Modelling ship collision risk based on the statistical analysis of historical data: A case study in Hong Kong waters

Yan-Fu WANG[1, 3], Long-Ting WANG[1], Jun-Cheng JIANG[2], Jin WANG[3], Zai-Li YANG[3]

[1]Department of Safety Science and Engineering, College of Mechanical and Electronic Engineering, China University of Petroleum, Qingdao, China

[2]School of Environment & Safety Engineering, Changzhou University, Changzhou, China

[3]Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Department of Maritime and Mechanical Engineering, Liverpool John Moores University, Liverpool, UK

Abstract: Collision, as a common type of ship accidents, leads to serious property loss and personal injury. In this paper, a new framework of quantitative risk assessment is proposed by quantifying the probability and the corresponding consequence based on the historical accident data. Firstly, the consequences of ship collisions are quantified and classified using an equivalent consequence method. Secondly, a decision tree model is established to analyse the impact of ship attributes on the collision consequences. The main ship attributes contributing to collision are determined, based on which, a BP neural network model is developed to estimate the probabilities of the different consequences. Thirdly, the collision risk is predicted by integrating the collision probabilities with the corresponding consequences. Fourthly, a case study in Hong Kong waters is investigated and the results are compared with the available references to validate the proposed framework. The new model can be used to assess present risks to plan preventive measures for the potential collision accidents.

Keywords: Quantitative risk assessment, Ship collision, Statistical analysis, Decision tree model, BP neural network

1. Introduction

With the rapid development of the water transportation industry, the number of traffic accidents has increased (Dong, 2010). Collisions between ships dominate in all kinds of water traffic accidents and have caused serious consequences. The Port of Hong Kong has always been a hub port serving the South Asian Pacific region. It is one of the busiest container ports in the world. In terms of vessel arrivals and departures,

1. Corresponding author: Yan Fu Wang (Tel.: +86-13698651195), E-mail: wangyanfu@upc.edu.cn.

and cargo and passenger throughput, it is also one of the major ports of the world (Hong Kong Marine Department, 2019). In Hong Kong waters, the ship collision accidents account for 48.7% of the total accidents in the past ten years. They account for 84% of the total injuries and 83% of the total deaths in maritime accidents (Hong Kong Marine Department, 2019). The disastrous consequences of ship collisions necessitate the development of a collision risk assessment framework that ensures safety and functionality of maritime transportation systems (MARPOL, 2005; Dong and Frangopol, 2015). Therefore, it is important and necessary to study how to reduce collision risk.

Many researchers have studied the collision risk. Some authors discussed theoretical and methodological frameworks (Mou et al., 2005; Morel and Chauvin, 2006; Mentes et al., 2015; Chen et al., 2018) and others focused on the frequency of ship collisions. The collision frequency or probability was often modelled based on the work of Fujii et al. (1974) and Macduff et al. (1974). Examples include ship domain models (Fowler and Sorgard, 2000; Wang, 2010; Chai et al., 2017) and ship collision formula (Qu et al., 2011; Wrobel et al., 2016). In addition, Bayesian Network was also used to study the collision probability by several authors (Montewka, 2014; Goerlandt and Montewka, 2015; Goerlandt et al., 2015; Sotiralis, 2016; and Trucco, 2008). A methodology based on the Hough Transform algorithm and Monte Carlo simulations were proposed to assess the collision probability (Christian and Kang, 2017).

Compared with the collision probability, the references about the collision consequence are limited. It is observed from the literatures that some authors use theoretical approach to simulate ship collision consequences, where, finite element methods were used to analyse the damage of the ship (Ozgur 2019). On the other hand, other researchers used empirical formula. For example, the collision consequence equations were proposed to determine the longitudinal and transversal damage extents of tankers (Van de Wiel and Van Dorp, 2011).

There are also a few scholars who have investigated the collision risk by evaluating the probability and the consequence. For example, Dong and Frangopol (2015) assessed the collision risk by computing the collision probability using the model proposed by Cowi (2008) and evaluating the consequences using economic loss. The risk of ship collisions was evaluated using the Frequency–Number of Facility (F/N) curve (Chai et al., 2017). Two types of accident consequence including human life loss and oil pollution are estimated using empirical formula. However, those approaches cannot be used to predict the probability of different types of collision consequences. Moreover, the proposed risk curves were not validated by the historical data. Therefore, this paper proposes a new framework to assess comprehensively the collision risk by quantifying the probability and the collisions consequence based on the historical data.

This paper evolves through four stages elucidated as follows. Section 2 proposes a new framework of quantitative risk analysis and presents the methods used. In Section 3, the case study in Hong Kong waters is researched to analyse the causes and consequences of ship collisions. A decision tree model is used to investigate into the effects of ship attributes on collision consequences. The BP Network model is used to predict the probability of different consequences. Section 4 predicts the collision risk of different types of ship by quantifying the collision probability and the corresponding consequences.

2. Methods and contributions

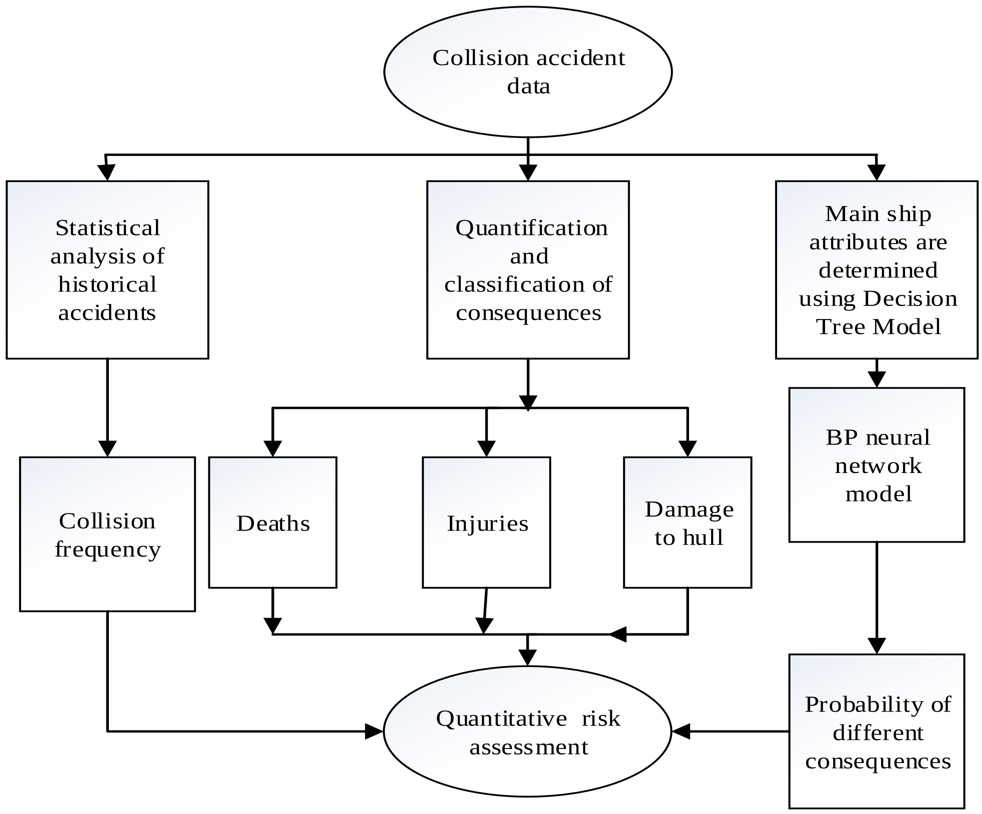


Fig. 1 The framework of quantitative risk assessment

As shown in Fig. 1, a new framework is proposed for quantitative risk assessment of ship collisions. Firstly, the collision accident data is statistically analysed to determine the collision frequency, the causes of ship collisions, the attributes of the ships involved in the collisions and the consequences (the number of deaths, injuries and the damage to the ship). Secondly, consequences are quantified and classified using an equivalent consequence method. Thirdly, the ship attributes and the corresponding consequences are input into a decision tree model to analyze the impact of ship attributes on the collision consequences. Fourthly, the main ship attributes, which are determined using the proposed decision tree model, are input into the BP neural network to estimate the probabilities of different consequences. At last, the collision risk is assessed by integrating the collision probabilities with the corresponding consequences.

