Zhou, K, Xing, W, Wang, J, Li, H and Yang, Z

A data-driven risk model for maritime casualty analysis: A global perspective

Reliability Engineering and System Safety 244 (2024) 109925

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

A data-driven risk model for maritime casualty analysis: A global perspective

Kaiwen Zhou [a,1], Wenbin Xing [a,1], Jingbo Wang [a,1], Huanhuan Li [b,*], Zaili Yang [b,*]

[a] School of Computer Science and Artificial Intelligence, Wuhan University of Technology, Wuhan, China
[b] Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool, UK

A R T I C L E I N F O

Keywords:
Maritime safety
Maritime accidents
Maritime casualty analysis
Data-driven bayesian network
Risk prediction

A B S T R A C T

Maritime casualty analysis needs to be addressed given the increasing safety demand in the field due to the accidents' low-frequency and high-consequence features. This paper aims to delve deeper into the factors that affect maritime accident casualties by establishing a new database and conducting an accident casualty evolution analysis. Based on the refined dataset, a pure data-driven Bayesian Network (BN) model is developed to conduct the casualty analysis of maritime accidents that occurred under different ship operational conditions. Methodologically, it introduces new risk factors to improve maritime casualty analysis accuracy through the enriched updated maritime accident database. Furthermore, the new database is categorised into five new datasets based on temporal development trends to better analyse the evolution of the casualty. Five risk analysis models are individually constructed based on different timeframes to illustrate the dynamics of the casualties and compared by seven evaluation indexes to demonstrate the effectiveness of the proposed data-driven BN model. It, for the first time, investigates the changing roles of different risk factors on maritime casualties with time. The insights gained from this model are invaluable, contributing to improved risk prediction and maritime safety strategies by acknowledging the changing patterns of maritime accidents.

1. Introduction

Followed by the widespread adoption of digital technologies, information technologies, and the manufacturing industry, the trend of large-scale, high-speed, and intelligent ships is becoming increasingly prominent [1,2]. However, busy maritime activities inevitably result in complicated maritime traffic situations [3,4], complex encounter situations [5], and intricate traffic patterns [6,7], thus increasing the risk of maritime accidents [8]. Maritime accidents are characterised by low probability and high consequences and often lead to significant casualties and property damage once they occur [9,10]. As an illustration, in 2021, the 'Ever Given' (i.e., a large container ship) was grounded on the banks of the Suez Canal, resulting in a blockage that caused trade losses of approximately £7 billion per day [11,12]. Hence, it is imperative to mitigate navigation risks, prevent maritime accidents, and improve maritime safety.

Many scholars have analysed maritime accidents and the associated risks to assess the emergency hazards of maritime waterways. For instance, Wan et al. [13] established a new model for evaluating maritime supply chain risk factors using Bayesian Networks (BN). Fan and Yang [14] developed a new model based on the Least Absolute Shrinkage and Selection Operator (LASSO) and BN to assess the impact of human fatigue factors on maritime risk. Xin et al. [15] proposed a multi-scale collision risk estimation model based on a graph-based clustering framework to identify the optimal spatial scale for risk assessment. Li and Yang [16] constructed a novel spatiotemporal pattern mining framework based on a new pirate incident database, analysing the influence of various risk factors on pirate attacks. Their work contributes to the rational formulation of anti-piracy measures to ensure maritime safety. Amin et al. [17] put forward a dynamic availability assessment method for the safety of maritime transport operations based on DBN. Cao et al. [18] investigated the accident severity based on the dataset from 2000 to 2019 using a data-driven BN to yield valuable insights. However, this approach incorporates subjective data and relies on experts' subjective experiences, and the accident dataset does not encompass the most recent accidents. Yip et al. [19] used Poisson regression to investigate the determinants of the degree of casualties in passenger ship accidents. Wang et al. [20] applied the Fuzzy Analytic

* Corresponding authors.
E-mail addresses: H.Li2@ljmu.ac.uk (H. Li), Z.Yang@ljmu.ac.uk (Z. Yang).
[1] Equal contribution.

https://doi.org/10.1016/j.ress.2023.109925
Received 14 March 2023; Received in revised form 26 November 2023; Accepted 29 December 2023
Available online 30 December 2023

Hierarchy Process (FAHP) to assess the navigational environmental risk in the shipping lanes of the South China Sea. They pointed out that the maritime ecological risk decreases in a V-shaped spatial distribution from north to south. Wu et al. [21] utilised BN to develop a probabilistic model to estimate the consequences of ship collisions along the Yangtze River. Lan et al. [22] proposed a data-driven approach that combines Association Rule Mining (ARM), Complex Network (CN), and Random Forest (RF) to identify the key risk factors for predicting the severity of ship collision accidents. However, the aforementioned studies reveal significant modelling limitations on single-directional analysis and the inability to expand the established models when new input parameters become necessary and available to add on. To solve this problem, Wang et al. [23] analysed the connection between maritime accident severity and the influential factors using an ordered logistic regression model. Moreover, Wang et al. [24] applied a Zero-Inflated Ordered Probability (ZIOP) model to classify maritime accident severity into two states and explore each factor's impact on the seriousness of maritime accidents. These studies in the field reveal that maritime accident/risk analysis predominantly focuses on identifying risk factors and assessing their combined impact rather than giving strong emphasis to casualty analysis, prediction, and diagnosis. Consequently, they do not effectively address the low-likelihood-high-consequence nature of maritime accidents.

A maritime accident is an initial incident or event, while maritime casualty indicates the specific consequences or outcomes that result from maritime accidents. Due to the global nature of shipping, a lack of comprehensive and up-to-date understanding of maritime casualties on an international scale can hinder the development of effective control measures suitable for international freight shipping. This understanding is crucial to prevent recurring maritime casualties of a similar nature. It highlights a substantial research gap in conducting maritime casualty analysis from a global and up-to-date perspective, involving a full set of all relevant Risk Influencing Factors (RIFs). However, the success of any solution to the gap requires a huge effort in developing a new maritime accident database to address the data limitations exposed by the existing studies and constructing new analysis models that can accommodate both increasing datasets and new RIFs.

The International Maritime Organization (IMO) oversees global shipping safety, security, and environmental protection. The IMO's Global Integrated Shipping Information System (GISIS) is a renowned dataset for maritime accident investigations but has faced criticism for lacking static ship information. To enhance the dataset, it is necessary and beneficial to incorporate complementary information from sources like Lloyd's Register Fairplay (LRF). In this paper, maritime accident data from 2017–2021, along with corresponding accident records, are collected from GISIS. The data with missing ship static information is then supplemented from the LRF database. Based on this comprehensive database, this paper has yielded several valuable contributions to both methodological and applied research, as detailed below.

(1) Development of a novel data-driven maritime casualty analysis model based on the BN method to analyse the causes of accident casualty.
(2) Analysis of maritime casualties from a global and multi-dimensional perspective, involving the features and impact of different regions, ship types, accident types and so on.
(3) Incorporation of new RIFs in maritime casualty analysis for the first time, such as 'information', and 'breadth', as well as new accident type 'occupational accident' and new ship type 'dredger'.
(4) Comparison of five maritime casualty analysis models, generated in line with temporal development trends and accessed using seven evaluation indexes, demonstrates the prediction performance of the proposed data-driven BN model.
(5) Provision of new managerial implications for preventing maritime accident prevention and reducing casualties through comparative analysis with previous studies and the evolution results of five models in this field.

The paper is structured as follows. A comprehensive literature review is presented in Section 2, summarising prior research on maritime casualty analysis and highlighting the current state of the field. Section 3 details the methodology and the comparison of five models. Section 4 presents the model validation content to demonstrate the effectiveness of the data-driven proposed BN model. Section 5 assesses the maritime casualty under different scenarios by scenario simulation and unveils the novel implications. The comparative evolution analysis of five models is listed in Section 6 to provide valuable findings and insights. Lastly, Section 7 concludes with new findings, discusses the limitations, and proposes suggestions for future research.

2. literature review

2.1. Review of maritime casualty analysis

Scholars have investigated maritime casualty in recent years from different perspectives. Wang et al. [23] used an ordered logistic regression model to analyse the global accident investigation reports from 2010 to 2019 and explored the connection between maritime accident severity and RIFs. Cakir et al. [25] utilised a Decision Tree (DT) and a data-driven BN model to analyse 1468 historical ship oil spill accidents to forecast the severity of future ship oil spills. The study revealed that ship type and accident type were the main factors that impacted the severity of ship oil spills. Wang et al. [24] applied a ZIOP model in 1128 maritime accident investigation reports from 2000 to 2019 to classify maritime accident severity into two states and explore the impact of each influential factor. Wang et al. [26] applied density and clustering analysis methods to visualise maritime accident frequency and spatial severity patterns by Geographic Information System (GIS). A zero-truncated binomial regression and a binary logistic regression model were developed by Weng and Yang [27] to predict the likelihood of fatal ship accidents and the corresponding mortality rates. The findings showed that fire/explosion, collision, grounding, sinking, and contact accidents occurring under darkness and adverse weather had higher mortality rates.

Although showing some attractiveness, they initiated the input data from accident reports and took little consideration of the impact of RIFs, hence suffering from the associated constraints in risk data (e.g., ship's static information) and parameters. In addition, the literature review also indicates that the previous maritime accident casualty analysis was conducted against such criteria as ship types, accident types, and areas accident locations (sea areas). The comparative results for these three criteria are presented in Table 1. According to the information in Table 1, 11 studies focus on the impact of ship types on accident severity, 8 studies centre around accident types and their influence on accident severity, and 10 studies explore accident severity concerning locations.

The influence of different ship types on maritime accident casualties exhibits notable variation, encompassing specific ship types and their respective effects on accident severity. While some studies have investigated the accident severity of certain vessel types, many have relied on expert knowledge to compensate for the lack of objective failure data, raising concerns about potential subjective bias in the results. Additionally, due to a lack of sufficient data, researchers have been unable to thoroughly explore the root causes of all accident types, which has hindered the development of comprehensive safety strategies.

Regarding accident types, much of the literature has concentrated on examining the impact of collisions and groundings on accident casualty. It is therefore challenging to explore the underlying causal factors and corresponding countermeasures for various accident types based on the abovementioned studies. The severity and frequency of maritime accidents vary depending on the complexity of traffic and environmental

Table 1
Comparison review of maritime casualty analysis.

Review focus	Research targets	Refs	Research methods
Impact of ship types on maritime casualty analysis	Passenger ships	[19]	Poisson regression
	Passenger ships	[28]	A generic evidence-based framework
	Passenger ships	[29]	Poisson regression and negative binomial
	Tankers	[30]	Statistical analysis method
	Tankers	[31]	A stochastic model
	Ferry ships	[32]	An ordered Probit model and Tobit regression
	Ferry accidents	[33]	Poisson regression
	LNG carriers	[34]	Fault Tree Analysis (FTA) and Fuzzy Evidential Reasoning (FER)
	Cargo vessels	[35]	Binary logistic regression and expert knowledge
	Fishing vessels	[36]	An ordered probit model
	Crew fatalities and ship failures in container transport environments	[37]	Questionnaire survey and sampling technique
Impact of accident types on maritime casualty analysis	Collision accidents	[21]	BN and a probabilistic model
	Collision accidents	[38]	BN
	Collision accidents	[39]	Probability model
	Grounding	[40]	A probabilistic approach
	Collision accidents	[41]	A quantitative procedure
	Fire accidents	[42]	A fire dynamics simulator
	Overboard	[43]	the Poisson model
	Fisheries safety	[44]	Object-Oriented Bayesian Networks (OOBN)
Impact of locations on maritime accident severity	The global area	[18]	BN
	The South China Sea waterway	[20]	FAHP and GIS
	The Istanbul Strait	[45]	Generalised Fuzzy Analytic Hierarchy Process (GF-AHP)
	The Arctic region	[46]	An Improved Fuzzy Analytic Hierarchy Process (IF-AHP)
	Chinese coastal waters	[47]	A data-driven BN
	Tianjin port	[48]	BN
	The Yangtze River	[49]	Formal Safety Assessment (FSA) and BN
	The Yangtze River	[50]	A congestion risk assessment method
	The lower Mississippi River	[51]	A clustering analysis
	Finnish waters	[52]	A hazard identification model
	Hong Kong waters	[53]	A negative binomial regression model

conditions in different seas. When it comes to analysing research areas related to the severity of maritime accidents, a noticeable gap is evident in recent global research. Consequently, the existing literature on maritime accident casualty analysis is often confined to specific regions, waterways, ship types, or particular accident types.

