



RESEARCH ARTICLE

# Automatic identification system data-driven model for analysis of ship domain near bridge-waters

Lei Jinyu,<sup>1,2</sup> Liu Lei,<sup>3\*</sup> Chu Xiumin,<sup>1,2</sup> He Wei,<sup>1,2</sup> Liu Xinglong,<sup>1,2</sup> and Liu Cong<sup>4</sup>

<sup>1</sup>National Engineering Research Center for Water Transport Safety, Wuhan University of Technology, Wuhan, China.

<sup>2</sup>Fujian Engineering Research Center of Safety Control for Ship Intelligent Navigation, Minjiang University, Fuzhou, China.

<sup>3</sup>School of Transportation, Southeast University, Nanjing, China.

<sup>4</sup>Department of Mechanical Engineering, Aalto University, Espoo, Finland.

\*Corresponding author. E-mail: lei1992@seu.edu.cn

**Received:** 24 November 2020; **Accepted:** 9 May 2021; **First published online:** 18 June 2021

**Keywords:** ship domain, automatic identification system, bridge-waters, ship collision

## Abstract

The ship safety domain plays a significant role in collision risk assessment. However, few studies take the practical considerations of implementing this method in the vicinity of bridge-waters into account. Therefore, historical automatic identification system data is utilised to construct and analyse ship domains considering ship–ship and ship–bridge collisions. A method for determining the closest boundary is proposed, and the boundary of the ship domain is fitted by the least squares method. The ship domains near bridge-waters are constructed as ellipse models, the characteristics of which are discussed. Novel fuzzy quaternion ship domain models are established respectively for inland ships and bridge piers, which would assist in the construction of a risk quantification model and the calculation of a grid ship collision index. A case study is carried out on the multi-bridge waterway of the Yangtze River in Wuhan, China. The results show that the size of the ship domain is highly correlated with the ship's speed and length, and analysis of collision risk can reflect the real situation near bridge-waters, which is helpful to demonstrate the application of the ship domain in quantifying the collision risk and to characterise the collision risk distribution near bridge-waters.

## 1. Introduction

The ship domain represents the surrounding effective waters which the navigator would like to keep free with respect to other ships and fixed objects (Goodwin, 1975; Rawson and Brito, 2020). Great efforts have been made in the research of ship domain models. The ship domain model is of great significance to the safety of inland shipping and the comprehensive utilisation of water resources (Liu et al., 2011). Meanwhile, the concept of ship domain has been widely applied in the study of navigation safety, especially in open waters, since the 1970s (Fujii and Tanaka, 1971; Goodwin, 1975; Wang et al., 2017).

In the literature, the ship domain models are divided into three types according to their shapes: the circular ship domain, the elliptical ship domain and the polygonal ship domain. Relevant methods for ship domain modelling are divided into three categories: statistical-based methods (Coldwell, 1983), analytical expression-based methods (Chen and Guo, 2008), and intelligent technology-based methods (Zhu et al., 2001; Chen et al., 2020). In statistical-based methods, using the ship domain calculation method based on statistics combined with the navigation characteristics of navigable waters and water collision avoidance rules, traffic observation data are analysed to determine the shape and size of the ship domain (Liu et al., 2011). Statistical-based methods have been used in several ship encounter situations. Therefore, these models have been simulated and analysed. Based on the dynamic model, the polygonal

ship domain was obtained focusing on different encounter situations in open waterways, poor visibility and narrow waterways (Pietrzykowski and Uriasz, 2004, 2006, 2009). Statistical research on inland ship domains has also been performed to investigate their characteristics. The method based on the analytical expression for calculating ship domain includes the speed, heading and position of the target ship and other ships as inputs using the ship motion equation. The deceleration model or the steering equation is utilised to calculate the minimum distance which the ship could maintain in all directions. Thus, the ship domain model is finally derived. Concerning analytical expression-based methods, the quaternion ship domain model is proposed to calculate the minimum range in four directions around the ship based on the ship's length and speed (Zhang et al., 2021). The methods based on intelligent technology generally include fuzzy theory and neural networks, among which the fuzzy theory is usually adopted for establishing the mapping relationship between the ship distance, the heading and relative orientation of the ship and the navigational risk of the ship. Thus one can proceed to determine the fuzzy boundaries of the ship domain (Pietrzykowski, 2008).

The dense vessel traffic flow near bridge-waters, which refer to the waters within a certain distance along the bridge axis, exert high pressure on the inland maritime administration. Although ship domain models have been accomplished, there are few studies in the literature focusing on ship domain models for inland waterways (Szlapczynski and Szlapczynska, 2017), especially the area near bridge-waters where collisions between ships and bridge piers often occur due to human error (Zhang et al., 2020). Because of the special navigation rules and conditions near bridge-waters, the ship domain model is expected to be different from the conventional models, and its characteristics need to be further explored.

AIS (automatic identification system) data are crucial for research on navigation safety. They provide not only guidance for ship collision avoidance, but also an effective way of retrieving ship trajectories. The frequencies of ship collisions and possible near-miss ship collisions have been widely investigated by researchers (Montewka et al., 2010; Weng et al., 2012; Zhang et al., 2015). Shipping risks were analysed by Wang et al. (2017) and Felski and Jaskólski (2013); analysis of ship traffic behaviour and a traffic simulation model were presented by Xiao et al. (2015) and Azevedo et al. (2015). In other research, situational awareness of inland navigation was achieved through vessel trajectory visualisation and traffic density mapping (Willems et al., 2013). Therefore, AIS data could be helpful for research on ship domain in the vicinity of bridge-waters since massive historical AIS data can reflect the actual navigation conditions to a certain extent (Pietrzykowski and Magaj, 2017; Zhang et al., 2017; Zhang and Meng, 2019; Rawson and Brito, 2021).

In this paper, the ship domain model for inland navigation on multi-bridge waterways is established by using AIS data. Ship domain is defined as the area free from ships and fixed objects, which mainly refers to bridge piers in multi-bridge waterways. Firstly, the inland ships are classified by their speed and length, and the trajectory distribution of surrounding ships is determined. The relative distance between ships and bridge piers is calculated when ships are passing through bridges. Secondly, the nearest point boundary determination method is adopted, and the elliptical ship domain is established. The shape of the ship domain (i.e., the longitudinal and lateral radii) is computed by the least squares method. Its relationship with the ship's length and speed is thus discussed. Fuzzy quaternion ship models are also presented for inland ships and bridge piers. A case study is carried out in the Wuhan reach of the Yangtze River, China. The model results are noted as consistent with field observations and the literature.

The main part of this study is organised as follows: the methods for construction of ship domain based on AIS data and calculation of the regional collision risk are introduced in Section 2. In Section 3, a case study is conducted in the Yangtze River, China, including the construction of the ship domains near bridge-waters, comparison between different domains, and analysis of the characteristics of collision risk distribution. In Section 4, the conclusions are presented.

## 2. Methodology

This section includes two parts, the first part describes the process of construction of ship domain in detail. The second introduces the risk assessment based on the fuzzy quaternion ship domain. The aim

**Table 1.** Description of variables.

Variable	Description
$\theta$	Ship heading
$(a, b)$	UTM coordinates of target ship
$(x, y)$	UTM coordinates of neighbouring ship
$(x', y')$	Relative coordinates of neighbouring ship
Dis	Distance between target ship and neighbouring ship
$\delta$	Angle of other ship in the clockwise direction of the target ship
$L$	Ship's length
$v$	Ship's speed over ground
Lx	Projected distance of Dis on the X axis
Ly	Projected distance of Dis on the Y axis
$a$	Longitudinal radii of the shipping domain
$b$	Lateral radii of the shipping domain
$K$	Positive scale factor in function expression for longitudinal radii
$K_0$	Constant in function expression for lateral radii
$\eta$	Coefficient in function expression for lateral radii
FQSD	Fuzzy quaternion ship domain
LonCR <sub>k</sub>	Longitudinal collision risk quantitative model
LatCR <sub>k</sub>	Lateral collision risk quantitative model
SCR	Spatial collision risk quantitative model
GSCI	Grid ship collision index

is to analyse ship collision risk by establishing the fuzzy quaternion ship domain for bridge-waters. The nomenclature table (Table 1) explains the meaning of the variables in this paper.

### 2.1. Construction of ship domain

The ship domain is determined by the relative positions of ships near bridge-waters. In order to determine the parameters of the ship domain model, the ships are classified according to their length and speed, and the position and direction of the neighbouring ships are superimposed to obtain the scope of the specific type of ships. In some research, the type of ship is supposed to be related to ship domain size (Szlapeczynski and Szlapeczynska, 2017). In narrow and crowded waters, however, the ship's length and speed are mainly considered. In this paper, 'target ships' refers to ships sailing in the research domain.

After preprocessing of AIS data, the moving ships are selected as target ships. The position and direction of the neighbouring ship are calculated during the current trajectory time of the target ship; multiple data are superimposed to obtain the ship domain of a single ship. According to the ship's length and speed, the data is classified and superimposed to obtain the ship domain of a specific ship type (Xiang et al., 2015). The algorithm flow of statistical analysis is shown in Figure 1.

As shown in Figure 1, the flowchart is composed of five critical processes: preprocessing of AIS data, selection of target ship, superimposition of single ship data, superimposition of data of the same type of ship and obtaining the ship domain of the specific type of ship. The specific steps are as follows:

- AIS data preprocessing
- Ships whose length or width do not conform the actual static data are excluded. The dynamic AIS data are cleansed by frequency and maritime mobile service identity (MMSI). The latitude and longitude drift data are filtered out. Ship trajectories are restored. The AIS data preprocessing is generally demonstrated as the following algorithm 1 (Zhang et al., 2018), and the results are presented in Figure 2. Although some points are not eliminated due to the method based on the spatial distance between adjacent points, they have no impact on subsequent content about ship domain statistics.