There is no model available for accurately predicting collision consequences. Therefore, a conservative prediction for the risk of different consequences can provide reliable suggestions for consequence mitigation and risk reduction. This paper introduces a framework for risk assessment, demonstrating its ability for efficient reasoning and instantaneous updating in the light of new data. The approach has been applied to a case study in “Hong Kong waters”. Then the results are compared with the available references to validate the new framework. It can be used to predict the risk of collisions and different consequences according to the ship attributes. Learning the risk of different consequences is significant and is capable of generating useful insights in collision risk analysis.

2.1 Classification and Quantification of Ship Attributes

The main ship attributes include ship age, tonnage, navigation speed, ship type, collision position and angles between the involved ships (Wang and Yang, 2018). Since, the ship attributes are widely distributed and some are not quantitative, it is necessary to classify and quantify all the ship attributes data. According to the standards of International Maritime Organization (IMO) and China Maritime Safety Administration (2019), the ship attributes can be grouped as shown in Table 1 (Wang and Yang, 2018; IMO, 2002; IMO, 2007a; IMO, 2007b; IMO, 2011).

Table 1 Classification of ship attributes

Ship attributes	Unit	Classification
Ship Age (A)	Years	1= [0,10）
		2= [10,20）
		3= [20, ∞）
Ship Tonnage (W)	Tons	1= [0,500）
		2= [500,3000）
		3= [3000, ∞）
Navigation Speed(V)	Knots	1= [0,5)
		2= [5,10)
		3= [10,15)
		4= [15,20)
		5= [20, ∞)
Ship Type (T)	________	1= {Small-sized ship}
		2= {Passenger ship}
		3= {Conventional cargo ship}
		4= {Liquid cargo ship}
		5= {Container ship}

Collision Position (L)	________	1= {Bow}
		2={Amidships}
		3= {Stern}
Conflict Types (K)	________	1= {Overtaking conflict}
		2= {Crossing conflict}
		3= {Head-on conflicts}
Crew members	________	1= [0,10）
		2= [10,20）
		3= [20,50）
		4= [50, +∞）

2.2 Classification and Quantification of Collision Consequences

Consequences are mainly described with respect to the number of deaths, the number of injuries and hull damage. According to the IMO standards, the collision consequences can be classified “Very serious”, “Serious” and “Minor” as shown in Table 2 (Wang and Yang, 2018; IMO, 2002; IMO, 2007a; IMO, 2007b; IMO, 2008; IMO, 2011).

Table 2 Classification of collision consequences

Classification	Collision consequences
Very serious	Involves total loss of the ship, loss of life or severe pollution (the case of pollution produces a major deleterious effect upon the environment).
Serious	Does not qualify as "Very serious casualties" but involves a fire, explosion, collision, grounding, contact, heavy weather damage, ice damage, hull cracking, or suspected hull defect, etc.
Minor	Does not qualify as "Very serious casualties" or "serious casualties" and includes marine incidents which themselves include "hazardous incidents" and "near misses".

Based on the available references, a measure of equivalent consequence can be defined as (IMO, 2002; IMO, 2007a; IMO, 2007b; IMO, 2008; IMO, 2011; IMO, 2019a; IMO, 2019b): One-person death, Ten injuries and Critical damage to the hull.

In order to facilitate the quantification of accident consequences, the equivalent consequence method is adopted to analyse the collision consequences. The hull damage can be divided into four types: undamaged, minor damage, critical damage and sunk. The degree of hull damage is expressed as a number between 0 and 100, where 0 means that the hull is undamaged, 40 means that the hull is slightly damaged, 80 means that the hull is critically damaged, and 100 means that the hull is sunk (Dai, 2003; Kim et al., 2015). The equivalent consequence (F) of collision accidents is calculated as:

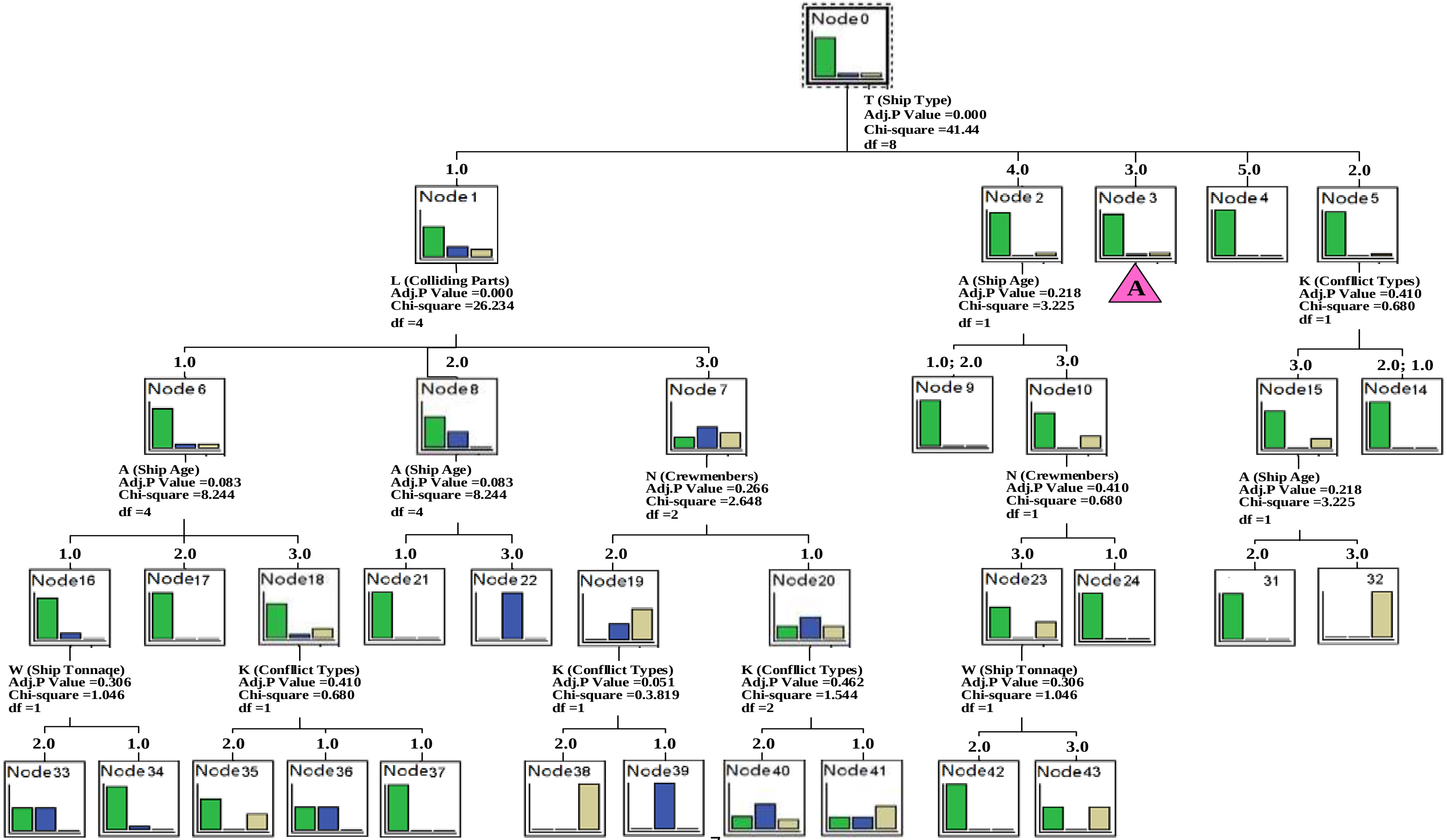
$$\mathrm{F} = \frac{D}{1} + \frac{I}{10} + \frac{B}{B_0}$$

where, D is the number of fatalities, I is the number of injuries, B is the damage degree to the ship, and B_0 is one equivalent damage degree to the ship (B_0= 80).

