IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 8, 2021, accepted July 2, 2021, date of publication July 26, 2021, date of current version August 9, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3099831

A New Risk-Based Early-Warning Method for Ship Collision Avoidance

YINGFAN CHEN[1], CHENG XIE[1], SHUZHE CHEN[1], AND LIWEN HUANG[1,2]

[1]School of Navigation, Wuhan University of Technology, Wuhan 430070, China
[2]Hubei Key Laboratory of Inland Shipping Technology, Wuhan 430070, China

Corresponding author: Cheng Xie (cxie@whut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51379170, and in part by the Fundamental Research Funds for the Central Universities under Grant WUT: 2021IVA009B.

ABSTRACT The ship collision accident (SCA) risk for any ship approaching any other change from the causation probability (CP) to the geometric probability (GP) in regime. Because ship operators may not be aware of the environmental factors (EFs) related potential risks in high CP during the initial stage of the GP analysis process, it is likely that higher-grade SCA measures will not be taken. However, if any EF-related CP is told to ship operators, they can take more effective and intentional measures in time; moreover, if the CP corresponding to navigation-related EFs is no less than the risk early warning critical value (REWCV) calculated based on historical SCA data, SCAs will be in a high-risk level. A new method was put forward here based on a quantitative analysis of EFs and previous SCA statistics to provide early warning of any SCA risk. On this basis, a REWCV is obtained based on quantified EFs by means of such method which is relatively simple but high operational and practical. A case study of Three Gorges Reservoir in China indicates that the range of EF values whose probability of a SCA grows rapidly is consistent with environmental limits defined by Chinese maritime standards. Moreover, the modified critical value of the EF-related CP shall be further refined to act as the REWCV for CAs. In addition, the relationship (REWCV vs. the number of previous SCAs) was clarified.

INDEX TERMS Causation probability, early warning, environmental factor, ship collision accident, risk.

I. INTRODUCTION

As the risk is a complex concept involving several uncertainties and risk impact factors, ship navigational risk is difficult to assess considering the involved probability of occurrence, consequence of loss and injury under an unpredictable circumstance. Thereby, a well-established model for assessing SCAs is necessary and it shall primarily consider the related risk factors such as EFs, and human and ship factors.

Based on the statements of International Regulations for Preventing Collisions at Sea (COLREGs), the following factors should be clarified in case of evaluating the SCA either under micro level and those safety factors are given as follows:

(1) Visibility;

(2) Traffic density including the intensity of fishing boats or any other types of vessels;

(3) Ship maneuverability (especially stroke and turning under the circumstances);

(4) Night background light (such as shore light or backscatter light of the light of the ship);

(5) Those situations such as wind, waves and currents or situations near any dangerous object; and

(6) The relationship between the draft and the available water depth.

As the selection of risk impact factors are not consensus, the SCA models may select various risk factors under a specific practice so as to cause a question that most of SCA models are complicated and the interaction among these factors are not clear. Therefore, it would be difficult to determine the basic CP under various EFs.

This study aims to present an early warning model where the risk impact under various EFs is taken into account. It assumes that probability of SCAs can be determined if the EFs for a certain scenario can be reconstructed to provide detailed information, while the potential risk can be assessed if the REWCV can be determined in the model.

The associate editor coordinating the review of this manuscript and approving it for publication was P. Venkata Krishna.

Thus, the potential SCA for local waters can be lowered. Moreover, this study utilizes different REWCVs to verify our SCA model when detecting potential collision candidates in an early warning stage.

This paper is organized as follows:

Section II: SCA-related literature review;

Section III: presentation of a quantitative calculation method for SCA;

Section IV: development of an early warning method;

Section V: a case testing our method;

Section VI: verification of our method; and

Section VII: Summary and conclusions.

II. LITERATURE REVIEW

The probability (P) of SCA is divided into two independent parts (namely CP (P_C) and GP (P_G)) [14], which are contrastive. P_C as a macroscopic concept means the probability of accidents which always occur but are not directed against any concrete ship, driver or pilot in a specific water area. As for a certain accident determined by multiple factors, its total P will rise while one or more factors may be worsened. Unfortunately, those accident ships may not be defined in a water area at the moment. P_C can be determined based on the contribution of each EF to the human-ship-environment system and is unrelated to geometry [9]. Alternatively, P_G as one microscopic description refers to the probability of accidents of specific objectives (primarily including ships, navigation marks and marine constructions) and drivers. On the other hand, those factors such as sizes, trajectories, speeds and courses of specific ships are defined, which focus on scales of ships and their motions to calculate P for ships and other obstacles. Thus, the microscopic field of each single ship is targeted.

Constant CPs were assumed for some certain waterbodies in a few studies so that their CPs may be used for analysis of accident cases. While the CP for a certain part of the waterbody is determined based on historical accident data, it can be applied to the entire waterbody [11]. The GP may depend on geometric parameters (such as the area of the waterbody, ship size, traffic volume, ship speed and course) and would be relevant only in particular accident cases. CPs correspond to management-level decision making and are suitable for performance of large-scale analysis. Rather, GPs correspond to operation-level decision making and are more accurate in a small scale level. P (the probability of collision accidents) is expressed as [15]:

$$P = P_G \times P_C \tag{1}$$

CP or GP may be generally focused interactively depending on the scale of any sailing ship. Typically, CPs are initially assumed for a large scale. However, an accurate geometric model is necessary for SCA of any ship close to any other one [6].

As the distance between the ship in question (hereafter referred to as the ship) and other ships changes, P also changes so that P shall depend on EFs related to the human-ship-environment system. While ships are spaced in a large distance and even though there might be various EFs for the ship, operators may not notice potential navigation environment risks due to absence of SCA objects. The risk is mostly determined for a collision case based on experiences of operators. While they had not been aware of the risk for each EF, necessary measures would not be taken to prevent SCA and guarantee safety. While human factors were not aware of existence of risks, their levels of defense would have been actively raised so that human factors shall not be dominant in such a situation. As mentioned in several accident samples below, environmental factors always exist while the combined effects of human and ship factors may be weakened while positive and negative effects of various factors are superimposed. Thus, the overall effects of human factors are also attenuated. While the ship tends to collide with any other ship, operators become aware of the risk and human factors are very useful in determining P based on their SCA strategy related decisions. However, EFs are also crucial for determination of P in such case. Thus, while ships are spaced in a large distance and even though EFs might not directly lead to a SCA, behaviors of operators may be influenced [9]. Additionally, P_G dominates P for any ship close to any other ship.

The human-ship-environment system constantly changes for a sailing ship. EFs also change with the waterbody and meteorological and hydrological conditions regardless of its tending to approach any other ship. However, EFs do not change significantly while two ships started to approach each other and a SCA occurs. Those factors related to the navigation environment do not typically change significantly during any local navigation period when ships cross a relatively short route or navigable waters so that they can be assumed as constants [2]. P_C may change for a SCA depending on them. Historical accident-related data show that P_C can be positively affected by human factors so that P_C shall fall. Conversely, while operating personnel actively respond to effects of external environmental factors on collision accidents and take gaming measures as much as possible, P_C of environmental factors will fall. For example, the captain shall be allowed to go to the bridge and constant VHF communications shall be performed to understand intentions of other ships in case of any severe environment such as poor visibility, strong wind or intensive currents; otherwise, if operating personnel were not sensitive to respond the environment at such case; good human responses were not taken in a harsh environment or even they neglected observation or were tired, P_C might be negatively affected so that P_C shall rise.

Occurrence of any single ship collision accident shall be due to interaction of people-ship-environment. In the existing maritime accidents, there are always certain navigational environment characteristics. Human and ship mistakes may be caused by vastly different reasons. While a certain number of accidents are used as research objects, some of them may come from too strongly oppressive environmental factors while human operation is correct. Rather, gaming between human and environmental factors would be out of balance in other accidents due to not only insufficiently oppressive environmental factors but also more negligent alert minds of operating personnel. However, there are conventional management procedures such as ship inspection, the probability of those maritime accidents due to failures of ship instruments shall be small. Thus, effects of the inherent features (such as inherent maneuverability and maneuverability reliability over time) shall be more concerned for the ship factors. Overall, the ship factors do not fundamentally change the trend of accident risks but their main values are to enlarge or reduce the magnitude of the accident risk due to gaming between human and navigation environment factors. In case that the accident risk is only studied, the following expression shall be satisfied:

$$\text{Accident risk} = \text{Ship factors} \times (\text{human factors} \oplus \text{environmental factors})$$

where: human and environmental factors are among their optimal and worst values, respectively.