**Algorithm 1:** AIS Data pre-processing.

```

Input: AIS dynamic data
Result: Matrix of Same Track Data
1 Sort data according to MMSI;
2 SameShipDataMat  $\leftarrow$  Create a mat to save data for the same ship;
3 SameTrackDataMat  $\leftarrow$  Create a mat to save data for the same track of the same ship;
4 while AIS dynamic data is not empty do
5   M  $\leftarrow$  Calculate the value of this MMSI minus the previous one;
6   if M equals 0 then
7     Save data of this line to this mat in SameShipDataMat;
8   else
9     Create a new mat in SameShipDataMat;
10    Save data of this line to the new mat;
11  end
12 end
13 for data in SameShipDataMat do
14   Sort data according to timestamp;
15   for each timestamp in data do
16     T  $\leftarrow$  Calculate the time interval between this timestamp and the previous one;
17     if T < t then
18       D1  $\leftarrow$  Calculate distance between this point and the previous one by longitude and latitude;
19       D2  $\leftarrow$  Calculate distance between this point and the previous one by T and ship speed;
20       if D1 < D2+d then
21         Save data of this line to this mat in SameTrackDataMat;
22       else
23         Delete data of this line;
24       end
25     else
26       Create a new mat in SameTrackDataMat;
27       Save data of this line to the new mat;
28     end
29   end
30 end
31 Return SameTrackDataMat;
  
```

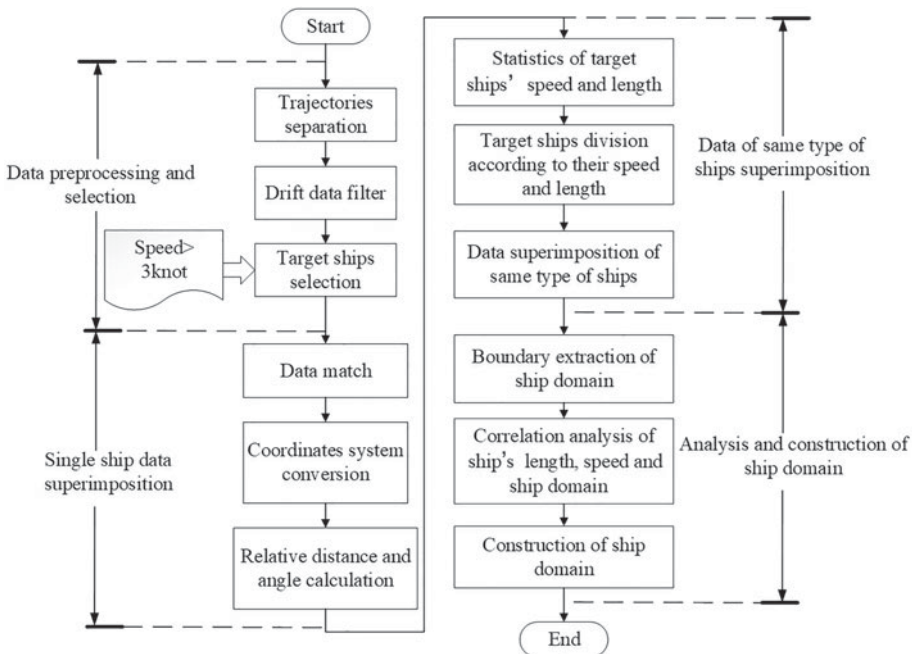
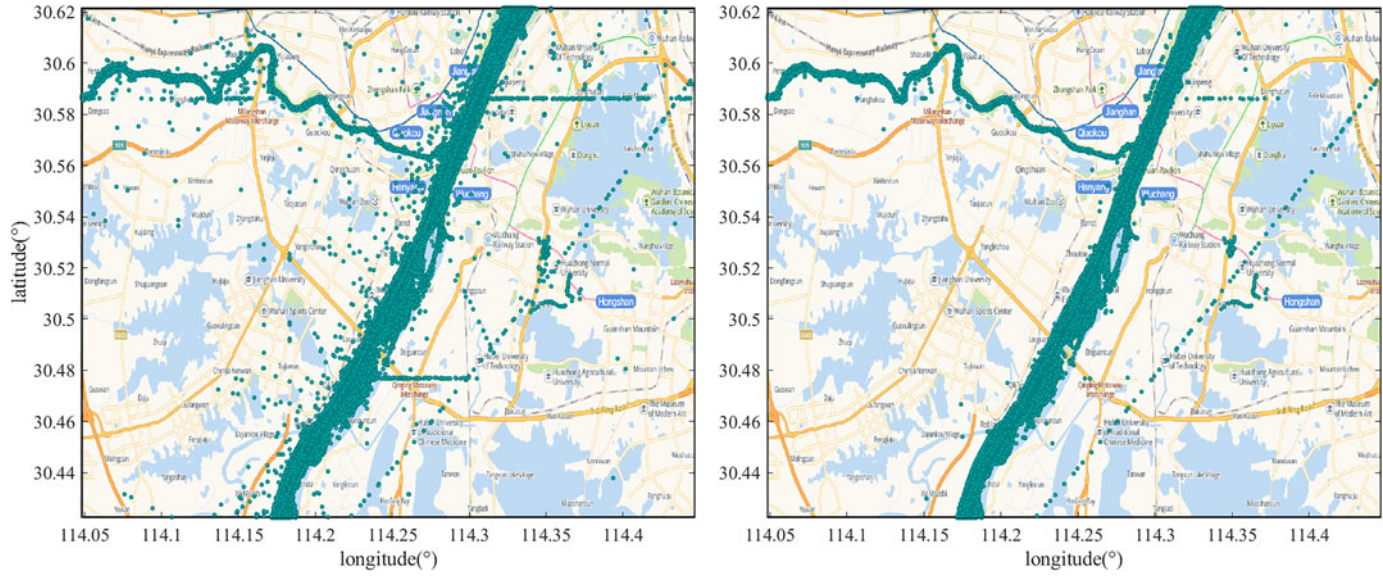


Figure 1. Flowchart of ship domain statistical algorithm.



*Figure 2. Comparison of inland AIS data before (left) and after (right) preprocessing.*

- Target ships selection
- Select various types of ships in the research domain with stable speeds of not less than 3 knots as target ships from AIS data according to the partition method of reporting intervals in ITU-R m.1371-4 Proposal (ITU, 2010).
- Data matching
- To reduce the calculation complexity and improve the efficiency of the algorithm, the matching process is divided into two steps: preliminary matching and advanced matching. The preliminary matching uses overlapping timestamps to determine that the ships appear in a certain area at the same time. Meanwhile, the advanced matching can find the ship's trajectory point within the target ship's range, which is prepared for ship domain analysis. When another ship's timestamp is different from the target ship, it is necessary to predict their movement at the same time. In this paper, piecewise cubic Hermite interpolation is adopted to obtain the prediction result.
- Coordinate transformation
- Firstly, because the accurate location of ship-borne AIS is not available at the moment, the AIS antenna position is regarded as the centre of the target ship. Furthermore, since the latitude and longitude data cannot be directly used for coordinate calculation, latitude and longitude are converted to Universal Transverse Mercator (UTM) coordinates. Let  $\theta$  denote the angle between ship heading and the north direction,  $(a, b)$  and  $(x, y)$  represent the UTM coordinates of target ship and neighbouring ship respectively, the coordinate transformation is performed for neighbouring ships (in Figure 3). The origin is set at the position of the target ship.

$$x' = (x - a) * \cos \theta - (y - b) * \sin \theta \tag{1}$$

$$y' = (x - a) * \sin \theta + (y - b) * \cos \theta \tag{2}$$

- Distance and angle calculation
- Through the above coordinate transformation, the distance between the target ship and neighbouring ship can be calculated as follows:

$$\text{Dis} = (a - x)^2 + (y - b)^2 \tag{3}$$

The clockwise angle between other ships and the target ship can be calculated as follows:

$$k = (\vec{m} \otimes \vec{n}) / (|\vec{m}| * |\vec{n}|) \tag{4}$$

$$\delta = \text{Arc cos}(k) \tag{5}$$

$$\delta = \begin{cases} \delta, & x' > 0 \\ 360 - \delta, & x' < 0 \end{cases} \tag{6}$$

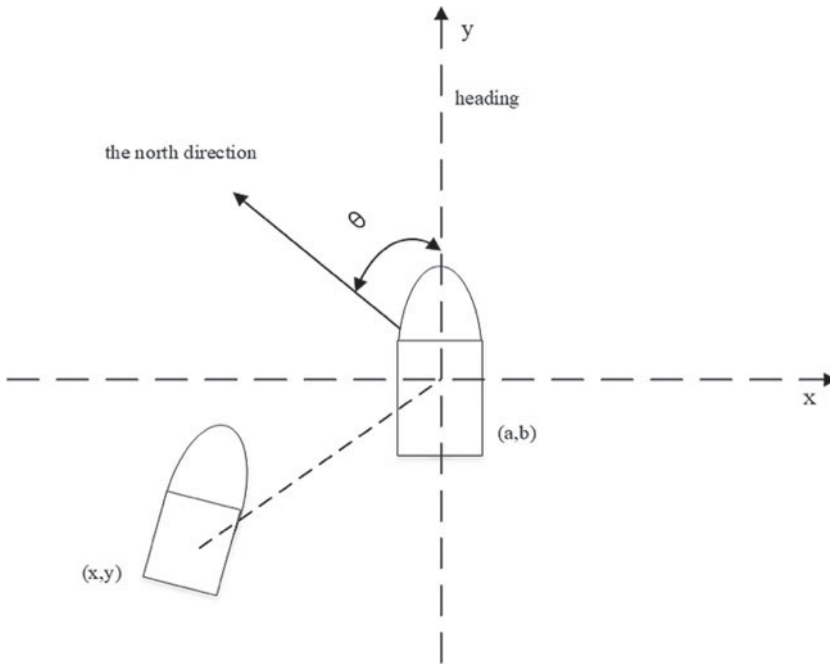
$\vec{m}$  represents the positive direction of the y-axis, which can be taken as (0,1),  $\vec{n}$  denotes the transformed coordinates of neighbouring ships  $(x', y')$ .

