

Article

A Novel Time-Frame Regional Collision Risk Model Based on Dynamic Time Warping

Zihao Liu *  and Peijun Yu

Navigation College, Dalian Maritime University, Dalian 116026, China; yupj@dlmu.edu.cn

* Correspondence: zihao.liu@dlmu.edu.cn

Abstract: The quantification of collision risk in the water area is crucial for marine traffic safety. However, most existing studies primarily adopt a statistical perspective or only consider collision risks at discrete time points, neglecting the dynamic nature and temporal characteristics of collision risk. To overcome the problem, this paper proposed a novel time-frame regional collision risk model based on dynamic time warping (DTW). The time series of the ships in the water area were generated by modeling the collision risk from the perspectives of geometric encounters and traffic situation pressure. Dynamic time warping was used to mine the eigenvalue of each sequence in representing collision risk. The time-frame regional collision risk can be obtained by synthesizing the collision risk eigenvalues through considering the contribution of each ship to the entire collision risk. To validate the proposed model, some experiments were carried out by using the real AIS data in the Bohai Strait. The results show the capability of the proposed model in reasonably and effectively identifying the time-frame regional collision risk and prove its merits compared with the traditional methods. Therefore, it can better assist the supervision and analysis of regional collision risk so as to further enhance marine traffic safety.

Keywords: collision risk; time series; dynamic time warping; geometry encounter; traffic situation pressure



Academic Editors: George Zaraphonitis, Sebastian Feuerstack, Ivica Pavic, Andrej Androjna, Maciej Gucma and Marko Perkovic

Received: 18 December 2024

Revised: 24 January 2025

Accepted: 29 January 2025

Published: 30 January 2025

Citation: Liu, Z.; Yu, P. A Novel Time-Frame Regional Collision Risk Model Based on Dynamic Time Warping. *J. Mar. Sci. Eng.* **2025**, *13*, 263. <https://doi.org/10.3390/jmse13020263>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the swift advancement of the global shipping industry, the prevention and management of ship collision accidents have emerged as crucial subjects in the field of maritime safety [1]. Ship collisions not only result in casualties and property damage but also have the potential to cause significant pollution in the marine environment, influencing international trade and economic activities [2]. Despite the gradual increase in the volume of ship traffic making significant contributions to global economic growth, it concurrently poses greater threats to maritime traffic safety and poses challenges to the supervision of ship traffic [3]. The growth in ship traffic leads to an increase in the density of ships in water areas, resulting in particularly congested conditions and elevating the likelihood of collision accidents. Meanwhile, environmental factors such as wave conditions and ocean currents may exacerbate the situation by intensifying the risk of collision. These natural conditions not only increase the complexity of navigation control but also potentially affect the ship's position and speed, thereby elevating the likelihood of accidents. In response to such a situation, conducting a comprehensive and quantitative analysis of collision risks in specific water areas during specific periods is essential for guaranteeing maritime traffic safety in that region.

To enhance maritime traffic safety, numerous scholars have conducted quantitative research on the collision risks of water areas.

In the early stages, marine traffic researchers typically employ the traffic flow probability model to evaluate collision risks within a specific water area. This involves constructing a mathematical model based on the relationship between two traffic flows and relevant parameters, enabling the estimation of ship collision frequency or probability. Initial studies by Fujii and Shiobara [4], as well as Macduff [5], lay the foundation for subsequent advancements made by scholars. For instance, Pedersen [6] substitutes traffic flow for ship density in Fujii and Shiobara's model to estimate geometry collision frequency, and subsequent studies have extensively applied this approach, such as Kujala et al. [7] employing it in marine safety analysis conducted in the Gulf of Finland, Silvera et al. [8] utilizing it to identify the collision candidate ships off the coast of Portugal, and Weng [9] and Chai [10] using the relationship between ship domains in the framework of collision probability to estimate the geometry collision risk, etc.

With advancements in information technology, scholars have increasingly adopted more sophisticated methodologies to comprehensively model collision risks in an area. Hu and Park [11] propose a novel ship navigational collision risk model based on the assessment of collision risk and the vulnerabilities of marine accidents. The factors of vulnerabilities are integrated into the fuzzy reasoning system to evaluate the navigational conditions and environment, which is useful for collision risk assessment—especially for coastal operating ships and fishing ships in practice. For special areas such as arctic waters, Zhang et al. [12] propose an integrated risk assessment model for analyzing the risks of getting stuck in the ice and ship–ice collision based on a Bayesian network, which is realized by calculating the probabilities and severities. The proposed model is useful for identifying the safe speed for arctic navigation. Zhang et al. [13] propose a regional multi-ship collision risk level identification model based on a convolutional neural network. The method can recognize ship collision risks by matching the images of a multi-ship encounter. Utilizing the mechanism of image recognition, the proposed model can judge the collision risk of a multi-ship encounter beyond a simple pairwise comparison of ship encounters quickly and is useful when organizational resources are limited. Chen et al. [14] propose a regional ship collision risk analysis method based on velocity obstacles, enabling the simultaneous identification of collision risks and the complexity of encounter situations. In addition, a density-based clustering method is used to identify the hotspots of marine traffic so as to assist the maritime safety operators. Zhang et al. [15] propose a machine-learning method for the collision risk evaluation, where K-means algorithm and DBSCAN algorithm are used. An avoidance behavior-based collision detection model is introduced to identify potential collision scenarios and the collision risk can be quantified when the evasion action is made. Rong et al. [16] conduct a spatial correlation analysis of near-collision hotspots based on the local maritime traffic characteristics off the coast of Portugal. They employ Moran's I and the Getis–Ord G_i^* spatial autocorrelation method to identify the spatial clustering pattern of near collisions from the local and global angle, which can provide a means of improving maritime safety. Weng et al. [17] conduct an analysis of the impact of ship domain shapes on ship collision risk identification, taking into account both collision frequency and severity. The overtaking encounter is found to have a low severity and the large-angle crossing encounter is found to have a low frequency and high severity. Based on the study, two high-risk areas with high frequency and severity are identified. Antão et al. [18] conduct a quantitative assessment of ship collision risk influencing factors on the collision probability. The method is inspired by the Cox proportional hazard regression model and integrates a Bayesian rule and least-squares method. The assessed factors are proven to have a great impact on ship collisions. Cheng et al. [19] propose a multi-

ship collision risk measurement method based on non-linear velocity obstacles and time proximity. The collision risk is measured from the aspects of both the difficulty and urgency of the encounter. The proposed model has been demonstrated to quantify changes in multi-ship collision risk in a more intuitive and effective manner.

As an important concept in interpreting collision risk, many scholars have been researching it. Fujii and Tanaka [20] first proposed the definition for the ship domain, describing it as the two-dimensional space surrounding a ship and asserting that other ships should avoid this area. Goodwin validated Fujii's theory through multiple simulation experiments and subsequently introduced a ship domain model comprising three fan-shaped regions [21]. Davis et al. [22] refined Goodwin's model by simulating it, resulting in a circular representation of the ship domain. Coldwell [23] investigated maritime traffic in restricted waters and confirmed Fujii's overtaking situation ship domain model. Pietrzykowski et al. [24] examined the impact of ship orientation and operator qualifications, integrating these factors with a ship collision risk model to derive a polygonal representation of the ship domain. Wang [25] proposed a four-dimensional ship domain model that incorporated fuzzy set theory to comprehensively account for various factors affecting the scale of the domain, leading to the development of a fuzzy four-dimensional ship domain model. Silveira et al. [26] defined the ship domain around each reference ship using derived traffic density matrices based on AIS data, proposing an adapted modified four-dimensional number ship domain model for improved applicability.

Analyzing the regional collision risk research, it can be found that the traditional methods usually identify the collision risk of a given period using a statistical method, which is limited in generating time series of collision risk and lacks an analysis of the process of change. For some studies that model collision risk from a microscopic perspective, the entire collision risk is modeled through the results of each isolated timestamp, which neglects consideration of the time continuity, and makes it difficult to mine the characteristics of the time series in influencing the regional collision risk during the specified time period. Therefore, this paper proposed a novel time-frame regional collision risk model based on dynamic time warping (DTW), which can model the time-frame regional collision risk in a more scientific manner. The proposed model can identify the regional collision risk for a given period based on time series of collision risk, and it is capable of considering the time continuity of collision risk and analyzing the dynamic changing process by extracting the eigenvalues from each sequence that represents the collision risk. The proposed model also takes into account the contribution of each ship to the entire collision risk. Meanwhile, the proposed model demonstrates enhanced capability and flexibility in addressing the collision risks with a nonlinear time difference and provide more solid data foundation for the predictions of the future collision risk. The reminders of the paper are arranged as follows: In Section 2, the dynamic time warping-based time-frame regional collision risk model is established with the generation of the collision risk faced by a ship from the perspectives of geometric encounters and traffic situation pressure. To validate the effectiveness of the proposed model, some experimental case studies were conducted as described in Section 3 by using real AIS data from the Bohai Strait. In Section 4, the merits of the proposed model are discussed by some additional comparison experiments. The reliability of the proposed model was systematically analyzed, and the limitations of the proposed model were also analyzed. At last, conclusions are drawn in Section 5, and the future directions of this study are summarized.

2. Methodology

Regional collision risk refers to the potential occurrence of collision accidents involving all ships or ship traffic within a specific water area. In order to quantitatively assess the

collision risk of a water area within a designated time frame, it is essential to examine the level of collision risk faced by each individual ship contributing to this risk in the water area. The collision risk faced by a ship in a given water area should be viewed as continuous over time, rather than based on isolated incidents. Therefore, the collision risk for a specific period of time in the water area should be an organic combination of the risks encountered by all ships navigating within that area during their respective voyages throughout that period.

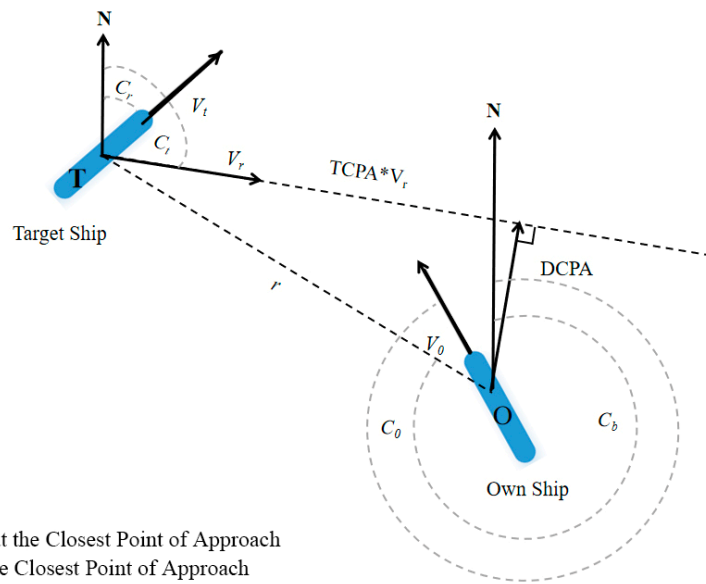
2.1. Collision Risk Faced by the Ship at an Individual Timestamp

In order to generate the time series of collision risk, the initial stage is to model the collision risks encountered by each ship at every individual timestamp in chronological order, which measures the likelihood of collision in the perspectives of geometric encounters and traffic situation pressure comprehensively.

2.1.1. The Collision Risk of Geometry Encounter

The movement trajectory of a transport ship sailing in a specific water area, following an established route and destination, typically involves continuous sailing after entering the water area and eventually exiting at a certain point. Furthermore, the navigation trajectory within the water area is also influenced by the size of said water area. If the water area is too large, the ship may remain in the water area most of the time, with no entry or exit. When conducting navigation in this water area, ships will encounter constantly changing collision risks over time as a result of alterations in the traffic environment they are exposed to. To be more specific, during the navigation process, ships will encounter different ships and varying numbers of ships at different moments, and the encounter states will be diverse. Sometimes ships will navigate independently with no other ships in the vicinity, while at other times they will encounter other ships and be required to perform collision avoidance maneuvers. The encounters can be classified as two-vessel encounters, multi-vessel encounters, etc.

Hence, for quantifying the collision risk faced by a ship during navigation, one of the relatively important factors is to describe the collision risk constituted by the surrounding ships from the perspective of geometric encounters through collision avoidance parameters. Commonly used geometric encounter collision avoidance parameters are often described in terms of both time and space, namely the Distance at the Closest Point of Approach (DCPA) and the Time to the Closest Point of Approach (TCPA) between ships. The DCPA and TCPA can, respectively, measure the collision risk between ships from the spatial and temporal dimensions. When jointly utilized, they can effectively conduct quantitative assessment of the collision risk between ships. Therefore, they are widely applied in collision avoidance maneuvers as key parameters for judging the timing and effect of collision avoidance. Based on AIS data, the DCPA and TCPA between two ships can be calculated through the following algorithm (Algorithm 1) in a relative motion triangle as shown in Figure 1.



DCPA: Distance at the Closest Point of Approach
TCPA: Time to the Closest Point of Approach

Figure 1. The relative motion between ships for calculating DCPA and TCPA.

Algorithm 1. The calculation of DCPA and TCPA.

Input: Ship1_position (x_0, y_0) , speed (V_0) , course (C_0) , Ship2_position (x_t, y_t) , speed (V_t) , course (C_t) ,

Output: DCPA TCPA

1. Convert the course to a radian
 2. $V_r = \sqrt{V_t^2 + V_0^2 - 2 * V_t * V_0 * \cos(C_t - C_0)}$
 3. $C_b = \arctan((x_t - x_0) / (y_t - y_0))$
 4. IF $C_b < 0$:
 5. $C_b = C_b + 2\pi$
 6. $R = \sqrt{(x_0 - x_t)^2 + (y_0 - y_t)^2}$
 7. FOR each ship i
 8. FOR each ship $i + 1$
 9. IF $V_r = 0$
 10. $TCPA = NaN$
 11. $DCPA = R$
 12. ELSE
 13. $DCPA = R * \sin(C_r - C_b - \pi)$
 14. $TCPA = R * \cos(C_r - C_b - \pi) / V_r$
 15. Continue to next ship pairs
-

Based on the calculated DCPA and TCPA, for a certain ship, the average spatial and temporal urgencies of the collision risks imposed by all other ships on it can be characterized by the following equations, and thereby describe the collision risks imposed by other ships on it in the studied water area from the spatio-temporal perspective of geometric encounters.

$$DCPA_{ship\ i} = \frac{\sum_{j=1}^n DCPA_{ship\ i - ship\ j}}{n - 1} \quad (1)$$

$$TCPA_{ship\ i} = \frac{\sum_{j=1}^n TCPA_{ship\ i - ship\ j}}{n - 1} \quad (2)$$

where $DCPA_{ship\ i-ship\ j}$ and $TCPA_{ship\ i-ship\ j}$ refer to the $DCPA$ and $TCPA$ between Ship i and Ship j , n refers to the number of ships, and $DCPA_{ship\ i}$ and $TCPA_{ship\ i}$ are the $DCPA$ and $TCPA$ faced by Ship i .

Utilizing negative exponential equation, the values obtained in Equations (1) and (2) can be standardized, thereby mapping their relationship with the collision risk. Referencing the collision risk index modeling in [27], the collision risk of a geometric encounter can be expressed as follows:

$$CR_{DCPA} = \alpha_D \cdot e^{-\beta_D \cdot DCPA_{ship}} \quad (3)$$

$$CR_{TCPA} = \alpha_T \cdot e^{-\beta_T \cdot TCPA_{ship}} \quad (4)$$

$$CR_{geometry} = \sqrt{CR_{DCPA}^2/2 + CR_{TCPA}^2/2} \quad (5)$$

where the coefficients α and β in Equations (3) and (4) are the adjustment coefficients of the negative exponential equation, which can be decided by presetting the values of the least and most dangerous situation. Specifically, according to the collision risk research [28], two sets of input and output can be preset, such as a very dangerous $DCPA$ value corresponding to a relatively high collision risk value and a very safe $DCPA$ value corresponding to a relatively low collision risk value, and the parameters can be solved by the two equations established by the two sets of input and output values.

2.1.2. Collision Risk of Traffic Situation Pressure

For a ship, during its navigation process, apart from the collision risks at the temporal and spatial encounter level from other individual ships, the overall marine traffic situation in its vicinity will likewise exert an influence on the collision risk [29]. The deck officers' environmental and situational awareness pressure will continue to intensify as the surrounding ship traffic becomes increasingly congested and intricate [30]. The increasing pressure will have an impact on the accuracy and timeliness of its maneuvering and collision avoidance decisions, rendering the ship more susceptible to collision accidents. Consequently, in order to obtain a more precise assessment of the collision risk encountered by a ship during its voyage, it is imperative to consider the influence from its surrounding traffic situation.

For the surrounding traffic situation, the deck officers of a ship can most intuitively feel the traffic intensity within a certain range around the ship, which is determined by the number of existing ships in its vicinity. A higher number of ships within this range indicates greater traffic intensity and a more complex traffic situation, leading to potential complications and increased difficulty for multiple ships to avoid collisions. Consequently, deck officers will face significant pressure, further elevating collision risks for the ship.