If the human factors are within the zone ([−1,1]), human factors between the accident samples may be positive or negative based on analysis of the sample set including a number of accidents to reflect the tact that the integrated effects of human factors of collision accidents may be attenuated or even eliminated. Human factors may become more stable or form a certain distribution (similar to a normal distribution or other modes) depending on effects of human sub-factors based on analysis of a large number of accident samples. It is generally thought that environmental factors are within the zone ([−1,1]) due to their objectiveness and statistical analysis of multiple accident samples that may ensure their being close to a zone or threshold. This is also the focus here. Additionally, the ship factors are within the zone ($(0, +\infty)$) including $(0, 1)$ and $(1, +\infty)$ reflecting amplification and reduction of integrated risk of human and environmental factors.

If the safety management of maritime jurisdiction is regarded as a macroscopic goal and there are a large number of accident samples, the corresponding accident risk macroscopically viewed as such fact that ship characteristics can be described by means of a certain linear or nonlinear function so that such linear or non-linear characteristics vary with time and space and remain constant at small spatial and temporal scales. Human factors are difficultly measured at present due to there being many characteristics such as physiological and psychological factors which are difficult to be quantified. On the whole, if there are big data characteristics for accident samples, the human factors with positive and negative operations shall be counterbalanced mutually or tend to a certain stable value or conform to a certain distribution in case of large samples.

In case of large SCAs, effects of the ship and human factors will act in a certain proportion or conform to a certain distribution [4]. P_C related to the ship and human factors may even be compensated (Figure 1). Irrespective of positive and negative effects of the ship and human factors cancelling each other out, P will increase while the same (or similar) levels of EFs occur as those for a SCA. Thus, an early warning system is necessarily developed for ships tending to collision if P_C related to EFs can be quantified based on previous SCA data. Namely, the safety situation and P can be correspondingly estimated while REWCV for EFs can be determined based on previous SCAs [3].

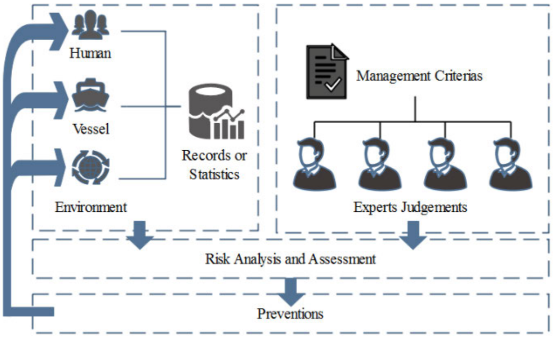

FIGURE 1. Positive and negative effects of factors related to human-ship-environment interaction.

III. METHOD FOR QUANTIFICATION OF P_C

The number of collision accidents is definite within a certain statistical period in a specific water body (channel). Whereas, the widths of channels where collision accidents occur may be wide ($\geq$ 600m) or narrow ($<$ 100m). Statistics is performed to lengths of channels in various widths to master effects of widths of channels on collision accidents. Subsequently, the ratio of lengths of channels whose widths are fixed and the total lengths of channels is extracted. The probability of effects of widths of channels on collision accidents may be determined by means of widths of channels where accidents occur at that time.

EFs may be quantified based on historical SCA data [16]. In case of assuming a certain number of SCAs within a given period and occurrence of every accident due to effects of some EFs, the effect per EF shall be quantified to determine P_C of a SCA due to EFs based on historical SCA data. The final CP can be determined as the sum of CPs related to individual EFs, some of which always take certain effects but vary in magnitude.

For example, visibility always acting as a navigation factor changes in probability from less than 50m to more than 1000m. In general, effects of visibility on the probability of accidents are smaller in case of greater visibility. In contrast, effects of visibility on the probability of accidents are greater in case of lower visibility. Effects of visibility on the probability of accidents are nonlinear as for different degrees of discretization. P_e can be divided into three independent parts:

1). Number of SCAs given a certain environmental factor (n, whose unit is in times);

2). Occurrence frequency of a certain environmental factor (F_z); and

3). Time interval for SCAs (n_0, whose unit is in day)

If the total number of SCAs are determined for a period and the level and rate of each EF, then P_e is expressed as:

$$P_e = \frac{n}{F_z \cdot n_0} \tag{2}$$

In fact, Eq. (2) has 2 functions:

1). Statistics of the number of collision accidents in channels in various widths within the period; and

2). Actual correction of the probability of effects in view of frequencies of collision accidents in channels in each width based on statistics of lengths of channels in various widths in a specific water body

Afterwards, this algorithm is replaced by means of application of Bayesian Theory to gain effects of each environmental factor on the accident probability. The equations are as follows:

$$P(B|A)P(A) = P(B)P(A|B) = P(AB) \tag{3}$$

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} \tag{4}$$

where: P(A|B) represents the probability of collision accidents in case of a certain factor (Grade B) within the statistical period; P(B|A) represents the probability of collision accidents in case of a certain factor (Grade A) within the statistical period; P(A) represents the probability of collision accidents within the statistical period; and P(B) represents the probability of existence of a certain factor (Grade B), which can be replaced by its frequency within the statistical period.

P_i presents the cumulative probability of a SCA occurring for all EFs. When it comes to independent EFs, their sum is one factor in case of occurrence of P_i is expressed as [13]:

$$P_i = \sum_{i=1}^{k} P_e(i) \tag{5}$$

Notably, both parameters (namely ship density and traffic volume) are not independent in a certain sense. It is generally difficult to describe their correspondence. Thus, it is regarded as that they are independent here. There is indeed such possibility ($\mathrm{P_i} < 1$). Our calculations of $\mathrm{P_i}$ are all no more than 1. Rigorously speaking, $\mathrm{P_i}$ based on Eq. (5) is only the total of $\mathrm{P_c}$s for ship accidents due to navigation environmental factors. Because the number of ships in actual waters in the statistical period is not taken into account, its representation significance lies in the relative magnitude of $\mathrm{P_c}$s environmental factors under different historical accidents. Eq. (5) can be improved in future to be applicable in different research fields.

IV. EARLY WARNING METHOD

Once the number of SCAs for a given period has been determined for a waterbody, the corresponding threshold (critical value) of ship collision accidents due to multiple environmental factors for a certain number of ships can also be determined. Each EF and CP changes continuously during any navigation period. When the CP related to all EFs is no less than that determined based on historical SCA data, the SCA probability related to the navigation environment can be regarded as the warning value [19].

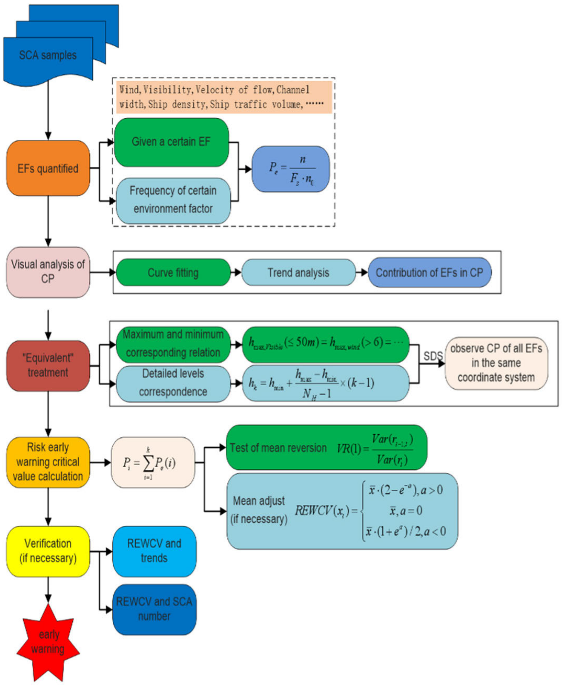

FIGURE 2. **Flow diagram of our early warning method.**

It is generally believed that the length of the ship domain in inland waters is 2-5 times of the length of the ship along its motion direction while those factors such as visibility and ship and current speeds take effect [23]. When the spacing between ships is five times of the length of the ship [5], an early warning should be issued prior to a collision incident so that a higher probability of SCA for the same or similar EFs of historical SCAs shall be presented in advance. Irrespective of any human or ship factor, the CP related to EFs that would be up to a threshold value suggests a SCA and ship operators shall be vigilant for issuing a warning [18]. This threshold value refers to the fact that the risk of accidents is mainly amplified or reduced by ship factors but the overall positive and negative effects of human factors on the probability of accidents are weakened mutually or become more stable or tends to a stable distribution based on statistical analysis of a large number of historical accident samples. Thus, the key navigation environment factors are sufficient to measure the risk of accidents to a certain extent. The method of measurement is the threshold (critical value) analysis in our study; namely, enough attentions shall be paid while the total of the probabilities affected by accidents in the navigation environment is more than that of the minimum influence probabilities of historical statistical accident data after analysis of historical data through the larger samples because ship collision accidents under this influence probability in history. The GP is evaluated subsequently and the early warning method is shown schematically in Figure 2.

