ISSN: 2710-0170

Research on spatiotemporal distribution characteristics of ship traffic flow in port waters based on AIS data

Tian Liu [1, a], Jianwen Ma[2, 3, b] and Yue Zhou[1, c]

[1]School of Shipping College, Shandong Jiaotong University, Weihai 264200, China;

[2]School of International Business, Shandong Jiaotong University, Weihai 264200, China.

[3]School of Merchant Marine, shanghai Maritime University, Shanghai 201200, China.

Abstract

With the gradual increase of ship traffic, water traffic management has become increasingly important. This paper is based on AIS data and uses the Analytic Hierarchy Process to comprehensively evaluate the relative hazard level of a specific area, in order to determine the relative consequences of regional hazards (RCORH). Next, we will use speed and traffic volume characteristics as indicators to construct a linear regression model for spatial relationship analysis. The research results indicate that the linear regression model proposed in this article can effectively identify areas with high collision risks, which has positive significance for improving the navigation safety management of port waters.

Keywords

Traffic flow, danger level, relative consequences, linear regression model.

1. Introduction

With the continuous increase in international trade and maritime transportation, the rapid expansion of the number of ships, and the increase in traffic density in port waters, water traffic safety issues have attracted widespread attention. As a special node in trade development, the navigable water area of ports is a crucial passage for waterway transportation. Various factors have led to the continuous compression of port navigable space and a decrease in ship safety performance. Therefore, the study of the characteristics of ship traffic flow and historical accidents in port waters is of great practical significance.

With the continuous development of AIS technology, more and more scholars are starting to apply AIS data to study ship traffic flow and maritime traffic characteristics [1-4]. Chen Renli, Li Zhenfu, and others based on AIS data and used GIS spatial analysis methods to analyze the time characteristics of ships in specific areas, constructing a new channel for ship traffic flow research [5-8]. On this basis, Hongchu Yu pointed out that there is a linear relationship between ship traffic flow and accident distribution To study the risk of ship accidents, scholars such as Shao Chengpu, Zihao Liu, Weibin Zhang et al. [10-12] constructed a ship collision risk assessment model and reconstructed the AIS database and conducted deep data mining. Fuzzy mathematics methods and Analytic Hierarchy Process were introduced to achieve scientific evaluation and prediction of ship collision risk. In order to effectively prevent traffic accidents, scholars such as Jiang Yujie, Wan Chengpeng, Li Yaling, and Hwang Taemin have studied ship trajectory prediction and the causes of frequent accidents based on the characteristics of maritime traffic, providing possibilities for monitoring and analyzing collision risks in water areas [13-16]. Other scholars have explored the risk evolution mechanism of water traffic accidents, identifying the key causal factors of water traffic accidents from the stages of accident incubation, diffusion, and occurrence, and conducting case studies [17-21].

ISSN: 2710-0170

At present, there is a lack of comprehensive statistical data on the spatial distribution and detailed consequences of accidents both domestically and internationally. In summary, this article comprehensively studies the relationship between accidents and influencing factors, further evaluates the risk level and relative consequences of specific areas, reveals the activity patterns of ship traffic flow, helps clarify the distribution of navigation risks, and provides theoretical support for the safety of navigation in Qingdao waters.

2. Data sources and research methods

2.1. Data sources

2.1.1. Sub-section Headings

AIS data includes ship static and dynamic information such as MMSI, ship size, speed, heading, and ship position. This article uses Python analysis and statistical techniques to construct an AIS database within the longitude range of 119.3 ° E to 121.3 ° E and the latitude range of 35.35 ° N to 37.09 ° N. The time range is from January 1, 2016 to November 1, 2022, focusing on general cargo ships, container ships, bulk carriers, and fishing vessels, without considering other ship types, such as government vessels, for the time being. Due to the large size of AIS data, sampling analysis was conducted every 4 seconds, resulting in 8160926 pieces of ship data information.

