$ according to their background information in Table 5. The basic probability masses for E_1 and E_2 can be modelled using Eqs. (2) - (5). Taking CPC_1 as an example, the basic probability masses are computed as Eqs. (23) and (24).

$$E_1: \begin{cases} m_{11}^1 = w_1\beta_{11}^1 = 0.6 \times 0.6 = 0.36 \\ m_{12}^1 = w_1\beta_{12}^1 = 0.6 \times 0.4 = 0.24 \\ m_{13}^1 = w_1\beta_{13}^1 = 0.6 \times 0 = 0 \\ m_{14}^1 = w_1\beta_{14}^1 = 0.6 \times 0 = 0 \\ \quad m_{L_1}^1 = 1 - \sum_{n=1}^{N} m_{1n}^1 = 1 - (0.36 + 0.24 + 0 + 0) = 0.4 \\ \quad \bar{m}_{L_1}^1 = 1 - w_1 = 1 - 0.6 = 0.4 \\ \tilde{m}_{L_1}^1 = w_1(1 - \sum_{n=1}^{N} \beta_{1n}^1) = 0.6 \times \left(1 - (0.6 + 0.4 + 0 + 0)\right) = 0 \end{cases} \quad (23)$$

$$E_2: \begin{cases} m_{11}^2 = w_2\beta_{11}^2 = 0.4 \times 0.5 = 0.20 \\ m_{12}^2 = w_2\beta_{12}^2 = 0.4 \times 0.3 = 0.12 \\ m_{13}^2 = w_2\beta_{13}^2 = 0.4 \times 0.2 = 0.08 \\ m_{14}^2 = w_2\beta_{14}^2 = 0.4 \times 0 = 0 \\ \quad m_{L_1}^2 = 1 - \sum_{n=1}^{N} m_{1n}^2 = 1 - (0.20 + 0.12 + 0.08 + 0) = 0.6 \\ \quad \bar{m}_{L_1}^2 = 1 - w_2 = 1 - 0.4 = 0.6 \\ \tilde{m}_{L_1}^2 = w_2(1 - \sum_{n=1}^{N} \beta_{1n}^2) = 0.4 \times \left(1 - (0.6 + 0.4 + 0 + 0)\right) = 0 \end{cases} \quad (24)$$

The basic probability masses m_{11}^1, m_{12}^1, m_{13}^1 and m_{14}^1 in Eq. (23), measure the DoBs exactly assigned to L_{11}, L_{12}, L_{13} and L_{14} by E_1. The probability mass of E_1 ($m_{L_1}^1 = 0.4$) unassigned to any individual L_{1n} is split into two parts, one caused by the relative importance of E_1 ($\bar{m}_{L_1}^1 = 0.4$), and the other due to the incompleteness of the belief degree assessment β_{1n}^1 ($\tilde{m}_{L_i}^g$=0). Similarly, the basic probability masses from E_2 can be modelled in Eq. (24). Using the same approach, the basic probability masses for the levels of other CPCs from E_1 and E_2 can be obtained.

Next, the combination of the basic probability masses from the two experts can be conducted using Eqs. (6) - (10).

$$\{L_{1n}\}: \begin{cases} \{L_{11}\}: m_{11}^{I(2)} = K_{I(1+1)}\left[m_{11}^{I(1)}m_{11}^2 + m_{11}^{I(1)}m_{L_1}^2 + m_{L_1}^{I(1)}m_{11}^2\right] \\ \qquad = 1.1617[0.36 \times 0.20 + 0.36 \times 0.6 + 0.4 \times 0.20] = 0.4275 \\ \{L_{12}\}: m_{12}^{I(2)} = K_{I(1+1)}\left[m_{12}^{I(1)}m_{12}^2 + m_{12}^{I(1)}m_{L_1}^2 + m_{L_1}^{I(1)}m_{12}^2\right] \\ \qquad = 1.1617[0.24 \times 0.12 + 0.24 \times 0.6 + 0.4 \times 0.12] = 0.2565 \\ \{L_{13}\}: m_{13}^{I(2)} = K_{I(1+1)}\left[m_{13}^{I(1)}m_{13}^2 + m_{13}^{I(1)}m_{L_1}^2 + m_{L_1}^{I(1)}m_{13}^2\right] \\ \qquad = 1.1617[0 \times 0.08 + 0 \times 0.6 + 0.4 \times 0.08] = 0.0372 \\ \{L_{14}\}: m_{14}^{I(2)} = K_{I(1+1)}\left[m_{14}^{I(1)}m_{14}^2 + m_{14}^{I(1)}m_{L_1}^2 + m_{L_1}^{I(1)}m_{14}^2\right] \\ \qquad = 1.1617[0 \times 0 + 0 \times 0.6 + 0.4 \times 0] = 0 \end{cases} \quad (25)$$