2.2. Comparative analysis of risk influencing factors in maritime accident casualty

Maritime accident casualty analysis is affected by various RIFs. A comparative literature review is carried out to investigate the effectiveness of the different RIFs and capture their current development. The keywords 'maritime accident' and 'severity' or 'casualty' are combined to search the relevant papers on the Web of Science (WoS). To ensure the high quality of the analysed results, only journal papers included in the Science Citation Index Expanded and Social Sciences Citation Index are selected. A total of 199 journal papers related to maritime accident research are obtained through this selection process. Subsequently, 19 papers about maritime accident severity are screened by analysing abstracts and filtering research content and conclusions. Moreover, all the RIFs appearing in the 19 retrieval results are counted quantitatively, and the same RIF is combined. Finally, 51 RIFs in these 19 retrieval results are extracted, and the frequency distribution is shown in Fig. 1. The results indicate that the most frequently appearing 8 RIFs in the 19 references are (1) time, (2) accident type, (3) location, (4) ship type, (5) wind, (6) accident injury severity, (7) visibility, and (8) ship age.

Table 2 lists all the 51 RIFs used in the 19 previous studies, along with the analysis methods employed. The RIFs are numbered from 1 to 51, and each entry is marked with a 'Y' or 'N' to indicate whether the associated RIFs are analysed interdependently through a BN method. The comparison result presented in Table 1 verifies the applicability of these RIFs in the related studies.

Table 3 presents the strengths and weaknesses of all references included in Table 2 from the dataset, the applied method, the research content, and the number of considered RIFs. Out of the 19 studies that are considered, 10 papers focused on examining the severity of accidents using various statistical models, listed in Table 2. It is notable that amongst the nine BN-based studies listed in Tables 2 and 3, only two studies [25,47] have made attempts to conduct data-driven analysis of accident severity without relying on expert knowledge. These two studies include 10 and 19 RIFs, respectively. Furthermore, the three other previous studies using BNs [21,38,48] suffer from such limitations as the use of restricted datasets and a lack of RIFs, potentially impacting the generality of the research findings.

Based on the comparative results, it is essential to develop a comprehensive global dataset and thoroughly investigate the influence of all the relevant RIFs on maritime accident severity from a global perspective.

2.3. The research gaps

The above literature review and comprehensive comparative analysis reveal a few research gaps that need to be addressed to prevent maritime casualties, including

(1) A shortage of effective and comprehensive accident casualty datasets, limited RIFs, or a narrow selection of accident types, rendering it challenging to conduct a comprehensive study on accident casualty analysis. Furthermore, research relying on expert knowledge and judgement could potentially be influenced by subjective biases.

Solution: This paper develops a new dataset combining multiple maritime accident data sources to support the 24 RIFs recommended by the IMO maritime casualty analysis. This dataset, the latest and most efficient for quantitative analysis, incorporates new RIFs and states for a thorough evaluation.

(2) Previous research concentrated on analysing specific criteria for maritime casualty analysis, such as ship types, sea areas, or accident types. Consequently, the findings may not be applicable to other contexts, limiting their practical relevance.

Solution: Leveraging the new maritime casualty dataset, this study undertook a comprehensive examination of the characteristics and nature of maritime accident casualty from a global perspective over the past five years. Furthermore, a comparative analysis with other pertinent literature offers practical guidance for maritime accident casualty analysis based on the obtained results and findings.

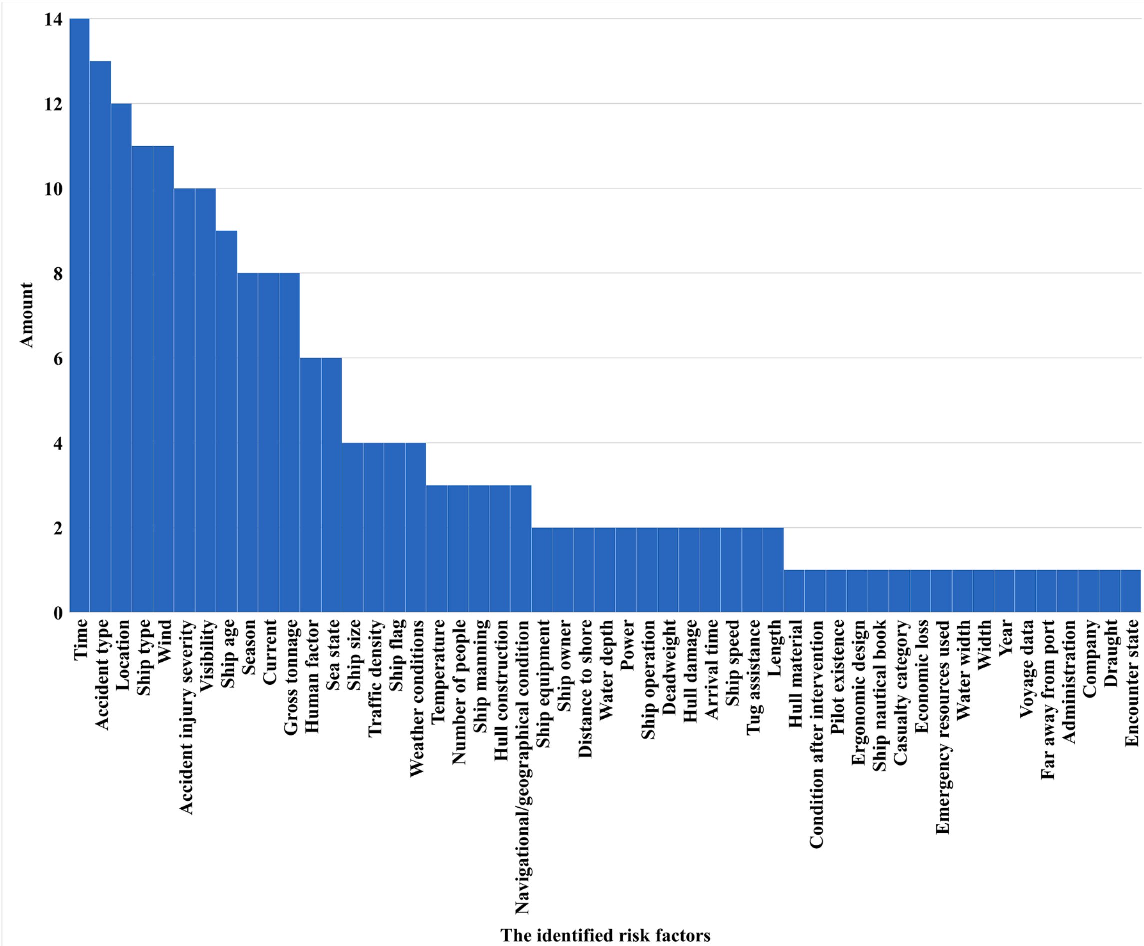

Fig. 1. The identified risk factors result from the literature.

(3) The factors affecting maritime accident casualties are not fully described, and the analysis of the causes of these casualties is not exhaustive.

Solution: This paper captures the features of comprehensive RIFs for maritime accident casualty from a global standpoint, addressing gaps identified in existing literature.

(4) Limited exploration of the effectiveness of influencing factors and the constructed risk analysis model.

Solution: Building upon the latest accident dataset, this paper analyses maritime accident casualties considering 24 influencing factors. The dataset is segmented into five datasets based on temporal development trends. A predictive model is constructed using the BN method, and its performance is evaluated using seven predictive indicators, thereby validating the identified influencing factors' impact.

(5) Insufficient research on the evolution of maritime accident casualties in the field of maritime risk analysis.

Solution: The new maritime accident casualty dataset is divided into five separate subsets according to their chronological development, encompassing the models using the data from 2021, 2020–2021, 2019–2021, 2018–2021, and 2017–2021, respectively. This thorough comparative analysis across the five models deepens the understanding of casualty evolution and brings to light significant findings.

3. Methodology

3.1. The proposed framework

This paper presents a novel framework for the data-driven BN model to perform maritime accident casualty analysis from a multi-dimensional and global perspective, highlighting its theoretical innovation. It is shown in Fig. 2, including five parts: dataset generation, model construction, comparison analysis of models, model validation, and model output. Compared to the traditional application of BN in maritime accident casualty analysis, this paper further demonstrates the efficacy of the constructed model and the identified factors in the comparison analysis of the models part. To do so, the original dataset spanning from 2017 to 2021 is categorised into five distinct datasets as follows:

(1) Accident records for the year 2021.
(2) Accident records covering the period from 2020 to 2021.
(3) Accident records spanning from 2019 to 2021.
(4) Accident records for the years 2018 to 2021.
(5) The comprehensive set of accident records from 2017 to 2021.

The prediction performance of the constructed models is evaluated by seven indexes to visualise the results. Furthermore, the comparative evolution analysis of these five models is listed clearly to reveal the invaluable insights.

Table 2
The RIFs comparison in different references.

Refs.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
[18]																		Y	Y							Y
[21]				Y				Y		Y			Y	Y			Y									
[23]												N								N						N
[24]			N																	N						N
[25]																					Y					
[27]																										
[29]																		N							N	
[35]							Y																Y		Y	
[36]	N																				N				N	
[38]				Y	Y			Y								Y										
[42]																										
[43]						N																				
[45]									N																	
[46]															N								N			
[47]		Y			Y						Y								Y					Y		
[48]										Y														Y		
[49]																						Y				
[50]																						Y				
[53]																										

Refs.	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
[18]				Y	Y		Y		Y	Y		Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y		
[21]		Y							Y				Y		Y	Y	Y	Y							
[23]							N	N		N		N	N		N	N			N						
[24]				N	N		N	N	N		N	N	N	N	N	N	N	N	N						
[25]				Y		Y					Y	Y		Y		Y	Y	Y	Y						
[27]	N	N												N		N	N	N	N						
[29]			N	N				N	N			N	N	N	N	N	N	N	N						
[35]						Y		Y		Y		Y	Y		Y		Y				Y				
[36]							N		N		N	N		N	N			N	N						
[38]	Y													Y				Y							
[42]		N	N			N											N	N							
[43]			N				N				N		N		N				N						
[45]					N	N				N															
[46]	N						N	N		N					N		N								
[47]	Y				Y			Y	Y		Y	Y	Y			Y		Y	Y		Y			Y	Y
[48]											Y			Y		Y	Y	Y	Y						
[49]									Y	Y	Y	Y	Y		Y	Y		Y	Y						
[50]									Y	Y	Y		Y	Y	Y		Y	Y	Y						
[53]														N		N	N		N						

Note: 'Y' and 'N' indicate the use of a BN method or not, respectively. Numbers 1–51 RIFs are 1. Year; 2. Width; 3. Water width; 4. Tug assistance; 5. Ship speed; 6. Casualty category; 7. Ship nautical book; 8. Arrival time; 9. Pilot existence; 10. Hull damage; 11. Hull material; 12. Far away from port; 13. Emergency resources used; 14. Economic loss; 15. Ergonomic design; 16. Deadweight; 17. Condition after intervention; 18. Ship operation; 19. Power; 20. Water depth; 21. Distance to shore; 22. Ship owner; 23. Ship equipment; 24. Length; 25. Hull construction; 26. Ship manning; 27. Weather conditions; 28. The number of people; 29. Temperature; 30. Ship flag; 31. Traffic density; 32. Ship size; 33. Sea state; 34. Human factor; 35. Gross tonnage; 36. Current; 37. Season; 38. Ship age; 39. Visibility; 40. Accident injury severity; 41. Wind; 42. Ship type; 43. Location; 44. Time; 45. Accident type; 46. Voyage data; 47. Navigational/geographical condition; 48. Administration; 49. Company; 50. Draught; 51. Encounter state.

Table 3
Comparative analysis of maritime accident severity studies.

Refs.	The method	Advantages	Disadvantages
[23]	An ordered logistic regression model	Identify influential factors from primary data	Data limitation (i.e., more on objective variables and less on human factors)
[21]	(1) NBN (2) Consider the expertise	Consider emergency management when predicting the consequences of maritime accidents	(1) Limited scope of applications (i.e., collision) (2) Data limitation (134 records)
[24]	A ZIOP model	Investigate injury severity in maritime accidents	Seven accident types only
[25]	(1) TAN (2) Purely data-driven	Analyse the severity of an oil spill at sea	(1) Limited attribute variables and data collection (2) A single case application (oil spills accident)
[27]	Binary logistic and zero-truncated binomial regression models	Propose a new model for the severity of injuries in shipping accidents	(1) Data limitation (no shipping traffic information) (2) The exclusion of shipping traffic information
[29]	Negative binomial and Poisson regression	Investigate passenger ship accidents	A few selected ship types
[35]	(1) NBN (2) Consider the expertise	Develop a risk assessment model using BN, logistic regression, and expert judgement.	A few selected ship types
[36]	An order probability model	Estimate the severity equations of crew injury and ship damage	Only in the northeastern United States
[38]	Bayesian belief network (BBN)	Introduce a systematic, transferable and proactive framework to estimate the risk for maritime transportation systems	(1) A single case (only the severity of collision) (2) Consider subjective input data
[42]	A probabilistic method	Estimate the death rate of ship fires	Only consider ship fires
[43]	A Poisson model	Analyse the impact of life jacket wear on the mortality rate of life-threatening events	A single case application
[45]	A FAHP model	Maritime risk perception while comparing it with 232 maritime accident statistics	Localisation analysis (only in the Istanbul Strait)
[46]	An IF-AHP model	Define navigational risk factors of the Arctic Ocean	Contain expert consultations and subjective input data
[47]	(1) TAN (2) Purely data-driven	Construct data-driven BN using advanced machine learning methods	(1) Data limitation (no human factors) (2) Localisation analysis (a single scenario)
[48]	(1) BBN (2) Consider the expertise	Combine statistical data and expert knowledge to evaluate probability distributions	Data limitation (234 accidents occurred in the waters of Tianjin Port between 2008 and 2013)
[49]	(1) NBN (2) Consider the expertise	Adopt a structured and systematic approach to risk assessment for FSA	Only in the Yangtze River
[50]	(1) BN (2) Consider the expertise	Propose a BN method based on an accident database to evaluate the congestion risk of inland waterways	(1) Localisation analysis (only in one region) (2) Data validity (adjust objective data using expert experience)

(continued on next page)

Table 3 (*continued*)

Refs.	The method	Advantages	Disadvantages
[53]	A negative binomial regression	Analyse traffic accident risks in Hong Kong ports	Localisation analysis (only one region)
[18]	(1) TAN (2) Purely data-driven	Apply a data-driven BN to analyse the relationship between accident severity and influencing factors	(1) Contain subjective data (2) Consider the experts' subjective experience

3.2. Data collection and cleaning

This study gathers information from two authoritative and reliable maritime accident databases, namely GISIS and LRF. GISIS provides comprehensive accident data records that include ship details, accident cause records, weather information, and accident reports. However, some static ship information is missing. To address this gap, the LRF database is utilised to supplement and complete the essential missing data. The detailed clean dataset generation steps are shown below.