2.2. Ship traffic density and flow calculation

Do Ship density refers to the number of ships per unit area in a certain water area at a certain time, which to some extent reflects the busyness and density of ships in that water area. The calculation formula is:

$$\rho = Q/h. \tag{1}$$

Ship flow refers to the number of ships passing through a certain water area and location within a unit time, and its value can reflect the traffic flow and busyness of the water area. The larger the ship flow, the more complex the ship traffic flow. Therefore, ship flow is an important indicator of maritime navigation safety and ship scheduling. The statistical model is:

$$\bar{Q} = \sum_{i=1}^{n} Q_i/n. \tag{2}$$

In the formula: $\bar{Q}$The average traffic volume of ships during a certain period of time; Q_i is the traffic volume of i at a certain moment; n represents time.

2.3. Analysis of point density in accident areas

The principle of density analysis is to calculate the number of data points per unit area range. Neighborhood method is one of the most widely used density calculation methods in GIS software. Neighborhood is defined as a circle, rectangle, or other shape surrounding the center of each grid. Add the number of points in the neighborhood and divide by the area of the neighborhood to obtain the density of point elements. This study divides Qingdao Port into several square cells with a side length of 1. Let the neighborhood radius be ρ， $N_k(\rho)$ Represents a circle centered around the k-th cell and with a neighborhood radius of ρ The number of accidents occurring within the circular range of. The accident density of the kth cell can be calculated using the formula:

$$Q_k = N_k(\rho)/(\Pi\rho^2). \tag{3}$$

Calculate the density of each cell to obtain the distribution of accident density. Considering the impact of ship traffic density, the density of maritime accidents per unit area of the sea area over a period of time cannot fully reflect the frequency of accidents occurring in that area. Therefore, it is necessary to calculate the relative accident density P_ k. It is defined as the accident density $Q_k$and ship traffic density Q_k The ratio of k, relative accident density P_k The calculation formula for k is shown in the formula:

ISSN: 2710-0170

$$P_k = Q_k / H_k. \tag{4}$$

2.4. Accident Hazard and Relative Consequence

The The core of accident hazard level is risk assessment. The indicators of accident hazard level in this article are based on the regulations of the Maritime Safety Administration, using three elements: casualties, economic losses, and oil spills for research.The accident hazard Q can be expressed as:

$$Q = W1 * X1 + W2 * X2 + W3 * X3. \tag{5}$$

The degree of impact of X1 event: This refers to the situation where a certain event occurs and causes casualties;

Severity of X2 risk: This refers to the serious economic losses caused by a certain risk to people;

X3 represents the degree of environmental pollution, which refers to the severity of a certain risk to the environment.

The relative consequences of water accidents in hazardous areas refer to the degree and impact of various potential consequences that may arise from water accidents occurring in specific water areas. This article mainly conducts research from three aspects: casualties, economic losses, and oil spills. The corresponding calculation formulas are as follows:

$$R = \sum_{i=1}^{m} \frac{Q_i}{\bar{Q}}. \tag{6}$$

Where "R" represents relative consequences, "Q"⁻ "Represents the relative danger level of a single accident, and" m "represents the frequency of accidents occurring in the grid.".

2.5. Model establishment

Model establishment

Using ArcGIS software, all designated areas were displayed as grids, with geographic grids as the basic unit. The statistical ship density and traffic flow of each grid were further analyzed, where ship traffic characteristics and accidents were synchronized based on geographical location. The ship traffic data is overlaid with the RCORH of each grid on ArcGIS to study the correlation or causal relationship between driving conditions and accident consequences. Finally, a regression model was established to link the relative consequences with the parameters of ship traffic.

3. Research area

Qingdao Port is located in Jiaozhou Bay on the south coast of the Shandong Peninsula, adjacent to the Yellow Sea. It consists of five major port areas: Dagang Port Area, Huangdao Oil Port Area, Qianwan Port Area, Dongjiakou Port Area, and Weihai Port Area. As an international trade port and transit hub along the Yellow River Basin and the western coast of the Pacific Rim in China, Qingdao Port has a high density of ships, dense port terminals, complex navigation conditions in some waterways, and a high probability of collision accidents, which poses hidden dangers to ship navigation safety, see Fig. 1.