The specific range of pressure brought by the surrounding ship is determined in this paper based on the navigation practice, focusing on the distance at which a ship typically initiates its first collision avoidance action. This area with this specific range was named as the perceptual field of traffic situation (PFTS).

In [31], Goodwin et al. determine the encounter rate by employing the ratio of the number of ships entering the ship domain to that entering a broader area. During the continuous approach of two ships, the encounter situation becomes increasingly perilous as the distance between them shortens. It is generally accepted that ships pass through three successive stages before a collision occurs: the stage of collision risk, the imminent situation, and the imminent danger. At the collision risk stage, the ships can mutually perceive the existence of danger. The ship's lights, serving as tools for mutual alerts, and their visible ranges are often utilized for delineating dangerous encounter situations. According to the requirements of the International Regulations for Preventing Collisions at Sea (COLREGs) [32], for ships with a length of no less than 50 m, the minimum visible

distance of the masthead light is 6 nautical miles, used to warn other ships of the presence of collision risk and reinforce their initiation of collision avoidance. Hence, the range of the minimum visible distance of the masthead light where the warning of collision risk emerges was taken as the range of the PTFS, thereby demarcating the range of the dangerous area around a ship that requires consideration. An illustration of the PTFS is shown in Figure 2.

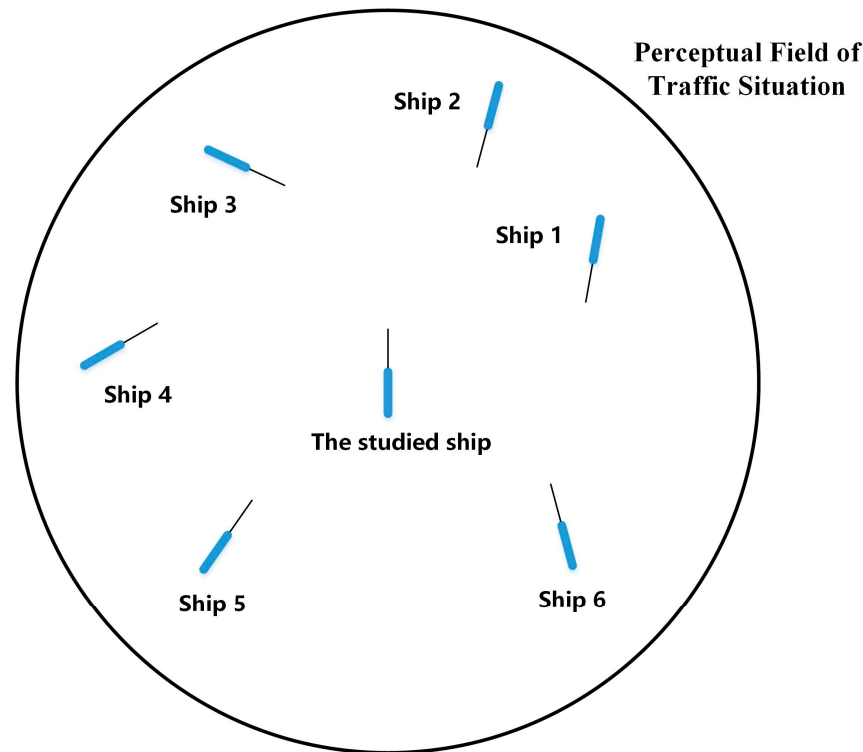


Figure 2. Illustration of the perceptual field of traffic situation.

In order to obtain the traffic intensity in PFTS, by incorporating the average density of the water area, an upper threshold for the number of ships surrounding a given ship was established. Subsequently, the collision risk based on the level of traffic situation pressure from the perspective of traffic intensity can be perceived by said ship using the following equation:

$$CR_{intensity} = \frac{N_{TI}}{N_{TI}^{max}} \quad (6)$$

where N_{TI} refer to the number of ships in PFTS; and N_{TI}^{max} refers to the max number of ships in PFTS, which can be obtained based on the average ship density of the studied water area for a relatively long period.

The perceived pressure faced by the deck officers is determined not only by the number of ships in a certain range around it, but also by the overall compact distribution of ships. The more closely ships are arranged, the less space there is for effective collision avoidance maneuvers for each ship, making it increasingly challenging to avoid collisions and escalating the risk of collision. In order to measure the distribution compactness of ships, the Ship Domain Overlapping Index (SDOI) in [33] was adopted, which is shown in Figure 3, and was further improved to evaluate in this paper. The compactness of the ship arrangement around a given ship was quantified by the average SDOI formed between the focal ship and other ships within its PFTS, and it was named as the Vicinity Ship Domain Overlapping Index (V-SDOI) in this paper, which is expressed as Figure 4. The SDOI, serving as an amplification or reduction factor of the ship domain, takes into account the compact spatial distribution among ships while considering collision avoidance

regulations. The dotted circles in Figures 3 and 4 refer to the amplified or reduced domains of ships. On this basis, V-SDOI can quantify the overall compact spatial distribution of the surrounding ships in relation to a specific ship. V-SDOI can be calculated based on the following equations:

$$SDOI_{ship\ i-ship\ j} = \frac{dis_{ij}}{r_{ship\ i} + r_{ship\ j}} \quad \text{if } i \neq j \quad (7)$$

$$V-SDOI_{ship\ i} = \frac{\sum_{j=1}^n SDOI_{ship\ i-ship\ j}}{n-1} \quad \text{if } i \neq j \quad (8)$$

where dis_{ij} refers to the distance between the centers of the ship domains of Ship i and Ship j . In order to consider the impact of COLREGS [32], the ship domain with a smooth unequal sector proposed by [22] was used, which can be considered as the improvement of the classical ship domain proposed by Goodwin [21]. In the relevant research of collision risk, the ship domain considering the collision avoidance strategy is widely used [34,35]. In this ship domain, the ship should be shifted to the port and stern side of center in order to consider the influences from the ships coming from different sectors from the perspective of collision avoidance responsibility. $r_{ship\ i}$ and $r_{ship\ j}$ refer to the radius of the ship domains of Ship i and Ship j . $SDOI_{ship\ i-ship\ j}$ refers to the SDOI between Ship i and Ship j .

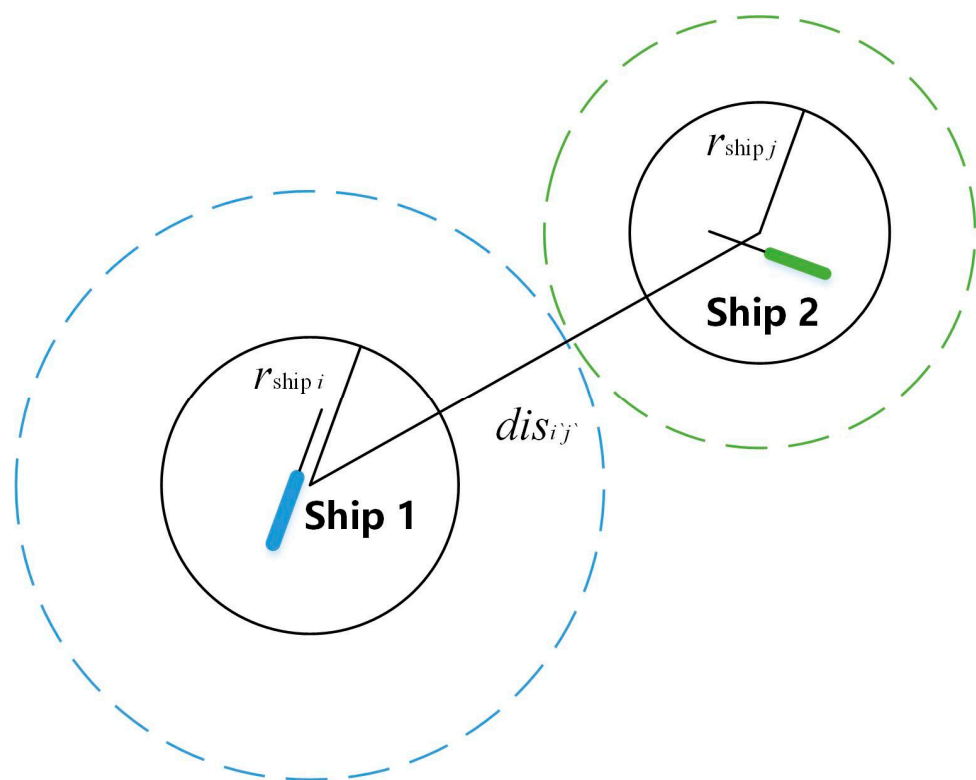


Figure 3. Illustration of SDOI.

Similarly, the result can be standardized by using a negative exponential equation:

$$CR_{compactness} = \alpha_S \cdot e^{-\beta_S \cdot SDOI_{ship}} \quad (9)$$

The adjustment coefficients can also be obtained by presetting the values of the extreme situations. The most critical situation occurs when ship domains align perfectly, while the safest situation arises when neighboring ships are positioned at the boundary of PFTS.

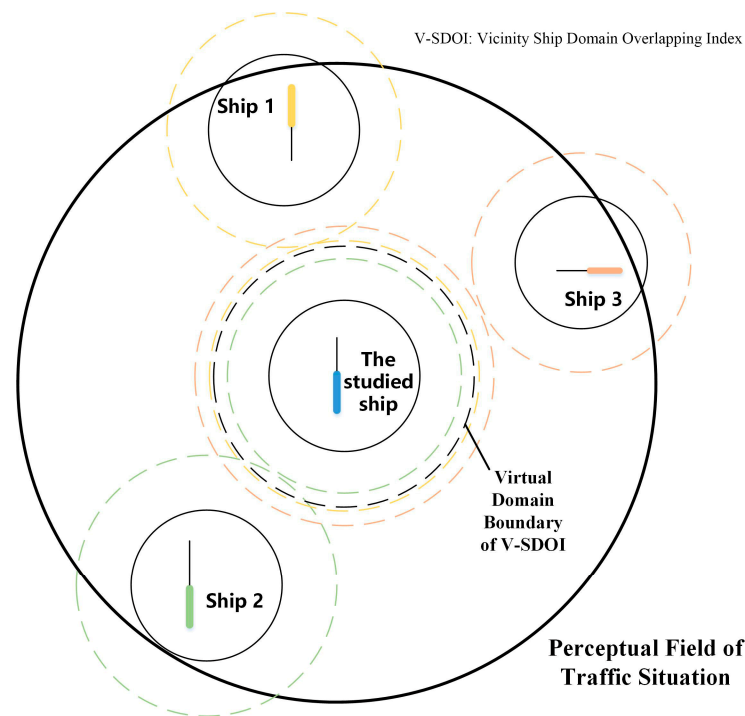


Figure 4. Illustration of V-SDOI.

The collision risk resulting from the overall traffic situation surrounding a ship can be calculated using the following equation. This, combined with the collision risk associated with the geometry encounter, allows for characterization of the perceived collision risk by a ship at any given moment.

$$CR_{situation} = \sqrt{CR_{intensity}^2/2 + CR_{compactness}^2/2} \quad (10)$$

$$CR_i^t = \sqrt{CR_{geometry}^2/2 + CR_{situation}^2/2} \quad (11)$$

where CR_i^t refers to the collision risk perceived by Ship i at timestamp t , namely the collision risk faced by the ship at an individual timestamp.

2.2. Time-Frame Regional Collision Risk Quantification

Based on the collision risk faced by the ship at an individual timestamp modeled in Section 2.1, the time series of collision risk perceived by the ship can be generated by combining these collision risk stamps in chronological order, which can express the collision risks faced by the ship over time:

$$CS_i = \{CR_i^1, CR_i^2, CR_i^3, \dots, CR_i^T\} \quad (12)$$

The time series represents the temporal evolution of collision risk faced by the ship while navigating in the studied water area. The overall collision risk within this period can be considered as an integrated combination of the aforementioned collision risks for all ships. The time series of collision risk for a specific ship can be analyzed not only to observe changes in collision risk over time from the perspective of the sequence itself, but also to further explore the intrinsic characteristics reflected in the representation of collision risk. This paper introduced a data-mining method, which is dynamic time warping, to further mine the characteristics of a time series by analyzing the fluctuation pattern of

the time series and identifying the similarity between it and dangerous sequence, thereby representing the collision risk by the cumulative distance between sequences.

For the similarity analysis of time series, a relatively simple method is used to calculate the Euclidean distance of two time series, which is a process of locking step measurement, where the time steps of the time series need to be aligned and the length of the series is consistent. However, the dynamic time warping algorithm can flexibly adjust the length of sequences by extending or compressing them based on the principle of finding the closest distance. It constructs correspondences between elements in two sequences of different lengths and evaluates their similarity. Therefore, dynamic time warping is an elastic measure, meaning it can assess the similarity between two time series CS_1 and CS_2 representing collision risk even if their time steps are misaligned and have different lengths.

The process of dynamic time warping mainly contains the establishment of a cumulative distance matrix and the finding of the shortest path, which is a process similar to dynamic programming. For two time series of collision risk CS_1 and CS_2 , the first step is to establish the distance matrix between them, where the column and the row correspond to the two series, respectively. The elements in the matrix represent the Euclidean distances between the collision risk stamps of CS_1 and CS_2 , corresponding to their respective rows and columns, as shown in Figure 5.

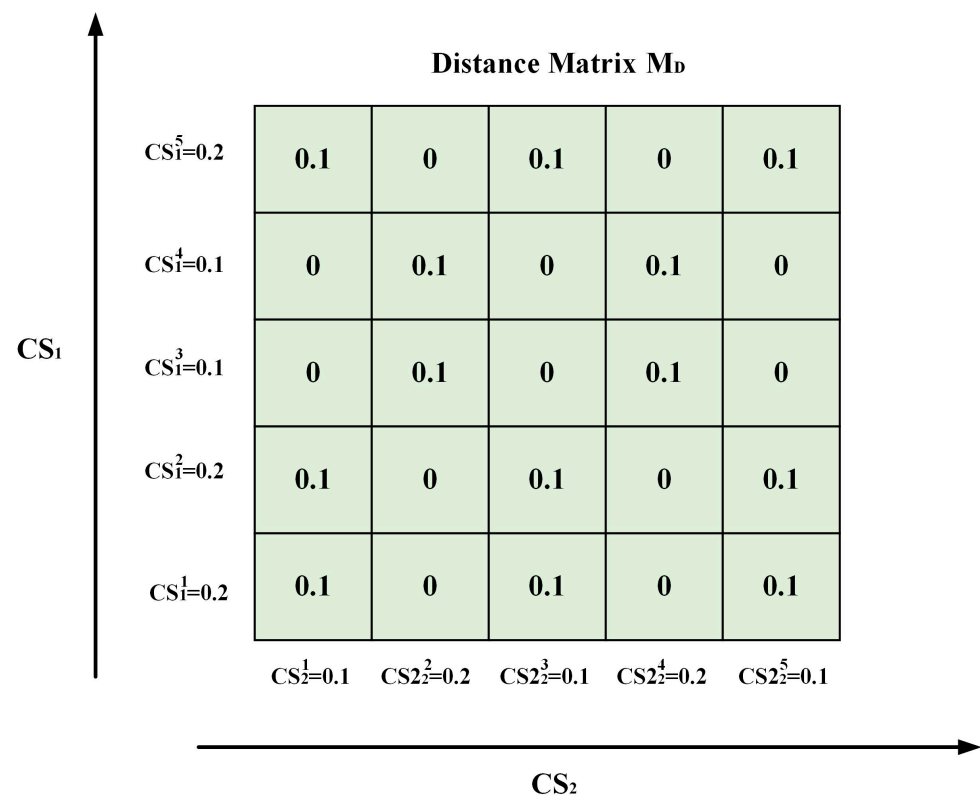


Figure 5. The distance matrix of dynamic time warping.

On the basis of the distance matrix, the cumulative distance matrix between the two collision risk time series CS_1 and CS_2 can be established, as shown in Figure 6. The cumulative matrix can represent the cumulative process starting from the first element between the two time series at the lower left corner of the matrix and ending at the last element of the two time series at the upper right corner. The initial conditions and recursive rules of the process can be expressed by the following equations.

$$M_C(1,1) = M_D(1,1) \quad (13)$$

$$M_C(p, q) = \min\{M_C(p, q-1), M_C(p-1, q), M_C(p-1, q-1)\} + M_D(1, 1) \quad (14)$$

where M_D refers to the distance matrix, M_C refers to the cumulative distance matrix, and p and q refer to the sequence number of the collision risk stamps in each time series. The recursive rule for calculating the cumulative distance is that the shortest path for each element is determined by adding its own distance to the shortest path of the previous element. The elements in the cumulative distance matrix generally have three adjacent elements, and the element to be added is the one with the shortest path, namely, the smallest value.