With reference to standards and literatures, the equivalent consequence value for a minor accident is [3, 10), for a serious accident [10, 30) and [30, +∞) for a very serious accident (Ministry of Transport of the People's Republic of China,2019; IMO, 2019a; IMO, 2019b; SOLAS, 1974).

2.3 Effect Analysis of Ship Attributes on Collision Consequences using Decision Tree Model

The equivalent consequences of each collision accident are calculated and are categorized into minor, serious and very serious accidents, respectively. The ship attributes are set as independent variables and the equivalent consequences are set as dependent variables. Based on the statistical analysis of historical data, a decision tree model is built as shown in Fig. 2. Since the variables are categorical variables, the chi-squared automatic interaction detector (CHAID) is chosen as the growth method of the decision tree. CHAID can be used for prediction, classification and detection of the interaction between variables. One of the CHAID's advantages is that its output is visual and easy to interpret. Another important advantage of CHAID over alternatives is that it is non-parametric.



Node 0
T (Ship Type)
Adj.P Value =0.000
Chi-square =41.44
df =8
1.0
4.0
3.0
5.0
2.0
Node 1
Node 2
Node 3
Node 4
Node 5
A
L (Colliding Parts)
Adj.P Value =0.000
Chi-square =26.234
df =4
A (Ship Age)
Adj.P Value =0.218
Chi-square =3.225
df =1
K (Conflict Types)
Adj.P Value =0.410
Chi-square =0.680
df =1
1.0
2.0
3.0
1.0; 2.0
3.0
3.0
2.0; 1.0
Node 6
Node 8
Node 7
Node 9
Node10
Node15
Node14
A (Ship Age)
Adj.P Value =0.083
Chi-square =8.244
df =4
A (Ship Age)
Adj.P Value =0.083
Chi-square =8.244
df =4
N (Crewmenbers)
Adj.P Value =0.266
Chi-square =2.648
df =2
N (Crewmenbers)
Adj.P Value =0.410
Chi-square =0.680
df =1
A (Ship Age)
Adj.P Value =0.218
Chi-square =3.225
df =1
1.0
2.0
3.0
1.0
3.0
2.0
1.0
3.0
1.0
2.0
3.0
Node16
Node17
Node18
Node21
Node 22
Node19
Node20
Node23
Node24
31
32
W (Ship Tonnaqe)
Adj.P Value =0.306
Chi-square =1.046
df =1
K (Conflict Types)
Adj.P Value =0.410
Chi-square =0.680
df =1
K (Conflict Types)
Adj.P Value =0.051
Chi-square =0.3.819
df =1
K (Conflict Types)
Adj.P Value =0.462
Chi-square =1.544
df =2
W (Ship Tonnaqe)
Adj.P Value =0.306
Chi-square =1.046
df =1
2.0
1.0
2.0
1.0
1.0
2.0
1.0
2.0
1.0
2.0
3.0
Node33
Node34
Node35
Node36
Node37
Node38
Node39
Node 40
Node 41
Node42
Node43

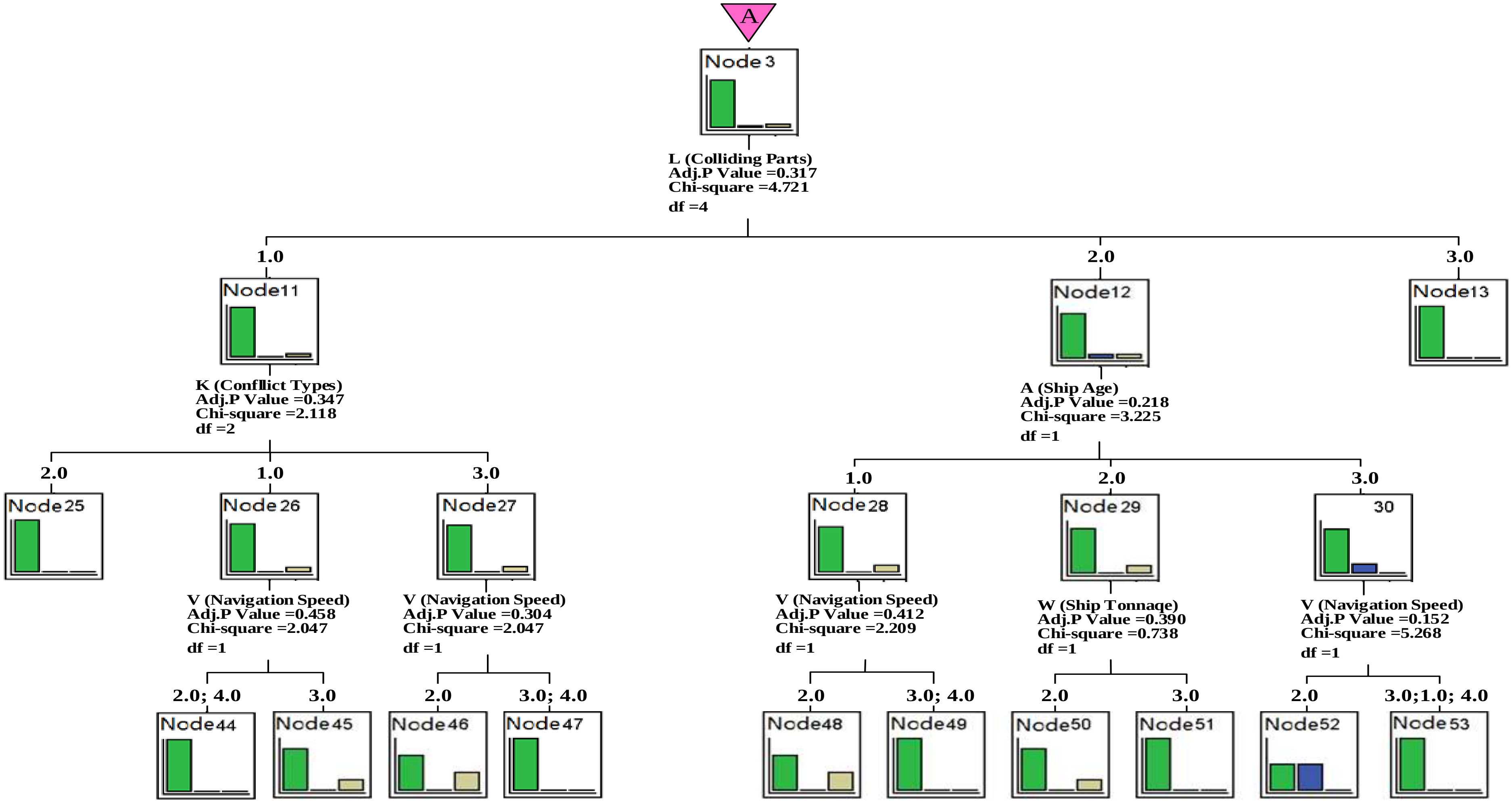


Fig. 2. Decision tree model of ship collision consequences

2.4 Predict the probability of different consequences using BP neural network model

The BP neural network is a hierarchical neural network with upper neurons fully associated with lower neurons. When samples are supplied to the network, the transferred values are propagated from the input layer through the middle layer to the output layer. Therefore, the neural network input response could be obtained from the neurons of the output layer.