According to the route map 1 of Qingdao Port, it can be seen that there are basically seven directions for ships entering Qingdao Port, namely: 1. Qingdao Port Main Channel 2. First Route 3. First Preparatory Route 4. Second Route (foreign vessels are only allowed to use the second and third routes) 5. Third Route (foreign vessels are only allowed to use the second and third routes) 6. Fourth Route 7. Laoshan Tourism Route, see Fig. 2.

ISSN: 2710-0170

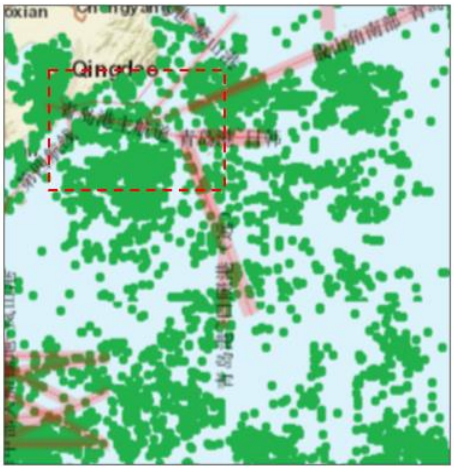

Fig. 1 Research area

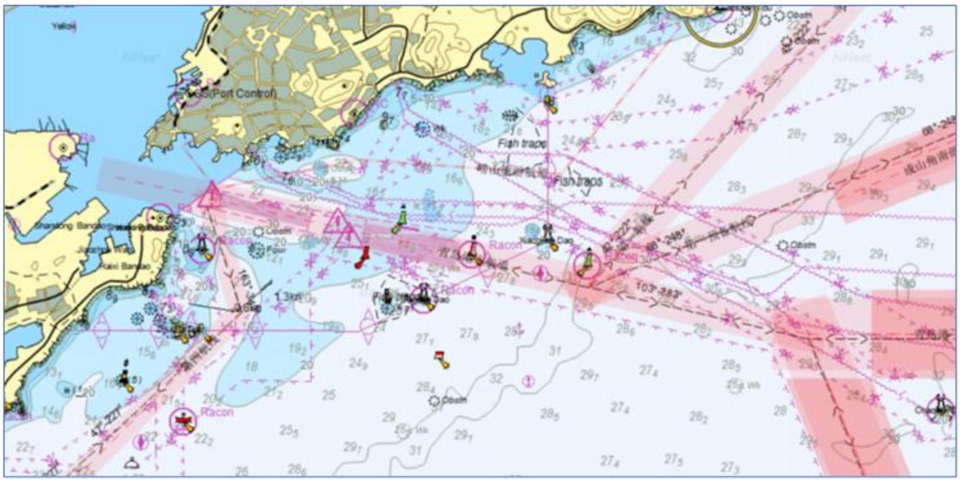

Fig. 2 Qingdao Port Route Map

4. Analysis of Water Traffic Accidents in Qingdao Port

The accident data in this paper is sourced from Shandong Maritime Safety Administration and covers the accident records from January 1, 2016 to January 1, 2022. This dataset contains 42 water traffic accidents, for which we conducted statistical analysis (data sourced from the official authoritative website of Shandong Maritime Safety Administration) http://www.sd.msa.gov.cn/ ）.

Water traffic accidents can be divided into the following types, including collision accidents, stranding accidents, reef accidents, collision accidents, wave damage accidents, fire and explosion accidents, wind accidents, and self sinking accidents. The types and proportions of various water traffic accidents that occurred between 2016 and 2022 are shown in the figure: collision accidents accounted for as high as 45.6%; Stranding accidents account for 13.1%; Fire and explosion account for 8.6%; Wind disasters account for 5.9% of accidents; Self sinking accounts for 2.1%; Reef accidents account for 1.6%; Wave damage accidents account for 6.3%, see Fig. 3.