$$\overline{m}_{L_1}^{I(2)} = K_{I(2)}\left[\overline{m}_{L_i}^{I(1)}\overline{m}_{L_i}^{2}\right] = 1.1617[0.4 \times 0.6] = 0.2788 \quad (26)$$

$$\{L_1\}:\ \widetilde{m}_{L_1}^{I(2)} = K_{I(2)}\left[\widetilde{m}_{L_1}^{I(1)}\widetilde{m}_{L_1}^{2} + \widetilde{m}_{L_1}^{I(1)}\overline{m}_{L_1}^{2} + \overline{m}_{L_1}^{I(1)}\widetilde{m}_{L_1}^{2}\right] = 1.1617[0 \times 0 + 0 \times 0.6 + 0.4 \times 0] = 0 \quad (27)$$

$$K_{I(2)} = \left[1 - \sum_{n=1}^{N}\sum_{\substack{h=1 \\ h\neq n}}^{N} m_{1n}^{I(1)} m_{1h}^{2}\right]^{-1}$$

$$= [1 - 0.36 \times 0.12 + 0.36 \times 0.08 + 0.36 \times 0 + 0.24 \times 0.20 + 0.24 \times 0.08 + 0.24 \times 0 + 0 \times 0.20 + 0 \times 0.12 + 0 \times 0 + 0 \times 0.20 + 0 \times 0.12 + 0 \times 0.08]^{-1} = 1.1617 \quad (28)$$

$$\{L_{1n}\}: \begin{cases} \{L_{11}\}: \beta_{11} = \dfrac{m_{11}^{I(2)}}{1-\overline{m}_{L_1}^{I(2)}} = \dfrac{0.4275}{1-0.2788} = 0.5928 \cong 0.59 \\ \{L_{12}\}: \beta_{12} = \dfrac{m_{12}^{I(2)}}{1-\overline{m}_{L_1}^{I(2)}} = \dfrac{0.2565}{1-0.2788} = 0.3557 \cong 0.36 \\ \{L_{13}\}: \beta_{13} = \dfrac{m_{13}^{I(2)}}{1-\overline{m}_{L_1}^{I(2)}} = \dfrac{0.0372}{1-0.2788} = 0.0516 \cong 0.05 \\ \{L_{14}\}: \beta_{14} = \dfrac{m_{14}^{I(2)}}{1-\overline{m}_{L_1}^{I(2)}} = \dfrac{0}{1-0.2788} = 0 \end{cases} \quad (29)$$

$$\{L_1\}: \beta_{L_1} = \frac{\widetilde{m}_{L_1}^{I(2)}}{1-\overline{m}_{L_1}^{I(2)}} = \frac{0}{1-0.2788} = 0 \quad (30)$$

As a result, the combined DoBs of L_{11}, L_{12}, L_{13} and L_{14} of CPC_1 are calculated as 0.5928, 0.3557, 0.0516 and 0 respectively. Similarly, the combined DoBs associated with other CPCs can be obtained and presented in Table 7.

Table 7
The combined DoBs of each level of 9 CPCs

CPCs	*DoBs*
CPC_1	(L_{11}, 0.59); (L_{12}, 0.36); (L_{13}, 0.05); (L_{14}, 0)
CPC_2	(L_{21}, 0); (L_{22}, 0.51); (L_{23}, 0.49)
CPC_3	(L_{31}, 0); (L_{32}, 0.77); (L_{33}, 0.21); (L_{34}, 0.02)
CPC_4	(L_{41}, 0.80); (L_{42}, 0.17); (L_{43}, 0.03)
CPC_5	(L_{51}, 0); (L_{52}, 1); (L_{53}, 0)
CPC_6	(L_{61}, 0.66); (L_{62}, 0.34); (L_{63}, 0)
CPC_7	(L_{71}, 0.47); (L_{72}, 0.53); (L_{73}, 0)
CPC_8	(L_{81}, 0.05); (L_{82}, 0.84); (L_{83}, 0.11)
CPC_9	(L_{91}, 0.06); (L_{92}, 0.77); (L_{93}, 0.17); (L_{94}, 0)

4.1.3. HEP of the whole task

Next, the couple ($\sum_{Improved}$, $\sum_{Reduced}$) can be calculated by summing up the DoBs of the improved and reduced effects. $\sum_{Improved}$=0.59+0+0+0.80+0.66+0.05+0.06=2.16 and $\sum_{Reduced}$=0.05+0+0.49+0.02+0.03+0+0+0.53+0+0.11+0=1.23. ($\sum_{Improved}$, $\sum_{Reduced}$) are obtained as (2.16, 1.23). For the specific anti-collision task in Shanghai coastal waters, the control mode is estimated as tactical and the corresponding probability interval is (0.001, 0.1). Furthermore, the value of context variable *x* is 2.16-1.23=0.93 in a general analysis.