In this section, the BP neural network model is established using MATLAB software. Seven ship attributes are used as input layer nodes. Three types of collision consequences are treated as the output layer nodes. The number of nodes in hidden layer is the twice of the nodes in the input layers plus one, so it is set to 15. Mapminmax is used to normalize the input node of the training samples and the test samples. The Sigmoid is selected as activation function. The corresponding MATLAB program encode is "logsig". "traingdx" is selected as the training function. The display period of training results is set to 50. The training number is 500 iterations, the minimum error of training target is 0.05. The learning speed is 0.01 (Ren et al., 2014). The training process of the proposed BP neural network is shown in Fig. 3. The detailed program is shown in Appendix 1.

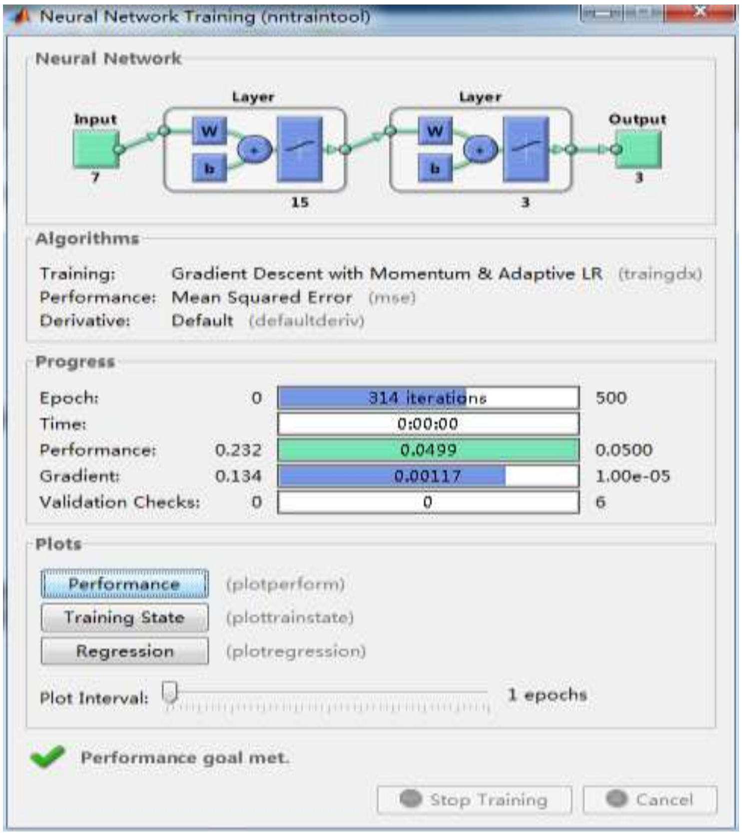


Fig. 3 Training process of neural network model

3. Application of the Case study in Hong Kong waters

3.1 Causes Analysis of the Ship Collisions

108 accident reports of ship collision with serious consequences were collected from the official sources of the Hong Kong Marine Department. Some very minor collision accidents are not included in this paper because the consequences are too minor to be considered. Each collision accident report includes the description of the ships, staff number, ship companies, accident consequence, accident process, causes analysis, collision angle and position as shown in Appendix 2.

The causes of collision accidents are summarized into four categories: human factors (Celik, 2009), ship and equipment factors, environmental factors and management factors (Sotiralis et al., 2016; Chai et al., 2017). The main causes of ship collision are identified through the statistical analysis of the historical accidents as shown in Table 3 and Fig. 4.

Table 3 Classification and the occurrence frequency of collision causes

Classification		Causes	Occurrence Number (N)	Nx100/108
Human factors	Perception stage	Misuse of navigation instruments	1	0.467
		Wrong diagnosis of the situation	25	11.682
		Fatigue	20	9.346
		Insufficient consideration of weather conditions	22	10.28
	Decision-making stage	Undetected signals	21	9.813
		Unsuitable route selection	41	19.159
		Not fully mastering environmental information	24	11.215
		Error estimation of collision risk	62	28.972
	Action stage	Improper ship handling	56	26.168
		Uncoordinated collision avoidance	30	14.109
		Untimely action	80	37.383
		Over-speed	35	16.355
Ship and equipment factors		Equipment failure	12	11.1111
		Communication system failure	7	6.4814
		Navigation system failure	8	7.4074
		Improper use of ship's signal lights	12	11.1111
Environmental factors		Bad channel environment	18	16.6667
		Bad weather	42	38.8889
		Poor visibility	38	35.1852
		High traffic density	10	9.2593
Management factors		Improper regulations for shifting of duty	13	12.037
		Lack of license	4	3.7037
		Inadequate communication	58	53.7037
		Inadequate training/ experience	41	37.963

