# Cause Analysis of Ship Accidents Based on Chi-Square Test

Xintong Liu, Qing Yu, Yutian Hu, Wenting Lu, Aipan Zhanga, Xujiang Bo*

* Corresponding author: xujianbo_0@hotmail.com

School of navigation, Jimei University, Xiamen, Fujian, China, 361000

**Abstract:** Studying historical accidents is of paramount significance to safeguard maritime transport. This study investigates the causation factors and their interactions from collected maritime accidents, which happened in the coastal water of China, to provide effective references for accident prevention. A chi-square test is used to assess the degree of association between two factors so that the potential causal relationships among major factors can be revealed. Then, the importance of each factor is ranked to identify the critical factors based on historical accident records. The research findings are expected to improve the understanding of the causation of the maritime accident while supporting maritime safety assessments and ship management.

**Keywords:** Chi-Square Test, Methods Use, Cause Analysis, Marine Accident.

## 1. INTRODUCTION

Nowadays, Waterborne transportation has been actively promoting the development of the national economy as an important mode of transportation, accounting for approximately 90% of the world's trade in volume [1]. However, catastrophic accidents are frequently noted during the past years with more and more challenging safety issues. For instance, in the accident of Suwatai 999, the ship sunk and brought catastrophic economic loss [2]. Maritime accidents are commonly understood as the results caused by complex technical, human, organizational, and environmental factors [3]. Thereby, each severe maritime accident shall be analyzed to take experience from the accidents and to form a basis for the conventions and contracts produced for the prevention [4]. However, previous studies only discussed the statistical characteristics of the accidents [5], without a deeper study of the association between each variable. To study the causes of maritime accidents, this study applies a chi-square test to analyze the historical accident records from the coastal water of China. The chi-square test is a versatile hypothesis testing method. The relationship between the two variables will be analyzed by the chi-square test to detect the importance of each causation factor.

The remainder of this paper is structured as follows. The methodology is presented in Section 2; Section 3 provides a case study and the conclusion is drawn in Section 4.

BDEDM 2023, January 06-08, Changsha, People's Republic of China

DOI 10.4108/eai.6-1-2023.2330244

## 2. METHODOLOGY

The chi-square test is used to analyze the interactions among all independent factors and the fitted distributions, which outputs the goodness of fit of a sample. The typical methods available for the chi-square tests include the Karl Pearson family, the Yates chi-square test, the Mantel–Haenszel chi-square, and the Maxwell–Stuart tests. One commonality of the above-mentioned methods is they use chi-square distribution as the reference distribution, which means that the chi-square test compares an observed set of data to what is expected [6].

Three parameters are used in the Karl Pearson family of chi-square tests: Goodness of fit, independence, and homogeneity. Each of these three tests has different interpretations and assumptions. The parameters can be calculated by using the following function:

$$\mathrm{x}^2 = \sum_{i=1}^{n} \frac{(O_I - E_I)^2}{E_i} \tag{1}$$

where O is the target factor and E is the compared factor, n is the number of cells in the table, in which $i \in n$. The difference and usages among the three chi-square tests relate to the applicable situations and specific problems. The chi-square goodness of fit test is used to compare a sample to a population with known parameters on the variable of interest, or, to test the independence and to detect the correlation between two factors. In addition, homogeneity is used to detect distributional variability in two or more independent samples on individual variables of interest. Its common usage is to compare two or more groups on classification results.

## 3. CASE STUDY

A total of 9 causation factors are selected and compared to study the interactions and their importance among them. They are ship type, accident type, day or night, visibility, extreme weather conditions, seasons, weather, ship density, and human factors. The accident records are collected from public resources, such as MSA and Marine Bureau's official website. Before processing the chi-square test, the obtained raw data need to be cleared by deleting the invalid data and errors. As result, a total of 108 records are used in this study.

The study selects the factor of "accident severity" as the target. Other factors of "ship type", "accident type", "day or night", "visibility", "extreme weather conditions", "seasons", "weather", "ship density", "human factors" and "count" are compared with the target factor.