- Distance conversion
- To obtain the ship domain, the spherical distance is projected into the transformed coordinate:

$$Lx = \text{Dis} * \sin \delta \tag{7}$$

$$Ly = \text{Dis} * \cos \delta \tag{8}$$

- Statistical results
- The ships are classified according to their length and speed, and superimposed on this basis to obtain the distribution map of other ships and bridge piers, so as to calculate the characteristic parameters of the ship domain. The ship domain boundary is fitted by the least squares method based on the nearest points around the target ship (Section 3.1).



**Figure 3.** Schematic diagram of coordinate transformation.

- Ship domain model construction
- Considering that ships are in a relatively crowded scenario when passing through river bridge-waters, the domain model of the target waters is constructed based on the model proposed in the papers (Jia, 1989; Fan et al., 2013) in which the ship domain model in congested river water areas is analysed, as follows:

$$a = \frac{1}{2}LKv \tag{9}$$

$$b = \frac{1}{2}BKv^{k_0} \tag{10}$$

where  $a$  and  $b$  indicate the longitudinal and lateral radii of the shipping domain respectively,  $L$  and  $B$  represent the length and width of the ship,  $v$  is the speed over ground of the target ship,  $K$  is a positive scale factor and  $K_0$  is a constant less than 1. Since the focus is on the influence of the ship’s length and speed on the ship domain, the ship width is replaced by ship’s length with a coefficient:

$$b = \frac{1}{2}\eta LKv^{k_0} \tag{11}$$

**2.2. Collision risk quantification model**

Based on the statistical analysis of inland ship domain, a spatial collision risk quantitative model combined with a novel inland fuzzy quaternion ship domain model is proposed to quantify collision risk (Szlupczynski and Szlupczynska, 2017). In this paper, collision risk refers to the probability of collision between ships or between ships and bridge piers. The key of collision risk assessment is the degree of invasion of one ship to another. Therefore, fuzzy quaternion ship domain is first introduced to distinguish different levels of intrusion, and then a collision risk quantification model is constructed. Then, the method of regional risk grid calculation is proposed to visualise the collision risk in a specific area, which is helpful to understand the risk pattern of ships passing through particular bridge-waters.

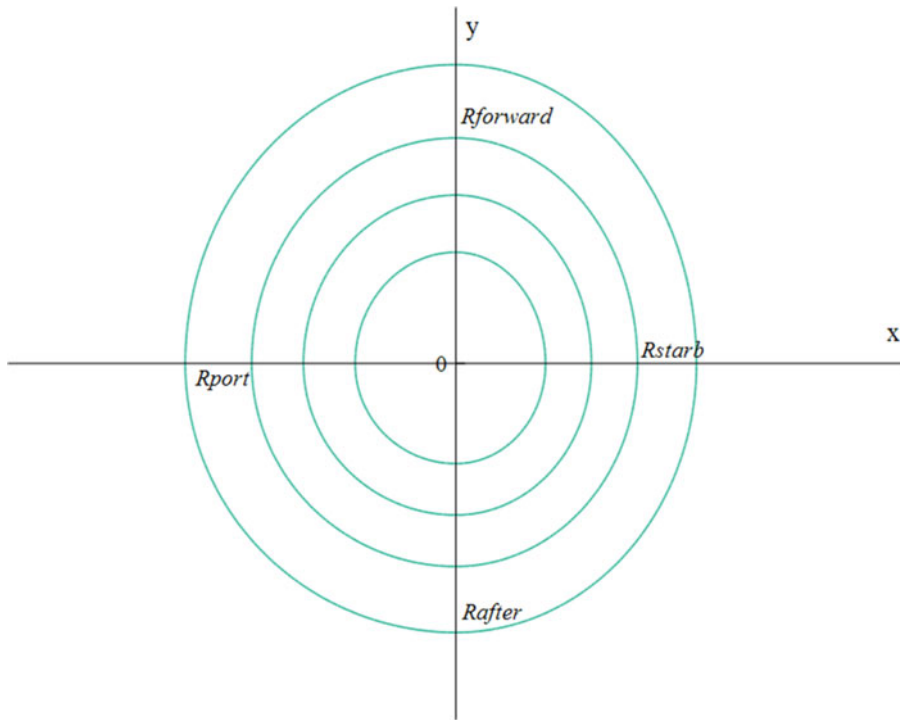


Figure 4. Schematic diagram of fuzzy quaternion ship domain model.

2.2.1. Fuzzy quaternion ship domain

In view of the fact that the longitudinal and lateral radii of the ship domain selected in this paper are determined by the nearest point of the target ship, the model of the ship domain can be further constructed by means of the fuzzy quaternary ship domain. Generally, a fuzzy quaternion ship domain model (Pietrzykowski, 2008; Szlapczynski and Szlapczynska, 2017) is defined as follows:

$$FQSD_{ce}(r) = \{(x, y) | f_{ce}(x, y : Q(r)) \leq 1\} \tag{12}$$

$$f_{ce}(x, y; Q(r)) = \left( \frac{2y}{(1 + \operatorname{sgn}x)R_{\text{before}}(r) - (1 - \operatorname{sgn}x)R_{\text{after}}(r)} \right)^2 + \left( \frac{2x}{(1 + \operatorname{sgn}y)R_{\text{starb}}(r) - (1 - \operatorname{sgn}y)R_{\text{port}}(r)} \right)^2 \tag{13}$$

$$Ri(r) = \left( \frac{\ln(1/r)}{\ln(1/r_0)} \right)^{1/2} \quad Ri, i \in \{R_{\text{before}}, R_{\text{after}}, R_{\text{starb}}, R_{\text{part}}\} \tag{14}$$

where  $\operatorname{sgn}x = \begin{cases} 1, x \geq 0 \\ -1, x < 0 \end{cases}$ ,  $R_{\text{before}}$ ,  $R_{\text{after}}$ ,  $R_{\text{starb}}$  and  $R_{\text{port}}$  represent the radius in the four directions of forward, after, port and starboard, respectively, as shown in Figure 4.  $r$  is the value corresponding to the boundary of a ship domain ( $r_0 = 0.8$  is adopted in the present study considering the innermost layer), thus the values of  $r$  from inside to outside are respectively 0.8, 0.6, 0.4, 0.2.

2.2.2. Collision risk quantification model

For the purpose of calculating collision risk, the fuzzy quaternion ship domain is applied to develop the longitudinal collision risk model and the lateral collision risk model of the ship (Zhang and Meng, 2019). The spatial collision risk quantitative model is then obtained. The longitudinal collision risk model is



mainly for the calculation of the risk in the fore and rear directions of the ship during navigation. The model is defined as follows.

$$\text{LonCR}_k(y) = \exp\left(-\left(\frac{2y}{(1 + \text{sgny})\sigma_{\text{before}} + (1 - \text{sgny})\sigma_{\text{after}}}\right)^2\right)$$

$$\sigma_i = \frac{R_i}{(\ln(1/r_0))^{1/2}}, i \in \{R_{\text{before}}, R_{\text{after}}\} \tag{15}$$

where  $\text{LonCR}_k$  represents the longitudinal collision risk, and  $x$  and  $y$  represent the relative position of the neighbouring ship (or pier) to the target ship shown in Figure 3, while other symbols have been explained in the previous section. The lateral collision risk model is used to calculate the risk value in the port and starboard. The model definition is shown as follows.

$$\text{LatCR}_k(x) = \exp\left(-\left(\frac{2x}{(1 + \text{sgnx})\sigma_{\text{starb}} + (1 - \text{sgnx})\sigma_{\text{port}}}\right)^2\right)$$

$$\sigma_i = \frac{R_i}{(\ln(1/r_0))^{1/2}}, i \in \{R_{\text{starb}}, R_{\text{port}}\} \tag{16}$$

where  $\text{LatCR}_k$  represents the lateral collision risk. Through the above two quantitative models, the spatial collision risk (SCR) quantitative model is established and defined in Equation (17).

$$\text{SCR}(x, y) = \text{LonCR}(y).\text{LatCR}(x)$$

$$= \exp\left(-\left(\left(\frac{2y}{(1 + \text{sgny})\sigma_{\text{before}} + (1 - \text{sgny})\sigma_{\text{after}}}\right)^2 + \left(\frac{2x}{(1 + \text{sgnx})\sigma_{\text{starb}} + (1 - \text{sgnx})\sigma_{\text{port}}}\right)^2\right)\right) \tag{17}$$