Consequently, the associated error probability (per anti-collision operation under the proposed context is computed as follows using Eq. (15).

$$HEP = 3.81\times10^{-3}\, exp\,(-0.619x) = 0.00381exp\,(-0.619\times0.93) = 2.14\times10^{-3}$$

This point value is located in the probability interval (0.001, 0.1) of the tactical control mode in CREAM.

4.2. Extended quantification of HEP

4.2.1. The weights of CPCs

The two experts were also invited to evaluate the relationship among CPCs in the defined context under investigation. The two 9×9 matrices containing the two experts' judgements on CPCs dependency are presented as follows.

$$A_1 = \begin{bmatrix} 0 & 2 & 2.5 & 2 & 0 & 0 & 0 & 2 & 1.5 \\ 0 & 0 & 0 & 0 & 2 & 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1.5 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2.5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},\ A_2 = \begin{bmatrix} 0 & 1.5 & 2 & 2 & 0 & 0 & 0 & 1.5 & 2 \\ 0 & 0 & 0 & 0 & 2.8 & 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 1.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.5 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The value of 2 in the first row and second column of matrix A_1 indicates that expert E_1 believes CPC_1 has "medium" influence on CPC_2 under the given scenario. This evaluation is also made by refereeing the influence relationship defined in CREAM (in Table 3). The numerical values at the diagonal of each matrix are 0 because a CPC cannot affect itself.

Using Eq. (16), the average matrix of A_1 and A_2 ($w_{1}= 0.6, w_{2}=0.4$) is obtained as follows.

$$M = \begin{bmatrix} 0 & 1.80 & 2.30 & 2.00 & 0 & 0 & 0 & 1.80 & 1.70 \\ 0 & 0 & 0 & 0 & 2.32 & 2.40 & 0 & 0 & 0 \\ 0 & 1.40 & 0 & 0 & 1.30 & 1.80 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1.50 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2.30 & 0 & 0 & 0 \\ 0 & 0.80 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.20 & 0 & 0 & 0 & 1.60 & 0 & 0 & 0 \\ 0 & 0.80 & 0 & 0 & 0 & 0 & 0 & 0 & 2.00 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Next the normalized direct matrix is calculated by Eq. (17).

$$X = \begin{bmatrix} 0 & 0.1875 & 0.2396 & 0.2083 & 0 & 0 & 0 & 0.1875 & 0.1771 \\ 0 & 0 & 0 & 0 & 0.2417 & 0.2500 & 0 & 0 & 0 \\ 0 & 0.1458 & 0 & 0 & 0.1354 & 0.1875 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.1042 & 0.1563 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.2396 & 0 & 0 & 0 \\ 0 & 0.0833 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.1250 & 0 & 0 & 0 & 0.1667 & 0 & 0 & 0 \\ 0 & 0.0833 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2083 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Consequently, the total relation matrix T is obtained as follows:

$$T = \begin{bmatrix} 0 & 0.2521 & 0.2396 & 0.2083 & 0.1151 & 0.1681 & 0 & 0.1875 & 0.2161 \\ 0 & 0.0263 & 0 & 0 & 0.2480 & 0.3160 & 0 & 0 & 0 \\ 0 & 0.1685 & 0 & 0 & 0.1761 & 0.2718 & 0 & 0 & 0 \\ 0 & 0.0155 & 0 & 0 & 0.1079 & 0.1860 & 0 & 0 & 0 \\ 0 & 0.0205 & 0 & 0 & 0.0050 & 0.2459 & 0 & 0 & 0 \\ 0 & 0.0855 & 0 & 0 & 0.0207 & 0.0263 & 0 & 0 & 0 \\ 0 & 0.1425 & 0 & 0 & 0.0344 & 0.2106 & 0 & 0 & 0 \\ 0 & 0.0855 & 0 & 0 & 0.0207 & 0.0263 & 0 & 0 & 0.2083 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