Incomplete navigational information	3	2.7778

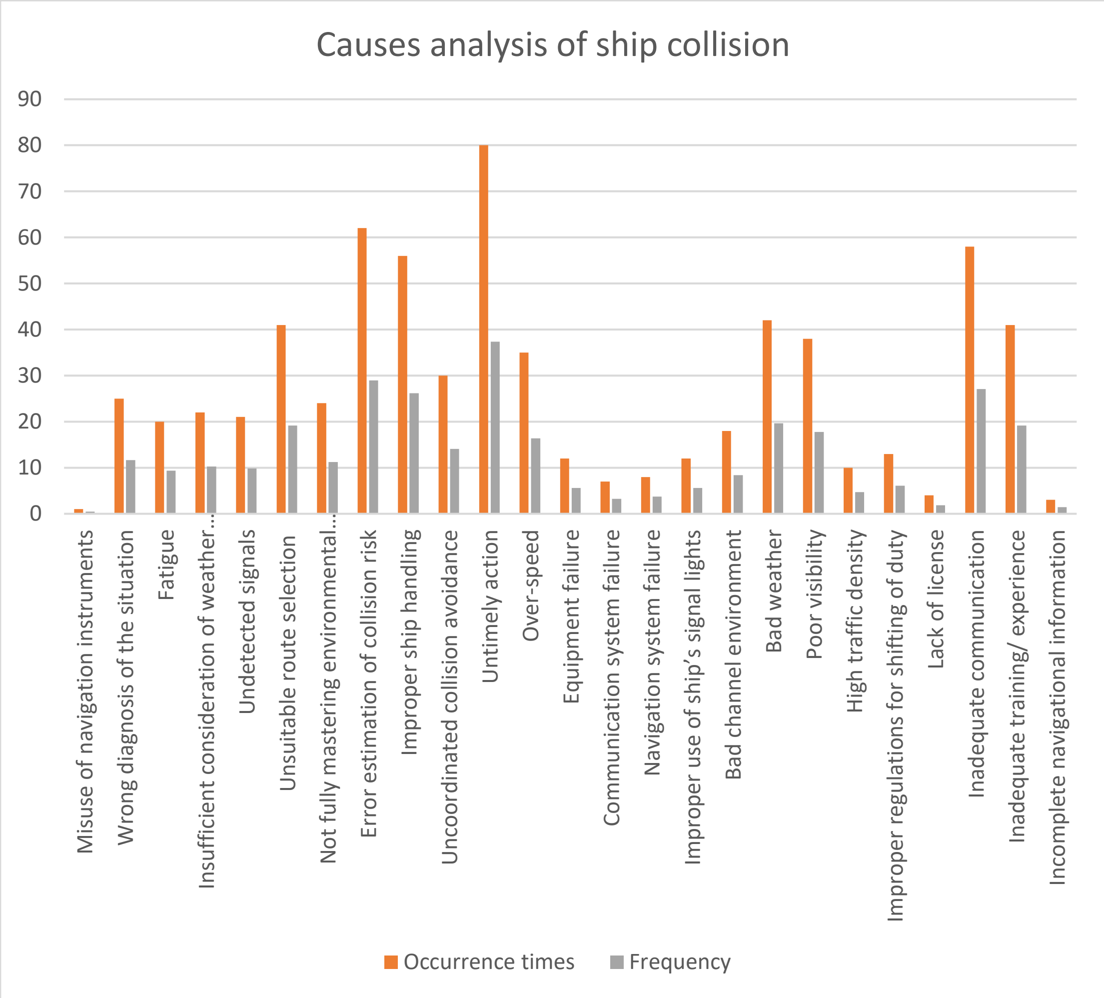


Fig. 4 Causes analysis of ship collision

As seen in Table 3 and Fig. 4, the most likely collision cause is Untimely action, followed by Error estimation of collision risk, Inadequate communication, Improper ship handling, Bad weather, Unsuitable route selection and Inadequate training/ experience in descending order. By comparing the results of this research with the literatures (Sotiralis et al., 2016; Chai et al., 2017), it can be found that Error estimation of collision risk, Inadequate communication, Improper ship handling, Bad weather, Inadequate training/ experience and Unsuitable route selection are the main contributing factors to ship collisions.

3.2 Effects of Ship Attributes on Collision Consequences

In this section, the minor accident is used for the case study. In order to avoid overfitting, the maximum number of layers is limited to 4. The corresponding information gain is shown in Table 4. Information gain measures how much “information” a feature gives us about the class, which is the main key that is used to construct a Decision Tree. An attribute with the highest information gain will be tested firstly.

Table 4 Information gain analysis of collision consequence decision tree

Node	Samples		Information gain		Confidence	Index
	N1	Percentage1	N2	Percentage2		
4	33	15.3%	33	17.9%	100.0%	117.4%
14	24	11.1%	24	13.0%	100.0%	117.4%
25	15	6.9%	15	8.2%	100.0%	117.4%
9	14	6.5%	14	7.6%	100.0%	117.4%
17	9	4.2%	9	4.9%	100.0%	117.4%
44	8	3.7%	8	4.3%	100.0%	117.4%
53	8	3.7%	8	4.3%	100.0%	117.4%
47	7	3.2%	7	3.8%	100.0%	117.4%
49	5	2.3%	5	2.7%	100.0%	117.4%
13	5	2.3%	5	2.7%	100.0%	117.4%
37	4	1.9%	4	2.2%	100.0%	117.4%
31	4	1.9%	4	2.2%	100.0%	117.4%
21	2	9%	2	1.1%	100.0%	117.4%
51	2	9%	2	1.1%	100.0%	117.4%
42	1	5%	1	5%	100.0%	117.4%
24	1	5%	1	5%	100.0%	117.4%
34	15	6.9%	14	7.6%	93.3%	109.6%
45	5	2.3%	4	2.2%	80.0%	93.9%
50	5	2.3%	4	2.2%	80.0%	93.9%
35	9	4.2%	6	3.3%	66.7%	78.3%
46	3	1.4%	2	1.1%	66.7%	78.3%
48	3	1.4%	2	1.1%	66.7%	78.3%
52	4	1.9%	2	1.1%	50.0%	58.7%
33	2	9%	1	5%	50.0%	58.7%
36	2	9%	1	5%	50.0%	58.7%
43	2	9%	1	5%	50.0%	58.7%
40	15	6.9%	4	2.2%	26.7%	31.3%
41	4	1.9%	1	.5%	25.0%	29.3%
38	2	9%	0	.0%	.0%	0%
39	1	5%	0	.0%	.0%	0%
22	1	5%	0	.0%	.0%	0%
32	1	5%	0	.0%	.0%	0%

In Table 4, N1 represents the number of samples contained in the node, Percentage1 is equal to N1 divided by the total number of all the samples, N2 indicates the number of colliding ships with minor consequence in this node, Percentage2 equals N2 divided by the total number of colliding ships with a minor consequence. Confidence equals N2 divided by the corresponding N1, which reflects the ratio of the researched category (in this case, minor accidents) to the total number of samples for this node. Index is the ratio of percentage 2 to percentage 1, which is used to judge the importance of the nodes. If the index of the node is larger than one, the contribution of this node is larger and attention should be paid to. The above decision tree model is analysed for the different types of consequences and the main rules, where the index is greater than 1, are summarized as shown in Table 5.

Table 5 the main rules of decision tree model

Rules	Ship Type	Ship Age	Colliding Parts	Conflict Types	Number	Tonnage	Speed	Consequence
1	Container ship	/	/	/	/	/	/	Minor
2	liquid cargo ship	<20	/	/	/	/	/	Minor
3	General cargo	/	Stern	/	/	/	/	Minor
4	Passenger ship	/	/	Overtaking conflict/ Crossing conflict	/	/	/	Minor
5	Small-sized ship	10-20	Bow	/	/	/	/	Minor
6	General cargo	/	Bow	Crossing conflict	/	/	/	Minor
7	Small-sized ship	<10	Bow	/	/	<500	/	Minor
8	Small-sized ship	>20	Bow	Crossing conflict	/	/	/	Minor
9	Small-sized ship	/	Amidships	Crossing conflict	10-20	/	/	Very Serious
10	Small-sized ship	/	Amidships	Crossing conflict	<10	/	/	Serious
11	Small-sized ship	/	Amidships	Overtaking conflict	<10	/	/	Very Serious
12	General cargo	/	Bow	Overtaking conflict	/	/	5-10，15-20	Minor
13	General cargo	/	Bow	Head-on conflicts	/	/	10-20	Minor
14	General cargo	<10	Amidships	/	/	/	10-20	Minor
15	General cargo	10-20	Amidships	/	/	500-3000	/	Minor
16	General cargo	>20	Amidships	/	/	/	5-10	Serious
17	General cargo	>20	Amidships	/	/	/	<5，10-20	Minor

By analysing the main rules generated by the above decision tree model, the following results are obtained:

(1) If the ship type is small-sized and the collision parts is bow, the consequences are mostly minor, and if the colliding parts are amidships, the consequences are serious or Very serious.

(2) For container ships, the mostly collision consequences are minor. If the ship type is small-sized, the mostly collision consequences are serious or very serious.

(3) The ship type, collision position and conflict type have a greater influence on collision consequences.

(4) The frequent attribute combinations leading to collision accidents can be determined by explaining and comparing the different rules of decision tree model.

(5) The collisions among the small-sized ships, passenger ships and general cargo ships cause more serious consequences than the other types of ships. Therefore, the consequences of these three-type ships will be further analysed in the next section.

3.3 Predict the Probability of Different Consequences using BP Neural Network

The main attributes of the above three types of ships, which are determined using the above decision tree model, are input into the proposed BP neural network. The first 80% of the statistical data is used as the training set and the last 20% is used as the verification set. The performance of the proposed BP neural network is shown in Fig. 5.