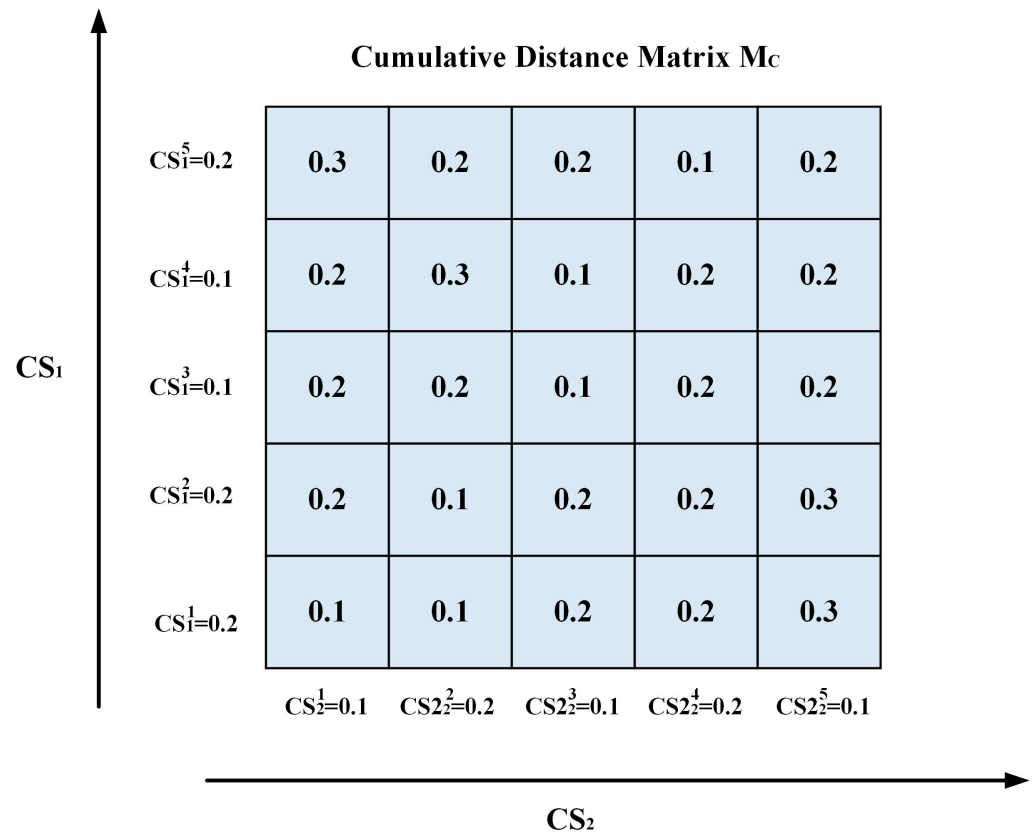


Figure 6. The cumulative distance matrix of dynamic time warping.

For some special cases, such as an element on the boundary of the cumulative distance matrix, since its previous element is only an adjacent element in the same column or row, the calculation of the cumulative distance can be performed according to the following rules:

$$M_C(p, 1) = M_C(p-1, 1) + M_D(p, 1) \quad (15)$$

$$M_C(1, q) = M_C(1, q-1) + M_D(1, q) \quad (16)$$

The final cumulative distance is in the upper right corner of the cumulative distance matrix, calculated according to the aforementioned rules. This value represents the similarity between two time series. A smaller value indicates a lower cost or distance between one series and another, indicating a higher level of similarity between the two time series, which can be expressed as the following equation. The process of using dynamic time warping to calculate the final cumulative distance is shown in Algorithm 2.

Algorithm 2. Dynamic time warping.

Input: The collision risk time series CS_1 and CS_2

Output: The final cumulative distance $distance_{cost}$

1. FOR each collision risk stamp CR_1^i in CS_1
2. FOR each collision risk stamp CR_2^j in CS_2
3. Calculate all distances: $dis_{1-2}^{i-j} = |CR_1^i - CR_2^j|$
4. Establish the distance matrix M_D with all distance dis_{1-2}^{i-j}
5. Initiate the cumulative distance matrix with the condition $M_C(1, 1) = M_D(1, 1)$
6. Fill the other elements $M_C(p, q)$ in M_C with the following rules
7. IF $p = 1$
8. $M_C(1, q) = M_C(1, q - 1) + M_D(1, q)$
9. ELSE IF $q = 1$
10. $M_C(p, 1) = M_C(p - 1, 1) + M_D(p, 1)$
11. ELSE
12. $M_C(p, q) = \min\{M_C(p, q - 1), M_C(p - 1, q), M_C(p - 1, q - 1)\} + M_D(1, 1)$
13. $distance_{cost}(CS_1, CS_2) = M_C(p_{max}, q_{max})$
14. Continue to next ship pair

$$distance_{cost}(CS_1, CS_2) = M_C(p_{max}, q_{max}) \quad (17)$$

Then, the most dangerous sequence was constructed to capture the characteristics of the collision risk time series, wherein the risk of collision remains consistently highest across all timestamps for ships navigating in these waters.

The time series of collision risk is constituted by the collision risks that each ship confronts at various time points. The most dangerous collision risk time series indicates that the ship is exposed to an extreme collision risk situation at every time point during its navigation in this water area, which is manifested in each sequence element; that is, the collision risk the ship encounters at each moment should be the maximum value of the collision risk it faces while navigating in this water area. The maximum value of the sequence elements needs to be comprehensively determined based on the range of the water area, the length of the study period, the traffic conditions in the water area, and the dimensions of the ships. For the studied water area, the longest line segment was taken as the navigation route of the ship of the most dangerous sequence, and the time unit is based on the time interval selected for the study. The parameters of the ship are obtained by conducting statistics on ships over a longer period of time, and the most dangerous sequence is formed by calculating the maximum collision risk the ship faces at each time point on the aforementioned route. These values will vary due to the traffic patterns at the ship's different location.

The dynamic time warping algorithm was then utilized to calculate the final cumulative distance between the specified collision risk time series and the most dangerous time series, which can be considered as its collision risk eigenvalue, and the negative exponential function was used to standardize the eigenvalue as follows:

$$CR_{ship\ i} = \alpha \cdot e^{-\beta \cdot distance_{cost}(CS_i, CS_D)} \quad (18)$$

where α and β are the adjustment coefficient, CS_i is the collision risk time series of Ship i , CS_D is the dangerous collision risk time series, and $CR_{ship\ i}$ refers to the collision risk faced by Ship i during the studied period.

The overall collision risk results in this period of time can be obtained by aggregating the collision risk eigenvalues based on the results in Equation (18). Given that each ship's

contribution to the overall collision risk varies within the water area during this period, it is necessary to assign different weights accordingly. As mentioned above, for each of the ships, as the durations of their presence in the water area are usually different, the contribution to the overall collision risk of the water area varies in terms of participation time, which was taken as the basis for calculating the weight of each ship in contributing to the overall collision risk. This consideration is expressed by Figure 7. Based on the weights and the collision risk eigenvalues calculated, the overall collision risk of the water area for a given period can be calculated by the following equations.

$$\omega_{ship\ i} = \frac{T_i^{end} - T_i^{initial}}{\sum_{i=1}^n (T_i^{end} - T_i^{initial})} \quad (19)$$

$$CR = \sum_{i=1}^n \omega_{ship\ i} \cdot CR_{ship\ i} \quad (20)$$

where $T_i^{initial}$ and T_i^{end} refer to the start time and end time of Ship i 's navigation within the designated study water area during the specified time frame; $\omega_{ship\ i}$ refers to the weight of Ship i on the overall collision risk, which can be obtained by dividing the time that the ship stays in the studied water area by the time that all the ships stay in the studied water area during the studied time period; and CR refers to the overall collision risk of the water area during the studied period.

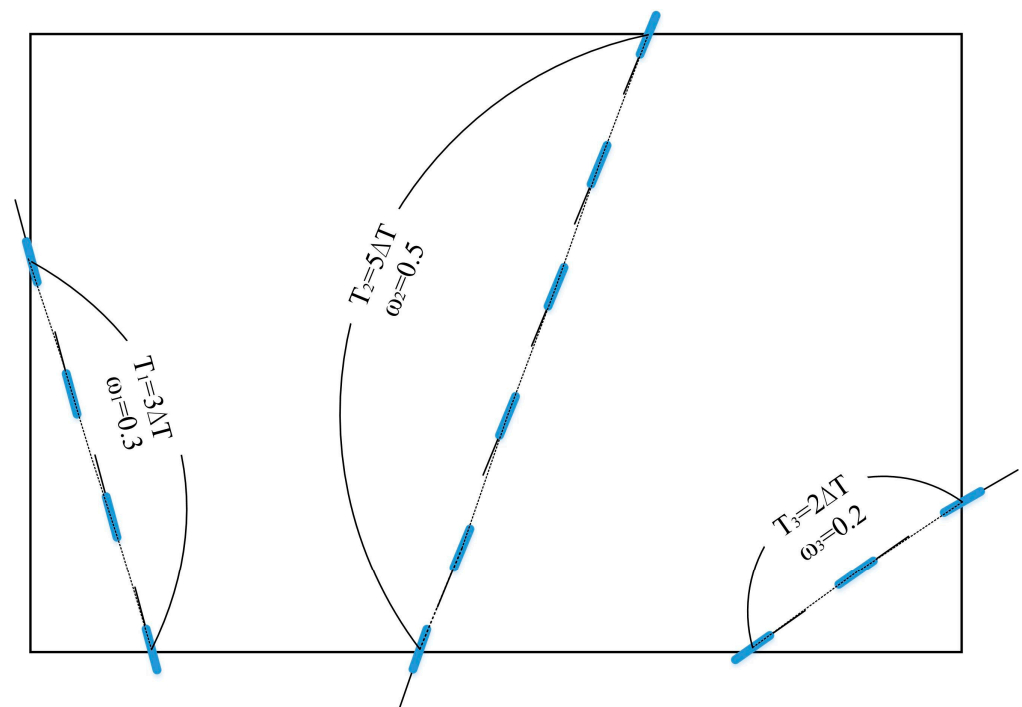


Figure 7. Illustration of determining the weights of ships.

The flow chart of the proposed model is shown in Figure 8. The results calculated by the proposed model enable effective characterization of the collision risk in the specified time period within the water area, providing a robust foundation for analyzing marine traffic safety in this region.

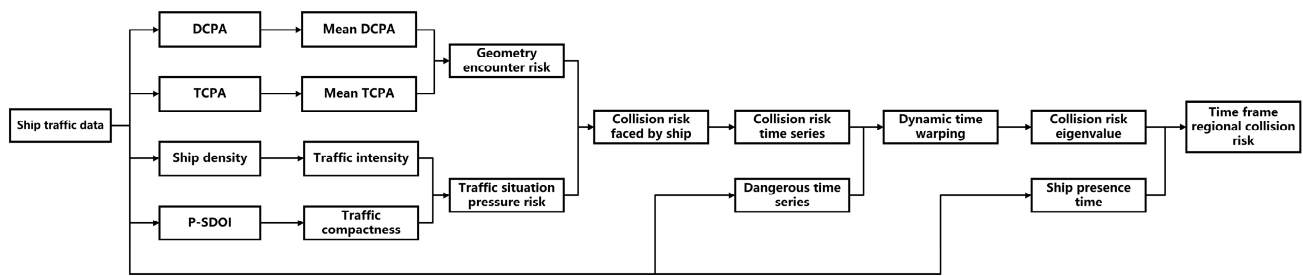


Figure 8. The flow chart of the proposed model.

3. Case Study

In order to validate the effectiveness of the proposed time-frame regional collision risk model, some experimental case studies were carried out by using the AIS data in the Bohai Strait. The studied water is located between the Bohai Sea and the Northern Yellow Sea of China, as shown in Figure 9. As the only way to enter or exit the Bohai Sea, the Bohai Strait was selected as the studied water area because the ship traffic intensity there is relatively high and the ship traffic is relatively complicated [36]; the traffic situation will lead to a more obvious collision risk between ships, which is suitable for the validation of the proposed model. Before experiments, AIS data should be processed first to remove anomalous information and align the timestamp.

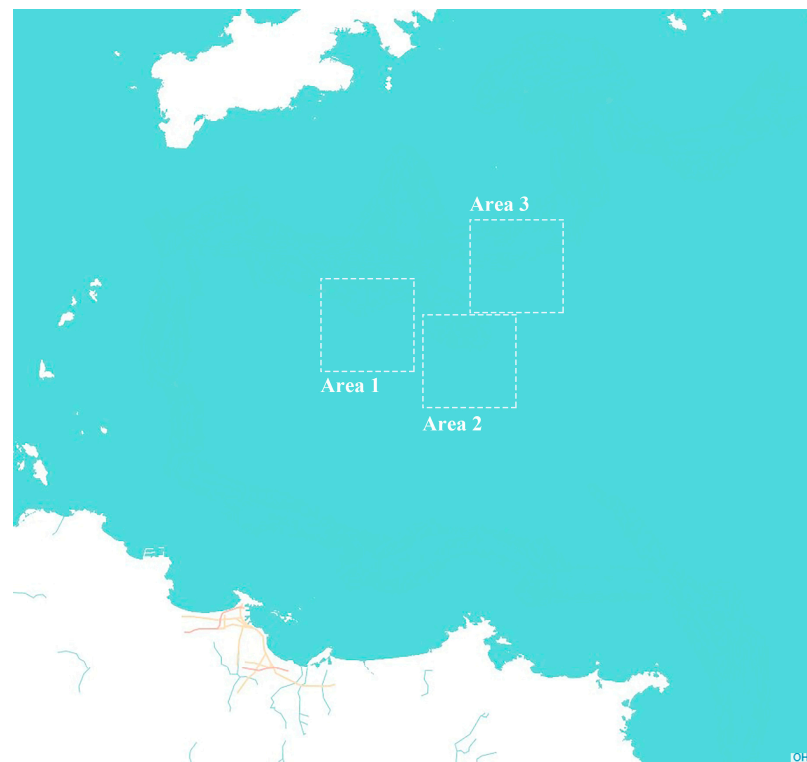


Figure 9. The studied water area [37].

Firstly, the proposed model was utilized to calculate the collision risk of Area 1 on 27 June 2022. Area 1 is an area with a size of $0.3^\circ \times 0.3^\circ$, which ranges from 121.45° E to 121.75° E and 38.15° N to 38.45° N. The whole time frame is from 0000 to 2400 on 27 June 2022. It is divided into 24 equal periods, and each period is divided by an interval of 300 timestamps. Area 1 is the intersection part of the two main traffic flows in the Bohai Strait. The area features relatively high traffic complexity because ships from different directions encounter others here with multiple encounter types. An example scenario

within the studied time frame is shown in Figure 10, which indicates the last interval of 0000–0100. The arrows refer to the speeds and courses of the ships. The collision risks of each hour in the studied period were calculated by using the proposed model, and the results are expressed in Figure 11.

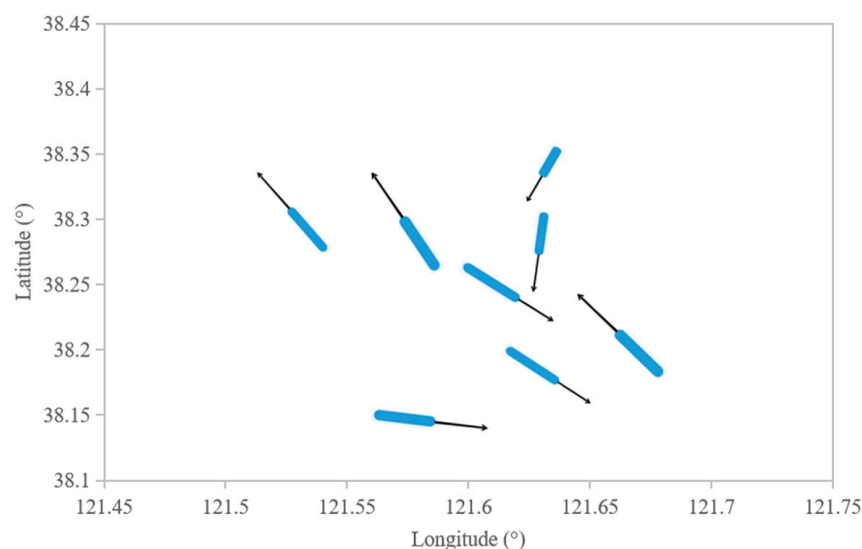


Figure 10. An example scenario within the studied time frame.

It can be found that the collision risks feature minor alteration, which means that the likelihood of the collision accident in this area is relatively stable. In order to further analyze the collision risk result, the index of traffic volume was introduced for comparison, because it can reflect the risk of marine traffic to some extent [38]. The traffic volumes of each hour in the studied period were obtained by calculating the number of ships passing by this area, and the results are shown in Figure 12.

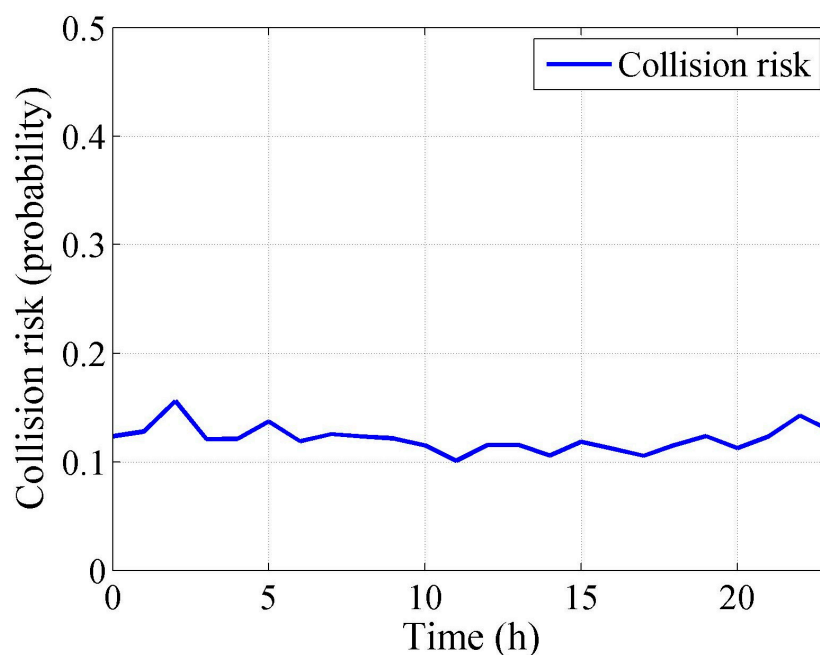


Figure 11. The collision risks of Area 1 on 27 June 2022.