According to Section 3.1.3, R_i is the sum of the *i*th row of matrix T and C_i is the sum of *i*th column of matrix T. Values of R_i, C_i, (R_i+C_i) and (R_i-C_i) are obtained and presented in Table 8 according to Eq. (18)

Table 8
The R_i , C_i , (R_i+C_i)and (R_i-C_i) values of CPCs

CPCs	CPC_1	CPC_2	CPC_3	CPC_4	CPC_5	CPC_6	CPC_7	CPC_8	CPC_9
R_i	1.3868	0.5904	0.6164	0.3094	0.2713	0.1325	0.3876	0.3409	0
C_i	0	0.7965	0.2396	0.2083	0.7279	1.4510	0	0.1875	0.4245
R_i+C_i	1.3868	1.3869	0.8560	0.5177	0.9992	1.5835	0.3876	0.5284	0.4245
R_i-C_i	1.3868	-0.2061	0.3769	0.1011	-0.4565	-1.3185	0.3876	0.1534	-0.4245

In the second column of Table 8, the value of 1.3868 in the second row indicates both the direct and indirect effects given by CPC_1. The value of 0 in the third row represents both direct and indirect effects received by CPC_1 from the other CPCs. The value of 1.3868 in the fourth row shows the total effects produced by CPC_1. The value of 1.3868 in the fifth row means the net effect that CPC_1 contributes to the system. Thus, the importance of the nine CPCs can be prioritized as C6>C2>C1>C5>C3>C8>C4>C9>C7 based on their individual (*R*+*C*) values. The weights (v_i) of CPCs (Table 9) can also be calculated using Eq. (19).

4.2.2. The adjusting index of CPC levels

Next the adjusting index (denoted as δ_{in}) of CPCs' levels can be determined according to the rules described in Section 3.2. For example, the weight of CPC_1 *AOO* is 1.55. The adjusting index of "*very efficient*", "*efficient*" and "*deficient*" can therefore be 1.55, 0 and -1.55 due to their individual most positive effect, neutral effect and most negative effect. For the level of "*inefficient*", its effect on human reliability is negative, but the degree of effect is less than "*deficient*". Its adjusting index is thus calculated as 0.78 using a linear distribution. In a similar way, all adjusting indices (Table 9) of the 9 CPCs' levels can be determined.

Table 9
The weights of CPCs and adjusting index of CPCs' levels

CPCs	CPC_1	CPC_2	CPC_3	CPC_4	CPC_5	CPC_6	CPC_7	CPC_8	CPC_9
v_i	1.55	1.55	0.95	0.58	1.11	1.77	0.43	0.59	0.47
δ_{i1}	1.55	1.55	0.95	0.58	0	1.77	0	0.59	0.47
δ_{i2}	0	0	0	0	-0.56	0	-0.22	0	0
δ_{i3}	-0.78	-1.55	0	-0.58	-1.11	-1.77	-0.43	-0.59	-0.24
δ_{i4}	-1.55	--	-0.95	--	--	--	--	--	-0.47

4.2.3. CFP of sub-tasks of ship collision avoidance

The specific effects of CPCs can be quantified using Eq. (20). For example,

$$x_1 = \sum_{n=1}^{N} \delta_{1n}\beta_{1n} = (1.55 \times 0.59) + (\ 0 \times 0.36) + (\ -0.78 \times 0.05) + (-1.55 \times 0) = 0.8755.$$

Similarly, the specific effects of all 9 CPCs are obtained as

$$[x_i] = [0.8755, -0.7595, -0.0190, 0.4466, -0.5600, 1.1682, -0.1166, -0.0354, -0.0126].$$

Consequently, the total context influence on CFP is obtained using Eq. (21).

$$x = \sum_{i=1}^{9} x_i = 0.99.$$

Compared to the context variable value of 0.93 in the general analysis, 0.99 in extended quantification analysis is slightly larger. The difference between two values is caused by the introduction of the different weights of CPCs and adjusting indices of their ranking levels when considering their dependency in specific task context.

In Table 10, the whole task of ship anti-collision is divided into sub-tasks and possible errors are identified. Cognitive functions of sub-tasks and generic failure types of possible errors are analyzed to determine the basic value of each possible error. Then the adjusted probabilities of possible errors can be calculated using Eq.22.