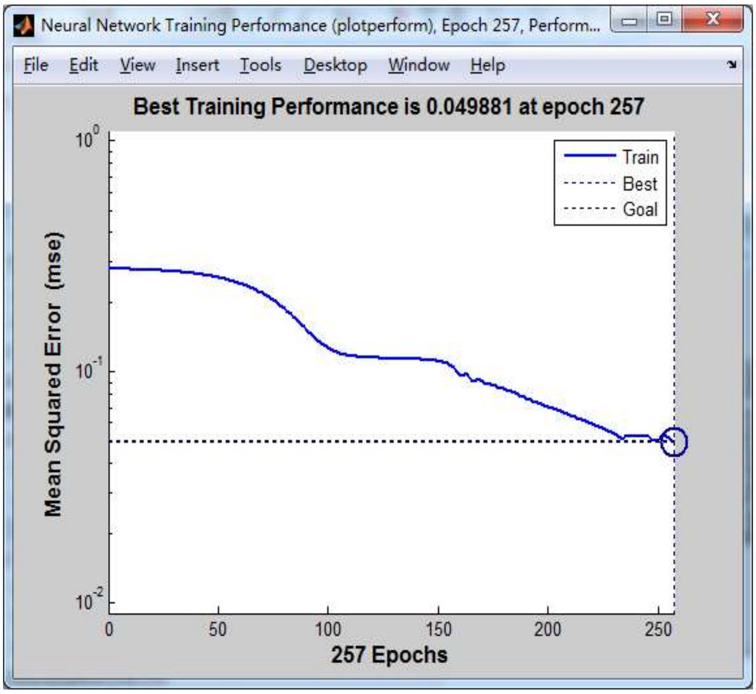


Fig. 5 Performance of BP neural network

The training results show that the accuracy rate of the model is larger than 90%, which verifies that the calculation results are at large consistent with the actual statistical data. The trained model is used to estimate the number of different consequences five times. The number of minor, serious and very serious consequences are statistically analysed as shown in Tables 6, 7 and 8 respectively for the three types of ships.

Table 6 Number of three-type collision consequences for small-sized ship

Times	Minor	Serious	Very Serious
1	43	27	2
2	44	21	7
3	48	21	3
4	41	21	10
5	41	26	5
Average	43.4	23.2	5.4
Proportion	0.6028	0.3223	0.075

Table 7 Number of three-type collision consequences for passenger ship

Times	Minor	Serious	Very Serious
1	54	0	0
2	54	0	0
3	53	0	1
4	54	0	0
5	53	1	0
Average	53.63	0.2	0.2
Proportion	0.9926	0.0037	0.0037

Table 8 Number of three-type collision consequences for general cargo ship

Times	Minor	Serious	Very Serious
1	68	2	2
2	69	1	2
3	68	2	2
4	66	3	3
5	65	1	6
Average	67.2	1.8	3
Proportion	0.9333	0.025	0.0417

The simulation results of the above three-type ships are compared with the results of the decision tree model and historical data as shown in Table 9.

Table 9 The proportion of different consequences for three-type ships

Ship Type	Data Category	Minor	Serious	Very Serious
Small-sized	Simulation results	0.6028	0.3222	0.075
	Decision tree results	0.636	0.212	0.152
	Statistical data	0.6571	0.1572	0.1857
Passenger	Simulation results	0.9926	0.0037	0.0037
	Decision tree results	0.932	0.012	0.056
	Statistical data	0.625	0.1562	0.2188

General cargo	Simulation results	0.9333	0.025	0.0417
	Decision tree results	0.914	0.029	0.057
	Statistical data	0.9018	0.0446	0.0536

From Table 9, it is found that:

(1) The simulation results of the BP neural network are slightly different from those of the decision tree model and historical data; however, they have shown a similar trend. The BP neural network can be used to predict the probability of different consequences if the conditions do not change drastically.

(2) The small-sized ships cause serious or very serious accidents more frequently than other types, which is consistent with the conclusion of the decision tree and historical data. Therefore, for the small-sized ships, the safety training and safety management should be especially strengthened to prevent the ship collisions.

(3) Most collisions of passenger ships and general cargo ships cause minor consequences. However, there is also a small amount of serious or very serious accidents, requiring safety awareness in addressing their anti-collision measures.

(4) For passenger ships, the probability of very serious accidents is greater than that of serious accidents, revealing that collision accidents involving passenger ships can result in large fatalities, although their frequencies are low.

4. Risk analysis of ship collision

In Table 10, the collision accidents number and the data of ship flow were collected for the small-sized ships, passenger ships and general cargo ships in Hong Kong waters from 2005 to 2015 (Hong Kong Marine Department, 2019).

Table10 Number of ships and collision accidents per year

	classification	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
collision accident number	small-sized	5	2	2	1	2	0	2	3	0	3	3
	passenger	0	2	0	5	2	0	4	2	1	0	1
	general cargo	1	2	3	4	1	2	3	1	4	1	1
ship flow	small-sized	680	690	640	430	462	438	452	555	545	557	523
	passenger	78510	79220	83450	83810	84943	90263	89895	84242	82852	82489	86722
	general cargo	32930	27950	23330	21470	21733	19150	17047	16486	19471	20185	19003

From Table 10, it can be calculated that the average collision frequency of the small-size ship is 3.8513×10^{-3} per ship visit, the collision frequency of a passenger ship is 1.8351×10^{-5} per ship visit and the general cargo ship's is 9.6333×10^{-5} per ship visit. By integrating the collision frequency with the ratio of different consequences in Table 9, the occurrence probabilities of different consequences are determined as shown in Table 11. For example, for small-sized ships, the proportion of minor consequence is 0.6571 as shown in Table 9 and the average collision probability is 0.003851 according to the calculation results of Table 10. Therefore, the occurrence probability of minor consequences for a small-sized ship is 0.002531.

Table 11 Occurrence probability for different consequences

Accident classification	Occurrence probability			Total
	Small-sized ship	Passenger ship	General cargo ship	
Minor	0.002531	1.14692E-05	8.68718E-05	0.002629
Serious	0.000605	2.8673E-06	4.30058E-06	0.000612
Very Serious	0.000715	4.01422E-06	5.1607E-06	0.000724
Total	0.003851	1.83507E-05	9.63331E-05	0.003966

The average equivalent consequences of collisions are statistically analysed for different types of ships as shown in Table 12. Some detailed historical data is shown in Appendix 2.

Table 12 Equivalent consequences of ship collisions

Accident classification	Average equivalent consequences			Total
	Small-sized ship	Passenger ship	General cargo ship	
Minor	0.878	0.585	0.478	1.941
Serious	2.186	2.01	2.02	6.216
Very Serious	7.042	15.607	15.817	38.466
Total	10.107	18.202	18.314	46.623

The risk of different consequences is evaluated for three types of ships as shown in Table 13. Taking the risk of minor accidents for small-sized ships as an example, the risk is equal to the probability (0.002531) in Table 11 multiplied by the equivalent consequence (0.878) in Table 12.

Table 13 The risk of different consequences for three types of ships

Accident classification	Risk			Total
	Small-sized ship	Passenger ship	General cargo ship	
Minor	0.002222	6.71E-06	4.152E-05	0.00227
Serious	0.001323	5.763E-06	8.687E-06	0.001334
Very Serious	0.005037	6.265E-05	8.163E-05	0.005181
Total	0.008582	7.51E-05	0.000132	0.008789

The F-N diagram is the most common form to illustrate the relationship between the accident consequence and the corresponding occurrence frequency. An F-N diagram of ship collisions in Hong Kong waters is depicted in Fig. 6. The equivalent consequence of the accident and the corresponding frequency are shown on the abscissa and ordinate, respectively.