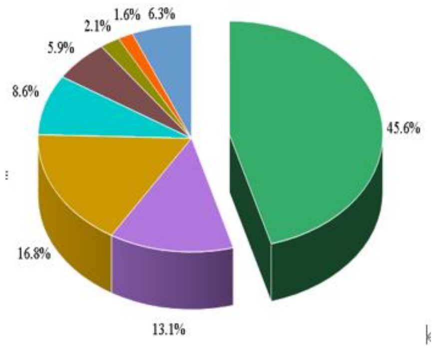

Fig. 3 Statistical chart of types of ship traffic accidents

A statistical analysis was conducted on 42 accidents in the waters of Qingdao Port according to time periods, and each time period showed different patterns and characteristics of accidents. Due to the frequent occurrence of sea fog weather from the latter half of the night to the morning, water traffic accidents mainly occur between 0am and 6am, accounting for up to 43%. Next is from 6am to 12pm, accounting for 29%. It can be seen that accidents are more likely to occur in the middle of the night. As shown in the figure, see Fig. 4.

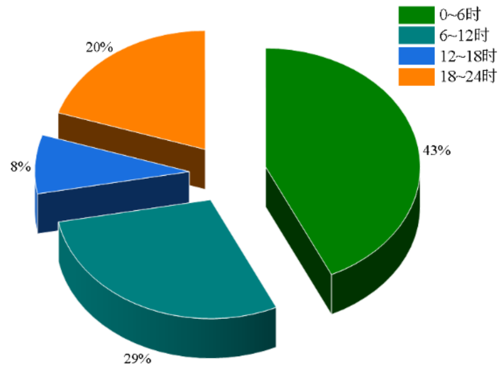

Fig. 4 Statistical chart of accident occurrence time period

According to the nuclear density analysis of 42 accidents in the waters of Qingdao Port by month, it was found that water traffic accidents are mainly concentrated in April May and September December. Among them, April to May is the peak period of sea fog weather, and it is also a busy period for fishing operations by offshore fishing vessels. A large number of fishing vessels are concentrated in offshore operations, making it easy for major and particularly serious accidents to occur, especially in the case of fishing vessel accidents and collisions between commercial and fishing vessels. With the start of the fishing ban in June, the number of maritime hazards remained at a relatively low level. However, with the end of the fishing ban in September, a large number of fishing boats returned to sea, and from October to December, they were affected by winter strong winds, leading to a high incidence of fishing related risks.

5. Model construction

This article uses AIS data to statistically analyze the traffic situation in the region, and establishes a 100 km/h model in the water area and its surrounding areas × 100, with a side length of 1n miles × A standardized statistical grid of 1n miles, with a total of 153 grids, see Fig. 5.