Table 10 Extended analysis and quantification for ship collision avoidance

Sub-task in ship collision avoidance		Possible errors	Cognitive function	Generic failure type	CFP_0	Adjusted CFP
1 Observation	1.1 Visual observation	Not maintaining Continuous visual lookout	Observation	O3 Observation not made	7.00E-02	3.79E-02
	1.2 Observation via navigational aids	Inadequate use of aids	Observation	O2 Wrong identification	7.00E-02	3.79E-02
	1.3 Communication with other ship	Communication interrupt	Execution	E5 Missed action	3.00E-02	1.63E-02
	1.4 Comparison of information	Comparison error	Interpretation	I3 Delayed interpretation	1.00E-02	5.42E-03
2 Judgment	2.1 Judgment of Risk of collision	Wrong judgment	Interpretation	I1 Faulty diagnosis	2.00E-01	1.08E-01
	2.2 Evaluation of situation	Wrong judgment	Interpretation	I1 Faulty diagnosis	2.00E-01	1.08E-01
	2.3 Choice of applicable items in COLREG 72	Select wrong items	Planning	P2 Inadequate plan	1.00E-02	5.42E-03
3 Decision-making	3.1 Decision of action to avoid collision	Wrong action	Interpretation	I2 Decision error	1.00E-02	5.42E-03
	3.2 Decision of time of action	Wrong time	Planning	P1 Priority error	1.00E-02	5.42E-03
4 Action	4.1 Alter course or/and speed	Improper implementation	Execution	E1 Action of wrong type	3.00E-03	1.63E-03

4.3. Model validation and result comparison

4.3.1. Validation using sensitivity analysis

Sensitivity analysis is a study on how the uncertainty in the output of a model can be apportioned to different sources of uncertainty in the model input. It is a method for testing the degree of sensitivity of a model's variables. The sensitivity analysis was carried out by changing the evaluation of CPC_1 by E_1. The original expert evaluation values associated with L_{11} were decreased from 0.6 to 0.1 at a step of 0.1. The assumption made is that the trend line of certain CFPs should increase steadily. The trend lines of selective CFPs are illustrated in Fig.2.

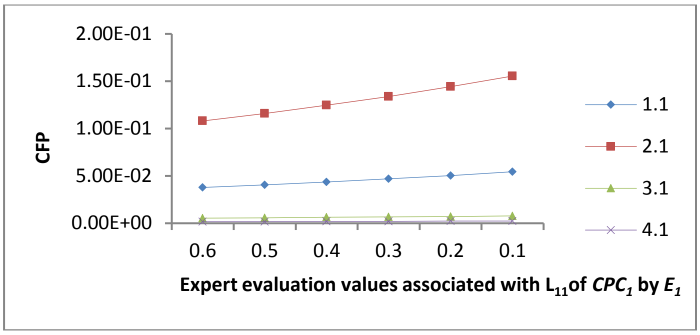

Fig.2 The trend lines of selective CFPs

A similar proposed approach is applied to determine if small changes of input variables will induce the corresponding changes in the output results of CFPs. It can be observed from the trend lines in Figure 2 that four selective CFPs have increased correspondingly with increment of evaluation values. Additionally, no abrupt large changes of CFPs were generated in the sensitivity analysis. Therefore, the model is able to illustrate the differences and has partially validated the model developed.

4.3.2. Result comparison

The cognitive model and framework of CREAM could be used in retrospective and prospective analysis. However, the quantification results provided are mainly based on the expert specification on CPCs. Lack of data, which makes it inevitable to adopt expert judgment, is the most important source of the uncertainty for many HRA methods (He et al., 2008) including the proposed method. The knowledge and experience imparity of experts leads to evaluation differences on CPCs and further on HRA analysis results. Therefore, the ER algorithm is used to deal with the differential judgments of experts. The quantification results derived from E_1, E_2 and ER algorithm are compared respectively. From Table 11, we can find that the CFP values derived from ER algorithm fall between the values resulted from two experts. It shows the strength of ER algorithm in the proposed method on dealing with the uncertainties from experts' judgments.