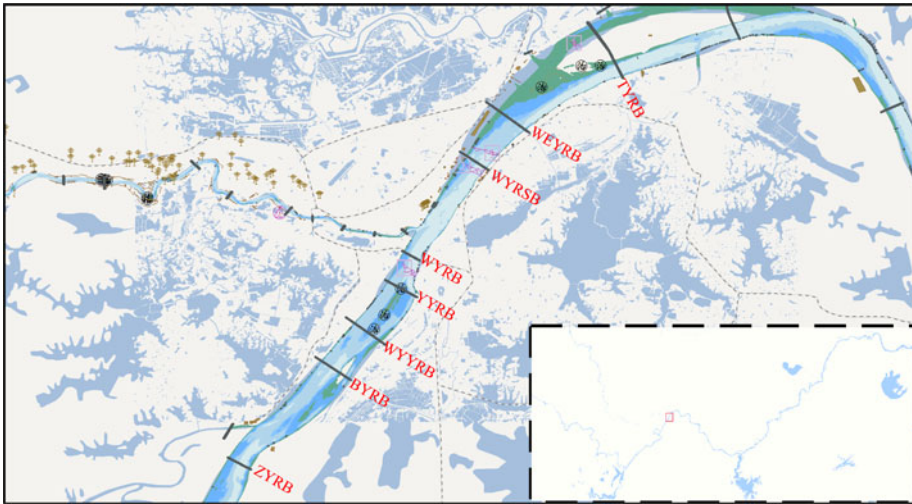
2.2.3. *Regional collision risk calculation*

Based on the SCR for the calculation of single ship spatial collision risk, the regional collision risk could be quantified, superimposed and visualised. The detailed process is shown as follows:

- 1 Area gridding: according to the longitude and latitude range of the area and the grid size, the water area is divided into grids.
- 2 Data selection: select ships navigating in the specified area within a specified time period.
- 3 Data matching: transforming coordinates and calculating distances are the same as the steps described in Section 2.1.
- 4 Ship risk evaluation: the risk of the target ship is calculated according to the distance measurement and spatial collision risk model aforementioned;
- 5 Grid risk accumulation: the position of the target ship is matched to the grid of the water area. Then the collision risk of each grid is superimposed to obtain the grid ship collision index (GSCI). Finally, we obtain the distribution of ship collision risk in the water area.

$$\text{GSCI}_{(x,y)} = \sum_i^n \text{SCR}^i_{(x,y)}, (x \in X) \cap (y \in Y) \tag{18}$$

where  $\text{SCR}_{(x,y)}$  refers to the collision risk of the corresponding grid and  $X, Y$  indicates the longitude and latitude coordinates of the grid.



**Figure 5.** The layout of bridges in the research domain.

### 3. Case study

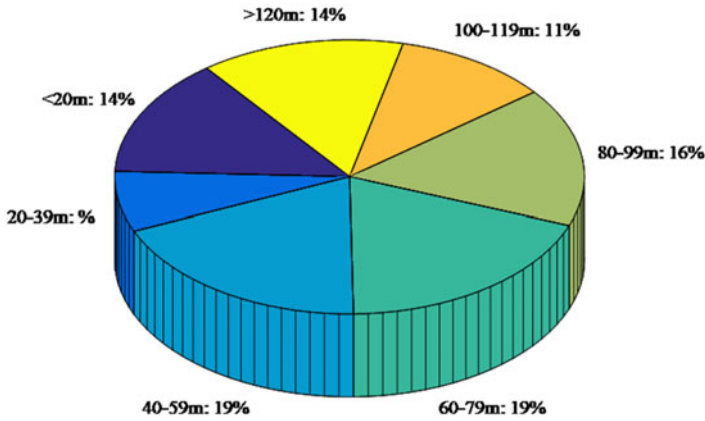
A case study has been conducted on bridge-waters in the Wuhan reach of the Yangtze River. The target ships are divided into different categories to analyse the relationship between ship domain and ship's length and speed. After obtaining the ship domain boundary through the construction process outlined in Section 2.1, the relationship between ship domain near bridge-waters and ship's speed and length is analysed, and then two ship domain models are constructed. In order to evaluate these models, comparison between the ship domains in this study and other popular domains is performed, and the model based on ship–pier data is applied to the actual on-going ship passing through the WYRB. Finally, the collision risk distribution for different bridges is characterised on the basis of GSCI when ships pass through.

#### 3.1. Statistical analysis of inland ship domain

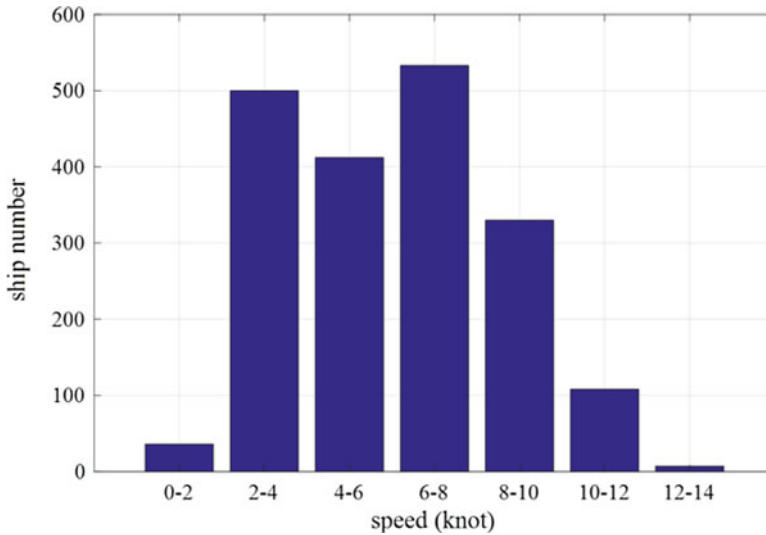
The Wuhan reach of the Yangtze River has been characterised as a multi-bridge waterway. There are eight bridges within the research domain: Zhuankou Yangtze River Bridge (ZYRB), Baishazhou Yangtze River Bridge (BYRB), Wuhan Yangsigang Yangtze River Bridge (WYYRB), Yingwuzhou Yangtze River Bridge (YYRB), Wuhan Yangtze River Bridge (WYRB), Wuhan Yangtze River Second Bridge (WYRSB), Wuhan Erqi Yangtze River Bridge (WEYRB) and Tianxingzhou Yangtze River Bridge (TYRB), from upstream to downstream (as presented in Figure 5). A total number of 5,994 trajectories are derived for statistical analysis in the present study. Navigation marks have been set up to ensure water transport safety in the research domain. Their positional information is obtained from the Yangtze River Maritime Bureau, Ministry of Transport, China.

Ships navigating on multi-bridge waterways should avoid collisions with neighbouring ships, piers and navigation marks. Therefore, the statistics of ship domain based on other ships, bridge piers and navigation marks are separately counted. Since the speed of inland ships varies with ship's length, the ships are divided into seven categories according to their length (m) as: <20, 20–39, 40–59, 60–79, 80–99, 100–119, and >120; the distribution of ship lengths is shown in Figure 6. Ships with a length of 40–59, 60–79, 80–99 and 100–119 were selected for the present study. The distribution of ship speeds is therefore provided for the corresponding ships (Figure 7).

According to the method described above, the spatial distribution of other ships and navigation marks around the target ship could be obtained. To determine the boundary of the ship domain, the distance corresponding to other ships and bridge piers in all directions around the target ship needs to