Then states are assigned to all factors. The states for "ship type" include: 1) "general cargo ship", 2) "oil tanker", 3) "fishing vessel", 4) "Passenger ship", 5) "Bulk carrier", 6) "Other". For "accident type", states assigned are: 1) "Fire", 2) "Sinking", 3) "Grounding", 4) " Allision", 5) "Collision", 6) "Other". Assigned states for "Day or Night" as 1) "Day" and 2) "Night". Assign the states of "visibility" as 1) "bad", 2) "good" and 3) "medium". Assign states for" Extreme weather conditions" as 1) "None" and 2) "Yes". The factors of "seasons" as 1) "Spring", 2) "Summer", 3) "Autumn", 4) "Winter" respectively. The states for "Weather" are 1) "Sunny", 2)

"Cloudy", 3) "Fog", and 4) "Rain". The states for "ship density" are 1) "Low" and 2) "High". The states for "Human Factors" are 1) "No" and 2) "Yes".

After assigning states to each factor, this paper uses the SPSS software program to perform the chi-square test as well as a correlation analysis among these factors. A total of nine chi-square tests are performed.

If the chi-square results satisfy the result of $P<0.05$, it indicates the interaction between two factors. A smaller chi-square value indicates a smaller deviation between the two factors, while the opposite indicates a larger deviation. The results are shown in Table 1.

**Table 1** The chi-square value

| Factors | Value |
|---|---|
| Day and night | 0.017 |
| The type of accident | 0.000033 |
| Vessel type | 0.027 |
| Visibility | 0.766 |
| Extreme weather | 0.747 |
| Season | 0.621 |
| Weather | 0.647 |
| Density | 0.668 |
| Human factor | 0.747 |

Table 1 reports the compared results of chi-square values among all factors. Three factors of day/night, type of ship accident, and ship type have significant interaction with"accident severity", with value less than the threshold value of 0.05 (0.017, 0.00003 and 0.027 respectively). The chi-square test results for other factors of visibility, extreme weather conditions, seasons, weather, ship density, and human factors are P=0.766, P=0.747, P=0.621, P=0.647, P=0.668, and P=0.747, respectively, in which the P-values of the above six factors are greater than 0.05, so there is no significant interaction between them and the severity of ship accidents.

In details, accident type, vessel type, and day or night are the three major influential factors of marine accidents in the coastal water of China, and the governments and the International Maritime Organization need to focus on these three factors and take effective measures.The frequency and proportions for all states of factors are reported in Table 2,3,4. There is the most significant correlation between the "types of accidents" and the "severity of the accident".

**Table 2** Type of accident as a percentage of accident severity

| **accident type** | **Frequency** | | | **Proportions** | | |
|---|---|---|---|---|---|---|
| | *Level 1 Minor accidents* | *Level 2 Serious accidents* | *Level 3 Catastrophic accidents* | *Level 1 Minor accidents* | *Level 2 Serious accidents* | *Level 3 Catastrophic accidents* |
| Fire | 4 | 0 | 0 | 4.80% | 0 | 0 |
| Sinking | 2 | 4 | 5 | 2.40% | 40% | 35.70% |

| | | | | | | |
|---|---|---|---|---|---|---|
| Grounding | 3 | 0 | 0 | 3.60% | 0 | 0 |
| Allision | 1 | 0 | 0 | 1.20% | 0 | 0 |
| Collision | 71 | 3 | 4 | 84.40% | 30% | 28.60% |
| Other | 3 | 3 | 5 | 3.60% | 30% | 35.70% |

**Table 3** Type of vessel as a percentage of accident severity

| **Vessel type** | **Frequency** | | | **Proportions** | | |
|---|---|---|---|---|---|---|
| | *Level 1 Minor accidents* | *Level 2 Serious accidents* | *Level 3 Catastrophic accidents* | *Level 1 Minor accidents* | *Level 2 Serious accidents* | *Level 3 Catastrophic accidents* |
| Grocery ships | 18 | 0 | 18 | 21.40% | 0% | 16.70% |
| Tanker | 5 | 0 | 6 | 6.00% | 0% | 5.60% |
| Fishing vessel | 7 | 2 | 11 | 8.30% | 20% | 10.20% |
| Passenger ship | 1 | 0 | 1 | 1.20% | 0% | 0.90% |
| Bulk carriers | 31 | 2 | 34 | 36.90% | 20% | 31.50% |
| Other | 22 | 6 | 38 | 26.20% | 60% | 35.20% |

**Table 4** Day/night as a percentage of accident severity

| **Day/night** | **Frequency** | | | **Proportions** | | |
|---|---|---|---|---|---|---|
| | *Level 1 Minor accidents* | *Level 2 Serious accidents* | *Level 3 Catastrophic accidents* | *Level 1 Minor accidents* | *Level 2 Serious accidents* | *Level 3 Catastrophic accidents* |
| Day | 28 | 7 | 2 | 33.30% | 70.00% | 14.30% |
| Night | 56 | 3 | 12 | 66.70% | 30.00% | 85.70% |