(1) Original data collection from the GISIS database.

The basic information on maritime casualty accidents is collected from the GISIS database, including the location, ship name, time of the accidents, accident type, and a brief description of their causes. Moreover, detailed accident information is received from the reporting forms in each record in GISIS, such as ship identification and particulars, data on very serious and serious casualties, fire casualty records, and life-saving appliance casualty records.

(2) Further investigation in the GISIS database.

The relevant maritime accident reports in the GISIS are further analysed to derive more comprehensive information like accident details, casualty descriptions, accident cause analysis and descriptions.

(3) Data cleaning.

A total of 1105 maritime accident records are compiled from the GISIS database from 2017 to 2021. To ensure the accuracy of the dataset, any records without necessary IMO or an Mobile Service Identification (MMSI) number or limited information are excluded from the study. After the initial stage of data cleaning, 462 accident records remained.

(4) Add missing static information from the LRF database.

Out of these, ships' basic information and physical parameters are added from the LRF database to generate 428 records using the ship's IMO and MMSI numbers to ensure the consistency of data from the two data sources, including details such as the ship's age, breadth, length,

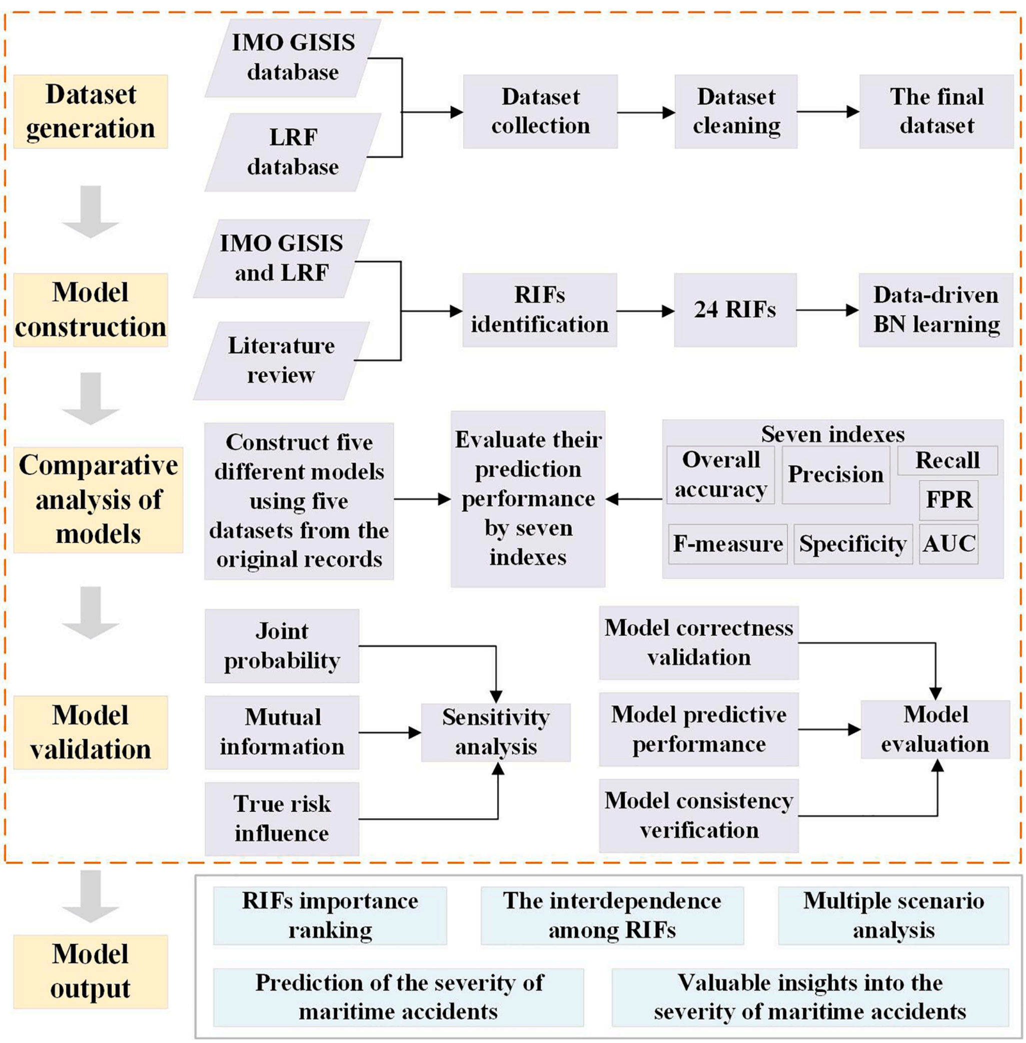

Fig. 2. The proposed framework.

type, deadweight, gross tonnage, and hull structure.

(5) Final filtering.

The dataset is further filtered to remove records with unclear information on the causes, environmental conditions, and ship equipment. This step is crucial in ensuring the validity and accuracy of the dataset. Ultimately, a total of 402 maritime accident records are used in this study.

3.3. RIF identification

The frequency of the RIFs presented in Table 2 and the comparative analysis in Table 3 highlight the innovation of this study, which develops a purely data-driven BN model for analysing maritime accident severity. This paper establishes and standardises a list of RIFs by analysing the most commonly used RIFs in literature and their relationship with the 24 risk factors that the IMO recommends as influential in maritime casualties. Additionally, a maritime accident severity database is developed using the GISIS and LRF databases to investigate the combined impact of various RIFs on a global scale. This model can guarantee that no relevant RIF is omitted due to data limitations. The use of accurate data and comprehensive RIFs ensures the robustness of the BN model.

Furthermore, previous studies simplified the state definitions of some involved RIF to avoid extensive requirements on data availability and computation when quantifying their interdependencies [38,45,46]. However, this approach may compromise the sensitivity of the BN model and result in inaccurate and insensitive outcomes. To address this problem, the current study redefines and standardises the states of RIFs based on the IMO standards. For example, the study defines ship types into 11 groups, including 'RORO', 'dredger', 'tug', 'container ship', 'cargo ship', 'bulk carrier', 'offshore ship', 'fishing ship', 'passenger ship', 'tanker' or 'chemical ship', and 'others'. It ensures that the accident severity analysis covers a wide range of ship types and enhances the comprehensiveness of the BN model. In addition, the detailed classification given by the IMO is applied for the voyage segment, containing 11 states (e.g., anchorage, archipelagos, at berth, etc.). Lastly, the definitions and status of all identified RIFs in this study are presented in Table 4, which are highly consistent with the IMO GISIS form, hence providing a standard for maritime accident casualty analysis in future. Therefore, the following 24 RIFs are adopted to explore their collective impact on maritime accident severity, generating new findings more compelling than previous studies based on the recommendation from

Table 4
Definition and states of 24 RIFs for maritime accident severity [54].

No.	Main factors	Classification	RIFs	Definitions	States
1	Accident	Accident type	Type of accident	Capsize, collision, contact/crush, fire/explosion, flooding, grounding, occupational accident, overboard, ship/equipment damage, sinking, others	1,2,3,4,5,6,7,8,9,10,11
2		Time	Time of day	(07:00,19:00], (19:00,07:00]	day, night
3	Environment	External environment	Wind (Beaufort scale)	>5, (0,5]	high, low
4			Visibility (nm)	>2, (0,2]	good, bad
5			Weather condition	Considering rain, visibility, wind, fog, and extreme weather	good, bad
6			Sea condition	Considering waves, currents, sea state, and falling or rising tide	good, bad
7	Ship	Ship details	Ship type	Bulk carrier, cargo ship, container ship, dredger, fishing ship, offshore ship, passenger ship, RORO, tanker or chemical ship, tug, others	1,2,3,4,5,6,7,8,9,10,11
8			Ship age (years)	(0,5], [6,20], >20, NA	1,2,3,4,5,6
9			Length (meters)	(0,100], (100,200], >200	1,2,3
10			Breadth (meters)	(0,20], (20,30], (30,40], >40	1,2,3,4
11			Hull type	Aluminium alloy, composite materials, GRP, light alloy, steel, wood, NA	1,2,3,4,5,6,7
12			Hull construction	Double bottom, double hull, single hull	1,2,3
13			Gross tonnage (GT)	(0,3000], (3000,10,000], (10,000,20,000], >20,000	1,2,3,4
14			Deadweight (DWT)	(0,5000], (5000,15,000], (15,000,30,000], >30,000	1,2,3,4
15			Draught (meters)	(0,6], (6,9], >9	1,2,3
16			Power (kW)	(0,3000], >3000	1,2
17			Vessel condition	Good state of ships or the accident is unrelated to the ship state; Bad state of ships (e.g., ship equipment failure and ship design errors)	good, bad
18	Navigation	Voyage particulars	Voyage segment	Anchorage, archipelagos, at berth, canal, channel, coastal waters, inland waters, open sea, port, port approach, river	1,2,3,4,5,6,7,8,9,10,11
19			Ship operation	At anchor, fishing, loading/unloading, on passage, manoeuvring, pilotage, towing, others	1,2,3,4,5,6,7,8
20			Ship speed (knots)	(0,6], (6,12], >12	low, middle, high
21			Information	Providing effective and updated information; Lack of effective and updated information (e.g., failure to send signals or respond appropriately, inadequacy of navigational equipment, unreliable and poor chart data, etc.)	good, bad
22		Equipment/ design	Equipment	The ship's equipment is in good state and is being operated correctly; Bad state involving malfunction or improper functioning (e.g., deactivation of alarm systems, unnoticed alarms, propulsion machinery failures, electrical installation failures, etc.)	good, bad
23			Ergonomic design	The ergonomic design of the ship is unrelated to the occurrence of accidents; Unfriendly ergonomic design (e.g., insufficient stability, poor bridge ergonomics, etc.)	good, bad
24	Human	Human factor	Human factor	The accident has nothing to do with human factors; Human errors or violations (e.g., stress, fatigue, inadequate familiarity or training, poor management, insufficient supervision, error in judgement, etc.)	no, yes

the IMO dataset.

3.4. Model construction

A BN was proposed by Pearl in 1988 [55] and is usually represented as a Directed Acyclic Graph (DAG) containing a set of nodes and directed edges connecting these nodes. Due to its powerful reasoning and learning capabilities, BN is widely used as a risk assessment method in maritime safety research [56]. It combines probabilistic knowledge with graph structures and visually demonstrates the interrelationships between variables. There are two main methods to configure a BN structure, a data-driven approach or a subjective approach based on the causal relationship between the variables/nodes. Comparatively, the data-driven approach is lately better used due to its objectivity and availability of relevant data.

The BN structure in this study is built based on a data-driven approach. Several types of BN classifiers have been used to train a data-driven BN structure, including Naive Bayesian Networks (NBN), Augmented Bayesian Networks (ABN), and Tree Augmented Naïve Bayesian (TAN). The NBN structure assumes conditional independence between each factor, but this is not always accurate in reality, which limits its practical applications of BN structures. The disadvantages of ABN include increased complexity, high data requirements, computationally intensive training, increased risk of overfitting, and reduced interpretability. The TAN structure is an improvement over NBN as it relaxes the independence assumption and better reflects feature dependencies. TAN enhances overall performance by allowing attribute nodes to be dependant on no more than one non-class node, simultaneously preserving the robustness and computational ease of NBN. Consequently, TAN is often used to construct reliable assessment models for maritime safety from large datasets.

In this study, TAN is chosen to train the BN for maritime accident severity analysis. A maritime accident severity BN model with 24 nodes is constructed based on the TAN mechanism. The Conditional Probability Table (CPT) for each node is obtained through parameter learning in this study, with reference to [57]. The constructed maritime accident severity BN model is displayed in Fig. 3.