A. DATA SORTING AND QUANTIFICATION

While SCA related to previous investigative reports are collected and sorted, the number of previous SCA s and the frequency of various EF levels can be determined. For a formal accident investigation report, locations of SCA shall typically be determined. Information regarding the water channel (including width, water depth and bending radius) shall also be gathered. The corresponding meteorological and hydrological conditions (including wind velocity, range of water velocity and visibility) shall be known. The density of the ships around the ship and the traffic volume shall be determined based on an automatic identification system (AIS) tracker or any other navigation apparatus.

In some cases, it is not easy to determine directly EF frequencies rather than calculate them indirectly. Relatively wind velocity and visibility level frequencies in a waterbody can be obtained directly from the meteorological authority. The year-round daily flow rates and water levels are necessary to determine the frequency of a certain water velocity. On this basis, boundary conditions are determined and the water velocity can be simulated. Subsequently, the frequencies of different water velocities in a waterbody are determined in a similar way based on the corresponding flow rates and water levels. Thus, the frequencies of various water velocities can be represented by the corresponding ones of different flow rates.

The length of a channel is first calculated while its width is fixed to calculate the frequency of a certain channel width by dividing the length of the entire water channel by the length while maintaining the constant width. Thus, the time (number of days) appears as a certain percentage of the whole day at a certain channel width if a SCA occurs in a waterbody.

The part of the waterbody within a certain distance from a previous SCA site shall be taken into account to determine the frequency of a certain ship density around the ship. The number of ships per hour and their corresponding distribution are determined for a certain period. The frequencies of various ship densities are then calculated by dividing the density frequencies of all ships. The frequencies of various traffic volumes provided by the maritime administration authority can be utilized to monitor traffic volumes in each region near the accident location and the number of ships going through each region per hour is counted. Thus, the frequency of a certain traffic volume can be calculated by dividing the frequency per traffic volume level by the frequency of a certain traffic volume.

B. VISUAL ANALYSIS OF CP

While historical accident data are collected and the frequencies of different EFs are determined, the CPs corresponding to different EF levels can be calculated by means of Eq. (2). By plotting the different levels of a given EF along the x-axis and the CP corresponding to the individual EF levels along the y-axis, discrete points shown in a graph were fitted by means of numerical methods of curve fitting (2012) to represent the CP trend corresponding to each particular EF. Any sharp change of the slope of a curve means a dramatic increase or decrease of the CP; namely, the corresponding EF level dominates the CP. Similarly, the CP trends can be determined for all EFs.

C. "EQUIVALENCE" TREATMENT

Various EFs correspond to its own dimensions. It is necessary to evaluate the subjective danger sensation (SDS) [17] for various EF levels to observe the CPs related to all EFs in the same coordinate system. Various EF levels correspond to different "grades" in any SDS classification system. Various types and levels of EFs can also be in the same SDS grade. However, it is difficult to determine the relationship between the different EFs and their levels. It is chiefly necessary to determine the relationships between the maximum quantitative data under a given index (h_{max}) and minimum quantitative data under a given index (h_{min}) EF values so that the quantitative data under a given index (h_k) can be presented. The index evaluation grade (N_H) classification process is performed in lights of the following equation:

$$h_k = h_{\min} + \frac{h_{\max} - h_{\min}}{N_H - 1} \times (k - 1)$$
$$(N_H > 1, N_H \in N; k = 1, 2 \ldots\ldots, N_H) \tag{6}$$

Then, the quantitative evaluation criteria for each EF level are expressed as:

$$\begin{aligned} h_k &= \{h_1, h_2, \ldots\ldots, h_{N_H}\} \\ &= \left\{h_{\min}, h_{\min} + \frac{h_{\max} - h_{\min}}{N_H - 1}, \ldots\ldots, h_{\max}\right\} \end{aligned} \tag{7}$$

As for definitions of $\mathrm{h}_{\max}$ and ranges of values mentioned, the extreme values of these environmental factors are defined based on historical statistics of accidents. A concept called as the maneuver danger degree is also introduced in view of these factors in different dimensions to perform comparison and calculation in the same system. As described in many Chinese literatures [13], experts and ship operators are invited to carry out subjective evaluation; and they think that effects of environmental factors which are below (or above) a certain degree on the maneuver danger degree are almost the same. This is the basis for determination of the equivalence of the extreme value (0 or 1). As a result, various environmental factors are classified after the correspondence between the extreme values of various environmental factors has been determined. It is believed that the various maneuver danger degrees are also put into a one-to-one relationship so that the correspondence between a certain navigational environmental factor and the danger of another navigation environment factor can be determined. Finally, various factors in different dimensions are unified under the unified system of the maneuver danger.

For example, if the SDS corresponding to a visibility (< 50m) is equal to that for a water velocity(> 2.5m/s), the visibility (< 50m) and water velocity (> 2.5m/s) correspond to the maximum of these EFs; namely:

$$h_{\max,Visible}(< 50m) = h_{\max,watervelocity}(> 2.5m/s) = 1 \tag{8}$$

where: the number "1" represents the highest SDS grade.

If the SDS corresponding to a visibility ($>$ 1000m) is equal to that for a water velocity ($<$ 0.5m/s), the visibility ($>$ 1000m) and water velocity ($<$ 0.5m/s) correspond to the minimum of the EFs in a similar way; namely:

$$h_{\min,Visible}(>1000m)=h_{\min,water\ velocity}(\leq 0.5m/s)=0 \tag{9}$$

where: the number "0" represents the lowest SDS grade.

Thus, the corresponding relationships for various levels of both EFs can be determined by means of Eqs. (3) and (4). The appropriate calculation methods must be used to determine precisely the relationships between different levels. Next, the CP corresponding to all EFs can be observed in the same coordinate system.

D. CALCULATION OF REWCV

After the CP corresponding to each EF has been determined, the SCA probability can be determined by means of Eq. (2). The CP values for all SCAs in a waterbody must exhibit a reversion to the mean if the number of samples is sufficiently high. That is to say, the CP determined by each EF will fluctuate around the critical or warning value. Alternatively, this also means that the REWCV of EFs is independent of human and ship factors. Thus, the mean reversion method can be used to not only verify the CP trends of previous CAs but also eliminate effects of human and ship factors on SCAs. Then, the critical value of the risk due to EFs can be known based on the early warning means. The ratio variance method [21] can be used to verify whether the CP values for various accidents are sorted chronologically. The variance ratio ($VR(k)$) represents the proportional relationship between a large-scale the variance of $r_{t-k,t}$ ($Var(r_{t-k,t})$) and a small-scale the variance of x_t ($Var(r_t)$), where t is between 1 and the number of $x_t(n)$. For the time series $\{x_t\}_1^n$ of the CP, the ratio can be calculated by:

$$VR(k)=\frac{Var(r_{t-k,t})}{k\cdot Var(r_t)},\ (0<k<t;1<t\leq n) \tag{10}$$

$$r_{t-k,t}=\log(x_{t-k}/x_t) \tag{11}$$

$$\text{Then}: VR(1)=\frac{Var(r_{t-1,t})}{Var(r_t)} \tag{12}$$

While $VR(1)$ is less than 1, there is a negative short-term autocorrelation to mean that CP fluctuates excessively in the short term. Thus, the CP shall be reverted the mean in the long term. In contrast, while $VR(1)$ is more than 1, there is a positive short-term autocorrelation to signify that the CP does not vary significantly in the short term. Hence, the CP will not be reverted the mean in the long term. In the mean-averting phase, the CP will deviate from the mean so that the REWCV shall be adjusted based on the slope of the fitting curve (a) and the intercept of line (b) which are the parameters of the equation. The curve is assumed as the following expression [12] to fit CP data:

$$y(x)=ax+b \tag{13}$$

$\bar{x}$ is the mean of x_t and Euler Constant (e) is approximated as 2.718. The REWCV can be adjusted as:

$$REWCV(x_t)=\begin{cases}\bar{x}\cdot(2-e^{-a}), & a>0\\ \bar{x}, & a=0\\ \bar{x}\cdot(1+e^{a})/2, & a<0\end{cases} \tag{14}$$

This adjustment is to reduce the range of variations in the slope of the fitting curve from $(-\infty,+\infty)$ to [0.5, 2] so that the mean shall be preferably adjusted. In fact, the slope of the fitting curve is within $[-1, 1]$. Thus, the adjusted range of the critical value of the risk lies in the range of [0.68, 1.63].