**Figure 6.** *Distribution of ship lengths.*

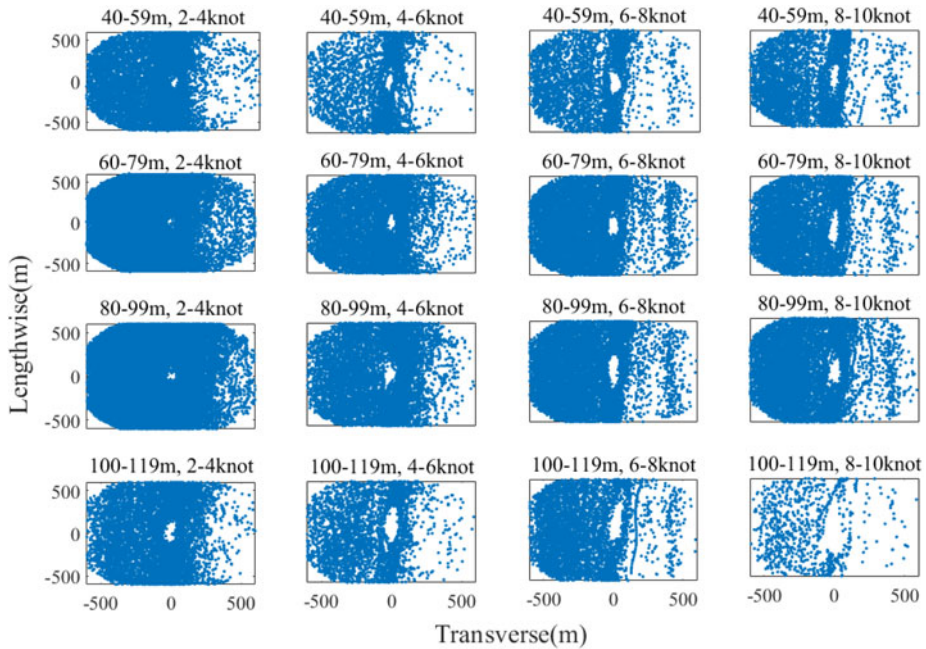


**Figure 7.** *Distribution of average ship speeds.*

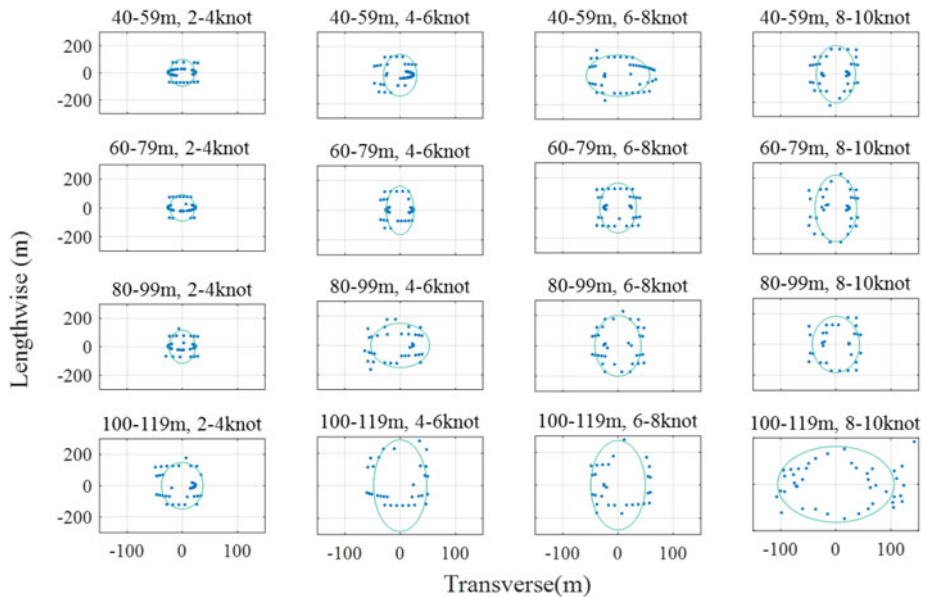
be computed. In view of the navigation rules and restrictions in multi-bridge waterways, ships are not allowed to surpass. Generally, both up-bound and down-bound ships keep to the right on the waterways. In other words, oncoming ships sail on the left side of the target ship. Combined with the least squares method, nearest boundary points of the target ship are used to determine the boundary of the ship domain. Specifically, it can be divided into two steps: firstly, the 600 m circular water area around the target ship is divided into 72 sectors with intervals of 5°, and the nearest boundary point is determined for each sector; secondly, the least squares method is used to fit the elliptical ship domain boundary after obtaining the 72 distance values in different directions. The selection of nearest boundary points is also based on the construction of a fuzzy ship domain model, and the boundary based on nearest boundary points can be the innermost boundary.

*3.1.1. Ship domain statistics based on other ships*

As stated before, ship trajectories have been classified and analysed according to the ship’s length and speed. The distribution of other ships around target ships is thus matched and spherical distances are calculated. The spatial distribution of other ships is presented in [Figure 8](#). The origin denotes the position of the target ship while the blue points represent surrounding ships.

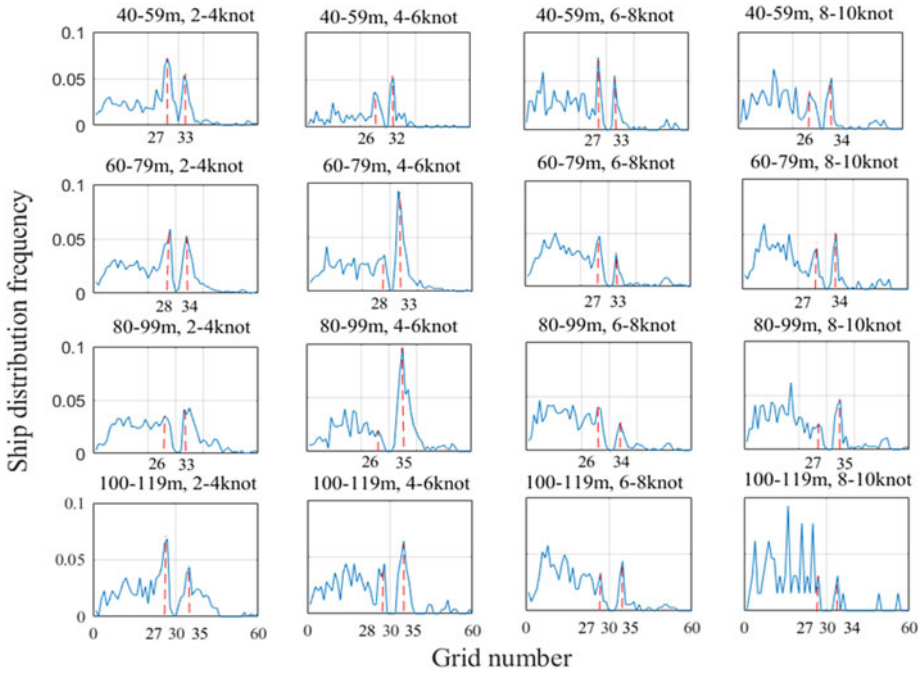


**Figure 8.** Spatial distribution of ships (blue points) around target ships of four different lengths and speeds.



**Figure 9.** Fitted elliptical ship boundary for ships with different speeds.

The analysis flow of visual exploration is schematically shown in Figure 2. First of all, the differential distance is calculated from the raw data and displayed in a scatterplot. According to the visualisation of distance in the scatterplot, the data quality problems are identified and their spatiotemporal distribution is displayed on figure by interacting with the scatterplot. Finally, the hidden values of dirty data are explored by means of interaction and observation with the distribution graph and the four-quadrant stacked graph.



**Figure 10.** The horizontal interface of the target ship domain for four ships of different lengths and speeds.

Some ships on the left travel in the opposite direction without characteristics of the ship domain boundary. Therefore, the relevant positional information should be excluded during the boundary determination. In the present study, the following rules have been defined and applied:

Rule I: the boundary points with positive abscissa values higher than the positive mean value or with negative abscissa lower than the negative mean value will be deleted;

Rule II: the boundary points with a distance from the centre greater than the threshold (X) will be deleted; the value of X is determined according to the actual situation.

By analysing the spatial distribution of other ships (in Figure 9), the threshold value of X is defined as 400. The fitting results of the elliptical ship domain boundary are illustrated in Figure 10. The major and minor axes are calculated for each elliptical ship domain and tabulated (in Table 2). The accuracy of elliptical ship boundary fitting is affected by the sample dataset. Therefore, part of the elliptical ship domain calculation has been excluded for analysis, e.g., the last row in Table 2.

The Pearson correlation coefficients of the major axis, minor axis and ship’s speed under the same ship’s length are 0.9375/0.9498/0.8721/0.9363, 0.5609/0.9818/0.4823/0.8627, which indicates that there is a strong linear relationship between shipping and the major and minor axes of ship domain. In other words, the ship domain will vary according to ship’s speed. This is consistent with the results of applying the density-in-cell method (Wen et al., 2018). The correlation coefficients of the major axis, minor axis and length of the ship at the same speed are 0.8901/0.7900/0.9667/0.7173, 0.8075/0.7765/−0.1942/0.8265 respectively, which also illustrates that the major and minor axes of the ship domain will vary with the ship’s length.

The sectional distribution patterns around the target ship are also explored. Take the minor axis section as an example (as shown in Figure 10), the water area near the target ship is divided by grids with the side length of 10 m, the x-axis represents the number of each grid where neighbouring ships in the port and starboard sides are located while the target ship is located in No.30 grid, and the Y coordinate indicates the frequency of the number of neighbouring ships in a corresponding grid divided

**Table 2.** Elliptical ship domain size for ships of different length and speed.

Ship type		Elliptical ship domain size	
Length (m)	Speed (knots)	Major axis (m)	Minor axis (m)
40–59	2–4	98.5	22.6
	4–6	145.5	30.2
	6–8	143.2	57.5
	8–10	203.7	35.5
60–79	2–4	92.1	22.5
	4–6	163.6	25.1
	6–8	166.2	33.3
	8–10	219.0	37.7
80–99	2–4	116.3	23.6
	4–6	150.7	53.2
	6–8	200.6	41.8
	8–10	183.6	42.7
100–119	2–4	148.6	37.6
	4–6	281.8	48.1
	6–8	272.2	49.5
	8–10	407.9	105.5

by the total number of ships in all the grids. The frequency of neighbouring ships in the grid between No. 0 and No. 30 is higher than that in the grid numbered larger than 30. This is because both up-bound and down-bound ships keep to the right on the waterways, which can be seen in [Figure 9](#). Although fluctuations are noted for the transverse frequency distribution, characteristics of the ship domain are found to be similar to [Wen et al. \(2018\)](#). However, compared with the relationship between the major and minor axis of the ship domain and the ship's length in [Wen et al. \(2018\)](#), the ratio in the present study is found to be smaller, especially on the minor axis. On the other hand, compared with the results in [Table 2](#) obtained by nearest boundary point and least squares method, the minor axis is relatively close, which verifies the effectiveness of the aforementioned method of obtaining the boundary. However, it does not show the relationship between ship domain and ship's speed and ship's length because it only considers peak values on the port and starboard sides instead of boundaries in all directions.

### 3.1.2. Ship domain statistics based on bridge piers

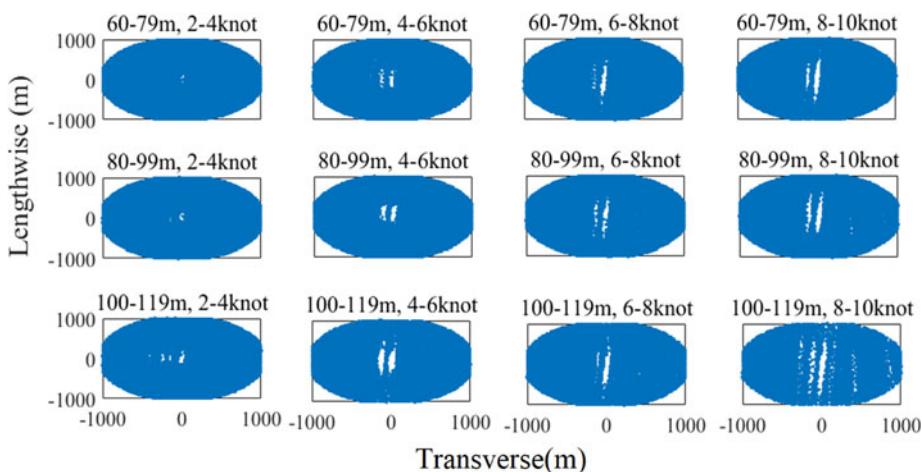
The position of the bridge pier relative to the target ship is overlapped and counted when the target ships pass by. [Figure 11](#) shows the spatial distribution of ships passing through specified bridge piers with different speeds. The origin denotes the target ships while the blue points represent the relative location of the bridge pier.