In Fig. 3, the 24 individual boxes around the edges represent the 24 RIFs, functioning as child nodes in the BN. The central box symbolises the research target, which is the type of casualty, serving as the parent node in the BN. On the left side of each box, the state divisions for the corresponding RIF are listed. For instance, 'type of casualty' is categorised into three states: 'less serious,' 'serious', and 'very serious'. On the right side of the boxes, the numbers and black bars indicate the probability values for each state. For RIFs with numerical states like 'draught', displayed as '1.8 ± 0.83', the first number 1.8 denotes the mean, and the second number 0.83 signifies the standard deviation, reflecting the dispersion of the RIF state values around the mean. These two numbers collectively express the uncertainty in the RIF state's probability distribution.

The marginal probability of each child node can be calculated based on the CPT of each node [57]. The calculation follows Bayesian rules and can be simulated using software packages such as Netica. Fig. 3 indicates that given the global maritime accident severity data from 2017 to 2021,

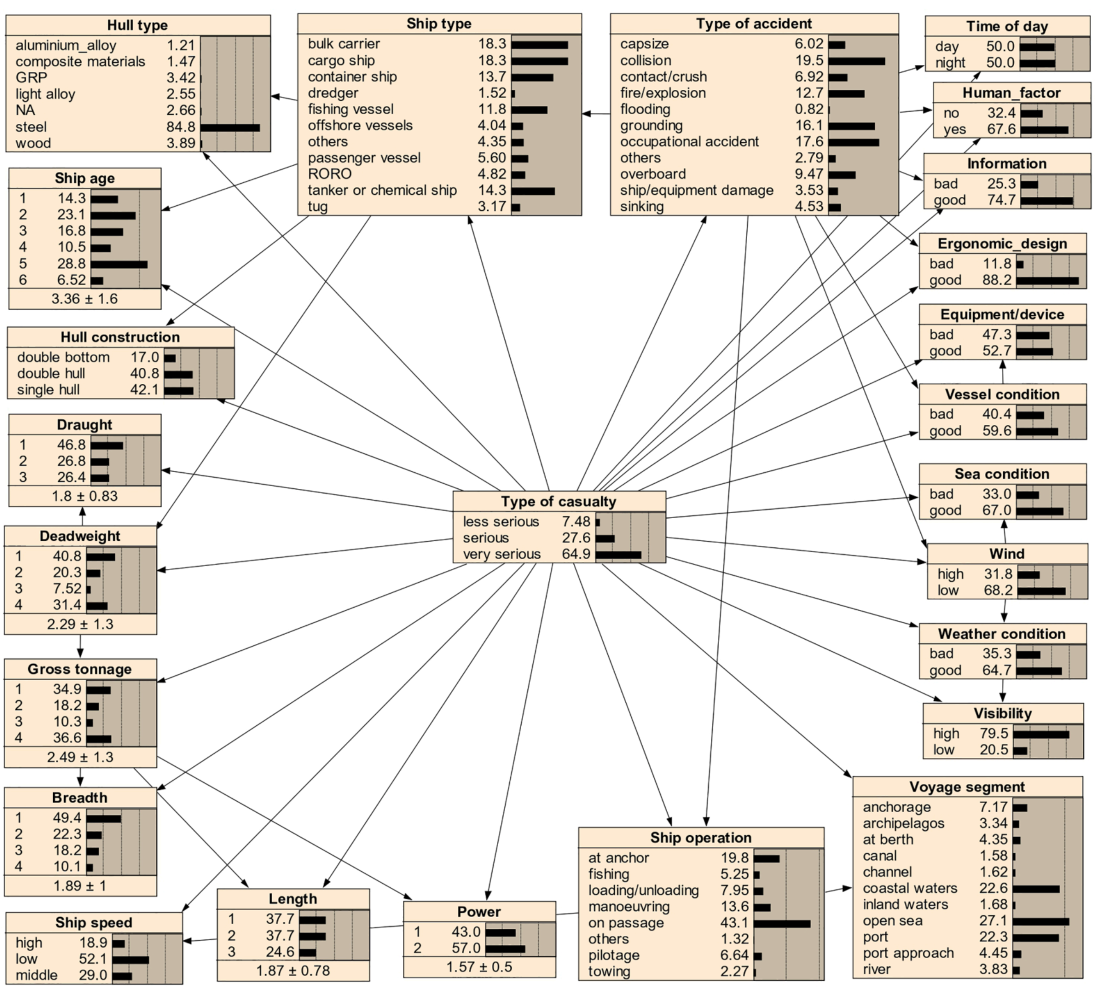

Fig. 3. The constructed BN model for maritime accident casualty.

the accident rates of ‘less serious’, ‘serious’, and ‘very serious’ were 7.48%, 27.6%, and 64.9%, respectively. To test the initial statical reliability of the model, the rates of each type of maritime casualty are directly computed using the raw data. The basic statistical analysis results are 7.46%, 27.61%, and 64.93% for ‘less serious’, ‘serious’, and ‘very serious’. The model’s initial test shows that the two sets of results keep a high-level consistency, revealing the validity of the model at this stage.

3.5. Comparison of five different models

To further demonstrate the effectiveness of the proposed model and the selected 24 RIFs, the five models by the temporal development trends (i.e., 2021, 2020–2021, 2019–2021, 2018–2021, and 2017–2021) are constructed and compared. This paper introduces a confusion matrix and seven indicators to evaluate the prediction performance of five different models comprehensively. The confusion matrix is displayed in Fig. 4, which is a standard format for representing accuracy evaluation, also known as the error matrix. The overall accuracy can be obtained by a confusion matrix, and generally, a higher accuracy indicates a better predictive classification. However, it is not scientific and rigorous to evaluate the model by accuracy alone.

80% of the accident records in the dataset are randomly selected to train the model, and the remaining 20% are utilised as the test dataset. To comprehensively assess the five models’ prediction performance, six evaluation metrics are also utilised based on the confusion matrix, including precision, recall, F-Measure, specificity, False Positive Rate (FPR), and Area Under Curve (AUC). The definitions of these seven indexes are listed in Table 5 to provide a clear understanding. To assess their performance, Table 6 displays the prediction results of five models based on these seven indices.

The prediction accuracy of five different models is assessed through seven distinct indexes, with the corresponding confusion matrices provided in Appendix I. Impressively, each model’s overall accuracy rate surpasses 90%, underscoring the effectiveness of the developed models and the 24 chosen RIFs. The models display consistent performance over various timeframes, indicating their stable predictive capabilities. Minor fluctuations in performance metrics over different years might indicate the models’ adaptability to varying data quantities and temporal shifts. The uniformly high scores on critical performance metrics confirm the robustness and dependability of the models for forecasting the severity of maritime accidents. The models used for maritime accident casualty prediction show excellent performance, with high accuracy and strong metrics across precision, recall, F-measure, specificity, FPR, and AUC. This suggests that the selected RIFs are effective predictors of maritime casualty severity.

To better present the results of the five-year model in our study, the confusion matrix and the accuracy rate are listed in Table 7. The prediction accuracy of the three types of casualty is 100%, 81.48% and 95.74%, respectively. Therefore, it is confirmed that the model can reasonably predict maritime accident severity levels.

4. Model validation

Along with the above marginal probability test and predictive performance analysis, the constructed model’s correctness is further

	Actual Positive	Actual Negative
Predicted Positive	True Positive (T_P)	False Positive (F_P)
Predicted Negative	False Negative (F_N)	True Negative (T_N)

Fig. 4. The confusion matrix.

validated by the following three methods [58,59].

4.1. Sensitivity analysis

Sensitivity analysis is a widely employed technique to assess uncertainty in maritime safety, aimed at pinpointing and measuring the impact of sensitive factors amongst various variables on key performance indicators. In this study, a combination of mutual information, joint probability distributions, and True Risk Influence (TRI) methods are employed to execute sensitivity analysis [60].

4.1.1. Mutual information

The degree of dependence between two random variables can be measured using mutual information, which can be represented as $I(X,Y)$. A higher mutual information value indicates a stronger correlation between the two random variables [61]. It signifies that the child nodes with higher mutual information values exert a greater influence on the parent node ‘Type of casualty’.

The calculated mutual information values, entropy reduction percentages, and belief variances are shown in Table 8. Based on Table 8, ‘Type of accident’ has the most significant influence on the accident severity, with a mutual information value of 0.2686. The mutual information values of each RIF and the trend of changes are shown in Fig. 5. The trend of entropy reduction is illustrated by the orange line, while the mutual information values are represented by the blue bars. Entropy reduction can measure the reduction in uncertainty within a system when an influential factor is added. A lower entropy reduction suggests a lesser impact of that variable on the target node in a BN. Variables are chosen based on a threshold criterion of 2% for entropy reduction, selecting those factors that exhibit an entropy reduction greater than 2%. As a result, the top 8 RIFs, including the ‘type of accident’ (22.4%), ‘ship operation’ (9.72%), ‘ship type’ (7.06%), ‘voyage segment’ (6.13%), ‘hull type’ (3.47%), ‘gross tonnage’ (2.46%), ‘length’ (2.12%), and ‘hull construction’ (2.11%), are identified to have significant impacts on accident severity.

4.1.2. Joint probability distribution

After the eight significant RIFs are screened from the mutual information calculation, the joint probability distribution is used to analyse the level of influence of different states and/or variables on the accident severity. To ensure the correctness of the probability distribution in the BN network, a normalisation condition is applied to the network. Simultaneously, this condition enables the computation of the posterior probability of any variable and facilitates Bayesian inference. Table 9 lists the results obtained by setting the probability of each state for each RIF independently to 100% to calculate the joint probabilities.

The probabilities for three states of accident casualty under the influence of different RIFs are presented in Table 9, with bolded values indicating the states where each RIF has the greatest and least impact on each type of accident severity. The first row depicts the initial probabilities before the RIFs’ states are adjusted individually. The rows in Table 9 illustrate the change in probability values for each type of casualty when a specific RIF is fixed in a particular state.

Table 9 reveals new findings and patterns. When the ‘type of accident’ is ‘occupational accident’, there is a high probability of a ‘very serious’(96.9%). Conversely, when the state ‘contact/crash’ in ‘type of accident’, the probability of a ‘very serious’ casualty is low (25.4%) while the probability of a ‘less serious’ casualty is high (35.1%). Regarding ship type, a ‘fishing vessel’ is most likely to cause a ‘very serious’ casualty (84.7%) and least likely to cause a ‘less serious’ casualty (2.5%). On the other hand, a ‘RORO’ has the lowest probability of causing a ‘very serious’ casualty (19.6%). When considering ‘hull construction’, a ship with a ‘single hull’ has a 69.0% chance of causing a ‘very serious’ casualty. Conversely, a ship with a ‘double bottom’ construction has the highest likelihood of causing a ‘less serious’ casualty, with a probability of 10.6%. It therefore provides empirical evidence of

Table 5
The seven indexes for evaluating the prediction performance of the constructed model.

Indexes	Equations	Definitions
Overall accuracy	$\frac{T_P + T_N}{T_P + F_P + T_N + F_N}$	The proportion of the entire sample that is accurately predicted.
Precision	$\frac{T_P}{T_P + F_P}$	The probability of a positive sample within the set of all samples predicted as positive.
Recall/ Sensitivity	$\frac{T_P}{T_P + F_N}$	The probability of receiving a positive prediction amongst the samples that are actually positive.
F-measure	$\frac{2 * precision * recall}{precision + recall}$	The overall average distribution.
Specificity	$TPR = \frac{T_N}{F_P + T_N}$	The ratio of correctly predicted negative samples to all actual negative samples.
FPR	$1 - TPR = \frac{F_P}{F_P + T_N}$	The proportion of samples predicted as positive amongst the actual negative samples.
AUC	It can be calculated from the Receiver Operating Characteristic (ROC) curve.	The area is encompassed by the ROC curve.

Table 6
The prediction performance comparison of five models based on seven indexes.

Datasets (data volume)	Overall accuracy	Indexes	Type of casualty		
			Less serious	Serious	Very serious
2021 (11)	100%	Precision	–	1	1
		Recall	–	1	1
		F-measure	–	1	1
		Specificity	–	1	1
		FPR	–	0	0
		AUC	–	1	1
2020–2021 (65)	92.86%	Precision	1	1	0.875
		Recall	1	0.8333	1
		F-measure	1	0.9167	0.9375
		Specificity	1	1	0.8571
		FPR	0	0	0.1429
		AUC	1	1	1
2019–2021 (140)	92.86%	Precision	1	0.8333	1
		Recall	1	1	0.8667
		F-measure	1	0.9167	0.9334
		Specificity	1	0.8889	1
		FPR	0	0.1111	0
		AUC	1	1	1
2018–2021 (258)	90.38%	Precision	1	0.8	0.9697
		Recall	1	0.9231	0.9143
		F-measure	1	0.8616	0.9420
		Specificity	1	0.9231	0.9412
		FPR	0	0.0769	0.0588
		AUC	1	0.9941	0.9891
2017–2021 (402)	91.25%	Precision	1	0.9167	0.9
		Recall	1	0.8148	0.9574
		F-measure	1	0.8627	0.9278
		Specificity	1	0.9623	0.8485
		FPR	0	0.0377	0.1515
		AUC	1	0.9476	0.9745

Table 7
The predicted results of different kinds of accident severity.