Since traffic flow can represent the degree of collision risk to a certain extent, the correlation between traffic flow and collision risk was examined through Pearson correlation

analysis, providing evidence for the effectiveness of the proposed model in identifying collision risk. The correlation analysis is shown in Figure 13.

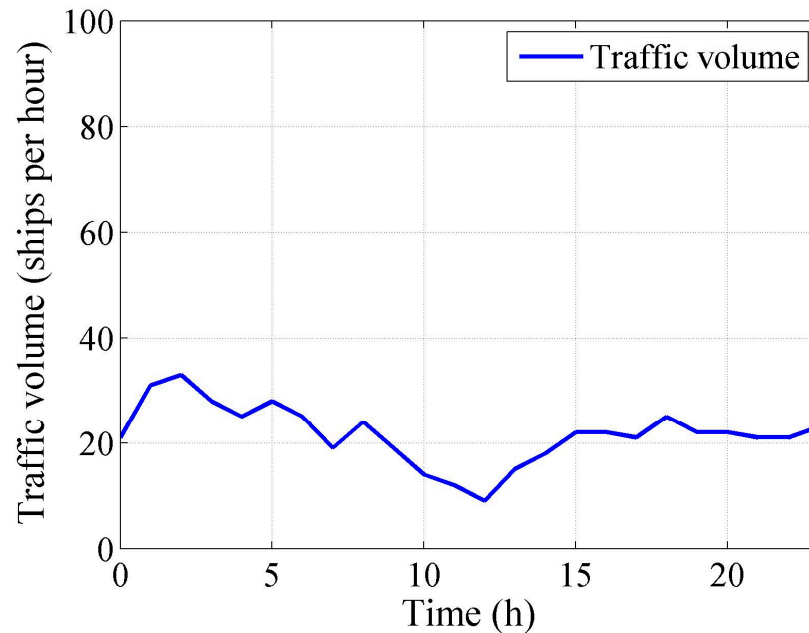


Figure 12. The traffic volumes of Area 1 on 27 June 2022.

It can be found that there is a positive correlation between the collision risks and the traffic volumes. The correlation coefficient was calculated to be 0.605 (relatively strong correlation), while the significance level yielded a result of 0.002. The correlation results show a certain correlation relationship between the collision risks and the traffic volumes at the significance level of 99%, which can prove the capability of the proposed model in identifying collision risk to some extent.

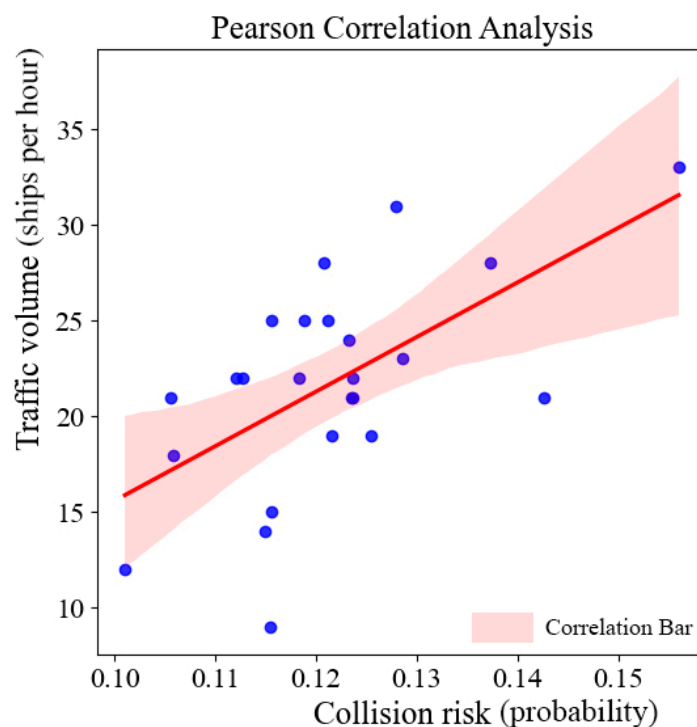


Figure 13. The Pearson correlation analysis between collision risks and traffic volume on 27 June 2022.

In addition, it can be found from the results that the risk of collision is higher during the night hours. Nighttime watchkeeping is more challenging compared to daytime watchkeeping due to the lack of natural light. Although artificial lighting equipment such as bow lights and navigation lights are used, their effective range is limited, leading to a significant reduction in visibility. In adverse weather conditions (such as fog, rain, or snow), nighttime visibility deteriorates further, increasing the difficulty of identifying other ships or obstacles [39]. While ship radars and other electronic devices can help to compensate for visual deficiencies, they may still encounter misjudgments or delayed responses in complex and dynamic maritime environments. Additionally, the human body clock is typically less active at night, resulting in slower reaction times and reduced attention spans compared to the daytime. Prolonged exposure to low-light conditions can lead to decreased alertness, further complicating watchkeeping efforts and increasing the risk of operational errors. This combination of factors significantly heightens the challenges of nighttime watchkeeping and increases the likelihood of operational mistakes.

In order to further demonstrate that the model proposed in this paper can effectively capture and quantify the potential collision risks in water areas, another index capable of depicting the likelihood of collision accidents occurring in a water area over a period of time was introduced to validate the results obtained from the model, which is the near miss. A near miss refers to instances where two ships will eventually pass each other at a very close distance. In such circumstances, although no collision occurs between the ships, the degree of collision risk between them is substantial. Therefore, near misses are also referred to as potential collisions and can reveal potential collision risks. Hence, the number of near misses during the research period in Area 1 as mentioned above was computed referring to the research in [40], and the results are presented in Figure 14.

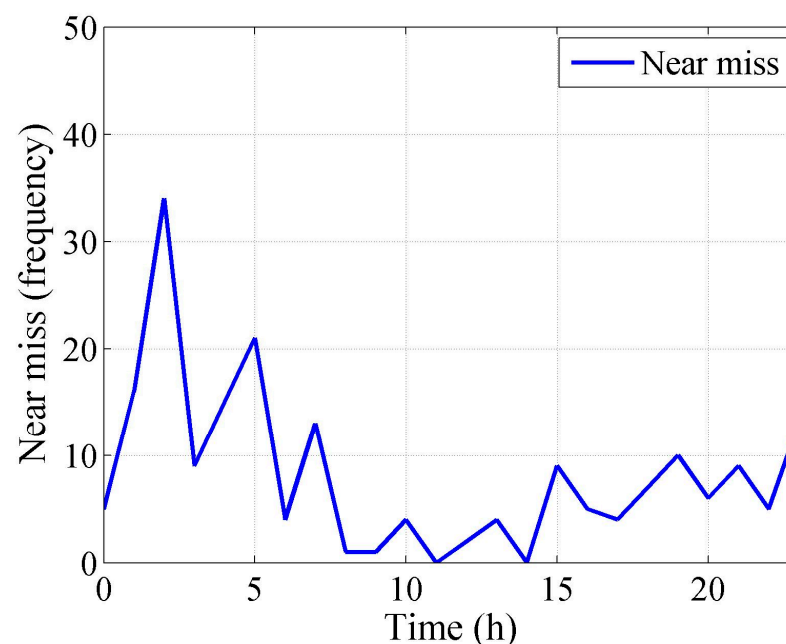


Figure 14. The near misses of Area 1 on 27 June 2022.

The correlation analysis between the near miss results and the collision risk results calculated by the model was also conducted, and the results are shown in Figure 15.

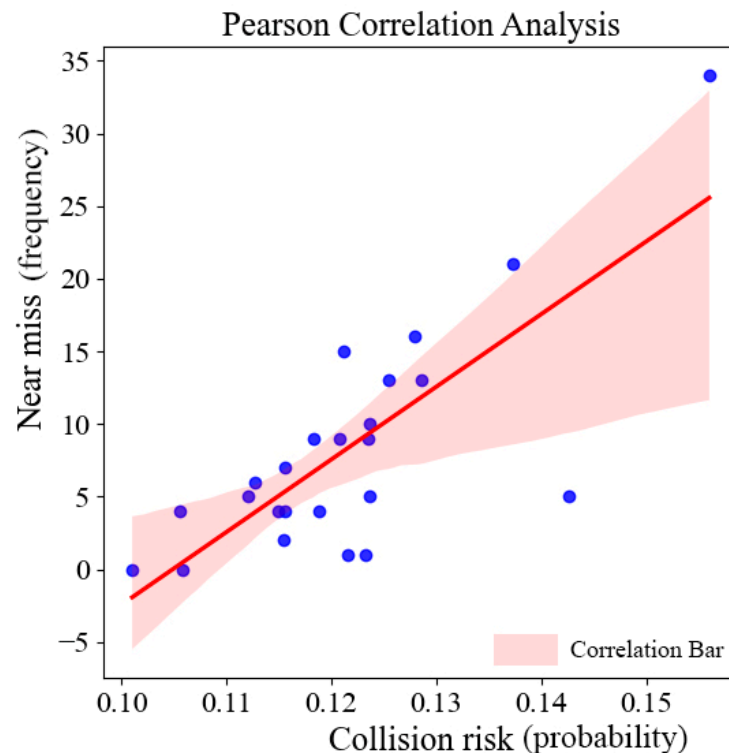


Figure 15. The Pearson correlation analysis between collision risks and near misses on 27 June 2022.

The results show that the collision risks are significantly correlated with the number of near misses at a 99% confidence level, with a correlation coefficient of 0.771 (strong correlation). The analysis results reveal a conspicuous positive correlation between the two variables, indicating that the model employed in this study effectively captures the actual potential collision risk in the water area and its validity is duly substantiated.

To further investigate the collision risks in Area 1, the experiment was carried out under the condition that the crossing flow of marine traffic was removed from the area. There would only be the traffic flow in the near east–west direction through the Bohai Strait in this water area, so the crossing encounters in the water area would be greatly reduced. As a result, the collision risk of this area is typically mitigated to a certain degree. In order to investigate the actual situation, the collision risk between 0200 am to 0300 am was calculated without the crossing flow based on the proposed model, and the result was 0.1341 with the traffic volume of 25 ships per hour. The collision risk value is obviously smaller than that when the crossing traffic flow exists, which is 0.1560. The results demonstrate that the model proposed in this paper effectively captures variations in collision risk resulting from changes in the encounter situation. Similarly, another experiment was conducted with removing the passing traffic flow. The collision risk of Area 1 under such conditions was recalculated using the model proposed in this paper, yielding a result of 0.1167 for a traffic volume of 15 ships per hour, which is smaller than the previous experimental outcome. This discrepancy can be attributed to the higher volume of ship traffic flows involved in passing rather than crossing traffic flow. By eliminating the flow, it is possible to further reduce collision risk within the water area while simultaneously decreasing the crossing encounters. The results obtained through the model presented in this paper align precisely with this objective reality, thus confirming its validity.

The following experiments were carried out in Area 2, which is part of the main passing traffic flow east to Area 1 in the Bohai Strait. Area 2 is an area with a size of $0.3^{\circ} \times 0.3^{\circ}$, which ranges from 121.85° E to 122.15° E and 38.00° N to 38.30° N. The time frame used is the same as the experiment of Area 1. Different from Area 1, Area 2 only

has the traffic flow in the near east–west direction, so the crossing encounters are very rare here and the traffic volume is also smaller. As a result, the collision risk in this area is typically lower compared to Area 1, and this is revealed by the collision risks calculated by the proposed model. The box charts for the collision risks and the traffic volumes in Area 1 and Area 2 are shown in Figures 16 and 17.

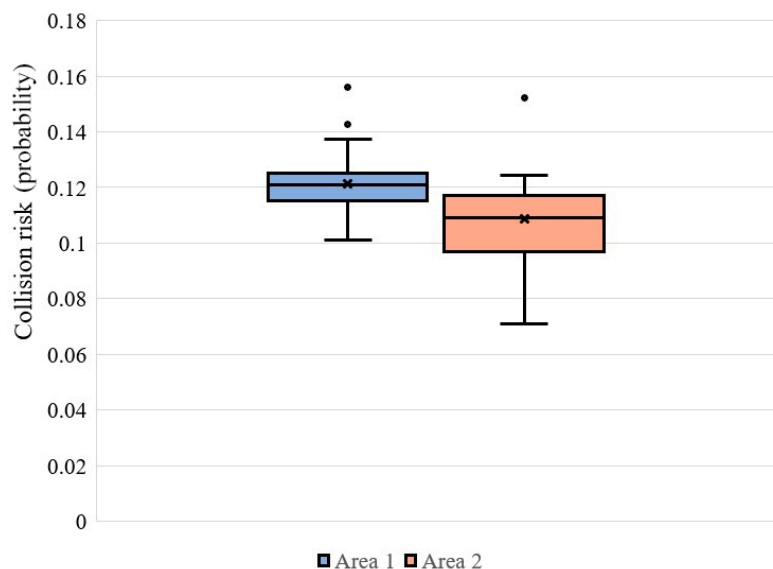


Figure 16. The box charts for the collision risks in Area 1 and Area 2.

It can be found that the collision risks of Area 2 are smaller than those of Area 1 on the whole, and the traffic volumes of Area 2 are also significantly smaller than those of Area 1. The results indicate that as the traffic flow decreases and the encounter situation becomes less complex, there is a corresponding decrease in collision risk. This observation aligns with the traffic reality, thus reaffirming the model's validity.

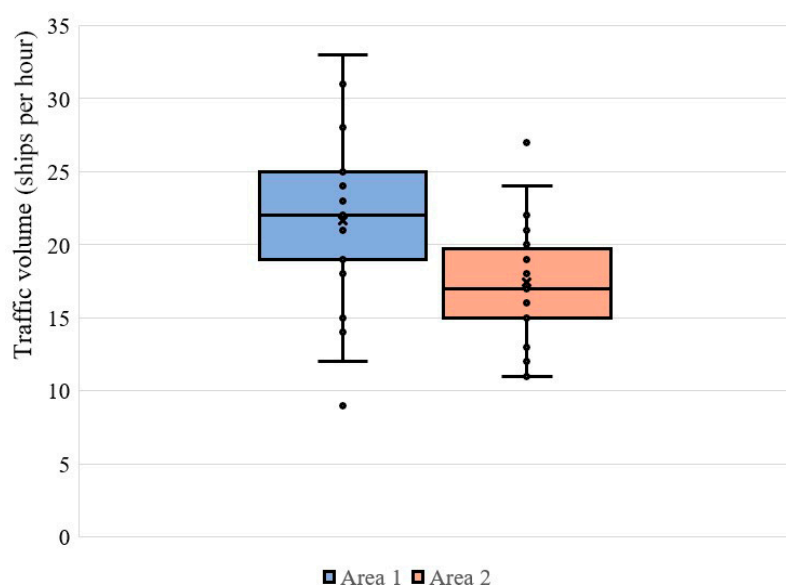


Figure 17. The box charts for the traffic volumes in Area 1 and Area 2.

In order to further investigate the collision risk in Area 2, the ship traffic in Area 2 was modified in a similar manner as that in Area 1 mentioned above. The eastbound and

westbound traffic flows were, respectively, eliminated within the area, and the collision risk for one hour within this area was recalculated. The results are presented in Table 1.

Table 1. The collision risks and traffic volumes of Area 2 in different traffic flow eliminating scenarios.

Situation	Collision Risk	Traffic Volume
With both flows	0.1230	27
Without eastbound flow	0.1082	17
Without westbound flow	0.0785	12

In addition, some random ships were added in the area for the experiment. The volume of ship traffic was initially increased from 18 to 23 per hour, and subsequently further increased to 26. The collision risk results calculated by the proposed model for the aforementioned scenarios are presented in Table 2.

Table 2. The collision risks of Area 2 with different traffic volumes.

Situation	Collision Risk	Traffic Volume
Scenario 0	0.1181	18
Scenario 1	0.1217	23
Scenario 2	0.1229	26

The above results demonstrate that the model is capable of providing an effective response to variations in ship traffic intensity within the studied area. When there is a reduction in both the volume and complexity of ship traffic, the risk of collision decreases accordingly. Conversely, with an increase in traffic due to additional ships and heightened complexity, the risk of collision also escalates. This indicates that the model proposed in this paper effectively and reasonably reflects changes in the collision risk based on objective conditions.