Table 11 Comparison of results

Sub-tasks		CFP_0	By E_1 CFP_1	By E_2 CFP_2	Combined CFP
1	1.1	7.00E-02	4.43E-02	3.72E-02	3.79E-02
	1.2	7.00E-02	4.43E-02	3.72E-02	3.79E-02
	1.3	3.00E-02	1.90E-02	1.60E-02	1.63E-02
	1.4	1.00E-02	6.33E-03	5.32E-03	5.42E-03
2	2.1	2.00E-01	1.27E-01	1.06E-01	1.08E-01
	2.2	2.00E-01	1.27E-01	1.06E-01	1.08E-01
	2.3	1.00E-02	6.33E-03	5.32E-03	5.42E-03
3	3.1	1.00E-02	6.33E-03	5.32E-03	5.42E-03
	3.2	1.00E-02	6.33E-03	5.32E-03	5.42E-03
4	4.1	3.00E-03	1.90E-03	1.60E-03	1.63E-03

Additionally, real data has been collected to validate the above result. Since Sep. 30th 2011 until Oct. 1st 2012, the operational data under the same context in Shanghai coastal waters has been collected and recorded. From the recorded data, it is revealed that from 2688 events that could lead to the error of "not maintaining continuous visual lookout or/and inadequate use of aids", 81 errors actually occurred (Fang et al., 2012). The error rate is 0.0301. The closeness of 0.0379 and 0.0301 shows a fairly good consistency between the estimated value and the real error rate, thus also provide evidence to support the research findings and proposed quantification method in this paper.

5. Conclusion

A new HEP quantification methodology is proposed based on the ER approach, DEMATEL and CREAM. It can be used for not only a general analysis but also the quantitative evaluation of HRA in maritime operations. The procedure of quantification is described and an illustrative case study is produced. A sensitivity analysis and comparison of quantification results is carried out for the validation of the newly proposed model.

The belief degree approach, instead of the deterministic one-or-zero way in specification of CPCs, can well model the uncertainty and be more practical when adopting experts' judgments. The ER algorithm can combine the different evaluations to improve the limitations of knowledge and experience of every expert in CPCs specification. The weights generated by DEMATEL and adjusted indices derived from the weights express the various influences of CPCs in a practical operation. Therefore, compared to the previous CREAM based HEP quantification studies, the new method has shown its superiority due to the fact that (1) the ER approach with the belief degree manner can deal with the uncertainties caused by insufficient information and data in the evaluation of CPCs and be capable of combining the DoBs from different sources; (2) setting of the adjusting indices based on the DEMATEL technique allows different importance of CPCs and more sensitive estimation of HEP.

This paper focuses on the development of the new modified CREAM method and its application in the maritime industry. The propagation of an event is made through the representation of a scenario of maritime operation. It is believed that the method provides useful insights on a systemic and structural cognitive task analysis of complex systems and hence will facilitate cognitive HRA in other sectors to enhance incorporation of human element in the safety management of critical systems. However, it only focuses on how seafarers' actions can fail, and concerns the human element in seafarers' reliability analysis. It has therefore revealed its applicability and limits in practice, with respect to Prof Eric Hollnagel's disclaimers on CREAM in his personal website.

Acknowledgements

This research was supported by an EU Marie Curie grant (ENRICH – 612546). The authors would like to thank Shanghai Pilot Station for its support in data provision. The authors would also like to thank the four anonymous reviewers for their constructive suggestions. The usual disclaimers apply.

References

1. Akyuz E. Quantification of human error probability towards the gas inerting process on-board crude oil tankers. Safety Science 2015; 80:77-86.
2. Akyuz E. Quantitative human error assessment during abandon ship procedures in maritime transportation. Ocean Engineering 2016; 120: 21-29.
3. Akyuz E, Celik E. A modified human reliability analysis for cargo operation in single point mooring (SPM) off-shore units. Applied Ocean Research 2016; 58:11-20.
4. Akyuz E, Celik M. Application of CREAM human reliability model to cargo loading process of LPG tankers. Journal of Loss Prevention in the Process Industries 2015; 34: 39-48.
5. Akyuz E. and Celik M. A methodological extension to human reliability analysis for cargo tank cleaning operation on board chemical tanker ships. Safety Science 2015; 75:146-155.