Predicted	less serious	serious	very serious	Actual total	Accuracy rate (100%)
less serious	6	0	0	6	100
serious	0	22	5	27	81.48
very serious	0	2	45	47	95.74
Predicted total	6	24	50	80	91.25

the effect of ‘double hull’ on maritime safety.

The latest research reveals how specific states of individual variables can affect maritime accident casualty, as well as the likelihood of different levels of casualties depending on the state of these variables. It is worth noting that the single variable state with the highest probability value under a given casualty warrants greater attention, as it denotes a higher likelihood of occurrence of the corresponding casualty.

Table 8
Different results between ‘Type of casualty’ and RIFs.

Variable	Mutual Information	Entropy Reduction Percent	Variance of Beliefs
Type of casualty	1.1973	100	0.2887
Type of accident	**0.2686**	**22.40**	0.0547
Ship operation	**0.1164**	**9.72**	0.0178
Ship type	**0.0846**	**7.06**	0.0163
Voyage segment	**0.0734**	**6.13**	0.0110
Hull type	**0.0416**	**3.47**	0.0068
Gross tonnage	**0.0295**	**2.46**	0.0047
Length	**0.0254**	**2.12**	0.0029
Hull construction	**0.0252**	**2.11**	0.0063
Human_factor	0.0234	1.95	0.0057
Power	0.0202	1.68	0.0037
Ship age	0.0187	1.56	0.0038
Deadweight	0.0184	1.53	0.0028
Ergonomic_design	0.0181	1.51	0.0043
Breadth	0.0169	1.42	0.0029
Draught	0.0144	1.20	0.0013
Ship speed	0.0131	1.10	0.0026
Information	0.0073	0.612	0.0019
Vessel condition	0.0036	0.303	0.0009
Visibility	0.0019	0.158	0.0002
Equipment/device	0.0019	0.157	0.0003
Weather condition	0.0017	0.146	0.0002
Time of day	0.0013	0.106	0.0001
Wind	0.0006	0.048	0.0001
Sea condition	0.0002	0.0193	0.00003

4.1.3. *True risk influence*

TRI is an effective analytical method for testing multivariate sensitivity proposed by Alyami et al. [60] and is popularly used in maritime safety research. As the RIFs in this study have multiple states, the TRI method is utilised to assess the influence of the significant RIFs on the severity of each accident. The TRI values for each significant RIF on accident severity can be calculated by the initial probability values and the bold values in Table.

Taking the TRI value of ‘type of accident’ on ‘less serious’ as an example, in the second column of the ‘type of accident’ section of Table 10, the maximum and minimum probability values are 35.101 and 0.400, respectively. The difference between the original probability values of 7.482 and 35.101 is the High-Risk Inference (HRI) (i.e., 27.619). The difference between the original probability value of 7.482 and 0.400 is the Low-Risk Inference (LRI) (i.e., 7.082). The TRI of the RIF for the type of casualty is 17.4. The remaining RIFs’ TRI on the type of casualty can be calculated similarly, and the results are shown in Table 10. According to the average column in Table 10, the eight important RIFs can be ranked by the magnitude of their impact on

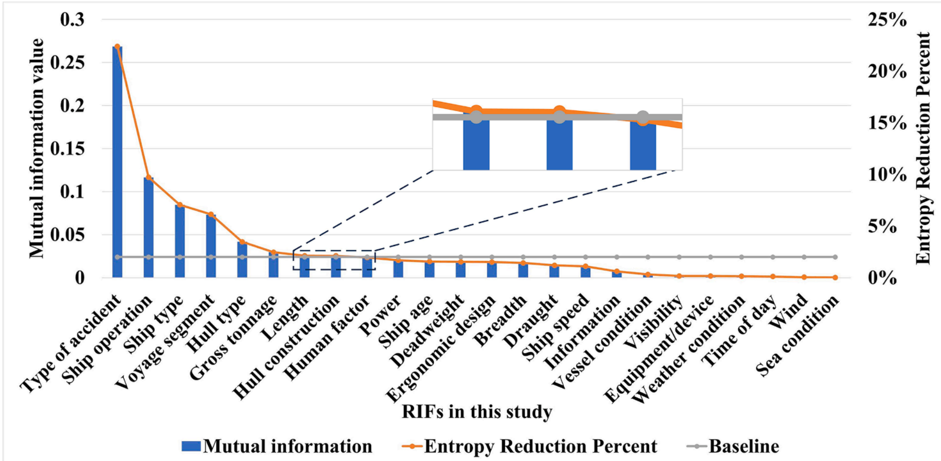

Fig. 5. Mutual information values and changes.

maritime accident casualty. Amongst the eight factors, the type of casualty is most significantly influenced by the 'type of accident', while the 'length' has the least impact. The ranking results are as follows:

Type of accident > Voyage segment > Ship type > Hull type > Ship operation > Gross tonnage > Hull construction > Length

Table 11 displays the descending order ranking of TRI values for each accident severity by the results in Table 10, ranging from 1 to 8. It can be observed that the RIFs have varying levels of influence on different types of casualty. For instance, 'ship type' has the most significant impact on 'serious' type and is less critical for other types of casualty. On the other hand, the 'voyage segment' has the greatest impact on the 'less serious' and 'very serious' casualty and is less critical for the 'serious' type.

4.2. Model correctness validation

To further validate the correctness of the constructed BN-based model, an additional sensitivity analysis is conducted for the given RIFs by analysing the model's results. The sensitivity analysis reasoning process should satisfy at least the following two axioms:

Axiom1: A minor increase or decrease in the prior probabilities of each RIF should contribute to the correspondence increase or decrease in the posterior probability of the target node.
Axiom 2: The total impact of integrating the probability variations of x parameters should be larger than the one from the set of y ($y \in x$) RIFs.

To prove that the model satisfies the two axioms, the eight significant RIFs are taken as a set of variables to examine the combined impact of multiple variables on the type of casualty. Given that the parent node has multiple states, this study computes the variation of probability values for each state. Taking the 'less serious' casualty as an example, 'hull construction' is chosen as the first node, and its prior probability is increased by a 2% step to reach two extreme states with the maximum and minimum influence on the 'less serious', respectively. Similarly, the process is then applied to the other seven RIFs. The cumulative probability change values are obtained in the order of 'hull construction', 'length', 'gross tonnage', 'hull type', 'voyage segment', 'ship type', 'ship operation', and 'accident type'. Finally, the above process is used for the two remaining accident casualty types to acquire the results shown in Table 12.

The second column of Table 12 displays the original probability values for each type of casualty in the TAN structure, with the updated cumulative change values presented in the remaining columns. The results demonstrate that an increase or decrease in the prior probability of the chosen RIF correspondingly leads to an increase or decrease in the posterior probability of the type of casualty, which verifies Axiom 1. In addition, the last three rows in Table 12 show that the cumulative probability change values of the parent node increase in sequence as the number of variables changing increases, testing Axiom 2. Therefore, the results prove the correctness of the constructed model.

4.3. Model consistency verification

The sample of maritime accident casualty in this study is unevenly distributed. The 'very serious' type accounts for 64.9% of accident severity, while the 'less serious' accounts for only 7.48%. Cohen's kappa coefficient is therefore introduced to test the consistency of the model in predicting the severity of various types of accidents [54].

The Kappa coefficient is calculated between [−1,1] but usually between [0,1]. When the kappa value falls in the range of [0.81,1], it can be inferred that the model predicts the maritime accident casualty with almost identical accuracy. Based on the above formula and the confusion matrix, the kappa coefficient is calculated as 0.83363. Therefore, the kappa coefficient falls between [0.81,1], proving the strong consistency of the model.

5. Results and implications

5.1. Scenario analysis

This study conducted a comprehensive risk analysis to investigate the impact of different RIFs on maritime accident casualty. The BN model constructed based on TAN can simulate various types of casualty and derive the corresponding RIF states. To illustrate this in detail, multiple scenario analyses are conducted to reveal the casualty of a particular accident occurring under a specific scenario with multiple RIFs combined.

5.1.1. Scenario one: the combined impact of important RIFs

The BN model enables the investigation of the combined influence of eight important RIFs (i.e., type of accident, voyage segment, ship type, hull type, ship operation, gross tonnage, hull construction, and length) on maritime accident casualty. Scenario one simulates the most likely accidents to cause a 'very serious' severity. By using the joint probability table of important RIFs, the state with the most significant impact on 'very serious' type is set to a 100% probability. The resulting change in the probability of 'very serious' casualty is shown in Fig. 6, where the probability value increases significantly from 64.9% to 99.7% compared to the initial state. This variation indicates that the casualty of an occupational accident is most likely to be 'very serious' (e.g., crew casualty) when a wooden, single-hull fishing vessel is engaged in loading/

Table 9
The joint probability of each variable and accident severity.

original	less serious 7.482	serious 27.616	very serious 64.902
Type of accident			
capsize	**0.400**	4.505	95.095
collision	8.737	36.670	54.592
contact/crush	**35.101**	39.502	**25.397**
fire/explosion	5.895	44.980	49.126
flooding	2.952	33.248	63.800
grounding	9.124	**52.231**	38.644
occupational accident	1.500	1.536	**96.965**
others	0.862	36.174	62.965
overboard	5.335	**0.260**	94.404
ship/equipment damage	7.506	42.625	49.868
sinking	0.531	5.983	93.486
Ship operation			
at anchor	**1.865**	25.444	72.691
fishing	6.199	15.815	77.986
loading/unloading	5.522	**5.544**	**88.934**
manoeuvring	17.439	41.738	40.823
on passage	4.158	26.613	69.229
others	10.964	15.244	73.793
pilotage	**25.208**	**56.343**	**18.449**
towing	16.037	8.886	75.078
Ship type			
bulk carrier	4.813	29.435	65.752
cargo ship	8.173	23.721	68.106
container ship	3.797	34.724	61.478
dredger	8.673	12.396	78.931
fishing vessel	**2.500**	12.760	**84.740**
offshore vessels	13.804	33.529	52.667
others	12.650	15.361	71.989
passenger vessel	**25.647**	32.030	42.323
RORO	10.879	**69.460**	**19.662**
tanker or chemical ship	5.172	28.352	66.476
tug	10.822	**5.957**	83.220
bulk carrier	4.813	29.435	65.752
Voyage segment			
anchorage	**1.946**	24.887	73.166
archipelagos	10.587	33.217	56.195
at berth	3.206	**9.176**	**87.618**
canal	**47.437**	40.406	**12.156**
channel	34.101	39.210	26.689
coastal waters	4.796	31.551	63.652
inland waters	21.065	10.196	68.740
open sea	3.486	20.353	76.161
port	10.130	31.798	58.072
port approach	17.216	**40.674**	42.110
river	3.642	34.870	61.487
Hull type			
aluminium_alloy	17.830	**59.900**	22.271
composite materials	14.715	33.319	51.966
GRP	6.306	7.712	85.982
light alloy	**31.736**	48.131	**20.133**
NA	8.104	27.389	64.507
steel	6.596	28.207	65.198
wood	**5.541**	**6.777**	**87.681**
Gross tonnage			
1	8.937	**20.692**	**70.371**
2	6.861	24.611	68.529
3	**15.341**	**42.085**	**42.574**
4	**4.192**	31.652	64.156
Length			
1	**9.599**	**19.224**	**71.177**
2	8.706	32.354	**58.940**
3	**2.369**	**33.211**	64.420
Hull construction			
double bottom	**10.639**	**44.181**	**45.179**
double hull	**6.003**	25.125	68.871
single hull	7.639	**23.334**	**69.027**

unloading operations at berth. Therefore, maritime fisheries authorities should take specific action based on these critical factors, such as timely vessel maintenance and improved crew protection measures during loading/unloading to prevent casualties.

Table 10
TRI of RIFs for casualty type (100%).

Node	TRI less serious	serious	very serious	average
Type of accident	17.4	26.0	35.8	26.4
Ship operation	11.7	25.4	35.2	24.1
Ship type	11.6	31.8	32.5	25.3
Voyage segment	22.7	15.7	37.7	25.4
Hull type	13.1	26.6	33.8	24.5
Gross tonnage	5.6	10.7	13.9	10.1
Length	3.6	7.0	6.1	5.6
Hull construction	2.3	10.4	11.9	8.2

Table 11
The most important RIFs for casualty type.

	less serious	serious	very serious
Type of accident	2	3	2
Ship operation	4	4	3
Ship type	5	1	5
Voyage segment	1	5	1
Hull type	3	2	4
Gross tonnage	6	6	6
Length	7	8	8
Hull construction	8	7	7

5.1.2. *Scenario two: the combined impact of environment-related factors*

Scenario two simulates the changes in maritime accident casualty under the combined influence of environment-related RIFs. These RIFs contain 'time of day', 'sea condition', 'visibility', 'weather condition', 'wind', 'voyage segment', and 'ship operation'. Depending on different environmental conditions, maritime accident casualty may exhibit different types.