V. CASE STUDY

Based on the key EFs for SCAs in Three Gorges Reservoir in China during 2003-2007 (Table 1) and their investigation and analysis, it is concluded that the maximum and minimum values of key EFs in the SDS system are related, which are presented as follows:

$$\begin{aligned}
&h_{\max,Visible}(\leq 50m) && h_{\min,Visible}(>1000m)\\
&=h_{\max,wind}(>6) && =h_{\min,wind}(\leq 1)\\
&=h_{\max,water\ velocity}(>2.5m/s) && =h_{\min,water\ velocity}(\leq 0.5m/s)\\
&=h_{\max,channel\ width}(\leq 100m) && =h_{\min,channel\ width}(>800m)\\
&=h_{\max,ship\ \text{density}}(>30) && =h_{\min,ship\ \text{density}}(\leq 5)\\
&=h_{\max,traffic\ \text{volume}}(>80)=1 && =h_{\min,traffic\ \text{volume}}(\leq 10)=0
\end{aligned}$$

Subsequently, the "grading" and "equivalence" can be performed by means of Eqs. (6) and (7). After the frequencies of each EF level are determined, the CP for each EF level can be calculated by means of Eq. (2). Their corresponding results are shown in Tables 2-7.

In fact, no any accident shall occur in the extreme environment including navigational environmental factors as for the current number of accident samples. Rational assumptions are also made in case of there being no accident samples in our analysis here. If effects of P_G is taken into account, the collision risk of the ship or other specific ships will be predicted more rigorously on a micro basis to form the complete logic of macro causal risk + micro-geometric scale risk. Division of specific ships can also be taken into account by means of the ship factors in our collision risk model (Accident risk = Ship factors × (human factors ⊕ environmental factors)). These are being at the point in our research to take above advantages and overcome those deficiencies. Another paper is being hatching for international publication and exchange with global counterparts after successful modeling, simulation and experimental verification.

The discrete points corresponding to different EF levels are fitted into curves (Figures 3-8) by plotting the SDS grade on the horizontal axis and the CP for a SCA on the y-axis. These curves clearly reveal the relationships between different EF levels and the CP. For example, the visibility curve shows that the CP grows with SDS ($<$ 0.8). The CP also increases sharply while SDS exceeds 0.6. More attentions shall be paid by the operators while the SDS of visibility is more than 0.6

TABLE 1. Key EFs for historical SCAs in the study area.

Serial No.	Visibility (m)	Wind velocity (Beaufort scale)	Water velocity (m/s)	Channel width (m)	Ship density* (ships)	Traffic volume** (ships)
1	(0,50]	1	1.5	800	16	24
2	(50,100]	1	0.5	600	14	10
3	(500,1000]	2	0.5	400	7	6
4	(500,1000]	3	3	300	8	7
5	(200,300]	1	0.5	600	23	50
6	(500,1000]	4	0.8	400	7	11
7	>1000	3	0.8	450	8	10
8	(50,100]	1	0.4	700	21	42
9	>1000	4	0.6	150	7	14
10	(200,300]	3	1	400	13	28
11	(100,200]	1	2.5	150	4	8
12	>1000	6	1.2	800	28	56
13	(500,1000]	3	0.6	600	15	32
14	(100,200]	1	0.8	600	5	16
15	>1000	4	0.8	750	14	34
16	>1000	5	1	550	13	38
17	(300,500]	3	2	600	31	64
18	(300,500]	2	0.4	400	12	31
19	(100,200]	2	2.2	300	4	10

Notes: * Ship density is the number of ships in the upper and lower reaches within a 1000m range around the ship.

** Traffic volume is the number of ships going per hour through a ship traffic flow statistics region near the location of the SCA.

*** The type of ship is not taken into account at all here because of insufficient accident samples so far.

TABLE 2. Visibility results.

Visibility (m)	Number of times observed	Frequency	CP	SDS
≤50	1	0.02	2.74E-02	1.0
(50,100]	2	0.03	3.65E-02	0.8
(100,200]	3	0.08	2.05E-02	0.7
(200,300]	2	0.10	1.10E-02	0.6
(300,500]	2	0.12	9.13E-03	0.4
(500,1000]	4	0.30	7.31E-03	0.2
>1000	5	0.35	7.83E-03	0.0

TABLE 3. Wind velocity results.

Beaufort scale	Number of times observed	Frequency	CP	SDS
>6	0	0.03	0.00E+00	1.0
=6	1	0.04	1.37E-02	0.8
=5	1	0.08	6.85E-03	0.7
=4	3	0.15	1.10E-02	0.6
=3	5	0.18	1.52E-02	0.4
=2	3	0.20	8.22E-03	0.2
≤1	6	0.30	1.10E-02	0.0

TABLE 4. Water velocity results.

Water velocity(m/s)	Number of times observed	Frequency	CP	SDS
>2.5	1	0.03	1.83E-02	1.0
(2.0,2.5]	2	0.05	2.19E-02	0.8
(1.5,2.0]	1	0.08	6.85E-03	0.7
(1.0,1.5]	2	0.12	9.13E-03	0.6
(0.8,1.0]	2	0.18	6.09E-03	0.4
(0.5,0.8]	6	0.26	1.26E-02	0.2
≤0.5	5	0.28	9.78E-03	0.0

TABLE 5. Channel width results.

Channel width (m)	Number of times observed	Frequency	CP	SDS
≤100	0	0.05	0.00E+00	1.0
(100,200]	2	0.05	2.19E-02	0.8
(200,400]	6	0.15	2.19E-02	0.7
(400,500]	1	0.25	2.19E-03	0.6
(500,600]	6	0.20	1.64E-02	0.4
(600,800]	4	0.20	1.10E-02	0.2
>800	0	0.10	0.00E+00	0.0

TABLE 6. Ship density results.

Ship number density(ships)	Number of times observed	Frequency	CP	SDS
>30	1	0.04	1.37E-02	1.0
(25,30]	1	0.05	1.10E-02	0.8
(20,25]	2	0.08	1.37E-02	0.7
(15,20]	1	0.28	1.96E-03	0.6
(10,15]	6	0.25	1.32E-02	0.4
(5,10]	5	0.20	1.37E-02	0.2
≤5	3	0.10	1.64E-02	0.0

TABLE 7. Traffic volume results.

Traffic volume (ships)	Number of times observed	Frequency	CP	SDS
>60	1	0.04	1.37E-02	1.0
(50,60]	1	0.08	6.85E-03	0.8
(40,50]	2	0.15	7.31E-03	0.7
(30,40]	4	0.20	1.10E-02	0.6
(20,30]	2	0.20	5.48E-03	0.4
(10,20]	3	0.18	9.13E-03	0.2
≤10	6	0.15	2.19E-02	0.0

(its corresponding actual visibility: < 300m). The region of the curve where the CP falls along with SDS (namely SDS: > 0.8) can be attributable to an error due to there being only a few samples. If the number of previous SCAs corresponding to a visibility (< 50m) rises, it is expected the CP would increase continuously with SDS (Figure 3).

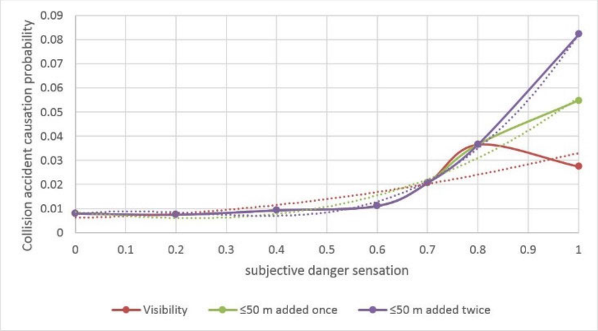

FIGURE 3. Fitted curves (the number of SCAs vs. visibility (< 50m)).