[Figure 11](#) shows that most of the ships do not go straight ahead when they pass through the vicinity of bridge-waters. Instead, they either deflect to the right and then to the left or keep skewing to the left. The ship domain boundary based on the bridge pier is obtained by the method for the nearest boundary point as described above. The fitted elliptical ship boundary for ships passing through bridge piers is presented in [Figure 12](#). Meanwhile, the results of major and minor axes calculations are shown in [Table 3](#).

The ship domain model of navigation marks is also investigated in the present study. However, the results indicate no obvious evidence of its existence. Since the navigation marks are relatively small compared with the ships and their location is not fixed, ship domain based on navigation marks is not obvious. This will not be considered and discussed hereinafter.

**Table 3.** Elliptical ship domain size for ships of different length and speed passing through piers.

Ship type	Elliptical ship domain size			
	Length (m)	Speed (knots)	Major axis (m)	Minor axis (m)
40–59		2–4	230.4	29.5
		4–6	321.3	40.5
		6–8	559.8	44.2
		8–10	474.3	47.5
60–79		2–4	255.7	23.5
		4–6	261.9	40.7
		6–8	447.4	41.4
		8–10	437.3	39.9
80–99		2–4	204.4	24.5
		4–6	306.9	47.3
		6–8	407.4	38.6
		8–10	461.2	43.5
100–119		2–4	245.5	42.1
		4–6	346.6	54.6
		6–8	452.1	43.8
		8–10	499.4	58.8



**Figure 11.** Spatial distribution of ships passing through bridge piers.

### 3.2. Ship domain model in WYRB waters

Based on the statistical results in Section 3.1, the ship domain was constructed based on other ships and piers respectively.

#### 3.2.1. Fuzzy elliptical ship domain model based on ship–ship data

According to the above-mentioned quaternion ship domain model, given  $R_l = R_{\text{before}} = R_{\text{after}}$  and  $R_s = R_{\text{strab}} = R_{\text{port}}$ , the function fitting the longitudinal and lateral radii with the ship’s speed and the

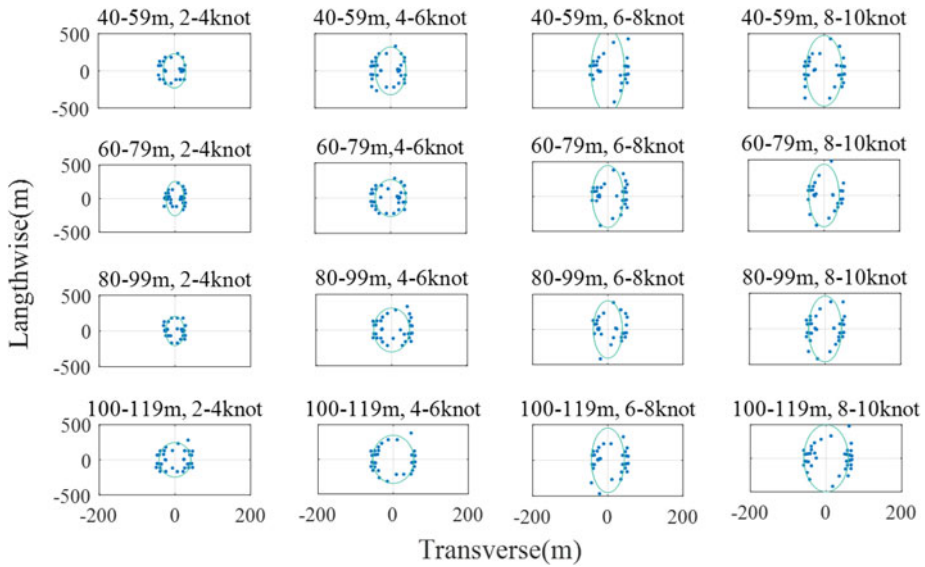


Figure 12. Fitted elliptical ship boundary for ships passing through bridge piers.

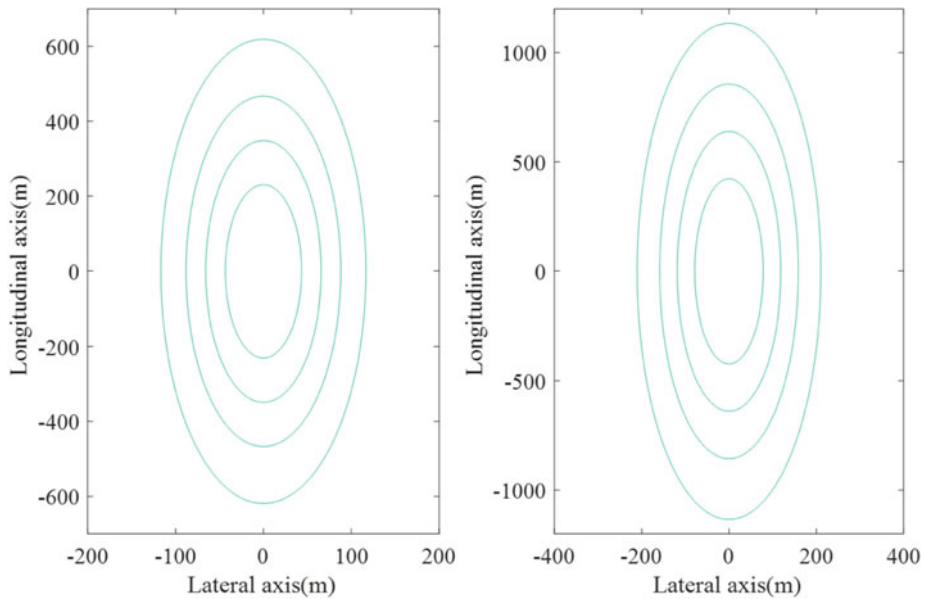


Figure 13. Fuzzy ship field model based on the neighbouring ships (left) and bridge piers (right).

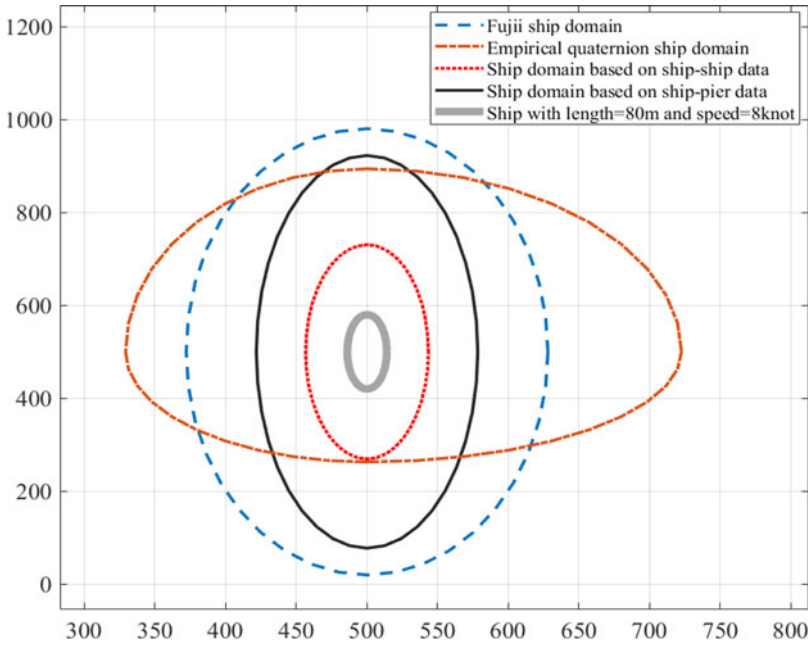
ship’s length using least squares is obtained, as shown in the following equation:

$$R_l = \frac{1}{2} \cdot 0.7202 \cdot l \cdot v \tag{19}$$

$$R_s = \frac{1}{2} \cdot 0.6138 \cdot l \cdot 0.7202 \cdot v^{0.432} \tag{20}$$

where  $R_l$  and  $R_s$  represent the longitudinal and lateral radii respectively,  $v$  and  $l$  represent the speed and length of target ships. The ship domain model is shown on the right of Figure 13.