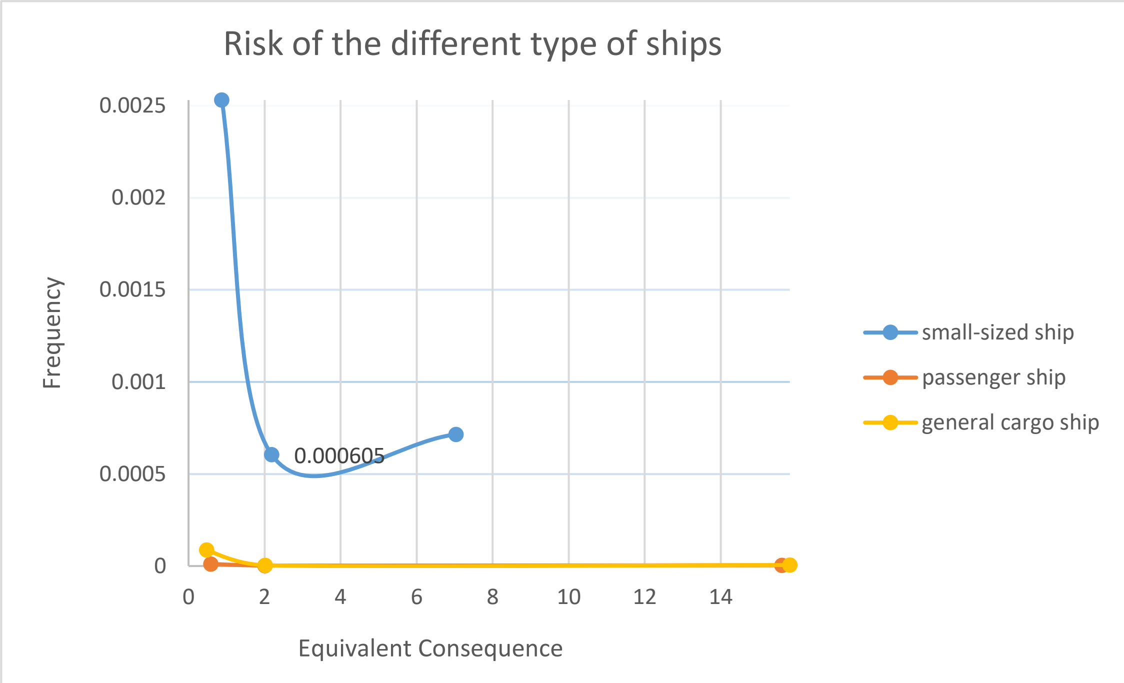


Fig. 6 F-N diagram of ship collisions in Hong Kong waters

From Table 13 and Fig. 6, it is seen that firstly the collision risk of small-sized ships is significantly higher than that of the other ships. Secondly, comparing all the results, the risk of very serious accidents is the largest, while the risk of serious accidents is the smallest. Thirdly, the collision frequency of a ship in Hong Kong waters is 0.003966 per ship visit. The average collision accident frequency of ships is about 0.00129 collisions per ship year (Montewka et al., 2014) and 0.0078 collisions per ship year from the EU research project Goal-Based Damage Stability (Zaraphonitis et al.,2012). The overall collision risk of a ship in Hong Kong waters is about 0.008789 per ship year. The research results of this paper are in the same level with the available references (Zaraphonitis et al., 2012, Montewka et al., 2014).

5. Conclusions

(1) According to the cause analysis of collision accidents, it can be seen that Error estimation of collision risk, Inadequate communication, Improper ship handling, Bad weather, Inadequate training/ experience and Unsuitable route selection contribute more to ship collisions. These results are in line with the available references.

(2) The equivalent consequence decision tree was built to determine the frequent attributes, which significantly contribute to the ship collisions. The main ship attributes that have great impact on collision consequences are separately the ship type, the collision position and conflict types. The collisions among the small-sized ships,

passenger ships and general cargo ships cause more serious consequences than the other types of ships.

(3) The BP neural network is proposed to predict the probabilities of different consequences. By comparing the simulation results with those of the decision tree model and historical data, it can be concluded that the proposed model can be used to predict the probability of different consequences if the conditions do not change drastically.

(4) The collision risk of different types of ships are calculated based on the statistical analysis of historical data. The collision risk of small-sized ships is significantly higher than that of other ships. The collision frequency of ships in Hong Kong waters is 0.003966 per ship visit. The overall collision risk of a ship is around 0.008789 per ship year. The proposed framework is validated by comparing prediction results with the available references.

This paper proposes a new framework to predict the collision risk of different types of ships based on historical data. Limitations are the deficiency of historical accidents data from different waters. In the future, more studies will be carried out to synchronize the traffic data and accident data by integrating AIS data and historical accidents to model the relationship between vessel traffic and accidents. It may be necessary to predict the dynamic collision risk in order to propose a collision risk alerting system.

Acknowledgments

This work is supported by Shandong Provincial Natural Science Foundation (Project No. ZR2019MEE080), “Key R&D projects in Shandong Province (Project No.2018GSF120021) and “the Fundamental Research Funds for the Central Universities” (Project No. 17CX02062 and No.19CX02025A)”. This research has received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Individual Fellowships– 840425. This research is also partially supported by EU H2020 RISE 2016 RESET – 730888.

References

Celik M., Cebi S., 2009. Analytical HFACS for investigating human errors in shipping accidents. Accident Analysis and Prevention. 41(1), pp. 66-75.

Chai T., Weng J., Xiong D Q., 2017. Development of a quantitative risk assessment model for ship collisions in fairways. Safety Science. 91, pp. 71-83.

Chen L J., Yan X P., Huang L W., et al, 2018. A systematic simulation methodology for LNG ship operations in port waters: a case study in Meizhou Bay. Journal of Marine engineering and Technology. 17, pp.12-32.

China Maritime Safety Administration, [online], URL. http://en.msa.gov.cn/, Accessed date: 1 May 2019

Christian R., Kang H G., 2017. Probabilistic risk assessment on maritime spent nuclear fuel transportation (Part II: Ship collision probability). Reliability Engineering & System Safety. 164, pp.136-149.

COWI ,2008. Risk Analysis for Sea Traffic in the Area around Bornholm. Danish Maritime Authority, Kongens, Lyngby.

Dai T Y., 2003. Ship Impact against Bridge and Its Risk Assessment (Doctoral dissertation). Harbin Engineering University.

Dong C L., 2010. The new impetus for the revival and prosperity of the Greek shipping model engine. Market Research. 11, pp.19-22.

Dong Y., Frangopol D. M., 2015. Risk and resilience assessment of bridges under mainshock and aftershocks incorporating uncertainties. Engineering Structures. 83(15):198-208.

Fowler T.G., Sørgård E., 2000, Modeling Ship Transportation Risk. Risk Analysis 20 (2):225-244

Fujii Y., Oshima R., Yamanouchi H., Mizuki N., 1974. Some Factors Affecting the Frequency of Accidents in Marine Traffic: I- The Diameter of Evasion for Crossing Encounters, II- The probability of Stranding, III- The Effect of Darkness of the Probability of Collision and Stranding. The Journal of Navigation 27 (2):239-247.

Goerlandt F., Montewka J., 2015. A framework for risk analysis of maritime transportation systems: A case study for oil spill from tankers in a ship–ship collision. Safety Science. 76, pp.42-66.

Goerlandt F., Montewka J., Kuzmin V., et al, 2015. A risk-informed ship collision alert system: Framework and application. Safety Science. 77(1), pp.182-204.