ISSN: 2710-0170

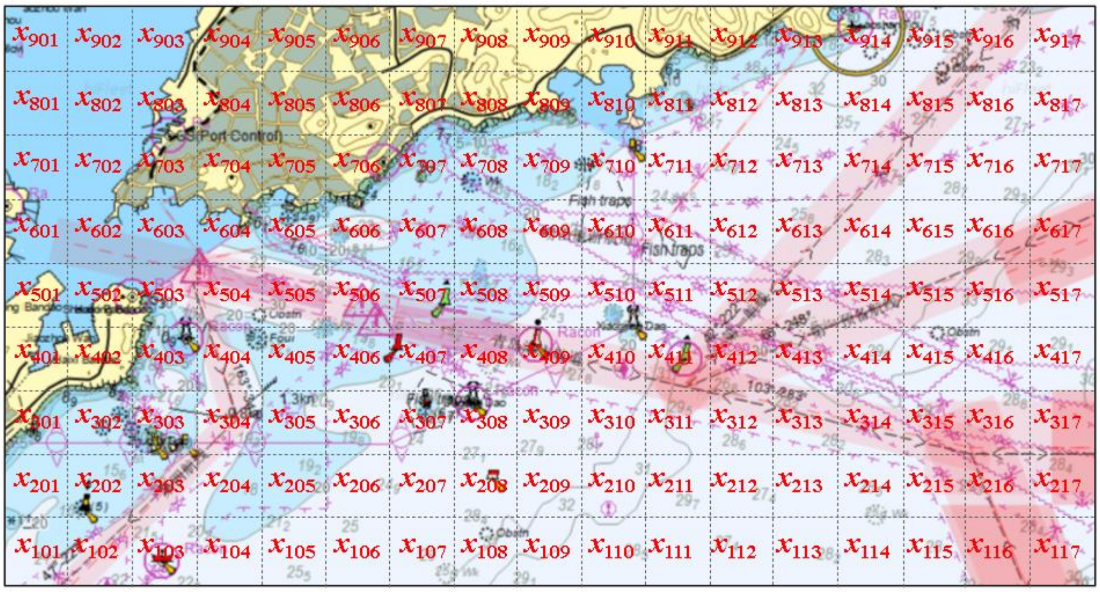

Fig. 5 Mesh

This article categorizes water traffic accidents into the following levels: particularly serious accidents, major accidents, major accidents, and general accidents (represented by numbers 4, 3, 2, and 1). According to the criteria for accident classification, 10 serious injuries are usually equivalent to 3 deaths. Therefore, it is converted to 0.3 deaths (not integers), and so on, see Table 1.

Table 1 Accident level

Accident level	4	3	2
death toll	>3	3-1	1-0.3
Economic losses	>3	3-1	1-0.2
Oil spill	>1	1-0.5	0.5-0.1

Using Analytic Hierarchy Process (AHP) to analyze accident variables, including oil spills, fatalities, and economic losses, in order to select the best solution and determine the various factors to be evaluated and their hierarchical structure. For each pair of sub factors, a judgment matrix was constructed, where 9, 7, 5, 3, and 1 respectively represent the degree of decreasing importance from extremely important to equally important, see Table 2.

Normalize the matrix by column and calculate the weights using the arithmetic mean method.

Wi is the weight of factor i;

λ I is the eigenvalue of the eigenvalue matrix, representing the relative importance of factor i;

N is the total number of factors.

Table 2 Judgment matrix summary results

	death toll	Economic losses	Oil spill	Eigenvector	Weight value (%
death toll	1	7	9	3.979	77.842
Economic losses	0.143	1	4	0.83	16.234
Oil spill	0.111	0.25	1	0.303	5.925

After calculation, W1=0 77842, W2=0 16234, W3=0.05925.

The table presents the weight calculation results of the Analytic Hierarchy Process, which have been validated through consistency testing.

ISSN: 2710-0170

5.1. Calculation of hazard values.

By calculating the distribution map of ship accident risk values in the waters of Qingdao Port, it can be observed that the distribution of ship accidents that have occurred in the past 5 years in the waters of Qingdao Port is basically consistent with the grid distribution of high ship collision risk calculated. In addition, water traffic accidents were numbered according to time and their accident hazards were calculated, and then the hazards were matched with the risk level. the basic agreement between accident hazard level and risk level verifies the feasibility of the model. This result emphasizes the effectiveness of the model and strengthens its potential application value in evaluating accident risks and managing water traffic safety, see Fig. 6, see Table 3.

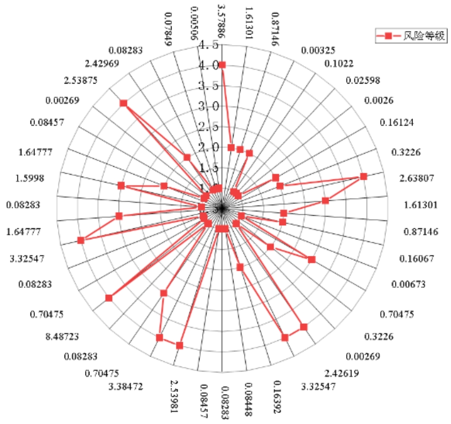

Fig. 6 Risk level

Table 3 Accident risk

Accident Number	Accident Hazard Level
1	3.578855
2	1.613014
3	0.871455
4	0.003247
…	…
42	0.005065