Tables 2 to 4 show the three factors that are correlated with the severity of the accident. the occurrence of minor accidents was mainly ship collisions, with numbers of 71 cases, accounting for 84.4% of the total number of accidents. Most of the vessels involved in minor accidents were general cargo ships, bulk carriers, and other vessels, which happened mainly at night, accounting for 21.4%, 36.9%, and 26.2% of the minor accidents respectively. Serious accidents occurred mainly due to ship sinking, ship allision, and collision, and the number of accidents related to these three factors was 4, 3, and 3 respectively, accounting for 40%, 30%, and 30% of the number of serious accidents. Most of the vessel types with serious accidents were fishing vessels, bulk carriers, and other vessels, accounting for 20%, 20%, and 60% of the number of serious accidents respectively, and the accidents occurred mainly in the daytime. Catastrophic accidents occurred mainly due to ship sinking, collision, and other causes, which mainly happened at night, there are 5, 4, and 5 catastrophic accidents related to these three factors, accounting for 35.7%, 28.6%, and 35.7% of the total catastrophic accidents, respectively. The majority of ship types with catastrophic accidents were cargo ships, bulk carriers, and other ships, accounting for 16.7%, 31.5%, and 35.2% respectively.

From the above three tables, the important factor influencing accident severity and their inner relationships can be seen. The result shows collision is the major accident type that occurs in coastal waters mainly happening on bulk carriers at night. Different state combinations of these three factors would lead to different accident severity. IMO should strengthen the cooperation among departments and accelerate the speed of salvage. Meanwhile, the ship safety inspection should be carried out to ensure that the ships are in condition, and enhanced communication between ships to reduce the occurrence of ship collisions. According to the data in this paper, the probability of accidents at night is higher than during the day because the crew is more likely to be tired at night, so the crew needs to maintain regular observation at all times of the night. Once a maritime accident is unavoidable relevant departments and rescuers should establish an effective plan to avoid secondary hazards.

## 4. CONCLUSION

This paper reports a study of causation analysis of ship accidents based on the chi-square test. The chi-square test is used to analyze the correlation between accident severity and other nine factors. As results, the day/night, type of accidents and ship type are three important factors. In addition, the states of those three factors are also analyzed. The test reveals some potential causal relations including collision is the major accident type that occurs in maritime bulk carriers; minor accidents often happen in bulk carriers. Catastrophic accidents usually happened at night. The study results are useful to provide effective references for accident prevention and to ensure safety sailing at sea. To the result would be more accurate, the future research database should be expanded to involve more data sources to increase the confidence level of experimental data.

## REFERENCES


[1] J. Chen, W. Bian, Z. Wan, Z. Yang, H. Zheng, and P. Wang, “Identifying factors influencing total-loss marine accidents in the world: Analysis and evaluation based on ship types and sea regions,” *Ocean Eng.*, vol. 191, Nov. 2019, doi: 10.1016/j.oceaneng.2019.106495.

[2] L. Wang and Z. Yang, “Bayesian network modeling and analysis of accident severity in waterborne transportation: A case study in China,” *Reliab. Eng. Syst. Saf.*, vol. 180, pp. 277–289, Dec. 2018, doi: 10.1016/j.ress.2018.07.021.

[3] L. Du, F. Goerlandt, and P. Kujala, “Review and analysis of methods for assessing maritime waterway risk based on non-accident critical events detected from AIS data,” *Reliability Engineering and System Safety*, vol. 200. Elsevier Ltd, Aug. 01, 2020. doi: 10.1016/j.ress.2020.106933.

[4] K. Liu, Q. Yu, Z. Yuan, Z. Yang, and Y. Shu, “A systematic analysis for maritime accidents causation in Chinese coastal waters using machine learning approaches,” *Ocean Coast. Manag.*, vol. 213, p. 105859, 2021.

[5] R. J. Bye and A. L. Aalberg, “Maritime navigation accidents and risk indicators: An exploratory statistical analysis using AIS data and accident reports,” *Reliab. Eng. Syst. Saf.*, vol. 176, pp. 174–186, Aug. 2018, doi: 10.1016/j.ress.2018.03.033.

[6] T. M. Franke, T. Ho, and C. A. Christie, “The Chi-Square Test: Often Used and More Often Misinterpreted,” *Am. J. Eval.*, vol. 33, no. 3, pp. 448–458, 2012, doi: 10.1177/1098214011426594.