Hong Kong Marine Department, [online], URL. https://www.mardep.gov.hk/en/home.html, Accessed date: 2 June 2019.

IMO. [online], URL. http://www.imo.org/en/Pages/Default.aspx, Accessed date: 8June 2019

IMO. [Online], URL. http://www.imo.org/en/OurWork/MSAS/Casualties/Pages/Default.aspx. , Accessed date: 8 June 2019

IMO, 2002. Formal Safety Assessment of Bulk Carrier Safety. International Maritime Organization, London, Document No.: MSC75/5/2, February 12, 2002.

IMO, 2007a. Formal Safety Assessment FSA-Container Vessels. International Maritime Organization, London, Document No.: MSC83/INF.8, July 3, 2007.

IMO, 2007b. Formal Safety Assessment FSA-Liquefied Natural Gas (LNG) Carriers. International Maritime Organization, London, Document No.: MSC83/INF.3, July 3, 2007.

IMO, 2008. Formal Safety Assessment FSA-Crude Oil Tankers. International Maritime Organization, NY.

IMO, 2011. Formal Safety Assessment FSA of General Cargo Ships. International Maritime Organization, London, Document No.: MSC2088/INF.8, 2011.

Kim S., Yang Y., Samy I., et al, 2015. Environmental Consequences Associated with Collisions Involving Double Hull Oil Tanker. Ships and Offshore Structures. pp.1-9.

Macduff T., 1974. The Probability of Vessel Collisions. Ocean Industry 9 (9). pp.144-148.

MARPOL, 2005. International Convention for the Prevention of Pollution from Ships, Annex VI.

Mentes A., Akyildiz H., Yetkin M., et al, 2015. A FSA based fuzzy DEMATEL approach for risk assessment of cargo ships at coasts and open seas of Turkey. Safety Science. 79, pp.1-10.

Ministry of Transport of the People's Republic of China, 2019. [Online], URL. http://xxgk.mot.gov.cn/jigou/fgs/201410/t20141020_2973305.html. , Accessed date: 28 July 2019

Montewka J., Ehlers S., Goerlandt F., et al, 2014. A framework for risk assessment for maritime transportation systems—a case study for open sea collisions involving RoPax vessels. Reliability Engineering & System Safety. 124(2), pp.142-157.

Morel G., Chauvin C., 2006. A socio-technical approach of risk management applied to collisions involving fishing vessels. Safety Science. 44(7), pp.599-619.

Mou J M., Zou Z J., Huang L W., et al, 2005. On Pattern of Vessel Accident. Journal of Wuhan University of Technology. 29(3), pp.489-492.

Ozgur O., 2019. Structural damage of ship-FPSO collisions. Journal of Marine Engineering and Technology. 18(1), pp. 1-35.

Qu X., Meng Q., Li S., 2011. Ship collision risk assessment for the Singapore Strait. Accident Analysis and Prevention. 43(6), pp.2030-2036.

Ren C., et al, 2014. Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting. Knowledge-Based Systems. 56, pp.226-239.

SOLAS, 1974. International Convention for Safety of Life at Sea.

Sotiralis P., Ventikos N P., Hamann R., et al, 2016. Incorporation of human factors into ship collision risk models focusing on human centered design aspects. Reliability Engineering & System Safety. 156, pp. 210-227.

Trucco P., Cagno E., Ruggeri F., et al, 2008. A Bayesian Belief Network modelling of organisational factors in risk analysis: A case study in maritime transportation. Reliability Engineering & System Safety. 93(6), pp.845-856.

Van de Wiel G., René van Dorp J. ,2011. An oil outflow model for tanker collisions and groundings. Annals of Operations Research.187(1), pp.279-304.

Wang L K., Yang Z L., 2018.Bayesian network modelling and analysis of accident severity in waterborne transportation: A case study in China. Reliability Engineering & System Safety. 180, pp.277–289.

Wrobel K., Krata P., Montewka J., et al, 2016. Towards the Development of a Risk Model for Unmanned Vessels Design and Operations. Transnav International Journal on Marine Navigation & Safety of Sea Transportation. 10(2), pp.267-274.

Zaraphonitis, G., Skoupas, S., Papanikolaou, A., Cardinale, M., 2012. Multi-objective optimization of ROPAX ships considering the SOALS 2009 and GOALDS damage stability formulations. In: Proc. 11th International Conference on Stability of Ships and Ocean Vehicles (STAB2012), Athens, September.

Appendix 1. BP neural network model

```
clear
[f1,f2,f3,f4,f5,f6,f7,class]=textread('train.txt','%f%f%f%f%f%f%f%f',300);
[input,PS]=mapminmax( [f1,f2,f3,f4,f5,f6,f7]');
s=length(class);
output=zeros(s,3);
for i=1:s
    output(i,class(i))=1;
end
net=newff(minmax(input),[15,3],{'logsig','logsig'},'traingdx');
net.trainparam.show=50;
net.trainparam.epochs=500;
net.trainparam.goal=0.05;
net.trainparam.lr=0.01;
net=train(net,input,output');
[t1,t2,t3,t4,t5,t6,t7,c]=textread('test.txt','%f%f%f%f%f%f%f%f',300);
testInput=mapminmax('apply',[t1,t2,t3,t4,t5,t6,t7]',PS);
Y=sim(net,testInput);
[s1,s2]=size(Y);
hitNum=0;
for i=1:s2
[m,Index]=max(Y(:,i));
if(Index==c(i))
  hitNum=hitNum+1;
end
end
sprintf(' the recognition rate is %3.3f%%',100*hitNum/s2)
```

Appendix 2. Part historical data of serious collisions accidents

Ship name	Death	Injures	Damage	B	F	Ship type	Tonnage	Age	Staff	Speed	Ship size	Angle	Position
Santa Maria	0	66	Structural damage to port bow	80	7.6	High speed passenger craft	267	35	255	42	23.93*8.53	40-60	Left port
Funchal	0	67	Structural damage to starboard bow	80	7.7	High speed passenger craft	267	30	255	42	23.93*8.53	40-60	Bow starboard
Neftegaz-67	18	6	Serious damage on starboard	80	19.6	Tugboat	2723	18	23	12	81.37*16.30	90	Bow starboard
No.3 Dae Kyung	7	2	Sank	100	8.45	Fishing boat		11	9	6	29.2*4.87	20	Central hull
Run Ze 001	8	6	Overturn and sink	100	9.85	Cargo ship	1978	5	17	7.5	75.75*14.8	150	Bow
Lu Rong Yu Shui 285	1	10	Serious damage to port bow	80	3	Fishing boat	145	3	11	7.5	32.37*6.20	30	Left port
Zhe Xiang Yu 27009	1	6	Serious damage to starboard bow	80	2.6	Fishing boat	113	1	7	2	35.31*6.30	30	Bow starboard
First Ferry III	0	40	Serious damage to bow	80	5	Passenger craft	451	13	393	20	33.39*10	120	Prow
Hai Bang Da 199	6	5	Disintegrated and sank	100	7.75	Bulk cargo ship	2966	9	11	6	96*15	90	Bow starboard
ANUGRAH 89	11	3	Broke in the middle and sank.	100	12.55	Fishing boat	29	/	14	10	22.1*4.15	90	Starboard amidships
Zhe Sheng Yu 05885	13	2	Broke in the middle and sank.	100	14.45	Fishing boat	240	1	15	9	48.8*6.8	90	Amidships on the port side
CM63963A	6	1	Disintegrated and sank	100	7.35	Fishing boat	227.06	24	7	9	37.8*7.3	60	Bow starboard