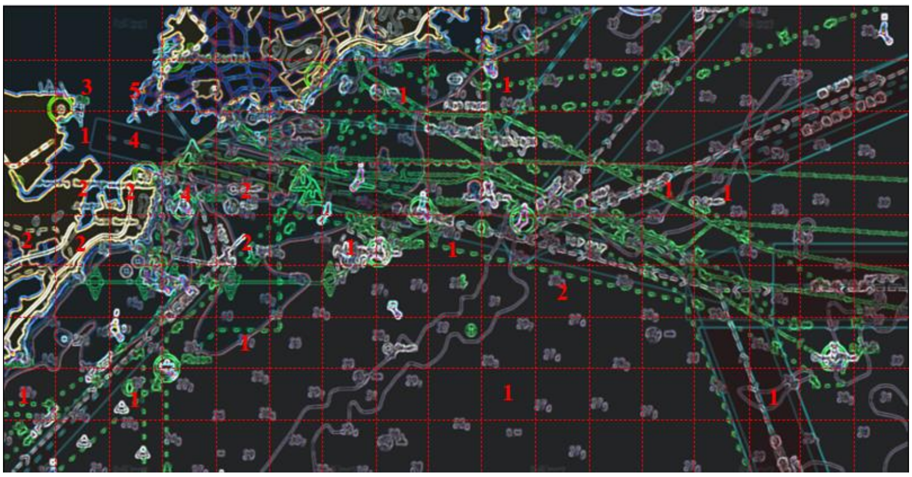

Fig. 7 Spatial distribution map of the number of ship accidents

ISSN: 2710-0170

By comparing the number of water traffic accidents with the grid distribution of ship collision risks calculated in the past 5 years in the waters of Qingdao Port, it can be observed that they are basically consistent. This comparison confirms the accuracy of the model, indicating a significant correlation between the distribution of high-risk areas in the waters of Qingdao Port and the actual occurrence of ship accidents, see Fig. 7.

Through a comprehensive assessment of the relative danger level of accidents occurring in the northern waters of Qingdao Port, we have determined the relative consequences of regional hazards (RCORH), see Table 4.

Table 4 Relative consequences of the accident

Accident number	RCORH
1	0.89471375
2	0.806507
3	0.4357275
4	0.0016237
...	...
42	0.005065

5.2. Regression model.

The final model of this study integrates both traffic and accident data. In order to deeply explore the impact of different risk factors on accidents, we selected two key indicators, namely ship speed and ship traffic flow, and fitted the relative consequences with the key indicators.

The fitting relationship between the relative consequences of hazardous areas and ship speed is shown in the following figure. We can observe a significant positive correlation between ship speed and the relative consequences of accidents. The formula of the model is expressed as follows: y=-0.272+0.118 * velocity. The R ^ 2 value of the fitting result is 0.153, indicating that the model can explain about 15.3% of the relative consequence changes, further confirming the existence of this positive correlation, see Fig. 8.

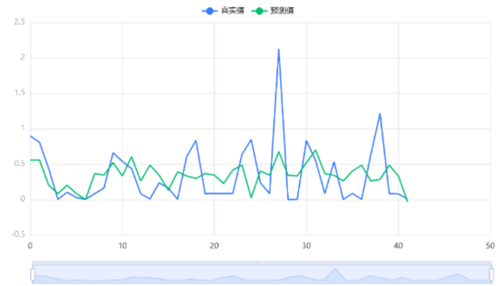

Fig. 8 Model diagram

The figure shows the fitting situation of the model, and the research results show that the overall fitting effect is good, with an adjusted R ^ 2 of 0.132. Meanwhile, the statistical value of the F-test is 7.239, with a significance level P-value of 5%. This result indicates that the independent variables of the model have statistical significance in explaining the dependent variable as a whole, supporting the effectiveness and credibility of the model.