When the above RIFs are set to the specific states, as shown in Fig. 7, the 'very serious' type has the highest probability of accident casualty, accounting for 98.8%, which is a significant increase of 33.9% compared to the initial probability. It indicates that if the ship sails on a river at night with low visibility and poor sea and weather conditions, the severity of an accident is likely to be 'very serious'.

If the 'sea condition' and 'weather condition' are set to 'good', 'voyage segment' to 'canal', 'time of day' to 'day', and the remaining RIFs are unchanged, the probability of 'less serious' casualty significantly increases to 98.9% compared to Fig. 7. This finding suggests that ships are safer sailing in the canals with good sea and weather conditions during daylight hours, and even if an accident occurs, the casualty is less serious and will cause insignificant damage. Therefore, shipowners should conduct adequate prior inspections before performing maritime operations, sail in good sea and weather conditions whenever possible, and pay attention to the choice of voyage segment. Additionally, relevant traffic authorities should improve preventive measures in accident-prone segments at night to reduce accidents' casualties.

5.1.3. *Scenario three: the combined impact of ship-related factors*

Scenario three investigates the impact of ship-related RIFs on maritime accident casualty by altering their status. These RIFs include 'ship type', 'hull type', 'ship age', 'deadweight', 'gross tonnage', 'length', 'breadth', 'power', 'draught', 'hull construction', and 'vessel condition'. By varying the status of these RIFs, the accident casualty can likewise be characterised differently.

When the ship-related RIFs are set to the specific state displayed in Fig. 8, the probability of a 'very serious' accident casualty increases significantly to 99.6% compared to the initial probability. This probability indicates that a fishing vessel in that particular state (e.g., 'hull type' is 'wood', 'length' is in (0,100], etc.) is most likely to suffer a 'very serious' casualty once an accident occurs while conducting maritime operations.

Furthermore, by setting the status of ship-related RIFs to 100% for

Table 12
The combined influence of multiple variables.

Hull construction		+2%	+2%	+2%	+2%	+2%	+2%	+2%	+2%
Length			+2%	+2%	+2%	+2%	+2%	+2%	+2%
Gross tonnage				+2%	+2%	+2%	+2%	+2%	+2%
Hull type					+2%	+2%	+2%	+2%	+2%
Voyage segment						+2%	+2%	+2%	+2%
Ship type							+2%	+2%	+2%
Ship operation								+2%	+2%
Accident type									+2%
Less serious	7.48	7.52	7.68	7.96	8.30	8.78	9.46	9.98	10.13
Serious	21.62	21.75	21.94	22.22	22.61	23.14	23.67	24.35	24.92
Very serious	64.90	65.13	65.48	65.86	66.39	66.97	67.58	68.10	68.78

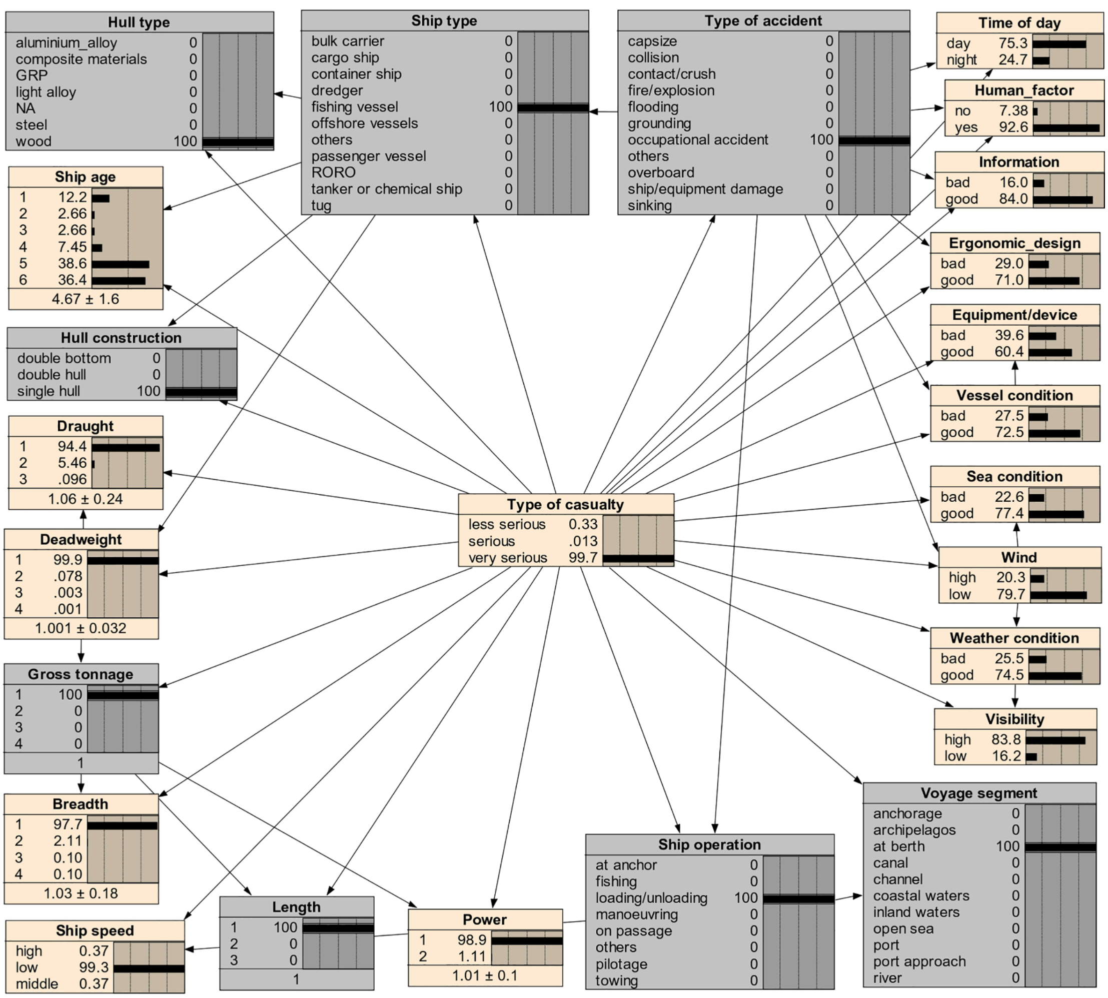

Fig. 6. The combined impact of important RIFs on 'very serious' casualty.

'offshore vessels', 'composite materials', '5' (ship age >20), '4' (deadweight > 30,000), '2' (gross tonnage in (3000,10,000]), '3' (length > 200), '3' (breadth in (30,40]), '1' (power in (0,3000]), '2' (draught in (6,9]), 'double bottom' and 'good', the probability of a 'less serious' accident casualty increases significantly to 99.9% compared to Fig. 8. The result demonstrates that the probability of a 'serious' or 'very serious' accident casualty is low when offshore vessels are in specific conditions (e.g., 'hull type' is 'composite materials', 'vessel condition' is 'good', etc.). Even if an accident does occur, it will result in low severity and not cause significant economic or human losses. This finding provides valuable insights for transport authorities and ship owners, helping them develop effective safety measures to reduce accident casualty during maritime operations based on the ship's factors.

5.1.4. *Scenario four: the most likely scenario for each severity type*

All three scenarios above are analysed in a forward manner using the BN model. Furthermore, the BN model enables reverse analysis by observing the effect on the variable node when the state of the parent node is adjusted. In Scenario four, the probability value of each state of the target node is set to 100% to observe the change in the variable node compared to the initial state.

As shown in Fig. 9, setting the probability of the 'less serious' to 100% reveals the most likely scenario for 'less serious' casualty. It can be found that the probability values of some node states increase significantly compared to their initial states, such as 'day' in 'time of day', 'good' in 'ergonomic design', 'port' in 'voyage segment', and 'manoeuvring' in 'ship operation'. This finding suggests that the accident

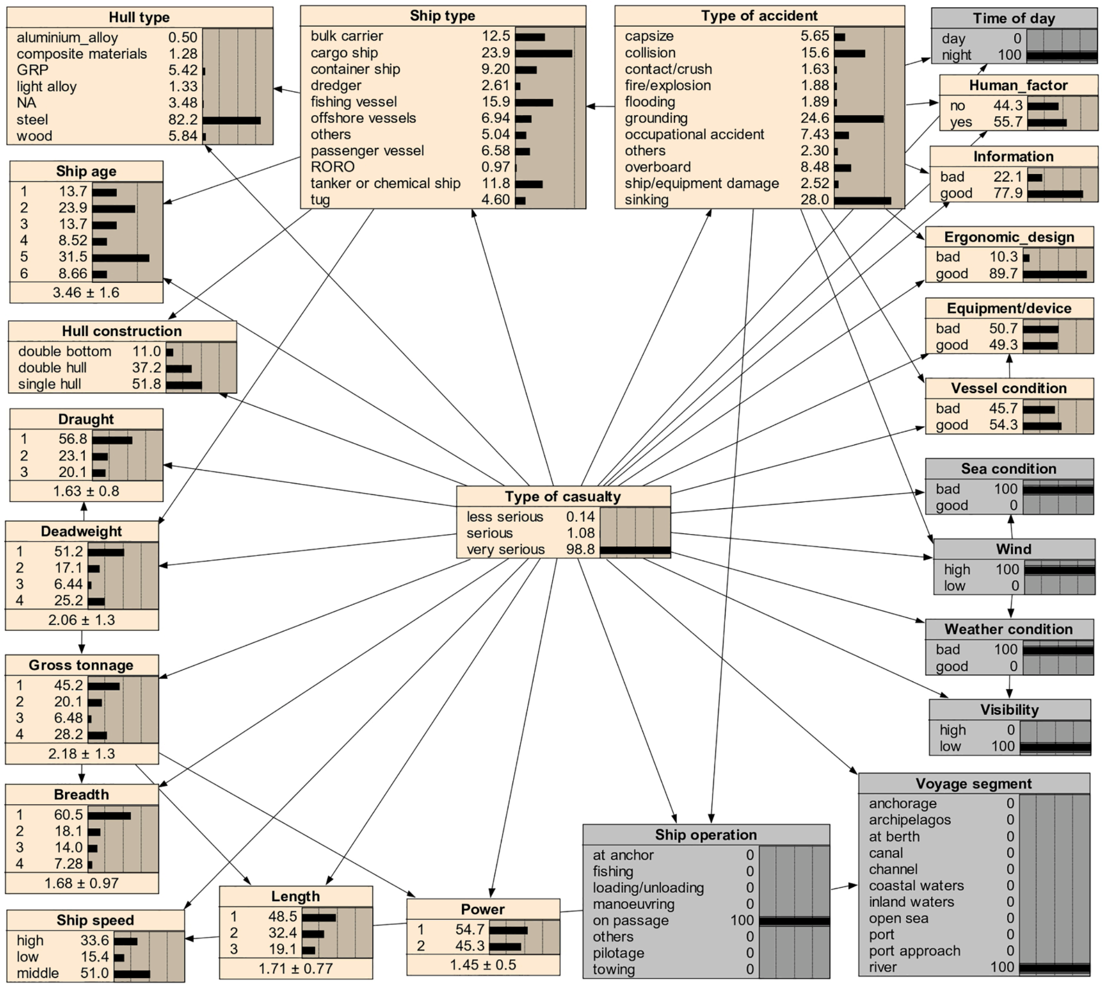

Fig. 7. Posterior probability analysis about 'very serious' type from environment-related factors.

casualty is most likely to be 'less serious' in a scenario consisting of the node states described above. Therefore, stakeholders such as transport authorities and shipowners can focus on these nodes when developing preventive measures to ensure safety and minimise accident damage caused by 'less serious' accident type.

When the probability of 'serious' is set to 100%, the most probable scenario corresponding to 'serious' accident casualty is revealed. Compared to the initial states, the probability of 'collision', 'fire/explosion' and 'grounding' in the 'type of accident' increased significantly, while the percentage of 'good' in 'information', 'equipment/device', and 'vessel condition' decreased. This result implies that 'collision', 'fire/explosion', and 'grounding' accidents are more likely to result in 'serious' casualty. Meanwhile, failure to ensure that the correct information is obtained promptly is more likely to result in a 'serious' accident during the voyage.

In maritime operations, stakeholders are keen to avoid 'very serious' accident casualty. Therefore, when the probability of 'very serious' is set to 100%, its corresponding most likely scenario is revealed. It can be found that the probabilities corresponding to 'fishing vessels' in 'ship type', 'single hull' in 'hull construction', 'open sea' in 'voyage segment', and 'bad' in 'weather condition' have all increased. Notably, there is a significant increase in the likelihood of 'occupational accident' in the 'type of accident'. This finding reveals that an occupational accident on a fishing vessel with a single-hull structure operating in the open sea is highly likely to have severe accidental consequences. Therefore, the relevant authorities should develop bespoke preventive measures, such as improving the hull construction of fishing vessels and crew protection measures, to avoid 'very serious' accident casualty.