In order to further validate the model, experiments in Area 3 were conducted. Area 3 is a part of the secondary traffic flow and exhibits sparser ship trajectories compared to Area 2, indicating lower congestion levels. Area 3 is an area with a size of $0.3^\circ \times 0.3^\circ$, which ranges from 122.00° E to 122.30° E and 38.30° N to 38.60° N. The time frame used is the same as the experiment of Area 1. Similarly, the collision risks for a duration of 24 h in Area 3 were computed, with the corresponding results depicted in Figures 18 and 19.

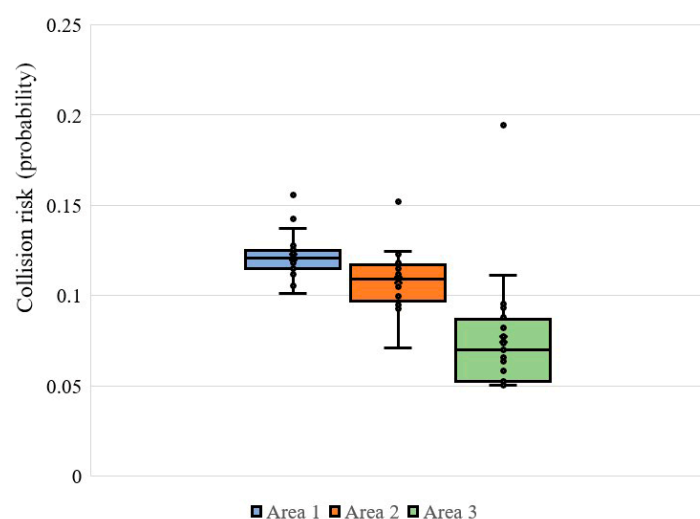


Figure 18. The comparison of the collision risks for the three studied areas.

The results indicate a reduction in collision risk in Area 3. This can be attributed to a significant decrease in traffic flow, resulting in reduced congestion and complexity within the area. Consequently, the collision risk is further minimized, reaffirming the effectiveness of the model. Additionally, the Pearson correlation analysis was performed by combining the results of Area 2 and Area 3 with those of Area 1. The correlation analysis is shown in Figure 20.

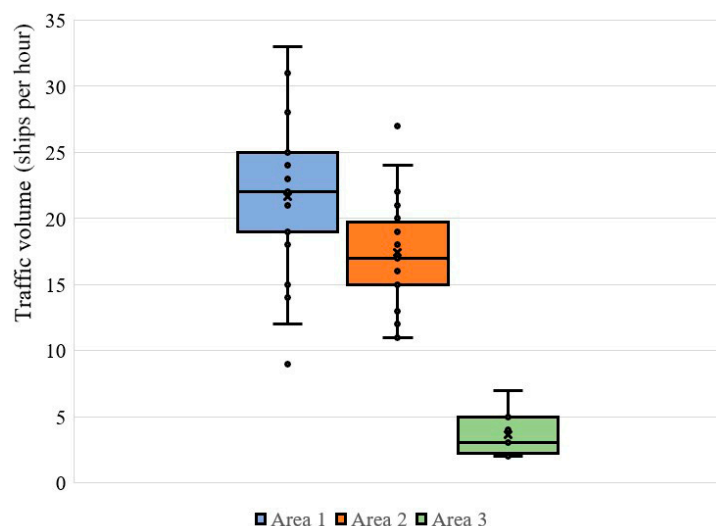


Figure 19. The comparison of the traffic volumes for the three studied areas.

It can be found that there is also a positive correlation between the collision risks and the traffic volumes. The correlation coefficient was calculated to be 0.677 (relatively strong correlation), while the confidence level is 99%. The inclusion of the results from Area 2 and Area 3 reveals a stronger correlation between the collision risk results and the traffic flow results, thereby demonstrating the model's capability to effectively differentiate collision risks in areas with varying degrees of traffic intensity and complexity. This further substantiates the efficacy of the proposed model.

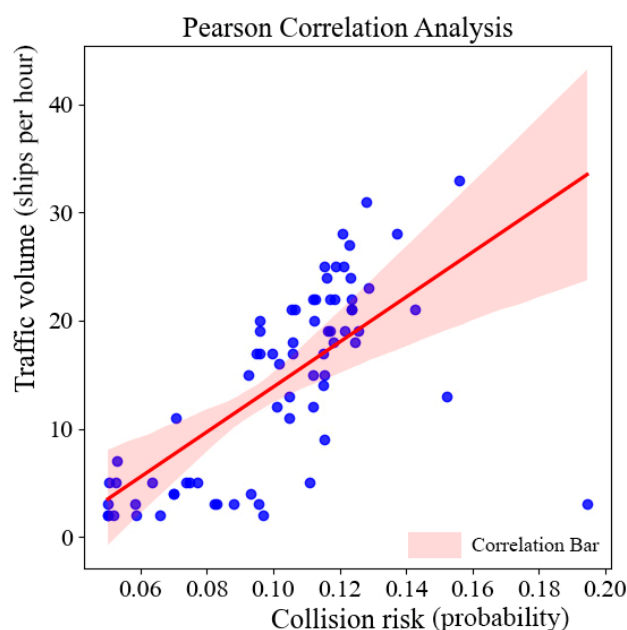


Figure 20. The Pearson correlation analysis between collision risks and traffic volumes of the combination results of the three studied areas.

A correlation analysis between the collision risk results and near miss results in the three areas was also conducted, and the findings are presented in Figure 21.

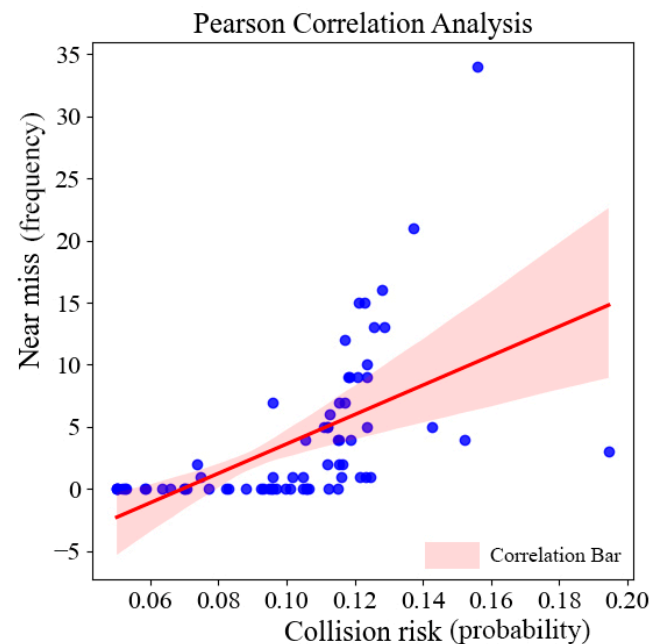


Figure 21. The Pearson correlation analysis between collision risks and near misses of the combination results of the three studied areas.

The results show that the collision risks are significantly correlated with the number of near misses at a 99% confidence level, with a correlation coefficient of 0.558 (relatively strong correlation). The presence of a positive correlation between the two variables remains evident, although affected by a lot of zero near misses in Area 3 to some extent. The results further substantiate the efficacy of the proposed model in identifying potential collision risk in the water area.

To further validate the effectiveness of the proposed model, the studied water area was extended to the majority of the Bohai Strait, with the area ranging from 121.0° E to 123.0° E and 37.9° N to 38.6° N. Within such a large region, the collision risk for each hourly period from 0000 to 2400 on 27 June 2022 was calculated, and the results are presented in Figure 22. To demonstrate the validity of the results, the traffic volume and the number of near misses during this period were also computed, and the correlation between them and the collision risk results was analyzed. The results are shown in Figures 23 and 24.

The results indicate that the collision risk results calculated by the proposed model in this water area are significantly correlated with the number of near misses at a 99% confidence level, with a correlation coefficient of 0.600. Moreover, the collision risk results are significantly correlated with the traffic volume at a 95% confidence level, with a correlation coefficient of 0.405. It can be observed that although both the traffic volume and the number of near misses present a certain positive correlation with the collision risk results, the correlation between the collision risk and the traffic volume is relatively weaker than that between the number of near misses. This is because after extending the water area to the majority of the Bohai Strait, the overall traffic volume in the water area is rather stable, remaining at approximately 110 ships per hour, with a relatively small variation on the overall scale. In contrast, the number of near misses does not tend to stabilize with the expansion of the water area and still exhibits heterogeneity at different times based on the varying encounter scenarios of ships, with a more pronounced variation range. Nevertheless, the positive correlation relationship manifested by the above results still

validates that the proposed model is effective even when the study water area is expanded to a larger extent.

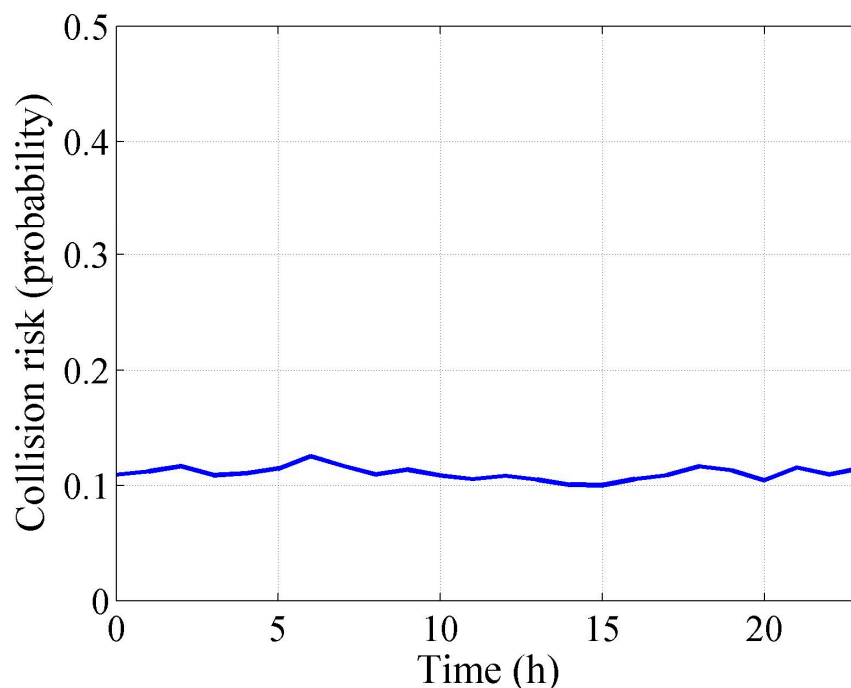


Figure 22. The collision risks of the extended area on 27 June 2022.

Furthermore, the studied water area of the Bohai Strait was divided into three sub-areas in accordance with their longitude. The specific division is depicted in the ship trajectory map in Figure 25, and the three sub-areas are A1, A2, and A3, respectively. Employing the model proposed in this paper, the 24 h collision risk results of the three sub-areas were calculated separately and are presented in Figure 26. Meanwhile, the 24 h traffic volume and the number of near misses in each sub-area were computed. Their average values as well as the average value of the collision risk are shown in Table 3.

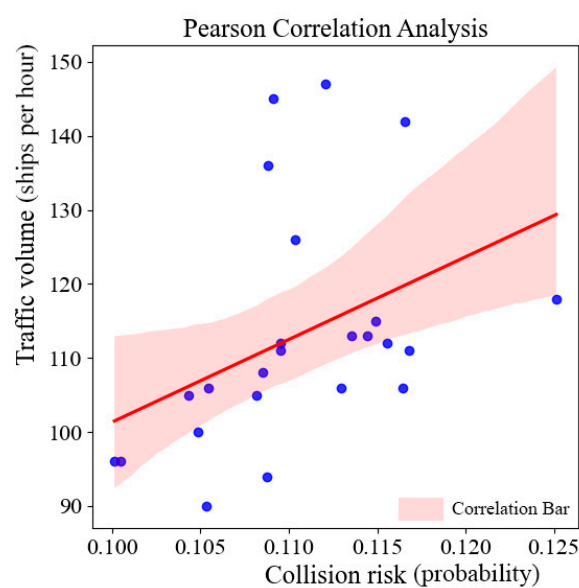


Figure 23. The Pearson correlation analysis between collision risks and traffic volumes of the extended area on 27 June 2022.

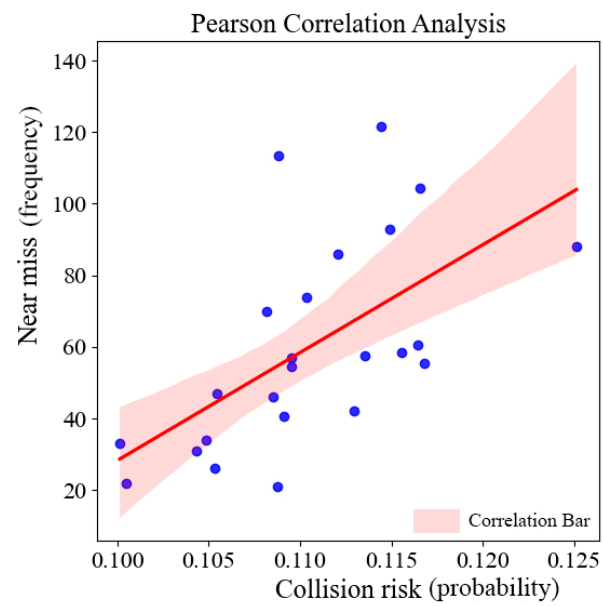


Figure 24. The Pearson correlation analysis between collision risks and near misses of the extended area on 27 June 2022.

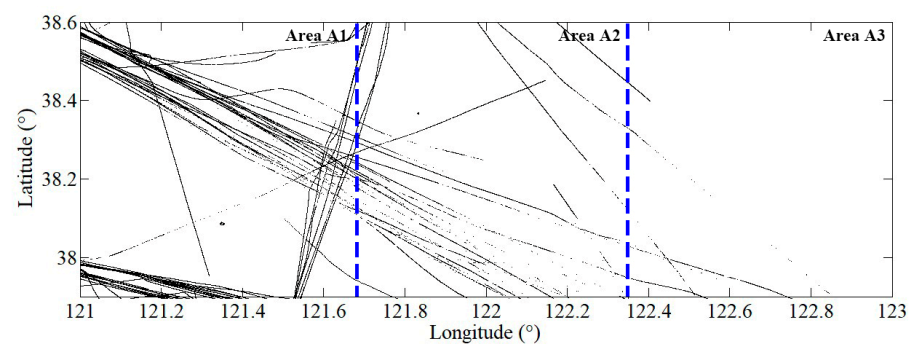


Figure 25. The ship trajectory map on 27 June 2022 and the sub-area divisions.

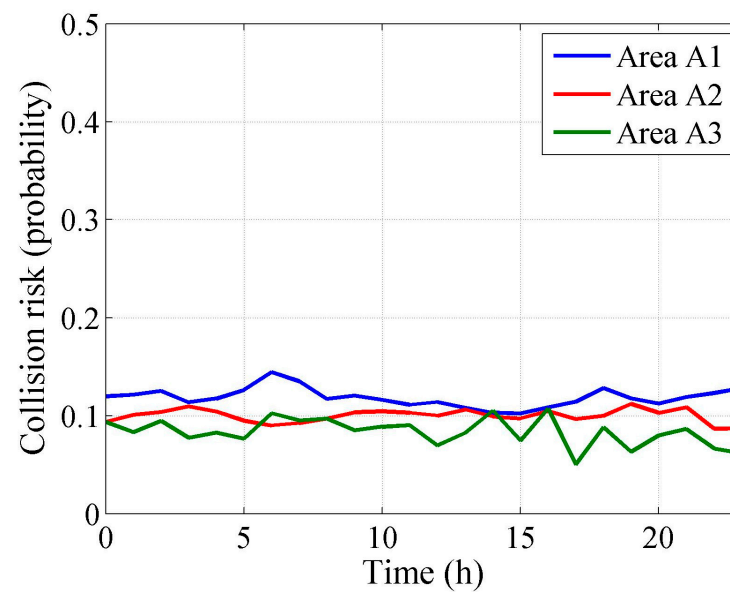


Figure 26. The collision risks for the three sub-areas A1, A2, and A3.

Table 3. The average values of collision risks, traffic volumes, and the numbers of near misses for the three sub-areas.

Area	Collision Risk	Traffic Volume	Near Miss
A1	0.1186	68.8333	41.6875
A2	0.0999	42.0000	10.9792
A3	0.0834	11.5833	1.1875

Firstly, it can be discerned from Figure 26 that the collision risk in Area A1 is conspicuously higher than that in Areas A2 and A3. Through analysis of the trajectory map in Figure 25, it is revealed that Area A1 encompasses the principal traffic flow in the northwest–southeast direction of the Bohai Strait and the crossing traffic flow approximating the south–north direction, with a dense array of ship trajectories and a substantial volume of ship traffic. Area A2 merely contains the principal traffic flow in the northwest–southeast direction and a portion of another traffic flow close to the south–north direction, featuring relatively sparse ship trajectories and a comparatively decreased volume of ship traffic. Area A3 merely encompasses a minor portion of the principal traffic flow and the traffic flow close to the south–north direction, with extremely sparse ship trajectories and an extremely small volume of ship traffic. The aforementioned facts can also be observed from Table 3. Whether it pertains to the traffic volume or the number of near misses, Area A1 is significantly higher than Areas A2 and A3, with Area A3 being the minimum. Commonly, the larger the volume of ship traffic and the denser the ship trajectories, the greater the risk of ship collision. This is because under such traffic conditions, the distance between ships is relatively small, and the situation is highly prone to the formation of complex and dangerous traffic outcomes. Under such circumstances, the free space for ships to avoid collision is diminished, and the difficulty of collision avoidance escalates. The results computed by the model proposed in this paper indicate that the collision risk in Area A1 is greater than that in Area A2, which is greater than that in Area A3, conforming to the abovementioned objective facts. Hence, it can further substantiate the effectiveness of the model. Additionally, in this study, a correlation analysis was carried out for the collision risk at all time points of the three sub-areas with the traffic volume and near misses. The correlation results are depicted in Figures 27 and 28. The results show that the collision risks are significantly correlated with the traffic volumes at a 99% confidence level, with a correlation coefficient of 0.835 (very strong correlation), while the collision risks and the near misses have a correlation coefficient of 0.753 at a 99% confidence level, indicating a strong correlation. Given that the traffic volume and near misses can represent the collision risk to some extent, the aforementioned experiments can further validate the effectiveness of the proposed model for wide areas.