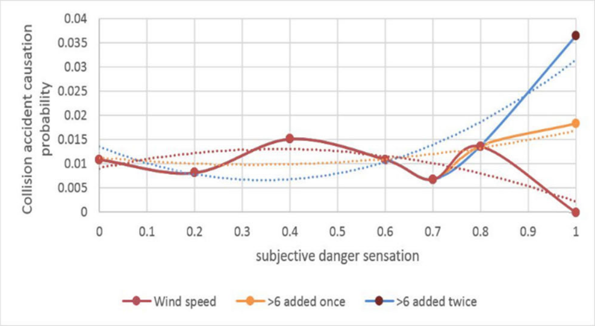

FIGURE 4. Fitted curves (the number of SCAs vs. the wind velocity (level: ≥ 6)).

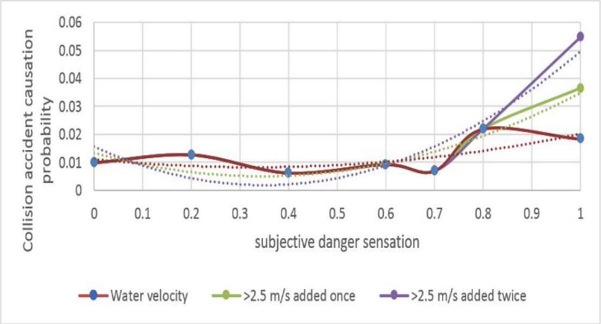

FIGURE 5. Fitted curves (the number of SCAs vs. the water velocity (> 2.5m/s)).

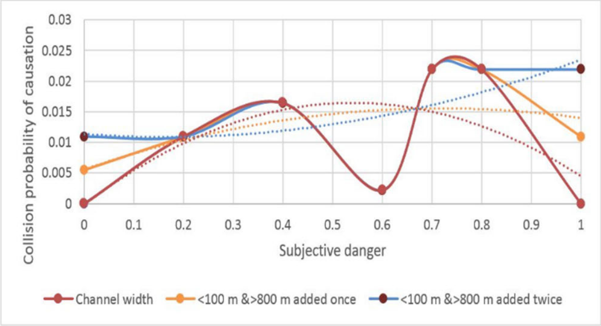

FIGURE 6. Fitted curves (the number of SCAs vs. the channel width (<100m or >800m)).

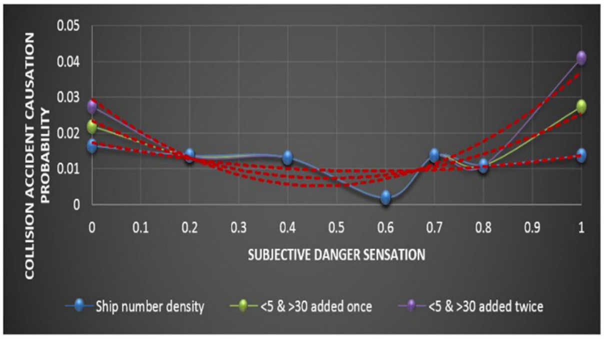

FIGURE 7. Fitted curves (the number of SCAs vs. the ship density (<5 or >30)).

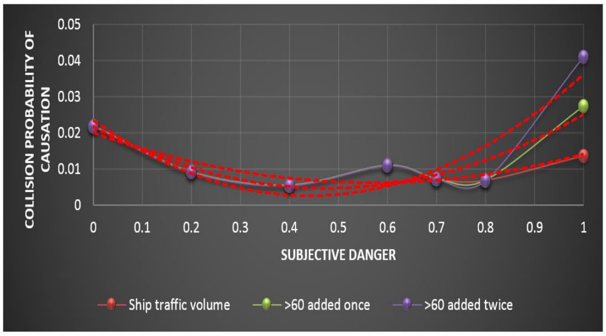

FIGURE 8. Fitted curves (the number of SCAs vs. the ship traffic volume (> 60)).

The wind velocity curve (Figure 4) shows that the CP does not change consistently with SDS. However, no previous SCA corresponds to a wind velocity level (≥ 6). If the wind velocity level (≥ 6) is assumed as the occurrence threshold for a ship collision, the CP increases with SDS but the curve fluctuates by a small margin. When SDS is greater than 0.72 and the wind velocity level is above 5, the probability of a collision increases remarkably so that more attentions shall be paid in these cases.

The fitted curve for water velocity is similar to a combination of those for visibility and wind velocity. The CP increases with SDS (< 0.8) but smaller fluctuations are observed. In contrast, the CP increases rapidly while SDS exceeds 0.68 which corresponds to a water velocity (≥ 1.8m/s). The part of the curve where CP falls with SDS (> 0.8) may attribute to there being only several samples. If the number of SCAs corresponding to water velocity (> 2.5m/s) rises, the CP increases with SDS without limit (Figure 5).

Absence of historical SCA data corresponding to channel width (< 100m or > 800m) results in occurrence of a CP peak while SDS is within [0.4, 0.75]. Even though the actual occurrence of any collision for any such channel width had been taken into account, the CP trend would not change considerably (Figure 6). For instance, CPs (@ SDS=0.6) are lower than those (@ SDS=0.4) and CPs (@ SDS>0.6) would be no more than those (@ SDS=0.75). A possible reason is that the channel width has little effect on CP in case of SDS (< 0.4) corresponding to a channel width (> 600m). Thus, effects of the channel width on the CP are attenuated quickly. Conversely, a channel width (< 200m) corresponding to SDS (> 0.75) significantly influences the CP so that the human factors may become more important. Thus, the CP decreases.

The fitted curve for the ship density only shows that the CP changes significantly in case of SDS=0.6; and the CP is very

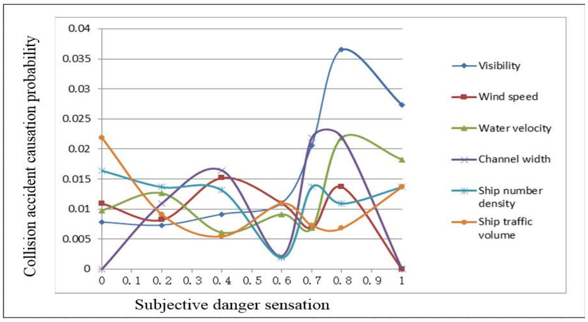

FIGURE 9. CP vs. EF.

TABLE 8. Results of variance ratio test.

Serial NO.	CP	Mean of x_t	Variance of x_t	$\log(p_t/p_{t-k})$	Variance of $\log(p_t/p_{t-k})$	$VR(1)$
1	6.59E-02	8.01E-02	2.67E-04	-	1.62E-02	6.04E+01
2	1.09E-01			2.18E-01		
3	8.28E-02			-1.18E-01		
4	9.83E-02			7.44E-02		
5	6.91E-02			-1.53E-01		
6	7.57E-02			3.91E-02		
7	7.35E-02			-1.26E-02		
8	7.89E-02			3.07E-02		
9	7.62E-02			-1.52E-02		
10	7.28E-02			-1.96E-02		
11	1.14E-01			1.94E-01		
12	5.94E-02			-2.82E-01		
13	7.57E-02			1.05E-01		
14	8.62E-02			5.61E-02		
15	6.65E-02			-1.12E-01		
16	6.13E-02			-3.53E-02		
17	7.35E-02			7.88E-02		
18	7.32E-02			-2.09E-03		
19	1.11E-01			1.81E-01		

TABLE 9. Adjusted results and REWCV.

CP fitting curve	*Slope*	$(1+e^a)/2$	*Mean of* x_t	$REWCV(x_t)$
$y = -0.0002x + 0.0825$	-2.00E-04	0.9999	8.01E-02	8.01E-02

low. This suggests that the CP is independent of any change in the ship density except for that while SDS is close to 0.6 (the corresponding ship density: 15-20). If the number of SCAs corresponding to the ship density (< 5 or > 30) rises, a higher ship density leads to an increase of CP (Figure 7).

Irrespective of any change of the ship traffic volume, the CP will always increase. The CP is minimal in case of SDS=0.4 corresponding to a ship traffic volume (20-30). Higher ship traffic volumes lead to larger CP values along the growth of the number of SCAs related to ship traffic volumes (> 60; Figure 8). It can be concluded that the CP may fall while an appropriate ship traffic volume is maintained. The CP also increases while the ship traffic volume is lower than such threshold value. Thus, operators shall pay more attentions in such case.