5.2. Implications

According to the scenario analysis, the new findings, enabled by the most comprehensive and up-to-date dataset, suggest that (1) 'very serious' accident casualties have a higher likelihood of occurring when ships are in the 'open sea' compared to 'port' and 'coastal waters'; and (2) the study highlights the need to focus on 'occupational accidents' alongside the commonly occurring accidents to prevent casualties more effectively.

To improve understanding of the changing risks linked with maritime accidents and the various factors contributing to maritime accident casualty, this study further undertakes a comparative analysis between our new findings and the established ones from the existing literature. Amongst the six comparative studies, 4 references [18,25,47,55] employed a BN model, while 2 other references [23,24] utilised an ordered logistic regression model and ZIOP model, respectively. This comparative analysis result discloses the uniqueness of the multi-dimensional and global maritime accident casualty analysis.

The comparison focuses on two aspects: (1) the identified important RIFs; and (2) the variables and states that are most likely to result in severe casualties in maritime accidents, including 'type of accident', 'ship type', and 'voyage segment'. The results of the comparative analysis are presented in Table 13.

Based on the comparative analysis presented in Table 13, the following results can be drawn:

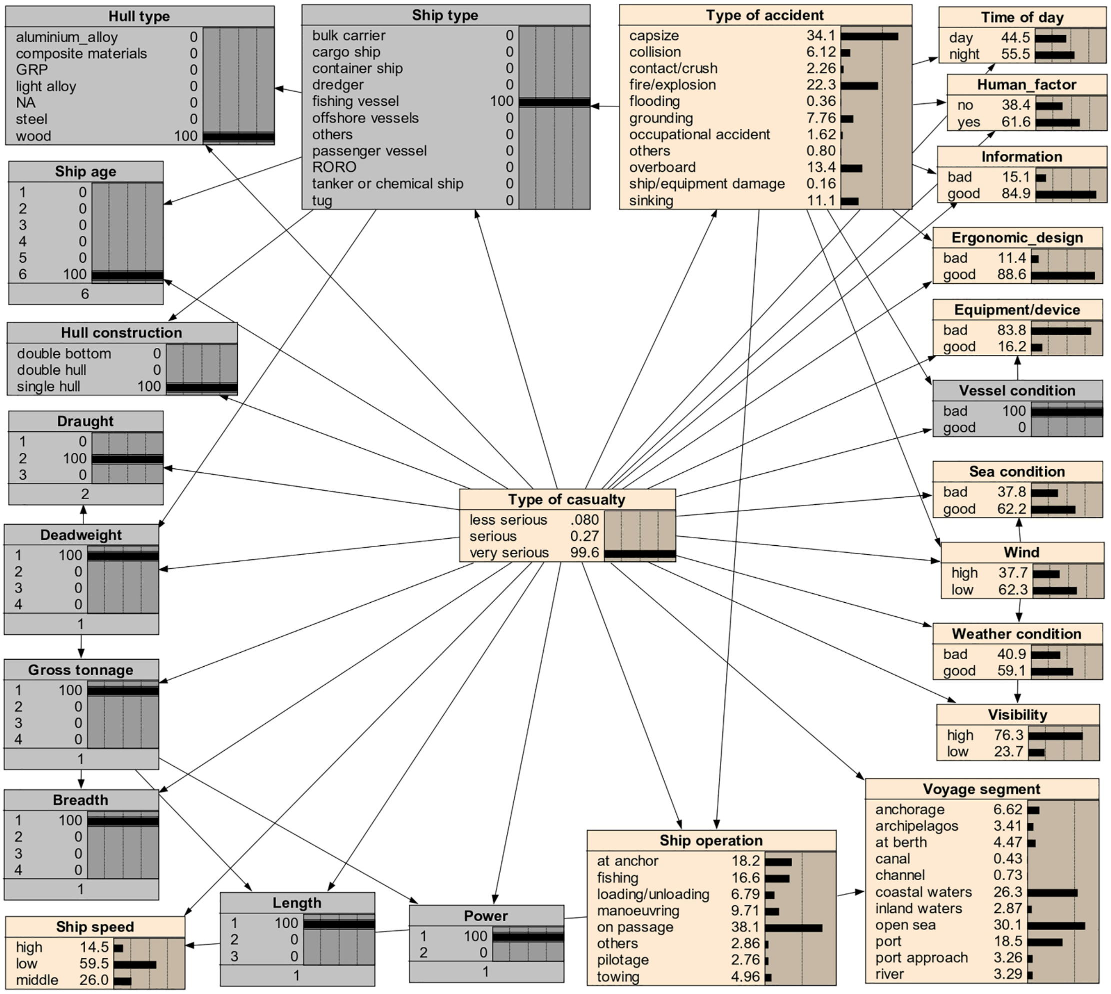

Fig. 8. Posterior probability analysis in 'very serious' type from ship-related factors.

(1) In terms of the most important RIFs influencing accident casualty, this study identifies 8 important RIFs. Specifically, this study identifies 'ship operation', 'hull type', 'hull construction', and 'length' as important RIFs for the first time, highlighting the importance of ship operations and characteristics for accident casualty from a global perspective. Improper ship operation (e.g., negligent navigation or manoeuvring) can lead to more severe consequences, resulting in loss of life and environmental pollution. Similarly, 'hull type', 'hull construction', and 'ship size' directly affect vessel safety and seaworthiness, thus impacting accident casualty.

(2) Compared to the identified 8 important RIFs, the frequency of a RIF appearance does not necessarily represent its importance. As a result, it raises an argument on the rationale for choosing the most frequently used RIFs in future maritime accident casualty studies.

(3) In contrast to previous research listed in Table 12, 'voyage segment' and 'hull type' are ranked as the top RIFs in this study. This finding suggests that the selection and evaluation of 'voyage segment' are important measures to prevent serious maritime accidents, including the development of reasonable navigation plans and the improvement of navigation safety management systems. Additionally, it highlights once again the significant impact of 'hull type' on vessel safety performance and the casualty of maritime accidents.

(4) Regarding the impact of accident types on casualty, 'capsize', 'sinking', 'occupational accident' are most likely to cause 'very serious' casualties, with 'occupational accident' having the greatest impact. Unlike the previous findings from the existing literature, this study includes 'occupational accident' as a state of accident types for the first time. It therefore generates a new finding, suggesting that personnel casualties caused by 'occupational accidents' can lead to erroneous ship operations and even cause the vessel to lose control, thus seriously endangering the safety of the vessel and other personnel. Without appropriate solutions, the shipping industry could suffer significant economic and reputational damage from 'occupational accidents', and hence adversely affecting the industry's sustainable development.

(5) In terms of ship type, 'dredger', 'fishing vessel', and 'tug' are more likely to cause 'very serious' accident consequences. Compared with the existing literature, this study first introduces 'dredger' in the state of 'ship type' and indicates that 'dredger' and 'tug' are more likely to cause serious accident consequences than 'passenger ship' and 'chemical ship'. This is due to the fact that 'dredger' and 'tug' relatively less follow international legislation compliance compared to other commercial ships and often appear in regions/waters of a high-level traffic density, increasing the probability of colliding with other vessels.

(6) Regarding voyage segments, 'open sea,' 'at berth,' and 'coastal waters' are more prone to 'very serious' maritime accident casualty. Unlike the previous relevant studies, this study proposes that vessels are more susceptible to experiencing severe maritime accidents when operating in 'open sea' and 'at berth'. This is

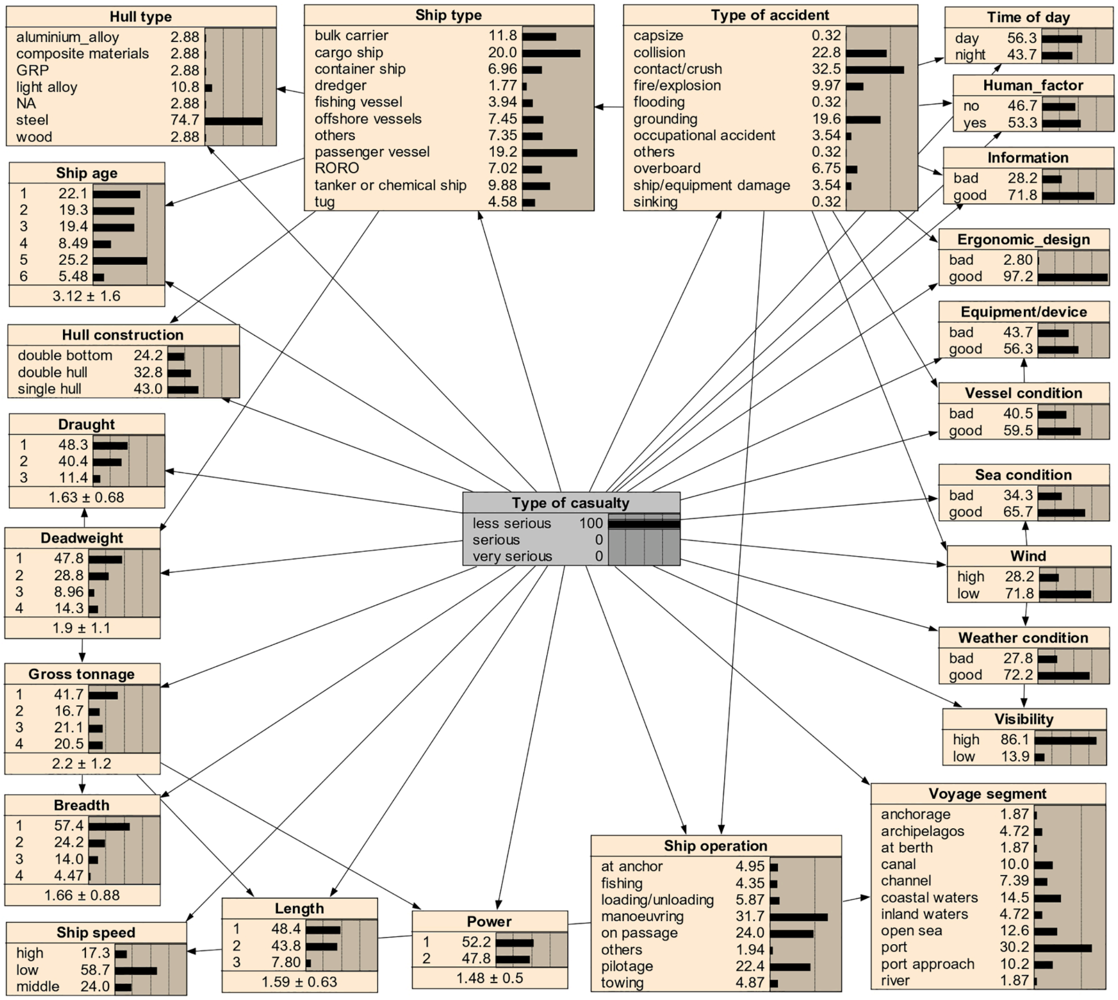

Fig. 9. Prior probability analysis in 'less serious' casualty type.

Table 13
A comparative analysis with the existing literature.

Refs.	The claimed important RIFs	Type of accident	Ship type	Voyage segment
[18]	accident type, ship type, engine power, gross tonnage, location (5 RIFs)	capsizing/sinking, hull/ machinery damage, collision	fishing ships, passenger ships, chemical ships	coastal areas
[23]	accident type, human element, ship type, ship condition, environment (5 RIFs)	sinking	fishing vessels, yachts, sailing vessels	far away from port
[24]	–	stranding/grounding, capsizing/ sinking	–	near the coast
[55]	accident type, location, ship type, ship age (4 RIFs)	sinking	fishing vessel	inland or coastal waterway
[47]	encounter situations, weather condition, and traffic density (3 RIFs)	collision	small general cargo ships	coastal waters
[25]	accident type, ship type (2 RIFs)	flooding, grounding	tanker, generaldrycargoVessel	strait, open sea
This study	type of accident, voyage segment, ship type, hull type, ship operation, gross tonnage, hull construction, length (8 RIFs)	capsize, occupational accident, sinking	dredger, fishing vessel, tug	open sea, at berth, coastal waters

because vessels are susceptible to adverse weather and sea conditions in 'open sea', and there are also many navigating vessels and hazardous obstacles, seriously affecting the safety of the navigation process. Furthermore, when vessels are at berth, they need to perform complex ship operations (e.g., anchoring, mooring, turning) in a small space, making it easy for equipment failures or ship collisions to occur.

The presented comparative analysis provides several valuable implications for the maritime industry and stakeholders, including:

(1) The identification of eight crucial RIFs by this study, including previously unexplored factors such as 'ship operation', 'hull type', 'hull construction', and 'length', enables new concerns when developing the countermeasures to reduce maritime accident casualty. This highlights the importance of ship characteristics and operations in determining the severity of accidents.
(2) The inclusion of occupational accidents as a new state of accident types underscores the need for targeted policies and prevention strategies to mitigate personnel casualties that can lead to serious vessel damage and economic losses.