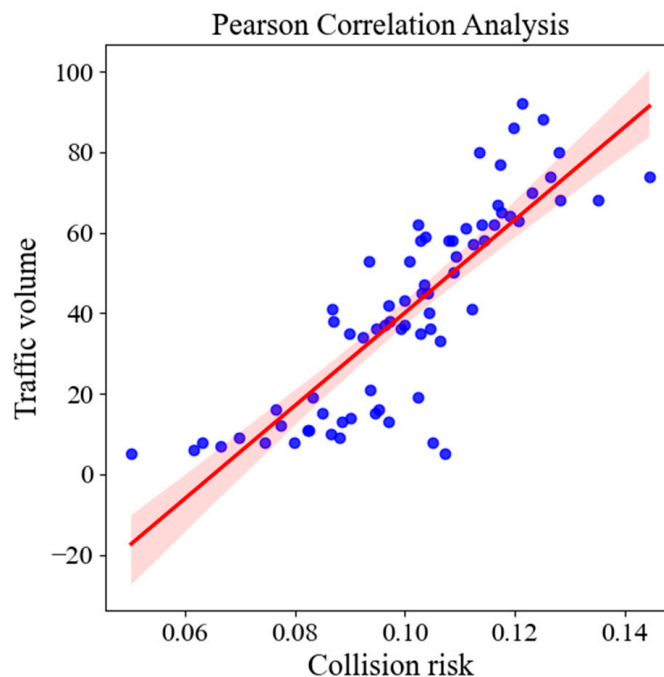


Figure 27. The Pearson correlation analysis between collision risks and traffic volumes of all three sub-areas on 27 June 2022.

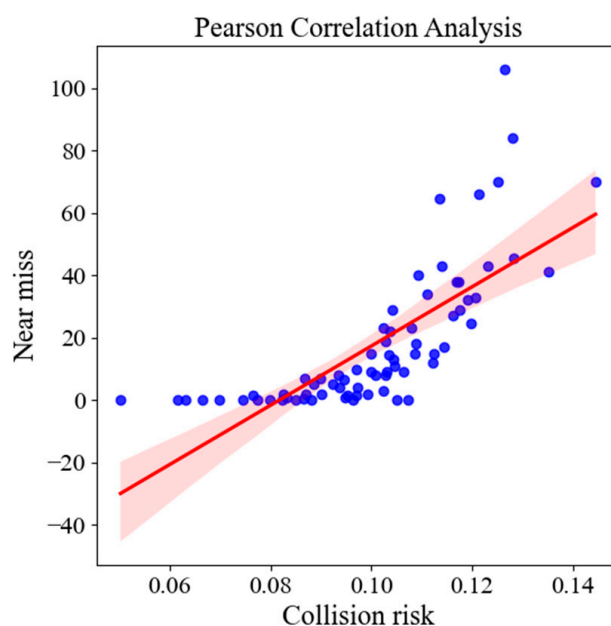


Figure 28. The Pearson correlation analysis between collision risks and near misses of all three sub-areas on 27 June 2022.

4. Discussion

This paper proposed a time-frame regional collision risk model based on dynamic time warping, which can quantify the collision risk of a water area for a certain time interval. Firstly, to generate the time series of collision risk, the collision risk of a ship at a timestamp was modeled from the perspectives of geometric encounters and traffic situation pressure. Then, the time series of collision risk was formed by combining these collision risk stamps in chronological order. After obtaining all of the time series of collision risk of a water area for a given period, dynamic time warping was used to extract their collision risk

eigenvalues, and the time-frame collision risk of the water area was reasonably represented by synthesizing these eigenvalues organically.

The collision risk time series of a ship is constituted by the diverse collision risks it encounters at each time point. Due to the varying durations of ships' presence in the water area, the collision risk time series of different ships possess different lengths. Hence, a ship's contribution to the time-frame regional collision risk is contingent upon both the length of its existence in the water and the magnitude of the collision risks it confronts at each time point when it is present in the water. This gives rise to differences in the sequence lengths and the numerical values of the sequence elements among different ships' collision risk time series, and such differences are taken into consideration in the establishment of the time-frame collision risk model. Ships with larger numerical values of the sequence elements will consume less cost to reach the end during the DTW calculation and have higher characteristic values of collision risk, which can be regarded as a consideration from the spatial perspective. Ships that exist for longer durations, that is, with longer sequence lengths, will be assigned higher weights in the combination of characteristic values of collision risk and can be regarded as a consideration from the temporal perspective. Therefore, despite the existence of collision risk time series with different lengths and magnitudes, they can be compared in order to describe the impact on the time-frame regional collision risk from the spatial and temporal perspectives.

In Section 3, some experimental case studies were carried out to validate that the proposed model can effectively identify the regional collision risk for a certain time period. In some of the experiments, the index of traffic volume was introduced to evaluate the collision risk results as it can reflect the collision risk to some extent. However, compared with the proposed model, traffic volume is limited in its ability to characterize collision risk and cannot distinguish the change in risk in some cases. For example, the collision risk for the 8th hour in Area 1 presented in Figure 9 is 0.1255 with a traffic volume of 19 ships per hour, and the collision risk for the 10th hour in Area 1 presented in Figure 9 is 0.1215 with the same traffic volume. In addition, for the 11th hour in Area 2, the traffic volume is also 19 ships per hour but the collision risk is 0.1172. It can be found that although the traffic volumes are same for the three cases mentioned above, their collision risks are different, because they present different risks in geometric encounters and situation pressure as expressed in Table 4.

Table 4. The collision risk derived from geometric encounters and situation pressure for different cases.

Case	Geometric Encounters	Situation Pressure
Case 1	0.1212	0.3828
Case 2	0.1312	0.3866
Case 3	0.1098	0.3409

The results explain that the identification of collision risk in waters cannot solely rely on the ship traffic volume, as it is limited in adequately accounting for the geometric encounters and environmental pressures faced by ships.

In Section 3, the results calculated by the model were validated through near miss, as near miss possesses the capability to identify potential collision risks. Nevertheless, through an in-depth investigation, it can be discerned that for the identical cases of zero near miss, the values of the collision risks are dissimilar, such as the 2nd, 7th, and 18th time intervals in Area 2. A zero near miss does not imply the absence of any collision risk, and the specific magnitude of the existing collision risk also varies depending on the specific traffic scenario. In some circumstances, although there is no current near miss

situation, under the existing traffic patterns, it is more prone to form a near miss in the future, thereby rendering the situation more dangerous. Simultaneously, through the investigation, it was discovered that in the scenarios presenting the same seven near misses, such as the 10th and 11th time intervals in Area 2, the results of the collision risks are also distinct. Through further exploration, it was found that there were obvious differences in the traffic patterns of the two scenarios. The more dangerous situation had a greater tendency towards ship conflicts, and this aspect could be captured by the model proposed in this paper, demonstrating its superiority in identifying potential collision risks.

To further validate the advantage of the proposed model, some comparison experiments were carried out between it and some traditional regional collision risk models, including the International Association of Lighthouse Authorities Waterway Risk Assessment Program (IWRAP) model, the DCPA/TCPA-based method, the ship domain-based method, the neural network method, and the Bayesian network method. The first compared model is a collision probability model which is widely used in assessing the collision risk of a specified water area for a time period, which is called the IWRAP model [41]. The IWRAP model can evaluate the collision risk of a specified water area using some crucial parameters of traffic flow and the ships sailing in the flow [42]. The collision frequency in the water area can be determined using the following formula when two traffic flows intersect.

$$N_{collision}^{crossing} = \sum_{i,j} \frac{Q_i^1 Q_j^2}{V_i^1 V_j^2} D_{ij} V_{ij} \frac{1}{\sin \theta} \quad (21)$$

where Q_i^1 and Q_j^2 refer to the traffic volume of the ship of type i and type j in traffic flows 1 and 2, respectively; V_i^1 and V_j^2 refer to the average speed of the ship of type i and type j in traffic flows 1 and 2, respectively; V_{ij} refers to the relative speed between the two flows and θ refers to the crossing angle of the two traffic flows; and D_{ij} refers to the diameter of the collision area, which can be calculated based on the ship dimension parameters and the parameters above.

To compare the proposed model with the IWRAP model, the two traffic flows in Area 1 were selected for experimentation, where the directions of the two traffic flows are approximately 120° and 194° . The two traffic flows intersect so that the IWRAP model can be used for collision frequency estimation. The time period of 0200–0300, which is featured with a relatively high traffic volume, was selected as the studied period. The IWRAP model and the proposed model were both used for the collision risk calculation, yielding results of 1.0279 and 0.1268, respectively. For this particular scenario, random perturbations of ship positions within a range of 0.2 nautical miles were performed. Although position is a consideration factor of the model, it is considered within the domain of the ships. It simultaneously takes into account factors such as the collision avoidance responsibilities between ships and the dimensions of the ships. The impact of the position of ships on the collision risk between them is not a simple linear relationship. Changes in the position of ships can cause many relative motion parameters between ships to change, thereby leading to variations in the collision risk. Therefore, the ability of the proposed model to accurately capture potential collision risks was examined through perturbations in the positions of ships to prove its advantage compared with the IWRAP model. Four iterations of such random perturbations were conducted and four additional scenarios were obtained. The collision risks associated with these four scenarios were evaluated using the IWRAP model and the proposed model. The results are shown in Table 5.

Table 5. The collision risks calculated by the traditional model and the proposed mode in different scenarios with position perturbations.

Scenario	1	2	3	4	5
Traditional	1.0279	1.0279	1.0279	1.0279	1.0279
Proposed	0.1268	0.1202	0.1207	0.1261	0.122

It can be observed that although the positions of ships were changed, the collision risks calculated by the IWRAP model remain the same. This is because the spatial relationship or geometric encounter relationship between ships were not considered in the traditional probabilistic model. However, the proposed model demonstrates its superiority over the traditional probabilistic model in assessing collision risk in water areas by effectively adapting to changes in ship position. Therefore, it can provide a more precise data basis for the analysis of marine collision risks and better supporting the enhancement of marine safety. For Scenario 1 and Scenario 3, the results of collision risk calculated through the IWRAP model are identical. From the perspective of the model's application, it ascribes the same collision risk to both scenarios. However, when an in-depth investigation of the ship traffic conditions in Scenario 1 and Scenario 3 was conducted, it was discovered that the ships in Scenario 1 were distributed more densely compared to those in Scenario 3, the tendency of ship conflicts was greater, and the number of near misses obtained through calculation was also marginally higher. This reflects a higher potential collision risk in this scenario from multiple aspects. This aspect can be captured by the model proposed in this paper, but the IWRAP model is incapable of achieving this, which proves the superiority of the proposed model compared with the traditional method.

To further substantiate the superiority of the model proposed herein in contrast to other commonly employed methods for quantifying collision risks, relevant experiments were carried out to compare the proposed approach with the DCPA/TCPA-based method [28,43] and the ship domain-based method [9].

The comparative experiments were initially conducted in the three sub-areas and the entire studied water area of the Bohai Strait as analyzed above. For the aforementioned four areas, the 24 h collision risks were calculated, respectively, by employing the DCPA/TCPA-based method, the ship domain-based method, and the proposed model. The average values for the collision risks obtained are presented in Table 6.

Table 6. The average collision risk calculated by the DCPA/TCPA method, the ship domain method, and the proposed model on 27 June 2022.

Method	Area 1	Area 2	Area 3	Entire Area
DCPA/TCPA	0.1095	0.0959	0.0584	0.0225
Ship domain	3.5556	1.1389	0.0590	24.0069
Proposed	0.1213	0.1086	0.0767	0.1105

It can be discerned from the results that all three approaches manifest the decrease in collision risk from Area 1 to Area 3. This is attributed to the fact that the ship traffic volume from Area 1 to Area 3 tends to decline, and the complexity of ship traffic simplifies, as previously analyzed. Nevertheless, with regard to the entire water area, it can be observed that the collision risk value obtained through the DCPA/TCPA-based method is extremely small, even smaller than that of Area 3. This is because when the DCPA/TCPA-based method computes the collision risk of the entire water area, it takes into account a considerable number of ship pairs with very large DCPA or TCPA, such as those pairs that are quite distant in the water area. These ship pairs have an extremely low collision

risk, thereby significantly diluting the final collision risk outcome. Simultaneously, it can also be noted from the results that the collision risk characterized by the number of ship domain conflicts in the entire water area is much greater than that in the three sub-areas. This is because this method merely characterizes the collision risk through the absolute quantity of domain conflicts, without considering the extent of the water area. In the entire water area, it accumulates the number of domain conflicts at all positions and regions, thus presenting a very large value. By contrast, the model proposed in this paper quantifies the surrounding collision risk from the perspective of the ships themselves to represent the collision risk of a region. It scientifically models the collision risk while taking into account the influence of the traffic range and is essentially not affected by extreme values. Hence, the results calculated by the proposed model can effectively represent the impact of traffic conditions on the collision risk and can still reasonably represent the collision risk on a unified scale when the range of water area changes.

Then, the comparative experiments were carried out in some specific scenarios. In the experiment comparing the DCPA/TCPA-based method and the proposed method, a three-ship crossing encounter scenario was simulated, as depicted in Figure 29a. In this scenario, the three ships have identical dimensions and speeds and are approaching one another. Taking the scene shown in Figure 29a as Timing 1, the collision risk was computed every three minutes using both the proposed model and the DCPA/TCPA method for a total of five timings, thereby obtaining the collision risk time series for the three-ship encounter scenario at five moments. On this basis, another scenario was simulated by eliminating the uppermost ship from the three-ship encounter scenario, resulting in a two-ship encounter scenario, as illustrated in Figure 29b. The collision risk time series for five moments were calculated for this scenario using both methods. The results of the collision risk calculated using the proposed model and the DCPA/TCPA method for these two scenarios are presented in Tables 7 and 8, respectively.

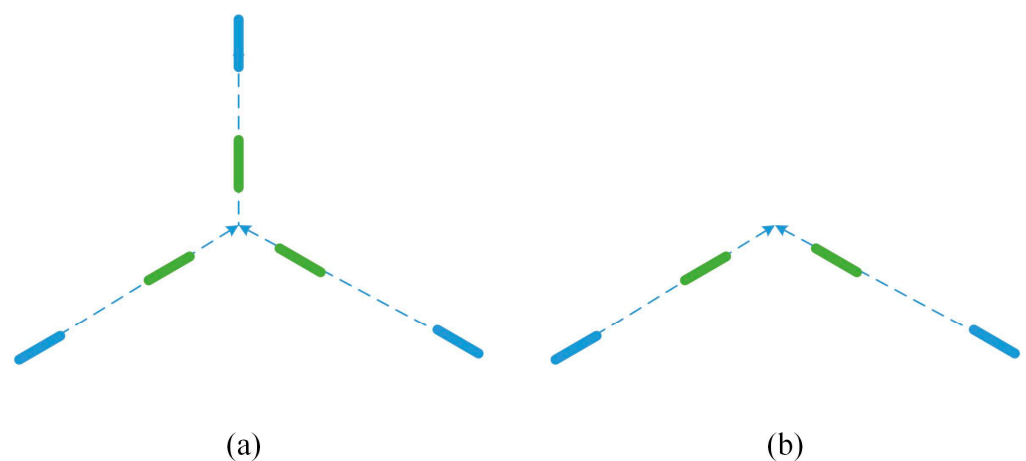


Figure 29. The two encountering scenarios for the comparative experiments (a) a three-ship encountering scenario (b) a two-ship encountering scenario.

Table 7. The collision risks calculated by the proposed model for the two three-ship encountering scenarios.