6. Alvarenga, MAB, Frutuoso e Melo PF, Fonseca RAA. Critical review of methods and models for evaluating organizational factors in human reliability analysis. Progress in Nuclear Energy 2014; 75: 25-41.
7. Apostolakis GE, Bier VM, Mosleh A. A critique of recent models for human error rate assessment. Reliability Engineering & System Safety 1988; 22:201-217.
8. Baraldi P, Podofillini L, Mkrtchyan L, Zio E, Dang VN. Comparing the treatment of uncertainty in Bayesian networks and fuzzy expert systems used for a human reliability analysis application. Reliability Engineering & System Safety 2015;138:176–193.
9. Calhoun J, Savoie C, Randolph-Gips M, Bozkurt I. Human reliability analysis in spaceflight applications. Quality and Reliability Engineering International 2013; 29(6): 869-882.
10. Calhoun J, Savoie C, Randolph-Gips M, Bozkurt I. Human reliability analysis in spaceflight applications, Part 2: modified CREAM for spaceflight. Quality and Reliability Engineering International 2014; 30(1): 3-12.
11. Chen ST, Wall A, Philip D, Yang Z, Wang J. A human and organisational factors (HOFs) analysis method for marine casualties using HFACS-maritime accidents (HFACS-MA). Safety Science 2013; 60: 105-114.
12. Cooper S E, Ramey-Smith A M, Wreathall J, et al. A Technique for Human Error Analysis (ATHEANA). NUREG/CR-6350, 1996.
13. Dempster AP. Upper and lower probabilities induced by a multi-valued mapping. Annals of Mathematical Statistics 1967; 38:325–339.
14. Embrey DE, Humphreys P, Rosa EA, Kirwan B, Rea K. SLIM-MADUD: an approach to assessing human error probabilities using structured expert judgment. NUREG/CR-3518, US Nuclear Regulatory Publication; 1984.
15. Fang Q.G., Hu S.P., Zhang J.P., Xi Y.T.. Report of ship-pilotage formal safety assessment (FSA) in Shanghai costal waters (2007-2012). Shanghai Maritime University / Shanghai pilot station; 2012.
16. Ford D, Serman JD. Expert knowledge elicitation to improve mental and formal models. Systems Dynamics Review 1998; 14: 309-340.
17. Fontela E, Gabus A. The DEMATEL observer, DEMATEL observer 1976 report. Geneva: Battelle Geneva Research Centre; 1976.
18. Fu C, Yang SL. The group consensus based evidential reasoning approach for multiple attribute group decision analysis. European Journal of Operational Research 2010; 206: 601–608.
19. Fujita Y, Hollnagel E. Failures without errors: quantification of context in HRA. Reliability Engineering & System Safety 2004; 83:145-151.
20. Gabus A, Fontella E. Perceptions of the world: Communication Procedure and Communicating with Those Bearing Collective Responsibility (DEMATEL Report No. 1). Geneva: Battelle Geneva Research Centre; 1973.
21. Gertman D, Blackman H, Marble J, Byers J, Smith C. The SPAR-H human reliability analysis method. Idaho National Laboratory; 2005.
22. Graziano A., Teixeira A.P., Guedes Soares C. Classification of human errors in grounding and collision accidents using the TRACEr taxonomy. Safety Science 2016; 86:245-257.
23. Hannaman GW, Spurgin AJ, Lukic YD. Human cognitive reliability model for PRA analysis. NUS-4531, Electric Power Research Institute; 1984.
24. He X, Wang Y, Shen Z, Huang X. A simplified CREAM prospective quantification process and its application. Reliability Engineering & System Safety 2008; 93: 298-306.