All EFs can be visualized under the same coordinate plane (Figure 9). In the context of the human-ship-environment system theory, any accident occurs due to interaction of various factors, which is only analyzed here in view of the navigation environment so that there shall be some certain congenital deficiencies. In future, the specific distribution of contribution of human and ship factors to accident risks will be analyzed in case of big data. The overall calculation will be finally performed in a human-ship-environment system. Its overall grasp will be carried out by means of the accident risk = ship factors × (man factors ⊕ environmental factors). Currently, environmental factors tend to deteriorate and the risk of accidents should be deteriorated in theory. Whereas, there will be "explicit" or "implicit" risk changes due to the fact that human and environmental factors are not fully taken into account. Human factors will hinder deterioration of accident risks in case that the risk of accidents is perceived. Thus, while the conditions are deteriorated, human factors will always avoid this deterioration of risks (this is a kind of human instinct). Thus, it is believed that the depiction of the curves shall conform to our actual cognition where effects of risk trends of a human-vessel-environment system shall be reflected only in the context of studying the navigational environmental factors.

The CP related to visibility is greater than that related to any factor in case of SDS (> 0.7). Thus, the visibility shall be focused for SCA early warning. The CP peaks for a channel width (SDS: > 0.65) or a water velocity (SDS: > 0.75). Notably, the CP remains high even in case of a ship traffic volume (SDS=0). The CPs per EF are combined by means of Eq. (5) so that the SCA incidents can be distributed chronologically. Then, reversion to the mean is verified by means of Eq. (12) and the results are listed in Table 8. Because VR (1) is larger than 1, SCAs are in the mean-averting phase. Thus, the mean is adjusted by means of Eq. (13). The adjusted results and the REWCV are shown in Table 9.

The REWCV means that a collision accident occurs and the probability of SCA is high in case that the CP related to the different EF levels is no less than 0.08012. Thus, such value can be used to issue early warnings. The REWCV does not represent the complete CP for a SCA. Rather, it means that there would be a higher probability of a SCA than usual if such EF is effective in spite of any human or ship factors.

VI. VERIFICATION OF RESULTS

Constraints are set by the Navigation Standards for Inland Waterways and Maritime Safety Administration in China. Based on comparison of these standards (Table 10), it can be concluded that our results are essentially consistent with those for visibility, wind velocity and water velocity in Chinese limitation standards to confirm the practical applicability of our study. SCA data for 2008 are collected to verify the applicability of the REWCV for the inland waterbody under study. The key EFs related to SCAs and the calculated CP values for the inland waterbody are presented in Table 11.

The mean CP value in 2008 is 0.101632 which is much larger than the REWCV (0.08012) calculated for SCAs for 2003-2007. This indicates that CPs related to EFs increases in 2008. The minimal CP for 2003-2008 (0.061329) is similar to that for 2003-2007 (0.059426). Such case indicates that

TABLE 10. Comparison of standards in China with the results of this study.

EF	Chinese standard	Conclusion of this study	Remark
Visibility	From ship to upper reaches: < 500m From ship to lower reaches: < 1000m	Probability increases sharply while visibility (< 300m).	Chinese standard is more stringent
Wind velocity	Level 6 and higher	Probability increases sharply while the wind velocity level (> 5).	Almost similar
Water velocity	> 2.0m/s	Probability increases sharply while the water velocity (> 1.8m/s).	Chinese standard is less stringent
Channel width	Not Applicable	CP peaks while the channel width ranges from 200m to 500m.	Not Applicable
Ship number density	Not Applicable	CP remains almost constant and its minimum value is among 15-20.	Not Applicable
Ship traffic volume	Not Applicable	Irrespective of any change of the ship traffic volume, the CP increases but its minimum value is among 20-30.	Not Applicable
CP of EFs	Not Applicable	≥ 8.01E-02	Not Applicable

TABLE 11. Key EFs related to SCAs in 2008 and calculated CP values.

Visibility (m)	velocity (Beaufort scale)	Water velocity (m/s)	Channel width (m)	Ship density (ships)	Traffic volume (ships)	CP
(500,1000]	6	1.2	300	27	31	1.18E-01
(300,500]	2	1.1	240	12	16	8.99E-02
(100,200]	≤1	2.2	370	21	27	1.22E-01
(100,200]	≤1	0.9	420	26	32	9.31E-02
>1000	3	1.2	620	31	27	8.55E-02

the minimum value can be used as the REWCV. On the other hand, it can be detrimental to use the minimum value for SCA early warning while its greater likelihood is taken into account. Thus, an appropriate warning value between the minimum and maximum CPs can be determined based on calculation collisions from 2003 to 2010. The longer study period changes CPs related to EFs. Comparison results (Table 12) indicate that CPs related to the same EF vary with the number of accidents per statistical period. EFs vary significantly for the studied waterbody in 2008 and 2009. CPs of SCAs increase substantially from 2008 to 2009.

Mean CP values related to EFs are minimal for the statistical period (2003-2007). The mean CP trends to grow with time for the statistical period (2003-2008) and increases slowly for the statistical period (2003-2009). The CP seems constant for the statistical period (2003-2010). Thus, the CP is larger for the same EFs from 2003 to 2008 and SCAs are more likely to occur. The high CP values related to EFs may be due to the fact that there are three times of the average number of accidents in 2009. The decrease of the mean CP value in 2010 is due to the number of accidents falling below the average in that year. The numbers of SCAs in 2003-2008 (Figure 10) indicates that the choice of statistical period greatly affects the CP related to a certain EF (Figure 11). For example, the number of SCAs in 2009 changes significantly.

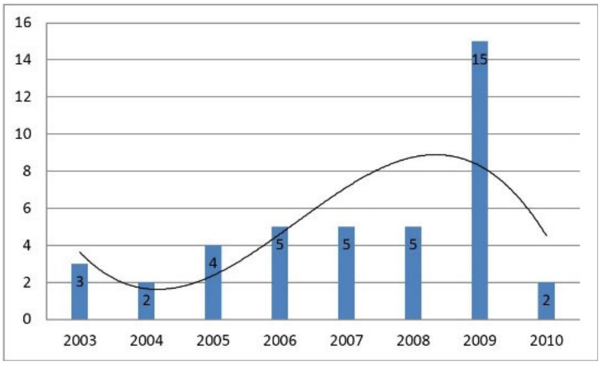

FIGURE 10. Number of SCAs (2003-2010).

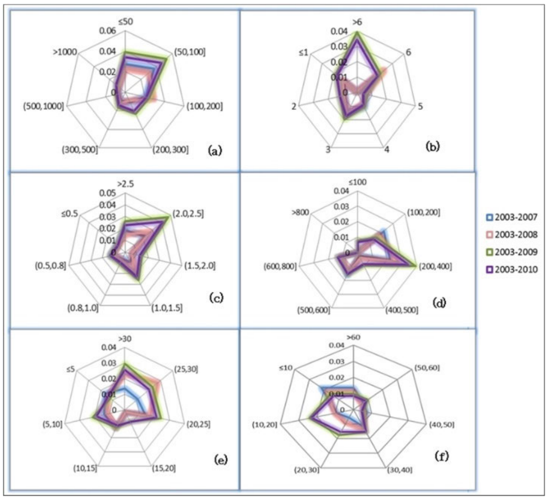

FIGURE 11. Changes in CP related to (a) visibility, (b) wind velocity, (c) water velocity, (d) channel width, (e) ship density, and (f) ship traffic volume according to the statistical period.

Based on a further analysis of changes of CPs related to EFs, SCA samples (2003-2010) are sorted chronologically. Then, the first 20 samples are used for calculation of CPs and then it is determined to replace the prior three samples with the following three ones. This method ensures that the number of SCAs shall be the same during each CP calculation period, which is based on the time between the first and last samples. The calculation results are shown in Figure 12.

Figure 12 shows a clear growth trend of CPs for the fourth SCA in 2008 (Line ②) and CPs for the fifth SCA in 2009 (Line ④), which becomes more stable. The CP trends indicate a rapid increase of CPs for the eighth SCA in 2009 (Line ⑤). CPs for the fourteenth SCA in 2009 (Line ⑦) shows that a decreasing trend becomes less tremendous than that in 2010. Understanding the upward trends in CPs between Lines ② and ⑥ can result in the effective control of SCAs. Our results

TABLE 12. Comparison results of CP values.