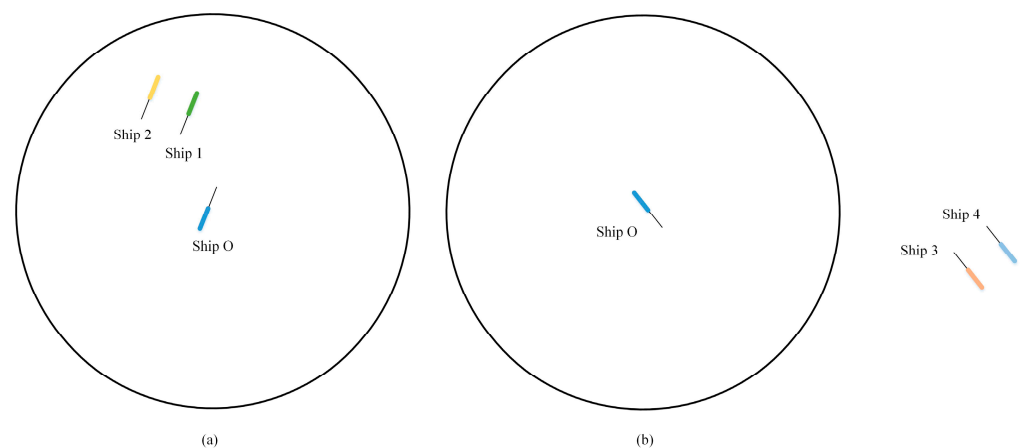
Timing	1	2	3	4	5
Scenario 1	0.4279	0.4661	0.5146	0.5763	0.6548
Scenario 2	0.4112	0.4494	0.4980	0.5597	0.6387

Table 8. The collision risks calculated by the DCPA/TCPA method for the two three-ship encountering scenarios.

Timing	1	2	3	4	5
Scenario 1	0.6435	0.6867	0.7429	0.8162	0.9115
Scenario 2	0.6435	0.6867	0.7429	0.8162	0.9115

It can be observed that although the collision risk computed through the DCPA/TCPA method can manifest the gradual escalation of collision risk during the continuous approach of encountering ships, the outcomes for three-ship encounters and two-ship encounters are identical, meaning it is incapable of identifying the alterations in collision risk brought about by a reduction or increase in the quantity of ships. Conversely, the model proposed in this paper, apart from being able to represent the augmentation of collision risk resulting from the continuous approach of encountering ships, can also discern the variations in collision risk induced by an increase or decrease in the number of ships. This is because the model proposed herein takes into additional consideration the ship traffic pressure in the encounter scenarios, thereby enabling a more objective and precise characterization of the collision risk circumstances.

In the experiments for comparison with the ship domain-based model, two scenarios of two ships in a head-on encounter were simulated, as depicted in Figure 30a. In both scenarios, the ships have the same dimensions and speeds. Ship O encountered Ship 1 and Ship 2, respectively. Nevertheless, the transverse distance between Ship 1 and Ship O was shorter than that between Ship 2 and Ship O. The collision risks of these two scenarios were calculated, respectively, using the model proposed in this paper and the method based on the ship domain. The results are presented in Table 9.

**Figure 30.** The two two-ship head-on encounter scenario set for the comparative experiments (a) Scenario Set 1 (b) Scenario Set 2.**Table 9.** The collision risks calculated by the proposed model and the ship domain method for the two encountering scenarios.

Scenario	1	2
The proposed model	0.5949	0.5332
Ship domain method	1	1

It can be observed that the traditional approach based on the ship domain, as it merely measures the collision risk from the aspect of the quantity of ship domain conflicts, can only conduct a relative characterization of the collision risk and has difficulty in identifying the

variations in the collision risk in two scenarios. In contrast, the proposed model, in addition to considering the conflicts within the ship domain, also takes into account the relative motion relationship between ships, thereby being capable of representing the reduction in ship collision risk more effectively.

In accordance with the aforementioned notions, another set of two-ship encounter scenarios was simulated, as shown in in Figure 30b. Distinct from the previous group of encounter scenarios, Ships 3 and 4 in this group did not encroach upon the domain range of Ship O. The transverse distance between Ship 3 and Ship O was nearer than that between Ship 4 and Ship O. Similarly, the collision risks of the two scenarios were computed through the proposed model and the ship domain-based method, and the results are presented in Table 10.

Table 10. The collision risks calculated by the proposed model and the ship domain method for the two encountering scenarios.

Scenario Set	1	2
The proposed model	0.4182	0.4032
Ship domain method	0	0

It can be discerned that in this set of scenarios, neither Ship 3 nor Ship 4 has encroached upon the domain of Ship O. Consequently, the number of ship domain conflicts obtained via the traditional ship domain-based method is all zero, indicating that the ship collision risk is non-existent and is zero, and the alteration of the collision risk resulting from the change in the relative motion relationship between the two ships cannot be identified. In contrast, the proposed model, in addition to taking into account the conflicts in the ship domain, also considers the relative motion relationship between ships and is likewise capable of representing the reduction in the ship collision risk more effectively. This also holds true for the other two major encounter types.

In the field of collision risk research, there exist several other approaches, such as neural network methods and Bayesian network methods.

Neural network methods are mostly used for the rapid calculation of collision risks. Based on ship traffic data, existing collision risk models are used to calculate the collision risk results, and the above data are used for neural network training, so that the trained neural network model can be used for the calculation of collision risks in the water area. The advantage of this method is that after training based on a large amount of data, the model can quickly identify collision risks in the water area. However, this method relies on existing collision risk models and has difficulty mining the time characteristics of collision risks by itself. The advantage of the proposed model is that it can independently calculate collision risks and identify the time characteristics of collision risks in the water area through the formed collision risk time series, thereby quantifying the collision risks in the water area during a certain period from an overall perspective. However, on the basis of the proposed model, neural networks can be further combined to achieve the rapid calculation of collision risks in the water area. The Bayesian network approach is mainly employed in the computation of the causation collision probability. Specifically, in the research of regional collision risk based on collision probability, the conditional probabilities that lead to ship collisions are analyzed and quantified from the aspects of marine traffic elements such as human, ship, and environment. It serves as a component of the collision probability and enables a more precise quantification of the probability of ship collisions on the foundation of geometric collision probability. Consequently, the research on collision risks based on the Bayesian network method primarily centers on the conditional probabilities of whether a collision accident can occur based on the likelihood of the accident in various elements of

marine traffic. Therefore, it is challenging for the Bayesian network method to provide a sufficient quantitative description of aspects such as the geometric encounter situation of ships and the ship traffic environment like the proposed model, and it is also difficult to conduct in-depth exploration of the temporal characteristics of collision risks. Nevertheless, as a research method of conditional probabilities, it can further enhance the accuracy of the results on the basis of the proposed model.

To further demonstrate the superiority of the proposed model, it is contrasted with the method in [33], which is capable of calculating the instantaneous collision risk and forming a time-continuous regional collision risk outcome, representing the change in collision risk in the studied water area over a certain period of time.

For comparison, a set of experiments was carried out, which encompassed two scenarios. The results of the continuous instantaneous regional collision risks for the two scenarios, computed by the method in [33], are presented in Table 11. It can be discerned that this approach can obtain the time series of collision risks, representing the temporal variations and extreme values of regional collision risks over time, yet it has difficulty undertaking further analyses of the overall pattern of regional collision risks during this period or can merely conduct simple processing and analyses based on the outcomes. As demonstrated in Table 11, by employing the average value of the collision risk time series to represent the overall collision risk level, it is discovered that the average collision risk of Scenario 1 is 0.2680, which exceeds that of Scenario 2, 0.2320. This implies that the overall collision risk level of Scenario 1 is greater than that of Scenario 2. Nevertheless, a meticulous investigation of the collision risk results of the two scenarios discloses that the collision risk of Scenario 2 remains at a relatively low level (approximately 0.1) for the majority of the time points, with a sharp escalation to 0.72 from the second to the third time point. In actual navigation, a sudden increase in collision risk indicates an extremely perilous situation, potentially attributed to the sudden formation of a complex multi-ship encounter situation. Although the overall average value of the collision risk level of Scenario 2 is lower than that of Scenario 1, from a practical perspective, it portends a more dangerous circumstance and deserves greater attention from marine traffic surveillance. If the DTW technology proposed in this paper is integrated, the overall collision risk levels of the two scenarios can be calculated as 0.1731 and 0.2294, respectively, signifying that the collision risk level of Scenario 1 is lower than that of Scenario 2, which is more congruent with the objective reality. This is because the proposed method can more comprehensively analyze the temporal characteristics and fluctuation patterns of collision risks during the study period via DTW technology, thereby enabling a more objective quantification of regional collision risks from the overall temporal dimension.

Table 11. The collision risks calculated by the method in [33] for the three scenarios.

Timing	1	2	3	4	5
Scenario 1	0.12	0.29	0.53	0.31	0.09
Scenario 2	0.12	0.11	0.72	0.09	0.12
Scenario 3	0.27	0.28	0.27	0.28	0.27

The circumstances featuring substantial fluctuations in collision risks deserve more attention from marine traffic surveillance than those in stable conditions, as they signify a sudden and sharp escalation or emergence of danger, hinting at greater potential collision risks. In this experiment, the magnitude of the collision risk time series in Scenario 2 was evidently greater than that in Scenario 1. Such a situation is hard to identify by traditional average processing approaches but can be recognized by the DTW-based model proposed in this paper. On this basis, Scenario 3 was added, where the collision risk exhibited almost no

fluctuation, ranging between 0.27 and 0.28. Through calculation, it was discovered that the average collision risk in Scenario 3 was 0.2740, slightly higher than that in Scenarios 1 and 2. Nevertheless, the collision risk calculated by the model proposed in this paper was 0.1136, which is obviously smaller than those in the first two scenarios. Such experimental results further illustrate that the DTW-based model proposed in this paper can better capture the fluctuations of collision risks during time periods and represent greater potential risks in cases with larger fluctuations.

Meanwhile, another set of experiments was conducted. In this set of experiments, there were also three scenarios, and the collision risk time series of the three scenarios are presented in Table 12. It can be observed that the collision risk time series of Scenario 1 did not exhibit a distinct changing trend and remained at approximately 0.2. However, the collision risk time series of Scenario 2 and Scenario 3 clearly demonstrated an increasing trend, among which the increasing rate of Scenario 3 was faster. Through the traditional averaging processing approach, it can be found that the average values of collision risk from Scenario 1 to Scenario 3 were all approximately 0.20. The collision risk results calculated by the DTW-based model proposed in this paper were 0.0957, 0.0968, and 0.1274, respectively. Such results are in accordance with the trend of danger manifested by the collision risk time series of the three scenarios. Compared with the stable trend of Scenario 1, from the overall perspective of the time period, the increasing trends of Scenario 2 and Scenario 3 imply greater potential risks. Furthermore, the faster increasing rate of Scenario 3 indicates a more rapid formation and development of collision risk, which is even more perilous.

Table 12. The collision risks calculated by the method in [33] for another three scenarios.

Timing	1	2	3	4	5
Scenario 1	0.2	0.2	0.18	0.23	0.21
Scenario 2	0.16	0.18	0.2	0.22	0.24
Scenario 3	0	0.1	0.2	0.3	0.4

To sum up, the DTW-based model proposed in this paper is capable of identifying the overall pattern of collision risk in water areas during a specific period. This pertains not only to the calculation of the overall collision risk during the period but also to the capture of the fluctuation amplitude and the increasing or decreasing trend of the collision risk time series, which is typically unachievable through traditional methods of instantaneous collision risk identification. Nevertheless, the proposed model herein can effectively identify the overall fluctuation pattern of collision risk and represent it through risk values.

To provide a more systematic illustration of the ability of the proposed model in identifying time-frame regional collision risk, a reliability analysis of the proposed model was carried out, where the performance of effectiveness, robustness, generalization, and sensitivity was evaluated.

In Section 3, the validity of the proposed model's calculation results was verified through correlation analysis with near misses in the water area. Since a near miss is a very dangerous state close to collision, the number of near miss occurrences within a certain period of time will reveal the degree of collision risk in the water area during that period. For the research period of Area 1, the number of near misses for each statistical period was calculated, and the results are shown in Figure 14. Through conducting a correlation analysis between the near miss results in Figure 14 and the calculated collision risk results in Figure 11, a strong positive correlation was discovered, thereby confirming the validity and accuracy of the proposed model in calculating the collision risk of the water area during a specific period. Meanwhile, in order to explain the ability of the proposed model to identify the time-frame regional collision risk when the conditions of ships in the water

area change significantly, another series of experiments were conducted. The size of the ship was enlarged in these experiments, being equivalent to one and a half times the original size. Longer ships typically require greater operational space to execute maneuvers such as turning or collision avoidance, which increases the likelihood of collisions with other ships and consequently may elevate the perceived pressure. The length of a ship significantly influences the dimensions of its ship domain, thereby exerting a notable impact on the calculation results of V-SDOI. The proposed model is utilized to calculate the regional collision risk by selecting five consecutive periods of 1 h duration each from Area 1, which is from 1300 to 1900. The results are shown in Table 8. After enlarging the size of the ship, the collision risks for the five periods were recalculated and the comparison between the collision risks and traffic volumes before and after the size enlarging process is expressed by Figure 31.

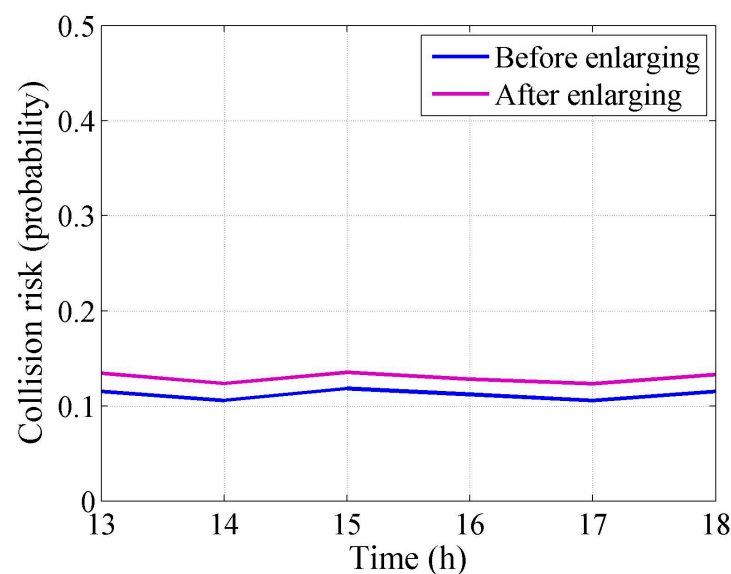


Figure 31. The comparison between the collision risks before and after the size enlargement.

It can be found that the collision risks elevate obviously after enlarging the ship size and the traffic volumes remain unchanged. This is because the enlargement of the ship's size will result in an expansion of its safety domain, thereby enhancing its compactness within a confined space. However, this augmentation also leads to a higher likelihood of urgent situations arising from the ship's spatial positioning relationship. Consequently, it imposes additional constraints on the deck officers' maneuvering space for collision avoidance and elevates the complexity associated with such avoidance maneuvers. Moreover, it intensifies the pressure on traffic situation awareness for deck officers and subsequently escalates collision risks. These facts corroborate the results derived from the model proposed in this paper, thereby further proving the effectiveness of the proposed model in identifying collision risk.

In terms of robustness, some random anomalies were generated in the data used for experiments, such as sudden changes in heading, excessive speed or negative values, and dramatic changes in ship position, etc. It can be found that despite the presence of these anomalies, the results calculated by the model were very similar to those obtained under normal data conditions. This is because the proposed model has a data cleaning step before use, which removes most of the abnormal data, leaving behind almost normal data, allowing the model to effectively calculate the collision risk results. At the same time, due to the frequent broadcasting of AIS information, the lack of a small number of abnormal data will not affect the final calculation results. We also removed some data to

simulate data loss, to check the effectiveness of the model in data loss situations. However, after removing these data, the results calculated by the model were still almost the same as before, because in the proposed model, AIS data are interpolated before use to obtain the data of different ships at the same time point, which allows us to obtain the missing data accurately even if some data are missing and accurately complete the calculation of collision risk. These facts show that the proposed model has a good robustness and can function under data abnormality and loss.

In terms of generalization performance, in addition to conducting experiments on Area 1 in Section 3, the experiments on the other two areas with obvious differences in traffic patterns from Area 1 were also conducted. Area 2 is located to the southeast of Area 1 and is also located in the main traffic flow, but without the crossing traffic flow, making the traffic pattern in the water area gentler and the traffic density lower, so the encounter situations are relatively less complex than in Area 1. Area 3 is located in the secondary traffic flow in the water area, with an even lower traffic density and simpler encounter situations. Therefore, the collision risk in the three research areas should gradually decrease. The collision risk results calculated by the proposed model are consistent with the above facts, which indicates the generalization performance of the proposed model in different traffic patterns in water areas; even if the traffic density, encounter patterns, and complexity of traffic in the water area change, the model can still effectively identify the time-frame regional collision risk.

Additionally, a sensitivity analysis on the weight parameter in the proposed time-frame regional collision risk model was conducted. In the proposed model, the weight is determined by the time each ship exists in the water area. The longer a ship exists in the water area, the greater its contribution to the time-frame regional collision risk, and therefore the higher its weight in the calculation of the collision risk. To test the sensitivity of the weight, the weight which was determined by the ship's existence time was adjusted to an equal weight situation, that is, each ship has the same contribution to the calculation of the time-frame regional collision risk. After adjusting the weight, the average values and average change rates of the time-frame regional collision risk in the three areas are shown in Table 13.

Table 13. The collision risks after adjusting weights and the change rates for the three studied areas.