25. Hollnagel E. Cognitive reliability and error analysis method (CREAM). Amsterdam: Elsevier; 1998.
26. Jang I, Kim AR, Al Harbi MAS, Lee SJ, Kang HG, Seong PH. An empirical study on the basic human error probabilities for NPP advanced main control room operation using soft control. Nuclear Engineering and Design 2013; 257: 79-87.
27. Kim M, Seong P, Hollnagel E. A probabilistic approach for determining the control mode in CREAM. Reliability Engineering & System Safety 2006; 91:191-199.
28. Konstandinidou M, Nivolianitou Z, Kiranoudis C, Markatos N. A fuzzy modeling application of CREAM methodology for human reliability analysis. Reliability Engineering & System Safety 2006; 91:706–716.
29. Lee SM, Ha JS, Seong PH. CREAM-based communication error analysis method (CEAM) for nuclear power plant operators' communication. Journal of Loss Prevention in the Process Industries 2011; 24:90-97.
30. Marseguerra M, Zio E, Librizzi M. Quantitative developments in the cognitive reliability and error analysis method (CREAM) for the assessment of human performance. Annals of Nuclear Energy 2006; 33:894–910.
31. Marseguerra M, Zio E, Librizzi M. Human Reliability Analysis by Fuzzy "CREAM". Risk Analysis 2007; 27(1): 137-154.
32. Nan C, Sansavini G. Developing an agent-based hierarchical modeling approach to assess human performance of infrastructure systems. International Journal of Industrial Ergonomics 2016; 53: 340-354.
33. Pennie, D., Brook-Carter, N., Gibson, W. Human factors guidance for maintenance. In: Human Factors in Ship Design, Safety and Operation Conference. The Royal Institution of Naval Architects, London, UK, 2007.
34. Ribeiro A.C., Sousa A.L., Duarte J.P., Frutuoso e Melo P.F. Human reliability analysis of the Tokai-Mura accident through a THERP–CREAM and expert opinion auditing approach. Safety Science 2016; 87:269-279.
35. Seyed-Hosseini SM, Safaei N, Asgharpour MJ. Reprioritisation of failures in a system failure mode and effects analysis by decision making trial and evaluation laboratory technique. Reliability Engineering & System Safety 2006; 91:872-881.
36. Shafer G. A Mathematical theory of evidence. Princeton: Princeton University Press, 1976.
37. Shieh JI, Wu HH, Huang KK. A DEMATEL method in identifying key success factors of hospital service quality. Knowledge-Based Systems 2010; 23: 277-282.
38. Sun Z, Li Z, Gong E, Xie H. Estimating human error probability using a modified CREAM. Reliability Engineering & System Safety 2012; 100:28-32.
39. Swain AD, Guttmann HE. Handbook of human reliability analysis with emphasis on nuclear power plant applications. Washington, DC: Nuclear Regulatory Commission; 1983.
40. Tamura H, Akazawa K. Stochastic DEMATEL for structural modelling of a complex problematique for realizing safe, secure and reliable society. Journal of Telecommunications & Information Technology 2005; 4:139-146.
41. Tseng M. Using the extension of DEMATEL to integrate hotel service quality perceptions into a cause–effect model in uncertainty. Expert Systems with Applications 2009; 36:9015-9023.
42. Tzeng G, Chiang C, Li C. Evaluating intertwined effects in e-learning programs: a novel hybrid MCDM model based on factor analysis and DEMATEL. Expert Systems with Applications 2007; 32:1028-1044.
43. Ung S.T. A weighted CREAM model for maritime human reliability analysis. Safety Science; 72: 144-152.

44. UK P&I Club. The United Kingdom Mutual Steam Ship Assurance Association (Bermuda) Limited. Analysis of Major Claims; 1992.
45. Wang A, Luo Y, Tu G, Liu P. Quantitative Evaluation of Human-Reliability Based on Fuzzy-Clonal Selection. IEEE Transactions on Reliability 2011, 60(3): 517-527.
46. Wang YF, Xie M, Ng KM, Habibullah MS. Probability analysis of offshore fire by incorporating human and organizational factor. Ocean Engineering 2011; 38: 2042-2055.
47. Wang YM, Yang JB, Xu DL. Environmental impact assessment using the evidential reasoning approach. European Journal of Operational Research 2006; 174(3): 1885–1913.
48. Williams JC. A data-based method for assessing and reducing human error to improve operational performance. In: Proceedings of IEEE 4th conference on human factor and power plants. 1988. Monterey, California: p. 436–53.
49. Wu W. Choosing knowledge management strategies by using a combined ANP and DEMATEL method. Expert Systems with Application 2008; 35: 828-835.
50. Wu W, Lee Y. Developing global managers' competencies using the fuzzy DEMATEL method. Expert Systems with Applications 2007; 32:499-507.
51. Xi Y, Fang Q, Hu S, Chen W. Evidential reasoning approach for CWA traffic safety assessment. Progress in Safety Science and Technology 2008; 7:84-91.
52. Xu DL, Yang JB, Wang YM. The evidential reasoning approach for multi-attribute decision analysis under interval uncertainty. European Journal of Operational Research 2006; 174:1914-1943.
53. Yang JB, Singh MG. An evidence reasoning approach for multiple attribute decision making with uncertainty. IEEE transactions on systems, man and cybernetics 1994; 24(1): 1–18.
54. Yang JB, Wang YM, Xu DL, Chin KS. The evidential reasoning approach for MADA under both probabilistic and fuzzy uncertainties. European Journal of Operational Research 2006; 171:309–343.
55. Yang Z, Ng AKY, Wang J. Incorporating quantitative risk analysis in port facility security assessment. Transportation Research Part A: Policy and Practice 2014; 59: 72-90.
56. Yang ZL, Usman M, Bonsall S. Wall A and Wang J. A Modified CREAM to Human Reliability Quantification in Marine Engineering. Ocean Engineering 2013; 58: 293-303.
57. Yang Z, Wang J. Use of fuzzy risk assessment in FMEA of offshore engineering. Ocean Engineering 2015; 95: 195-204.
58. Yang ZL, Wang J, Bonsall S, Fang QG. Use of fuzzy evidential reasoning in maritime security assessment. Risk Analysis 2009; 29: 95-120.