Year	Serial NO.	CP (2003-2007)	CP (2003-2008)	CP (2003-2009)	CP (2003-2010)
2003-2007	1	6.59E-02	7.62E-02	1.23E-01	1.12E-01
	2	1.09E-01	9.55E-02	1.23E-01	1.11E-01
	3	8.28E-02	8.20E-02	1.06E-01	9.53E-02
	4	9.83E-02	9.51E-02	1.31E-01	1.15E-01
	5	6.91E-02	6.64E-02	9.55E-02	8.76E-02
	6	7.57E-02	7.62E-02	1.18E-01	1.06E-01
	7	7.35E-02	6.69E-02	9.09E-02	8.28E-02
	8	7.89E-02	8.54E-02	1.22E-01	1.11E-01
	9	7.62E-02	6.73E-02	9.63E-02	8.94E-02
	10	7.28E-02	8.13E-02	1.23E-01	1.08E-01
	11	1.14E-01	1.18E-01	1.33E-01	1.21E-01
	12	5.94E-02	9.42E-02	1.04E-01	9.45E-02
	13	7.57E-02	7.36E-02	8.61E-02	8.01E-02
	14	8.62E-02	8.88E-02	1.08E-01	1.04E-01
	15	6.65E-02	6.54E-02	7.70E-02	7.41E-02
	16	6.13E-02	6.13E-02	7.08E-02	6.73E-02
	17	7.35E-02	8.03E-02	1.05E-01	9.39E-02
	18	7.32E-02	8.26E-02	1.04E-01	9.54E-02
	19	1.11E-01	1.24E-01	1.50E-01	1.36E-01
2008	1		1.18E-01	1.33E-01	1.18E-01
	2		8.99E-02	1.30E-01	1.17E-01
	3		1.22E-01	1.71E-01	1.51E-01
	4		9.31E-02	1.02E-01	9.22E-02
	5		8.55E-02	1.18E-01	1.06E-01
2009	1			1.70E-01	1.52E-01
	2			9.93E-02	9.14E-02
	3			1.12E-01	1.02E-01
	4			1.63E-01	1.47E-01
	5			1.31E-01	1.17E-01
	6			1.34E-01	1.19E-01
	7			1.30E-01	1.18E-01
	8			9.97E-02	9.19E-02
	9			1.53E-01	1.37E-01
	10			1.28E-01	1.13E-01
	11			1.41E-01	1.28E-01
	12			1.09E-01	9.82E-02
	13			1.34E-01	1.21E-01
	14			9.05E-02	8.53E-02
	15			1.00E-01	9.05E-02
2010	1				8.94E-02
	2				7.46E-02
Mean		8.01E-02	8.70E-02	8.71E-02	1.06E-01

TABLE 13. Comparison results of REWCV.

	2003-2007	2003-2008	2003-2009	2003-2010	Equation
Calculated value of REWCV	7.97 E-02	8.71 E-02	1.18 E-01	1.06 E-01	$REWCV(x_t) = \begin{cases} \bar{x} \cdot (2 - e^{-a}), a > 0 \\ \bar{x}, a = 0 \\ \bar{x} \cdot (1 + e^{a})/2, a < 0 \end{cases}$
Mean	8.01 E-02	8.71 E-02	1.18 E-01	1.06 E-01	
Slope of fitting line	2.00 E-04	9.00 E-04	5.00 E-04	2.00 E-04	
Intercept of fitting line	8.25 E-02	7.55 E-02	1.09 E-01	1.02 E-01	

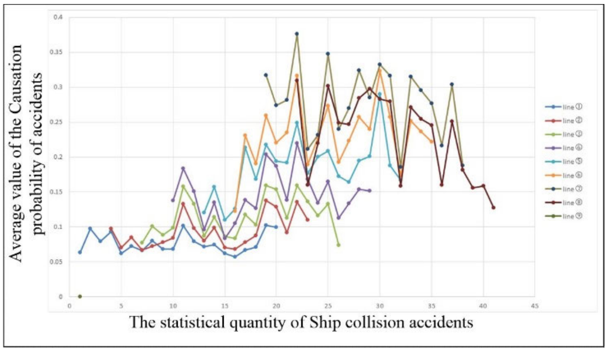

FIGURE 12. CP trends for SCAs (2003-2010) based on 20 samples per calculation period.

suggest that the following issues shall be necessarily solved in future.

1) REWCV Shall Be Accurately Determined to Enable Early Warnings: the mean value of all CPs related to all EFs is regarded as the REWCV and the slope of the linear fit is taken as the adjustment parameter here. The slope changes slightly so that the adjustment range shall be correspondingly small. Because the number of SCAs varies substantially each year, the linear fit also varies greatly. The calculated fit for one year cannot be applied to calculate the REWCV in the following year. A comparison of REWCV values for each statistical period and SCA based on application of linear fitting (Tables 13 and 14). If $e^{-a}(e^{a})$ in Eq. (15) is replaced by $e^{-100a}(e^{100a})$ under a given REWCV* and the slope is more than 0.001, the correction value will be close to the mean for the next calculation process. Thus, the REWCV can be calculated based on the following equation, where μ can be adjusted according to the calculation requirements.

$$REWCV(x_t) = \begin{cases} \bar{x} \cdot (2 - e^{\mu a}), & a > 0 \\ \bar{x}, & a = 0 \\ \bar{x} \cdot (1 + e^{\mu a})/2, & a < 0 \end{cases} \tag{15}$$

2) The Precise Relationship Between the Change of CPs and the Number of SCAs Shall Be Determined: it can conclude that the mean CP of SCAs will increase to result in a greater risk for the same EF in case that CPs related to EFs increase. If any increased risk is not recognized and processed, the number of SCAs will rise accordingly. Conversely, if the increased risk can be recognized, appropriate countermeasures will be taken according to human and ship factors so that the number of SCAs may decrease. In contrast, if CPs related to EFs decrease, a lower risk for the same EFs and an increase in the number of SCAs will come out. Additionally, if operators underestimate the risk related to

TABLE 14. Comparison results of REWCV*.

	①	②	③	④	⑤	⑥	⑦	⑧	Equation
Calculated value of REWCV	7.69E-02	9.38E-02	1.15E-01	1.47E-01	1.86E-01	2.33E-01	2.79E-01	2.29E-01	$REWCV(x_t)=\begin{cases}\bar{x}\cdot(2-e^{-a}),a>0\\ \bar{x},a=0\\ \bar{x}\cdot(1+e^{a})/2,a<0\end{cases}$
Calculated value of REWCV*	7.84E-02	1.11E-01	1.31E-01	1.60E-01	2.39E-01	2.74E-01	2.50E-01	1.81E-01	$REWCV(x_t)=\begin{cases}\bar{x}\cdot(2-e^{-100a}),a>0\\ \bar{x},a=0\\ \bar{x}\cdot(1+e^{100a})/2,a<0\end{cases}$
Mean	7.69E-02	9.36E-02	1.15E-01	1.46E-01	1.86E-01	2.32E-01	2.79E-01	2.30E-01	
Slope of fitting line	2.00E-04	2.00E-03	1.50E-03	1.00E-03	3.40E-03	2.00E-03	-2.40E-03	-5.60E-03	
Intercept of fitting line	7.47E-02	6.68E-02	8.99E-02	1.26E-01	1.09E-01	1.81E-01	3.47E-01	4.06E-01	

certain EFs, a negative response may occur and the number of SCAs can increase similarly.

As for broad acceptance of our method and results, shipping operations are directly forbidden in those cases (such as wind strength > Grade 6; the flood rate > 2.5m/s during the flood period; and visibility < 500m (in heavy fog) based on data collected by Yangtze River Maritime Affair Bureau of China Maritime Safety Administration (China MSA) in several extreme cases in its navigation restrictions. After occurrence of Eastern Star Tragedy, more specified administration provisions were issued to present more detailed regulations on navigation environment restrictions for sailing ships and those ready for sail. Our primary concern is that the risk of those accidents that may be perceived straightforwardly in any extreme navigation environment, which is called as the explicit risk of the navigation environment. Unfortunately, more navigation environment factors are acceptable (namely those extreme conditions stipulated by the authority are not be transcended) as for specific types of accidents. The probability of the risk of accidents is high (for example: foggy visibility > 800m or the flood rate > 2.0m/s). Whether the interaction of many environmental factors may intensifies extreme navigational factors shall be studied by means of a large number of accident samples. This is defined as the implicit risk which is to actually be determined here and can be identified to some extent. Thus, it is believed that a better promotion significance shall be provided after continuous development and improvement.