**Figure 14.** Comparison of ship domain models.

3.2.2. *Fuzzy elliptical ship domain model based on ship–pier data*

Using the above method to curve fit the ship domain based on bridge piers, the following formulas for the longitudinal and lateral radii are obtained:

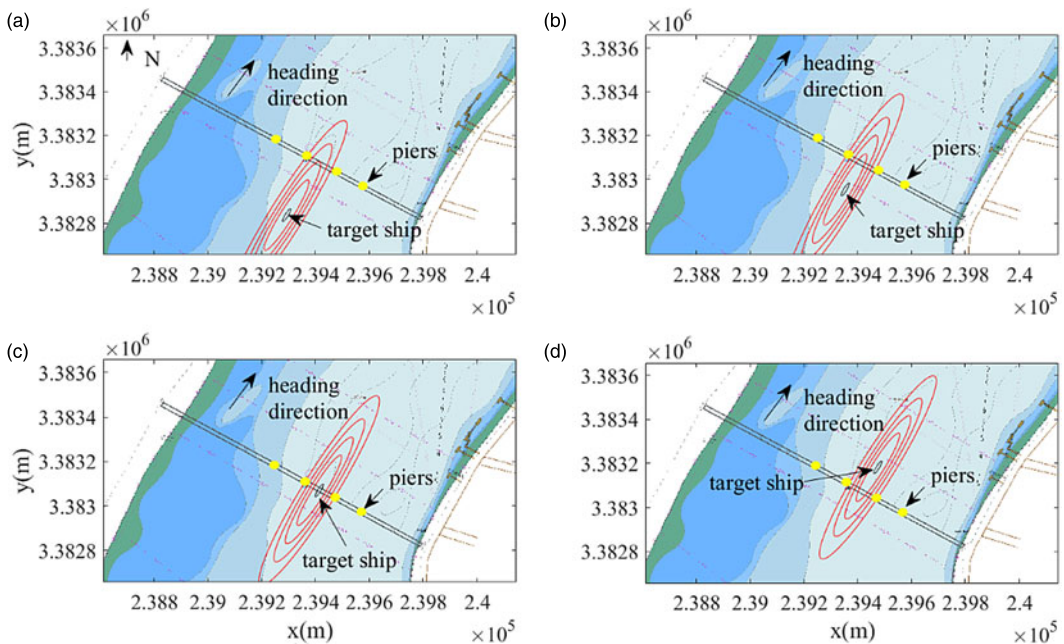
$$R_l = \frac{1}{2} \cdot 1.3198 \cdot l \cdot v \tag{21}$$

$$R_s = \frac{1}{2} \cdot 0.7875 \cdot l \cdot 1.3198 \cdot v^{0.3047} \tag{22}$$

where  $R_l$  and  $R_s$  represent the longitudinal and lateral radii respectively,  $v$  and  $l$  represent the speed and length of target ships. The ship domain model is shown on the right of Figure 13.

The ship domain is compared with the traditional ship domains, including Fujii’s domain (Fujii and Tanaka, 1971), a quaternion ship domain. The Goodwin’s domain (Goodwin, 1975) is not drawn in the figure because it is much bigger than other domains. Figure 14 shows the comparison results. From outside to inside, they are, respectively: Fujii’s domain, ship domain based on ship–pier data, and ship domain based on ship–ship data. Specifically, the rear boundary of the ship domain based on ship–ship data is basically equal to that of the quaternion ship domain. In terms of the fore boundary, the ship domain based on ship–pier data approximately matches the quaternion ship domain. The difference between the quaternion ship domain and other domains lies in the port and starboard, because the boundary distance in port and starboard is even greater than that in the fore and rear directions. It seems that the estimation method of longitudinal and lateral radii in the quaternion ship domain is not suitable for bridge-waters.

To evaluate further the applicability of the proposed ship domain, the ship domain based on ship–pier data is applied to the actual moving ship. In Figure 15, the scenario of target ship passing through the WYRB waters is illustrated. The piers are mainly located between the boundaries corresponding to  $r = 0.4$  and  $r = 0.2$  of the fuzzy ship domain when the ship passes through the bridge in the centre



**Figure 15.** Ship domain for ships passing through WYRB ((a). The first stage of target ship passing WYRB waters, the pier on the one side starts to enter the boundaries corresponding to  $r=0.2$ ; (b). The second stage of target ship passing WYRB waters, the pier on the other side starts to enter the boundaries corresponding to  $r=0.2$ ; (c). The third stage of target ship passing WYRB waters, the piers on the two side are located between the boundaries of  $r=0.2$  and  $r=0.4$ ; (d). The fourth stage of target ship passing WYRB waters, the piers on the two sides start to leave the ship domain boundaries).

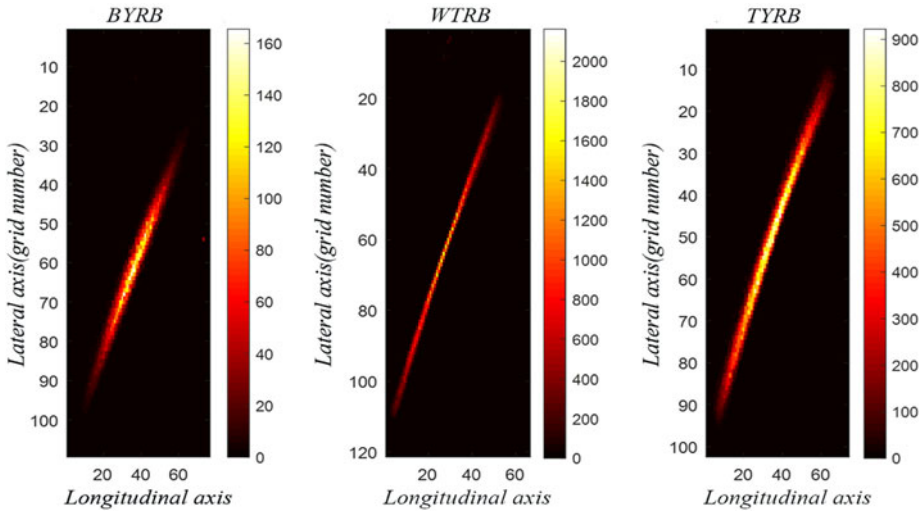
between the piers, which means that the collision risk is relatively small and verifies the effectiveness of the model.

### 3.3. Collision risk analysis in bridge area

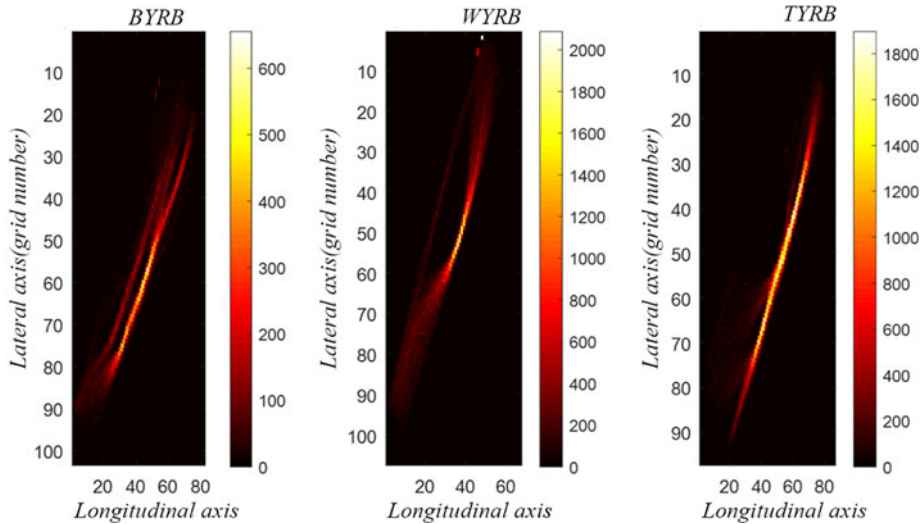
The concept of ship domain is broadly adopted for collision risk analysis in maritime traffic. Therefore, risk analysis for ships passing WYRB was conducted on the basis of the fuzzy quaternion ship domain model. In the beginning, the collision risk evaluation model was developed based on the fuzzy quaternion ship domain model mentioned above. Then, through the accumulation of regional collision risks, the risk analysis of ship navigation in the vicinity of bridge-waters was conducted.

In the present collision risk analysis, the aim is to measure and visualise the collision risk when ships navigating upstream and downstream pass through the bridge-waters. On the basis of the fuzzy quaternion ship domain based on bridge piers, analysis of the collision risk of ships passing the bridge is carried out. Figures 16–18 show the results of visualisation of collision risk when a number of ships pass through different bridges in turn. The colour bar shows the superimposed quantitative result of collision risk based on Equation (20) when ships passed the bridge.

It can be seen from Figures 16–18 that there are major differences in collision risk when ships pass through three particular bridges: WYRB, TYRB and BYRB. When a ship passes the WYRB, the risk zone is longer and thinner, and the maximum collision risk is also significantly higher than at the others. These patterns can be explained by the diameters of the navigation areas at WYRB, TYRB and BYRB, which are 128 m, 504 m and 618 m, respectively. As a result, affected by the bridge-waters' navigation conditions, when a ship passes the narrower channel between piles of WYRB, the collision risk is higher than that of TYRB and BYRB.



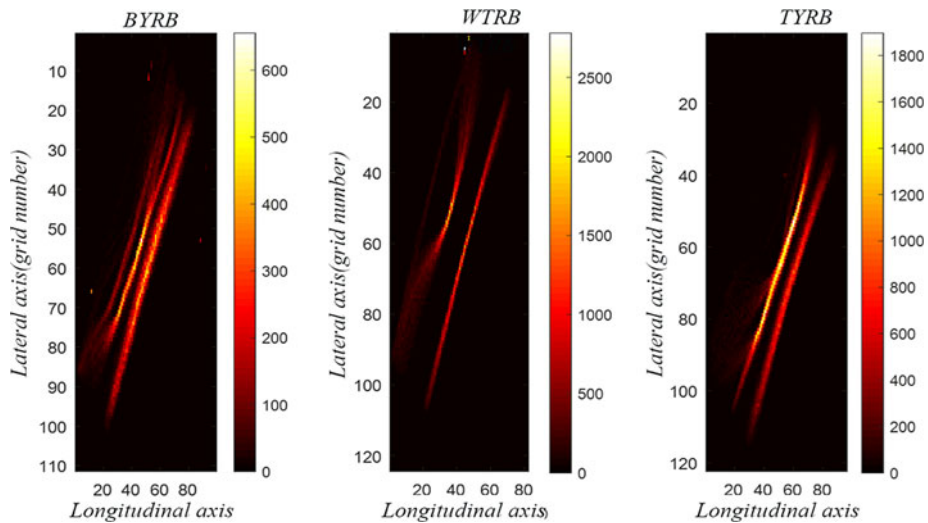
**Figure 16.** Collision risk heatmap of downstream ships passing through the bridge.