Area	Area 1	Area 2	Area 3
Risk after adjustment	0.1070	0.0800	0.0742
Change rate	−11.7%	−25.9%	−1.5%

It can be found that upon adjusting the weights to be equal, the collision risk outcomes in all three areas decreased. Specifically, Area 1 witnessed an average reduction of 11.7%, Area 2 an average reduction of 25.9%, and Area 3 an average reduction of 1.5%. Through in-depth investigation, it was discovered that before the weight adjustment, there were ships with an extremely short presence in the water area, thereby making minimal contributions to the collision risk in the water area, and they also faced negligible collision risks. However, with the assignment of the same weight to all ships, the contribution of these ships to the collision risk in the water area was enhanced, while the contribution of those ships that had a greater contribution to the collision risk in the water area and faced higher collision risks was decreased. This led to a reduction in the overall collision risk result. The aforementioned facts not only demonstrate the sensitivity of the proposed model to weight variations but also explain the rationality of weight settings in this paper. The weights in the experiment regarding the relationship between collision risks and near

misses in Area 1 depicted in Figures 14 and 15 were also adjusted. The weights of all ships were set to be equal and the collision risk results were recalculated. The correlation between the collision risk results after the adjustment to equal weights and the number of near misses were analyzed. The results indicated that after adjusting the weights, the correlation coefficient between collision risks and near misses decreased by 0.045, signifying a weakened correlation. This outcome suggests that when the weights are adjusted to be equal, that is, without considering the influence of weights, the validity of the collision risk results in representing the potential collision risk reduces. This once again proves that by scientifically setting the weights for each ship, this method can more rationally reflect the collision risk situation in the water area. Furthermore, with respect to the experiments depicted in Figure 22, a sensitivity analysis of the weights was also carried out. The weights of all ships were adjusted to be equal, and the regional collision risk results were recalculated. It was observed that after adjusting the weights of each ship to be equal, the regional collision risk decreased by 8% on average, and the correlation coefficient between the collision risk and near misses dropped by 0.012. The aforementioned experiments not only illustrate the sensitivity of the proposed model to weight variations but also reaffirm the rationality of the weight settings in this paper.

In sum, the proposed time-frame regional collision risk model based on DTW technology has the following advantages:

Firstly, the proposed model considers the geometric encounter relationship among ships and the pressure of the traffic environment surrounding them. It not only pays attention to the dynamic position and motion state of ships in the spatio-temporal dimension but also incorporates the impact of the surrounding traffic environment on the collision risk of ships. Such a multi-dimensional analysis can evaluate the collision risk faced by ships more comprehensively.

Secondly, the proposed model conducts an in-depth analysis of the dynamic evolution process of collision risks from the perspective of time series. It is capable of exploring the impact of the temporal attributes and characteristics of collision risks on the basis of considering the continuity and changing trend of the collision risks of individual ships over time.

Moreover, the proposed model commences from the perspective of the navigation of each ship, meticulously considering the contribution of each ship to the overall collision risk, and on this basis, organically combines them to form the regional collision risk. This modeling approach takes into account the influence of each ship on the regional collision risk, being more scientific and rational.

However, the proposed model still possesses certain limitations. For the collision risk derived from the traffic situation pressure, although the two crucial factors of traffic intensity and the traffic compactness surrounding the ship were incorporated, factors associated with the natural environment were not sufficiently considered. When the natural condition changes, the proposed model is limited in making a response. There are also some other factors influencing the perceived pressure on board. Due to differences in dimensions and ship type, the interaction patterns among ships exhibit distinct characteristics. Additionally, there are significant differences in the navigation performance of ships between their fully loaded and empty states, in addition to differences in maneuverability such as turning capability and acceleration/deceleration capabilities. In future research, key factors relevant to the natural environment that influence the collision risk, such as wind, wave, visibility, etc., and the ship static and dynamic parameters can be included in the model. Meanwhile, in modeling traffic compactness with V-SDOI, the ship domain with a smooth unequal sector was adopted to enhance the computing efficiency; although COLREGs were considered to some extent, the difference between the scopes of the domain in ship's beam

and fore–aft directions is still insufficient. The selection of ship domain models in future research can be tailored to the specific characteristics of different water areas. In addition, the proposed time-frame regional collision risk model was established by synthesizing the time series of the collision risks of different ships organically by the DTW algorithm; the relationship between different sequences or different trajectories deserves to be considered more deeply in order to further improve the accuracy of the model. The construction of the most dangerous time series in this paper also has limitations. It was constructed by calculating the maximum collision risk at each time point of the longest route of the ship in the studied waters. Although the influence of the traffic pattern was considered, we did not fully integrate the consideration of the traffic pattern into the route selection. In future research, the most dangerous route can be selected based on a detailed analysis of the traffic pattern in the studied waters, thereby making the most dangerous sequence obtained closer to the actual situation.

5. Conclusions

In this paper, a novel time-frame regional collision risk model was proposed based on dynamic time warping. The collision risk faced by the ship was established at first to generate the collision risk time series of the ships in the water area. The collision risk was modeled from the perspectives of a geometric encounter, which considers the spatial and temporal danger between ships, and the traffic situation pressure, which considers traffic intensity and compactness. Utilizing the generated time series of collision risk, dynamic time warping was used to extract the eigenvalue of each sequence and the time-frame collision risk of the water area could be obtained by organically combining these eigenvalues with the consideration of the contribution of each ship to the entire collision risk. For validating the proposed time-frame collision risk model, some experiments were carried out by using the real AIS data in the Bohai Strait. The results show that the proposed model possesses the ability to effectively identify the collision risk of a water area for a certain time period. Meanwhile, the proposed model was proven to have some merits compared with the traditional means by some comparison experiments. Therefore, the proposed model will provide assistance for the supervision and analysis of collision risk in marine traffic, contributing to future predictions of collision risk and thereby enhancing marine traffic safety.

The proposed model can also be improved in the following aspects. Firstly, the existing collision risk factors are considered as a basis, and the factors influencing collision risk in the natural environment can be taken into account to enhance the accuracy of the results. Secondly, the selection of the ship domain can prioritize the characteristic of the water area, leading to the choice of ship model with a more precise boundary. Thirdly, the time series-based collision risk modeling can incorporate the interaction between sequences, enabling the model to accurately depict more intricate traffic scenarios. In addition, with more abundant data, in-depth research on collision risks in complicated sea environments and the application of the proposed model in characterizing the spatial distribution of collision risk can be further conducted in the future.

Author Contributions: Conceptualization, Z.L.; methodology, Z.L.; software, Z.L.; validation, Z.L.; formal analysis, Z.L.; investigation, Z.L.; resources, Z.L.; data curation, Z.L. and P.Y.; writing—original draft preparation, Z.L. and P.Y.; writing—review and editing, Z.L. and P.Y.; visualization, Z.L. and P.Y.; supervision, Z.L.; project administration, Z.L.; funding acquisition, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Dalian Science and Technology Talent Innovation Support Policy Implementation Plan (Young Science and Technology Star), grant number

2023RQ002, the Fundamental Research Project of Liaoning Provincial Department of Education, grant number JYTQN2023041, and the Fundamental Research Funds for the Central Universities, grant number 3132024133.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are not publicly available due to privacy restrictions.

Acknowledgments: We would like to express our gratitude to the editors and anonymous reviewers whose valuable comments and suggestions have contributed to the improvements of the quality of this paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Maidana, R.G.; Kristensen, S.D.; Utne, I.B.; Sørensen, A.J. Risk-based path planning for preventing collisions and groundings of maritime autonomous surface ships. *Ocean Eng.* **2023**, *290*, 116417. [\[CrossRef\]](#)
2. Cao, R.; Rong, Z.; Chen, H.; Liu, Y.; Mu, L.; Lv, X. Modeling the long-term transport and fate of oil spilled from the 2021 A Symphony tanker collision in the Yellow Sea, China: Reliability of the stochastic simulation. *Ocean Eng.* **2023**, *186*, 102285. [\[CrossRef\]](#)
3. Xin, X.; Yang, Z.; Liu, K.; Zhang, J.; Wu, X. Multi-stage and multi-topology analysis of ship traffic complexity for probabilistic collision detection. *Expert Syst. Appl.* **2023**, *213*, 118890. [\[CrossRef\]](#)
4. Fujii, Y.; Shiobara, R. The analysis of traffic accidents. *J. Navig.* **1971**, *24*, 534–543. [\[CrossRef\]](#)
5. MacDuff, T. The probability of vessel collisions. *Ocean Ind.* **1974**, *9*, 144–148.
6. Pedersen, P.T. Collision and Grounding Mechanics. In Proceedings of the WEGEMT95, Copenhagen, Denmark; 1995.
7. Kujala, P.; Hänninen, M.; Arola, T.; Ylitalo, J. Analysis of the marine traffic safety in the Gulf of Finland. *Reliab. Eng. Syst. Saf.* **2009**, *94*, 1349–1357. [\[CrossRef\]](#)
8. Silveira, P.; Teixeira, A.P.; Soares, C.G. Use of AIS data to characterise marine traffic patterns and ship collision risk off the coast of Portugal. *J. Navig.* **2013**, *66*, 879–898. [\[CrossRef\]](#)
9. Weng, J.; Xue, S. Ship collision frequency estimation in port fairways: A case study. *J. Navig.* **2015**, *68*, 602–618. [\[CrossRef\]](#)
10. Chai, T.; Weng, J.; Xiong, D. Development of a quantitative risk assessment model for ship collisions in fairways. *Saf. Sci.* **2017**, *91*, 71–83. [\[CrossRef\]](#)
11. Hu, Y.; Park, G.K. Collision risk assessment based on the vulnerability of marine accidents using fuzzy logic. *Int. J. Nav. Archit. Ocean Eng.* **2020**, *12*, 541–551. [\[CrossRef\]](#)
12. Zhang, C.; Zhang, D.; Zhang, M.; Lang, X.; Mao, W. An integrated risk assessment model for safe Arctic navigation. *Transp. Res. Part A Policy Pract.* **2020**, *142*, 101–114. [\[CrossRef\]](#)
13. Zhang, W.; Feng, X.; Goerlandt, F.; Liu, Q. Towards a Convolutional Neural Network model for classifying regional ship collision risk levels for waterway risk analysis. *Reliab. Eng. Syst. Saf.* **2020**, *204*, 107127. [\[CrossRef\]](#)
14. Chen, P.; Li, M.; Mou, J. A velocity obstacle-based real-time regional ship collision risk analysis method. *J. Mar. Sci. Eng.* **2021**, *9*, 428. [\[CrossRef\]](#)
15. Zhang, M.; Montewka, J.; Manderbacka, T.; Kujala, P.; Hirdaris, S. A big data analytics method for the evaluation of ship-ship collision risk reflecting hydrometeorological conditions. *Reliab. Eng. Syst. Saf.* **2021**, *213*, 107674. [\[CrossRef\]](#)
16. Rong, H.; Teixeira, A.P.; Soares, C.G. Spatial correlation analysis of near ship collision hotspots with local maritime traffic characteristics. *Reliab. Eng. Syst. Saf.* **2021**, *209*, 107463. [\[CrossRef\]](#)
17. Weng, J.; Du, J.; Shi, K.; Liao, S. Effects of ship domain shapes on ship collision risk estimates considering collision frequency and severity. *Ocean Eng.* **2023**, *283*, 115070. [\[CrossRef\]](#)
18. Antão, P.; Sun, S.; Teixeira, A.P.; Soares, C.G. Quantitative assessment of ship collision risk influencing factors from worldwide accident and fleet data. *Reliab. Eng. Syst. Saf.* **2023**, *234*, 109166. [\[CrossRef\]](#)
19. Cheng, Z.; Chen, P.; Mou, J.; Chen, L. Novel collision risk measurement method for multi-ship encounters via velocity obstacles and temporal proximity. *Ocean Eng.* **2024**, *302*, 117585. [\[CrossRef\]](#)
20. Fujii, Y.; Tanaka, K. Traffic capacity. *J. Navig.* **1971**, *24*, 543–552. [\[CrossRef\]](#)
21. Goodwin, E.M. A Statistical Study of Ship Domains. *J. Navig.* **1975**, *28*, 328–341. [\[CrossRef\]](#)
22. Davis, P.V.; Dove, M.J.; Stockel, C.T. A Computer Simulation of Marine Traffic Using Domains and Arenas. *J. Navig.* **1980**, *33*, 215–222. [\[CrossRef\]](#)
23. Coldwell, T.G. Marine traffic behaviour in restricted waters. *J. Navig.* **1983**, *36*, 431–444. [\[CrossRef\]](#)

24. Pietrzykowski, Z.; Uriasz, J. The Ship Domain-A Criterion of Navigational Safety Assessment in an Open Sea Area. *J. Navig.* **2009**, *62*, 93–108. [\[CrossRef\]](#)
25. Wang, N. An Intelligent Spatial Collision Risk Based on the Quaternion Ship Domain. *J. Navig.* **2010**, *63*, 733–749. [\[CrossRef\]](#)
26. Silveira, P.; Teixeira, A.P.; Soares, C.G. A method to extract the Quaternion Ship Domain parameters from AIS data. *Ocean Eng.* **2022**, *267*, 111568. [\[CrossRef\]](#)
27. Szlapczynski, R. A unified measure of collision risk derived from the concept of a ship domain. *J. Navig.* **2006**, *59*, 477–490. [\[CrossRef\]](#)
28. Zhen, R.; Riveiro, M.; Jin, Y. A novel analytic framework of real-time multi-vessel collision risk assessment for maritime traffic surveillance. *Ocean Eng.* **2017**, *145*, 492–501. [\[CrossRef\]](#)
29. Zhang, F.; Liu, Y.; Du, L.; Goerlandt, F.; Sui, Z.; Wen, Y. A rule-based maritime traffic situation complex network approach for enhancing situation awareness of vessel traffic service operators. *Ocean Eng.* **2023**, *284*, 115203. [\[CrossRef\]](#)
30. Wen, Y.; Huang, Y.; Zhou, C.; Yang, J.; Xiao, C.; Wu, X. Modelling of marine traffic flow complexity. *Ocean Eng.* **2015**, *104*, 500–510. [\[CrossRef\]](#)
31. Goodwin, E.M.; Lamb, W.G.P.; Kemp, J.F. Quantitative measurements of navigational safety. *J. Navig.* **1983**, *36*, 418–429. [\[CrossRef\]](#)
32. International Maritime Organization. Convention on the International Regulations for Preventing Collisions at Sea, 1972 (COLREGs). Available online: <https://www.imo.org/en/OurWork/Safety/Pages/Preventing-Collisions.aspx/> (accessed on 14 November 2024).
33. Liu, Z.; Wu, Z.; Zheng, Z. A novel framework for regional collision risk identification based on AIS data. *Appl. Ocean Res.* **2019**, *89*, 261–272. [\[CrossRef\]](#)
34. Szlapczynski, R.; Szlapczynska, J. A ship domain-based model of collision risk for near-miss detection and Collision Alert Systems. *Reliab. Eng. Syst. Saf.* **2021**, *214*, 107766. [\[CrossRef\]](#)
35. Du, L.; Valdez Banda, O.A.; Huang, Y.; Goerlandt, F.; Kujala, P.; Zhang, W. An empirical ship domain based on evasive maneuver and perceived collision risk. *Reliab. Eng. Syst. Saf.* **2021**, *213*, 107752. [\[CrossRef\]](#)
36. Kang, R.; Zhou, S.; Chen, T.; Yin, H.; Si, L.; Deng, C.; Kaufmann, H. Extent and Sources of Heavy Metal Pollution from Discharging Rivers in the Bohai Region, China. *Water* **2024**, *16*, 982. [\[CrossRef\]](#)
37. OpenHistoricalMap. Available online: <https://www.openhistoricalmap.org/> (accessed on 14 November 2024).
38. Wang, X.; Du, D.; Peng, Y. Assessing the Importance of the Marine Chokepoint: Evidence from Tracking the Global Marine Traffic. *Sustainability* **2023**, *16*, 384. [\[CrossRef\]](#)
39. Liu, T.; Ma, J.; Zhou, Y. Research on Spatiotemporal Distribution Characteristics of Ship Traffic Flow in Port Waters Based on AIS Data. *Front. Humanit. Soc. Sci.* **2024**, *4*, 1–11. [\[CrossRef\]](#)
40. Yoo, S.L. Near-miss density map for safe navigation of ships. *Ocean Eng.* **2018**, *163*, 15–21. [\[CrossRef\]](#)
41. The International Association of Marine Aids to Navigation and Lighthouse Authorities. International Association of Lighthouse Authorities Waterway Risk Assessment Program. Available online: https://www.iala.int/wiki/iwrap/index.php/Main_Page/ (accessed on 14 November 2024).
42. Cucinotta, F.; Guglielmino, E.; Sfravara, F. Frequency of ship collisions in the strait of messina through regulatory and environmental constraints assessment. *J. Navig.* **2017**, *70*, 1002–1022. [\[CrossRef\]](#)
43. Zhou, M.J.; Chen, J.; Ge, Q.; Huang, X.; Liu, Y. AIS data based identification of systematic collision risk for maritime intelligent transport system. In Proceedings of the 2013 IEEE International Conference on Communications, Budapest, Hungary, 9–13 June 2013.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.