VII. CONCLUSION

Study of CPs of SCAs is a highly complicated task which is akin to using a "black box" in many cases. One knows input and output parameters rather than the actual process (namely, how to obtain outputs remains unknown). Even if there were a perfect model or method, it would be difficult to determine the basic CP corresponding to each EF. Our early warning method was put forward here based on effects of various EFs on ship navigation and assumptions of the facts that the future probability of SCAs can be determined if the EFs corresponding to previous SCAs can be reconstructed for a given waterbody and the hidden risks related to EFs can be recognized if changes of the REWCV can be determined. In such case, the SCA risk can be perceived in advance and the corresponding number of SCAs can be reduced accordingly. The REWCV for previous SCAs can be used to issue early warnings. Our method incorporating the advantages of the Bayesian inferring [20] and failure diagnosing [1], which are popular in current studies. Our method is more intuitive though the REWCV is only an estimate of the actual SCA probability. Additionally, our early warning method and calculations are relatively simple so that early warnings can be issued to prevent future collisions by continuously updating the data of previous SCAs. Thus, our method shall be of great applicable value.

APPENDIX

NOMENCLATURES

a	Slope of line
b	Intercept of line
e	Euler constant (≈2.718)
F_Z	Occurrence frequency of a certain environmental factor
h_k	Quantitative data under a given index
h_{max}	Maximum quantitative data under a given index
h_{min}	Minimum quantitative data under a given index
N_h	Index evaluation grade
n	Number of SCAs given a certain environmental factor; and Number of x_t
n_0	Time interval for ship SCAs, whose unit in day
P	Frequency of SCA (frequently used unit is 1/year)
P_C	Causation probability of failing to avoid collision
P_e	Causal probability due to a certain environmental factor
P_G	Frequency (unit: 1/year) of collision candidates which are cases where collision is unavoidable if no evasive action is taken.
P_i	Causation probability due to environmental factors
$VR(k)$	Variance ratio
$Var(r_{t-k,t})$	Variance of $r_{t-k,t}$
$Var(r_t)$	Variance of
$\bar{x}$	Mean of x_t

ACRONYMS AND ABBREVIATIONS

CP	Causation probability
GP	Geometric probability
EF	Environmental factor
REWCV	Risk early warning critical value
SCA	Ship collision accident
SDS	Subjective danger sensation

REFERENCES

[1] P. Antão and C. G. Soares, "Fault-tree models of accident scenarios of RoPax vessels," *Int. J. Autom. Comput.*, vol. 3, no. 2, pp. 107–116, Apr. 2006.

[2] J.-F. Balmat, F. Lafont, R. Maifret, and N. Pessel, "MAritime RISk assessment (MARISA), a fuzzy approach to define an individual ship risk factor," *Ocean Eng.*, vol. 36, nos. 15–16, pp. 1278–1286, Nov. 2009.

[3] C.-H. Chen, L. P. Khoo, Y. T. Chong, and X. F. Yin, "Knowledge discovery using genetic algorithm for maritime situational awareness," *Expert Syst. Appl.*, vol. 41, no. 6, pp. 2742–2753, May 2014.

[4] R.-M. Darbra and J. Casal, "Historical analysis of accidents in seaports," *Saf. Sci.*, vol. 42, no. 2, pp. 85–98, Feb. 2004.

[5] X. Fan, Q. Zhang, F. Zhou, Z. Tan, and M. Wang, "Model of ship domain in river water," *J. Dalian Maritime Univ.*, vol. 39, no. 1, pp. 46–48, 2013.

[6] F. Goerlandt and P. Kujala, "Traffic simulation based ship collision probability modeling," *Rel. Eng. Syst. Saf.*, vol. 96, no. 1, pp. 91–107, Jan. 2011.

[7] P. G. Guest, *Numerical Methods of Curve Fitting*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[8] M. Hanninen and P. Kujala, "The effects of causation probability on the ship collision statistics in the Gulf of Finland," in *Marine Navigation and Safety of Sea Transportation*, A. Weintrit, Ed. London, U.K.: Taylor & Francis, 2009, pp. 267–272.

[9] K. Hara and S. Nakamura, "A comprehensive assessment system for the maritime traffic environment," *Saf. Sci.*, vol. 19, nos. 2–3, pp. 203–215, Jun. 1995.

[10] C. Hetherington, R. Flin, and K. Mearns, "Safety in shipping: The human element," *J. Saf. Res.*, vol. 37, no. 4, pp. 401–411, Jan. 2006.

[11] F. Kaneko, "Methods for probabilistic safety assessments of ships," *J. Mar. Sci. Technol.*, vol. 7, no. 1, pp. 1–16, Jun. 2002.

[12] C. D. Lewis, *Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting*. London, U.K.: Butterworth-Heinemann, 1982.

[13] S. Li, Q. Meng, and X. Qu, "An overview of maritime waterway quantitative risk assessment models," *Risk Anal.*, vol. 32, no. 3, pp. 496–512, Mar. 2012.

[14] T. Macduff, "Probability of vessel collisions," *Ocean Ind.*, vol. 9, no. 9, pp. 144–148, 1974.

[15] J. Montewka, T. Hinz, P. Kujala, and J. Matusiak, "Probability modelling of vessel collisions," *Rel. Eng. Syst. Saf.*, vol. 95, no. 5, pp. 573–589, May 2010.

[16] A. Papanikolaou, K. Bitha, E. Eliopoulou, and N. P. Ventikos, "Statistical analysis of ship accidents that occurred in the period 1990–2012 and assessment of safety level of ship types," in *Proc. Maritime Technol. Eng. Conf. (MARTECH)*, 2014, pp. 227–233.

[17] T. Statheros, G. Howells, and K. M. Maier, "Autonomous ship collision avoidance navigation concepts, technologies and techniques," *J. Navigat.*, vol. 61, no. 1, pp. 129–142, Jan. 2008.

[18] A. Toffoli, J. M. Lefèvre, E. Bitner-Gregersen, and J. Monbaliu, "Towards the identification of warning criteria: Analysis of a ship accident database," *Appl. Ocean Res.*, vol. 27, no. 6, pp. 281–291, Dec. 2005.

[19] A. Toffoli, J. M. Lefevre, J. Monbaliu, H. Savina, and E. Bitner-Grgersen, "Freak waves: Clues for prediction in ship accidents?" in *Proc. 13th Int. Offshore Polar Eng. Conf.*, Honolulu, HI, USA, 2003, pp. 23–29.

[20] P. Trucco, E. Cagno, F. Ruggeri, and O. Grande, "A Bayesian belief network modelling of organisational factors in risk analysis: A case study in maritime transportation," *Rel. Eng. Syst. Saf.*, vol. 93, no. 6, pp. 845–856, Jun. 2008.

[21] J. H. Wright, "Alternative variance-ratio tests using ranks and signs," *J. Bus. Econ. Statist.*, vol. 18, no. 1, pp. 1–9, Jan. 2000.

[22] D.-L. Xu, J.-B. Yang, and Y.-M. Wang, "The evidential reasoning approach for multi-attribute decision analysis under interval uncertainty," *Eur. J. Oper. Res.*, vol. 174, no. 3, pp. 1914–1943, Nov. 2006.

[23] Z. Xu, J. Mu, and Y. Ji, "A study of 3D model of ship domain for inland waterway," *J. Wuhan Univ. Technol. Transp. Sci. Eng.*, vol. 28, no. 3, pp. 380–383, 2004.

YINGFAN CHEN received the M.Sc. degree from Wuhan University of Technology, where he is currently pursuing the Ph.D. degree in traffic control and information engineering. His research interests include autonomous navigation, smart intelligence, and information control.

CHENG XIE received the M.Sc. and Ph.D. degrees from Wuhan University of Technology, in 2015 and 2020, respectively. He is currently holding a postdoctoral position in traffic control and information engineering with Wuhan University of Technology. His research interests include safety engineering and shipping pollution prevention.

SHUZHE CHEN received the M.Sc. and Ph.D. degrees from Wuhan University of Technology, in 2006 and 2012, respectively. He is currently an Associate Professor with the School of Navigation, Wuhan University of Technology. He has published more than 30 scientific articles in international journals and conferences. His research interests include risk navigation and accident playback inversion.

LIWEN HUANG received the M.Sc. and Ph.D. degrees from the Ocean University of China (Qingdao), in 1993 and 1997, respectively. He is currently a Professor and a Ph.D. Supervisor with the School of Navigation, Wuhan University of Technology. He undertook more than ten national scientific research projects as the director. He gets the special allowance granted by State Council, since 2015. He has published three books and more than 100 scientific articles in international journals and conferences. His research interests include safety shipping, intelligent navigation, and ship pollution prevention technology. He is an Expert of the National Science and Technology Award. He is the Vice Chairman of the Inland Navigation Committee of Chinese Maritime Association.