**Figure 17.** Collision risk heatmap of upstream ships passing through the bridge.

**4. Conclusions**

The ship domain plays an important role in the assessment of collision risk. However, few research studies have considered the practical considerations of implementing this method with regard to collision risk near bridge-waters. In order to solve this problem, this paper has investigated the characteristics of the inland ship domain model in multi-bridge waterways considering ship–ship collision and ship–pier collision based on historical AIS data. Combined with the ship domain model, a grid ship collision index (GSCI) has been proposed, which realises the regional risk grid calculation and realises the visualisation and analysis of the collision risk near bridge-waters. Taking the multi-bridge waterways in the Wuhan waterway of Yangtze River as an example, the conclusions have been drawn as follows. Firstly, the ship domain based on navigation marks is not obvious, while the domains of neighbouring ships and bridge piers obviously exist. Secondly, the Pearson correlation coefficients of the longitudinal and lateral radii and ship’s speed and ship’s length are calculated after the ship domain boundary is determined by the nearest boundary point method, which indicates that there is a strong relationship between them. The



**Figure 18.** Collision risk heatmap of upstream and downstream ships.

models were also compared with others and it was found that the models are relatively small but move suitably for bridge-waters. Thirdly, the collision risk of ships passing through different bridges is quite different. The risk zone will be thinner and longer when the navigation area of the bridge is smaller.

Through the ship domain models constructed by the proposed method and the results of analysis of collision risk, this paper has demonstrated the application of ship domain in quantifying collision risk. In addition, a visual analysis of the collision risk near bridge-waters has been performed, which helps to characterise the collision risk distribution, provide guidance for maritime supervision and improve the safety of ships passing through bridge-waters. The accuracy of the model could be affected by the amount of sample data and the data for specific ship length and ship speed not being enough to obtain the corresponding boundaries of ship domain reliably. The ship domain may also be related to the ship type, which is not involved in this work due to the consideration of previous studies and the incompleteness of static data in AIS data. Therefore, the fuzzy quaternion ship domain would be further developed on the basis of more comprehensive consideration. On the other hand, future research can focus on the early warning of collision accident of ships passing the bridge-waters according to the distribution of collision risk in the area and the real-time status of ship navigation.

**Acknowledgement.** This work is supported by the Fujian Province Natural Science Foundation (No. 2018J01506, 2020J01860), and University-Industry Cooperation Programme of the Department of Science and Technology of Fujian Province (No.2019H6018) and The Key Project of Science and Technology of Wuhan (No. 201701021010132).

## References

- Azevedo, C. L., Ciuffo, B., Cardoso, J. L. and Ben-Akiva, M. E. (2015). Dealing with uncertainty in detailed calibration of traffic simulation models for safety assessment. *Transportation Research Part C: Emerging Technologies*, **58**, 395–412.
- Chen, H. Z. and Guo, G. P. (2008). Research of ship domain and the traffic capacity in paratactic bridge water area. *Ship & Ocean Engineering*, **37**, 113–116.
- Chen, Z., Chen, D., Zhang, Y., Cheng, X., Zhang, M. and Wu, C. (2020). Deep learning for autonomous ship-oriented small ship detection. *Safety Science*, **130**, 104812.
- Coldwell, T. (1983). Marine traffic behavior in restricted waters. *The Journal of Navigation*, **36**, 430–444.
- Fan, X. H., Zhang, Q. N., Zhou, F., Tan, Z. R. and Wang, M. Q. (2013). Model of ship domain in river water. *Journal of Dalian Maritime University*, **39**, 46–48.
- Felski, A. and Jaskólski, K. (2013). The integrity of information received by means of AIS during anti-collision manoeuvring. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, **7** (1).
- Fujii, Y. and Tanaka, K. (1971). Traffic capacity. *The Journal of Navigation*, **24**, 543–552.
- ITU. (2010). *Recommendation M.1371-4, Technical characteristics for an automatic identification system using time-division multiple access in the VHF maritime mobile band*. Geneva: International Telecommunication Union.

- Goodwin, E. M. (1975). A statistical study of ship domains. *The Journal of Navigation*, **28**, 328–344.
- Jia, C. Y. (1989). Ship domain in congested water area. *Journal of Dalian Marine College*, **15**, 15–19.
- Liu, S., Wang, N. and Wu, Z. (2011). Review of research on ship domain. *Journal of Dalian Maritime University*, **37**, 51–54.
- Montewka, J., Hinz, T., Kujala, P. and Matusiak, J. (2010). Probability modelling of vessel collisions. *Reliability Engineering & System Safety*, **95**, 573–589.
- Pietrzykowski, Z. (2008). Ship's fuzzy domain – a criterion for navigational safety in narrow fairways. *The Journal of Navigation*, **61**, 499–514.
- Pietrzykowski, Z. and Magaj, J. (2017). Ship domain as a safety criterion in a precautionary area of traffic separation scheme. *Transnav International Journal on Marine Navigation & Safety of Sea Transportation*, **11**(1), 93–98.
- Pietrzykowski, Z. and Uriasz, J. (2004). The Ship Domain in a Deep-Sea Area. *Proceeding of the 3rd International Conference on Computer and IT Applications in the Maritime Industries*, Siguenza, Spain.
- Pietrzykowski, Z. and Uriasz, J. (2006). Ship Domain in Navigational Situation Assessment in an Open Sea Area. *Proceeding of the 5th International Conference on Computer and IT Applications in the Maritime Industries*, Oegstgeest, Netherlands, pp. 135–141.
- Pietrzykowski, Z. and Uriasz, J. (2009). The ship domain – a criterion of navigational safety assessment in an open sea area. *The Journal of Navigation*, **62**, 93.
- Rawson, A. and Brito, M. (2020). A critique of the use of domain analysis for spatial collision risk assessment. *Ocean Engineering*, **219**, 108259.
- Rawson, A. and Brito, M. (2021). Developing contextually aware ship domains using machine learning. *Journal of Navigation*, **74**(3), 515–532.
- Szlapczynski, R. and Szlapczynska, J. (2017). Review of ship safety domains: Models and applications. *Ocean Engineering*, **145C**, 277–289.
- Wang, N. (2010). An intelligent spatial collision risk based on the quaternion ship domain. *Journal of Navigation*, **63**, 733–749.
- Wang, Y., Zhang, J., Chen, X., Chu, X. and Yan, X. (2013). A spatial–temporal forensic analysis for inland-water ship collisions using AIS data. *Safety Science*, **57**, 187–202.
- Wang, T. F., Yan, X. P., Wang, Y. and Wu, Q. (2017). Ship domain model for multi-ship collision avoidance decision-making with COLREGs based on artificial potential field. *Transnav. International Journal on Marine Navigation & Safety of Sea Transportation*, **11**(1), 85–92.
- Wen, Y., Li, T., Zheng, H., Huang, L., Zhou, C. and Xiao, C. (2018). Characteristics of ship domain in typical inland waters. *Navigating of China*, **41**, 43–47.
- Weng, J., Meng, Q. and Qu, X. (2012). Vessel collision frequency estimation in the Singapore Strait. *The Journal of Navigation*, **65**, 207–221.
- Willems, N., Scheepens, R., van de Wetering, H. and van Wijk, J. J. (2013). Visualization of vessel traffic. In van de Laar, P., Tretmans, J., and Borth, M. (eds.). *Situation Awareness with Systems of Systems*, New York: Springer, 73–87.
- Xiang, Z., Hu, Q., Shi, Z. and Yang, C. (2015). Computation method of ship domains in restricted waters based on AIS data. *Journal of Traffic and Transportation Engineering*, **15**, 110–117.
- Xiao, F., Ligteringen, H., Gulijk, C. V. and Ale, B. (2015). Comparison study on AIS data of ship traffic behavior. *Ocean Engineering*, **95**(feb.1), 84–93.
- Zhang, L. and Meng, Q. (2019). Probabilistic ship domain with applications to ship collision risk assessment. *Ocean Engineering*, **186**, 106130.
- Zhang, W., Goerlandt, F., Montewka, J. and Kujala, P. (2015). A method for detecting possible near miss ship collisions from AIS data. *Ocean Engineering*, **107**, 60–69.
- Zhang, M., Zhang, D., Fu, S., Yan, X. and Goncharov, V. (2017). Safety distance modeling for ship escort operations in Arctic ice-covered waters. *Ocean Engineering*, **146**, 202–216.
- Zhang, L., Meng, Q., Xiao, Z. and Fu, X. (2018). A novel ship trajectory reconstruction approach using AIS data. *Ocean Engineering*, **159**, 165–174.
- Zhang, M., Zhang, D., Yao, H. and Zhang, K. (2020). A probabilistic model of human error assessment for autonomous cargo ships focusing on human–autonomy collaboration. *Safety Science*, **130**, 104838.
- Zhang, M., Montewka, J., Manderbacka, T., Kujala, P. and Hirdaris, S. (2021). A big data analytics method for evaluation of ship–ship collision risk reflecting hydrometeorological conditions. *Reliability Engineering & System Safety*, **213**, 107674.
- Zhu, X., Xu, H. and Lin, J. (2001). Domain and its model based on neural networks. *The Journal of Navigation*, **54**, 97.