The fitting relationship between the relative consequences of hazardous areas and ship traffic flow is shown in the following figure. We can observe a clear positive correlation between traffic flow and the relative consequences of accidents, which means that as traffic flow increases, the relative consequences of accidents also show an upward trend. The specific fitting equation is y=-0.308+0.108 * traffic volume, and the R ^ 2 value of the fitting result is 0.843, indicating that the model can explain about 84.3% of the relative consequence changes, further confirming the significance of this positive correlation, see Fig. 9.

ISSN: 2710-0170

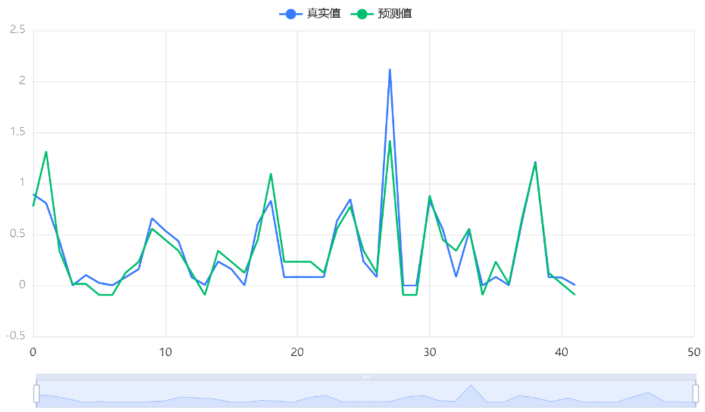

Fig. 9 Model diagram

The diagram clearly shows the fitting situation of the model, and the research results show that the overall fitting effect is very good. The adjusted R ^ 2 value of the model reached 0.839, which means that the model can explain most of the variance in the dependent variable. In addition, the statistical value of the F-test is 213.985, with a significance level P value of 10%. This result indicates that the independent variables of the model have a high degree of statistical significance in explaining the dependent variable as a whole, supporting the effectiveness and credibility of the model.

This result indicates a significant correlation between the independent variable and the dependent variable, rejecting the null hypothesis that the regression coefficient is zero. Therefore, it can be considered that the model is statistically reasonable. By comparing the fitting graphs of ship speed and ship traffic flow, it can be concluded that traffic flow has a dominant impact on RCORH. The R ^ 2 value of the fitting result is 0.843, which means that an increase in ship traffic flow will significantly increase the likelihood of accidents.

6. Conclusion

This article takes water traffic accidents as the research object, focusing on solving the statistical problem of accidents mentioned above, and delving into the characteristics of water traffic accidents. The core objective of this study is to explore in depth the correlation between ship traffic characteristics and accidents. To achieve this goal, we conducted a detailed analysis by combining historical accident data with AIS data of actual ship traffic. The results showed that our constructed model can highly fit these data, providing a powerful tool for in-depth research on the causes of ship accidents.

Specifically, we adopted a linear regression model in our study and divided the study area into multiple grids to achieve a research perspective from local to global. This enables us to have a clearer understanding of the ship collision risk values in different areas of the research object's water area, which helps to more accurately assess accident risks.The specific work of the study includes:

1) Based on AIS data, we analyzed the spatial distribution characteristics of water transportation vessels, identified accident hotspots through this analysis, and thus revealed areas with high accident incidence.

2) We then extracted information about the consequences of the accident from the accident report, taking into account factors such as the type of accident, time period, and type of consequences. Our model evaluates the relative consequences of accident areas using indicators such as fatalities, economic losses, and oil spills to better understand the relationship between different risk factors and collision risk. Two regression models were proposed, taking into account ship speed and traffic volume. Linear regression analysis was used to examine the factors influencing collision risk more comprehensively.

ISSN: 2710-0170

In summary, by combining historical accident data and real-time AIS data, as well as using advanced statistical models, this study aims to deeply explore the relationship between ship traffic characteristics and accidents, providing key insights and methods for improving navigation safety and traffic management. These research findings will provide valuable references and guidance for practitioners in academia and related fields.

Acknowledgements

Natural Science Foundation.

References

[1] Du L, Goerlandt F, Kujala P. Review and analysis of methods for assessing maritime waterway risk based on non-accident critical events detected from AIS data[J]. Reliability Engineering & System Safety, 2020, 200: 106933.

[2] Zhang L, Meng Q, Fwa T F. Big AIS data based spatial-temporal analyses of ship traffic in Singapore port waters[J]. Transportation Research Part E: Logistics and Transportation Review, 2019, 129: 287-304.

[3] Huang L, Wen Y, Geng X, et al. Estimation and spatio-temporal analysis of ship exhaust emission in a port area[J]. Ocean Engineering, 2017, 140: 401-411.

[4] Yu H, Fang Z, Lu F, et al. Impact of oil price fluctuations on tanker maritime network structure and traffic flow changes[J]. Applied Energy, 2019, 237: 390-403.

[5] Liu Z, Wu Z, Zheng Z. A novel framework for regional collision risk identification based on AIS data[J]. Applied Ocean Research, 2019, 89: 261-272.

[6] Zhang W, Feng X, Qi Y, et al. Towards a model of regional vessel near-miss collision risk assessment for open waters based on AIS data[J]. The Journal of Navigation, 2019, 72(6): 1449-1468.

[7] Li Y, Cheng Z, Yip T L, et al. Use of HFACS and Bayesian network for human and organizational factors analysis of ship collision accidents in the Yangtze River[J]. Maritime Policy & Management, 2022, 49(8): 1169-1183.

[8] J. Liu, E.L. Chen and Z.T. He: Journal of Shi Jia Zhuang Railway Institute (Natural Science), Vol. 22 (2009) No. 4, p.40-42.

[9] Hwang T, Youn I H. Latent-Cause Extraction Model in Maritime Collision Accidents Using Text Analytics on Korean Maritime Accident Verdicts[J]. Applied Sciences, 2022, 12(2): 914.

[10] Rentzeperis I, Van Leeuwen C. Adaptive rewiring in weighted networks shows specificity, robustness, and flexibility[J]. Frontiers in Systems Neuroscience, 2021, 15: 580569.

[11] Rentzeperis I, Van Leeuwen C. Adaptive rewiring in weighted networks shows specificity, robustness, and flexibility[J]. Frontiers in Systems Neuroscience, 2021, 15: 580569.

[12] Johansen T, Blindheim S, Torben T R, et al. Development and testing of a risk-based control system for autonomous ships[J]. Reliability Engineering & System Safety, 2023, 234: 109195.

[13] Luo X, He H, Zhang X, et al. Failure Mode Analysis of Intelligent Ship Positioning System Considering Correlations Based on Fixed-Weight FMECA[J]. Processes, 2022, 10(12): 2677.

[14] Fan C, Montewka J, Zhang D. A risk comparison framework for autonomous ships navigation[J]. Reliability Engineering & System Safety, 2022, 226: 108709.

[15] Hoem Å S, Veitch E, Vasstein K. Human-centred risk assessment for a land-based control interface for an autonomous vessel[J]. WMU Journal of Maritime Affairs, 2022, 21(2): 179-211.

[16] Liu C, Chu X, Wu W, et al. Human–machine cooperation research for navigation of maritime autonomous surface ships: A review and consideration[J]. Ocean Engineering, 2022, 246: 110555.

[17] Bao J, Yu Z, Li Y, et al. A novel approach to risk analysis of automooring operations on autonomous vessels[J]. Maritime Transport Research, 2022, 3: 100050.

[18] Yang X, Utne I B, Sandøy S S, et al. A systems-theoretic approach to hazard identification of marine systems with dynamic autonomy[J]. Ocean Engineering, 2020, 217: 107930.

ISSN: 2710-0170

[19] Fan C, Wróbel K, Montewka J, et al. A framework to identify factors influencing navigational risk for Maritime Autonomous Surface Ships[J]. Ocean Engineering, 2020, 202: 107188.

[20] Ramos M A, Thieme C A, Utne I B, et al. Human-system concurrent task analysis for maritime autonomous surface ship operation and safety[J]. Reliability Engineering & System Safety, 2020, 195: 106697.