(3) The study's introduction of dredgers in the state of ship type emphasises the higher likelihood of serious accidents for dredgers and tugs due to their current legislative frameworks and the traffic characteristics of their operating regions.
(4) The identification of open sea, at berth, and coastal waters as more prone to 'very serious' maritime accident casualty highlights the need for additional safety measures, such as better equipment and navigation systems, to minimise the impact of adverse weather and sea conditions, navigating vessels, and hazardous obstacles.

Overall, these insights from the analysis can inform the development of effective risk mitigation policies and practices to enhance the sustainability and safety of the maritime industry.

6. Comparative evolution analysis of five models

6.1. Comparative results of the basic content of the five models

Table 14 presents a comparative evolution analysis across five models, with the details outlined as follows:

The number of important RIFs varies across five models: 10 in 2021, 7 in 2020–2021, 7 in 2019–2021, 6 in 2018–2021, and 8 in 2017–2021. The models are based on different amounts of maritime accident data in the five different timeframes: 11 accidents in 2021, 65 in 2020–2021, 140 in 2019–2021, 258 in 2018–2021, and 402 in 2017–2021.

Each model shows varying statistics for common RIFs like gross tonnage, length, ship speed, weather conditions, sea conditions, voyage segments, and ship operations. The trends and percentages of these factors differ in each model, reflecting changes over time. The development trends of various factors in maritime accidents across different models are summarised as follows:

(1) Gross Tonnage ($<$ 10,000)

There is a fluctuating trend, starting at 53.1%, dropping to 50.4%, and then increasing to 59.4%. The fluctuating trend in accidents involving smaller ships (gross tonnage $<$ 10,000) suggests varying levels of risk exposure over time. This indicates the need for targeted safety measures and regulations for smaller vessels, considering their changing role in maritime incidents.

(2) Length (meters)

In terms of contributary probabilities, ships of less than 100 m show a general increase from 37.7% to 42.5%, while those over 100 m exhibit a slight decrease from 62.3% to 57.5% in the accident records. The increase in accidents involving ships of less than 100 m and a slight decrease for larger ships ($>$100 m) suggest that smaller vessels are becoming more prone to accidents. This could be due to increased traffic, operational challenges, or less stringent safety protocols for smaller ships.

(3) Ship Speed (knots) in (6, 12] and $>$12

The percentage of ships in the 6–12 knots range shows an increasing trend from 29% to 40.8%. Ships moving faster than 12 knots also increased from 18.9% to 23.6%. The rising trend of accidents at speeds of 6–12 knots and above 12 knots highlights the danger of high speed in navigation practices, especially in high-traffic areas or adverse conditions.

(4) Weather Conditions ('Poor')

Poor weather conditions rose from the original 35.3% to a peak of 46.5% in the latest accident records. The increase in accidents during poor weather conditions underscores the importance of enhanced weather forecasting, better preparation and training for adverse weather, and more robust ship design to withstand harsh conditions.

(5) Sea Condition ('Poor')

A general decreasing trend is seen from 33.0% to 25.2%, with a fluctuation by a value of 29.8%. The trend in accidents under poor sea conditions suggests the continuous improvement in the sea-readiness of vessels and crew training, especially in navigating rough seas.

(6) Voyage Segment ('Port')

There is a decreasing trend from 22.3% to 12.6% for accidents occurring in port segments. The decrease in accidents in port areas could indicate improvements in port safety and operations. However, continuous efforts are needed to maintain and enhance safety protocols, especially considering the complexity of port operations.

(7) Ship Operation ('At Anchor' and 'Manoeuvring')

The percentage changes of accidents 'at anchor' keep stable relatively, ranging from 19.2% to 19.8%. For 'manoeuvring', there is a slight decrease from 13.6% to 12.6%. The values with the changes indicate 'manoeuvring' aid to improve operational practices. However, these activities still represent a relatively high level of risk, requiring further attention on operational safety in these scenarios.

Each of these points reflects specific areas where maritime safety can be improved through targeted strategies, policy changes, and technological advancements.

6.2. Comparative results by various scenario analyses

The comparison of five models incorporates scenario analyses, focusing on the combined impact of important RIFs, ship-related factors, environment-related factors, and the most likely scenario. These scenarios illustrate the effect of varying RIFs configurations on maritime accident casualties. The overall comparison is provided below.

Most less serious maritime accident casualties that are due to 'contact or crush' involve cargo ships during manoeuvring or transit. This indicates a need for enhanced navigation and collision avoidance for cargo ships. Serious accident casualties often involve grounding and collision, especially with bulk carriers and container ships, pointing to a need for better route planning and situational awareness in coastal areas. Very serious accident casualties are frequently caused by occupational hazards on cargo and bulk carrier ships, emphasising the necessity for improved safety and training, especially at sea.

Although most accidents happen in good weather and sea conditions, these occurrences suggest a potential underestimation of risk and highlight the need for constant safety vigilance. The varied ages of ships involved in accidents suggest that maintenance and safety updates are crucial for all vessels, regardless of their age. The increase in coastal water accidents for very serious cases calls for heightened safety measures in these challenging environments.

An upward trend in less serious accident casualties caused by 'contact or crush' accidents underscores the need for better spatial awareness. The involvement of ships over 20 years is decreasing, yet mid-aged ships (11–15 years) are seeing increased incidents, which implies a continuous need for safety checks across all ship ages. The decline in good weather conditions as a factor in the fourth case further stresses the importance of preparedness for all conditions.

Overall, these points suggest that while environmental conditions are often favourable, improvements in human factors and operational practices are essential for reducing accident casualties.

Table 14
Summary comparison amongst the five models.

RIFs	2017–2021 24 RIFs in total, of which **8** are important RIFs	2018–2021 24 RIFs in total, of which **6** are important RIFs	2019–2021 24 RIFs in total, of which **7** are important RIFs	2020–2021 24 RIFs in total, of which **7** are important RIFs	2021 24 RIFs in total, of which **10** are important RIFs
Data records	402	258	140	65	11
The state change of common RIFs	(1) Gross tonnage < 10,000 (53.1%); (2) Length (meters) < 100 (37.7%), >100 (62.3%); (3) Ship speed (knots) in (6, 12] (29%), >12 (18.9%); (4) Weather conditions in 'poor' (35.3%); (5) Sea condition is 'poor' (33.0%); (6) Voyage segment in 'port' (22.3%); (7) Ship operation in 'at anchor' (19.8%), in 'manoeuvring' (13.6%).	(1) Gross tonnage < 10,000 (50.8%); (2) Length (meters) < 100 (32.9%), >100 (67.1%); (3) Ship speed (knots) in (6, 12] (29.4%), >12 (20.0%); (4) Weather conditions in 'poor' (31.3%); (5) Sea condition is 'poor' (27.8%); (6) Voyage segment in 'port' (18.6%); (7) Ship operation in 'at anchor' (19.3%), in 'manoeuvring' (11.2%).	(1) Gross tonnage < 10,000 (50.4%); (2) Length (meters) < 100 (34.2%), >100 (65.8%); (3) Ship speed (knots) in (6, 12] (33.4%), >12 (19.4%); (4) Weather conditions in 'poor' (36.9%); (5) Sea condition is 'poor' (32.7%); (6) Voyage segment in 'port' (18.7%); (7) Ship operation in 'at anchor' (18.2%), in 'manoeuvring' (11.5%).	(1) Gross tonnage < 10,000 (52.6%); (2) Length (meters) < 100 (37.7%), >100 (62.3%); (3) Ship speed (knots) in (6, 12] (37.3%), >12 (18.2%); (4) Weather conditions in 'poor' (38.8%); (5) Sea condition is 'poor' (25.2%); (6) Voyage segment in 'port' (13.5%); (7) Ship operation in 'at anchor' (17.7%), in 'manoeuvring' (6.00%).	(1) Gross tonnage < 10,000 (59.4%); (2) Length (meters) < 100 (42.5%), >100 (57.5%); (3) Ship speed (knots) in (6, 12] (40.8%), >12 (23.6%); (4) Weather conditions in 'poor' (46.5%); (5) Sea condition is 'poor' (29.8%); (6) Voyage segment in 'port' (12.6%); (7) Ship operation in 'at anchor' (19.2%), in 'manoeuvring' (12.6%).
Scenario one: The combined impact of important RIFs	**Important RIFs setting:** (1) Type of accident: occupational accident; (2) Ship operation: at anchor; (3) Ship type: container ship; (4) Voyage segment: anchorage; (5) Hull type: steel; (6) Gross tonnage: (0,3000]; (7) Length: (0,100]; (8) Hull construction: single hull. **The result:** (1) less serious: 0.010%; (2) serious: 0.11%; (3) very serious: **99.9%**	**Important RIFs setting:** (1) Type of accident: occupational accident; (2) Ship operation: at anchor; (3) Ship type: container ship; (4) Voyage segment: anchorage; (5) Hull construction: single hull; (6) Human factor: yes. **The result:** (1) less serious: 0.036%; (2) serious: 0.003%; (3) very serious: **100%**	**Important RIFs setting:** (1) Type of accident: occupational accident; (2) Ship operation: at anchor; (3) Hull construction: single hull; (4) Voyage segment: anchorage; (5) Ship type: container ship; (6) Deadweight: >30,000; (7) Gross tonnage: >20,000. **The result:** (1) less serious: 0.002%; (2) serious: 0.82%; (3) very serious: **99.2%**	**Important RIFs setting:** (1) Type of accident: occupational accident; (2) Ship type: container ship; (3) Ship operation: at anchor; (4) Voyage segment: anchorage; (5) Hull type: steel; (6) Ship age: [6,10]; (7) Deadweight: >30,000. **The result:** (1) less serious: 0.074%; (2) serious: 1.36%; (3) very serious: **98.6%**	**Important RIFs setting:** (1) Type of accident: occupational accident; (2) Ship type: cargo ship; (3) Ship age: [6,10]; (4) Sea condition: bad (5) Ship operation: at anchor; (6) Voyage segment: anchorage; (7) Gross tonnage: >20,000; (8) Length: >100; (9) Draught: (6,9]; (10) Breadth: >0. **The result:** (1) serious: 0%; (2) very serious: **100%**
Scenario two: The combined impact of ship-related factors	**The result:** (1) less serious: 0.080%; (2) serious: 0.27%; (3) very serious: **99.6%**	**The result:** (1) less serious: 0.013%; (2) serious: 0.092%; (3) very serious: **99.9%**	**The result:** (1) less serious: 0.18%; (2) serious: 3.31%; (3) very serious: **96.5%**	**The result:** (1) less serious: 0.022%; (2) serious: 0.059%; (3) very serious: **99.9%**	**The result:** (1) serious: 27.7%; (2) very serious: **72.3%**
Scenario three: The combined impact of environment-related factors	**The result:** (1) less serious: 0.14%; (2) serious: 1.08%; (3) very serious: **98.8%**	**The result:** (1) less serious: 0.44%; (2) serious: 1.27%; (3) very serious: **98.3%**	**The result:** (1) less serious: 0.65%; (2) serious: 1.48%; (3) very serious: **97.9%**	**The result:** (1) less serious: **77.2%**; (2) serious: 4.03%; (3) very serious: 18.8%	**The result:** (1) serious: 2.15%; (2) very serious: **97.9%**
Scenario four: The most likely scenario for each severity type	Note that 1) refers to 'Type of accident', 2) indicates 'Ship type', 3) represents 'Ship operation', 4) corresponds to 'Ship age', 5) pertains to 'Voyage segment', 6) is 'Sea condition', and 7) expresses 'Weather condition'.				
Less serious (100%)	1): contact/crush (32.5%); 2): cargo ship (20.0%); 3): manoeuvring (31.7%); 4): > 20 (25.2%); 5): port (30.2%); 6): good (65.7%); 7): good (72.2%);	1): contact/crush (27.3%); 2): cargo ship (25.4%); 3): on passage (37.7%); 4): (0,5] (29.9%); 5): coastal waters (23.1%); 6): good (77.7%); 7): good (77.5%)	1): contact/crush (35.0%); 2): cargo ship (22.1%); 3): on passage (38.9%); 4): [11,15] (27.6%); 5): coastal waters (26.4%); 6): good (72.9%); 7): good (73.9%);	1): contact/crush (52.5%); 2): cargo ship and passenger vessel (19.9%); 3): on passage (38.6%); 4): [11,15] (44.5%); 5): coastal waters (43.6%); 6): good (60.3%); 7): good (58.8%);	–
Serious (100%)	1): grounding (30.4%); 2): bulk carrier (19.5%); 3): on passage (41.6%); 4): > 20 (31.4%); 5): coastal waters (25.8%); 6): good (68.3%); 7): good (65.6%)	1): collision (27.9%); 2): container ship (23.5%); 3): on passage (42.3%); 4): >20 (28.8%); 5): open sea (23.1%); 6): good (74.3%); 7): good (69.9%)	1): grounding (30.5%); 2): container ship 21.1%); 3): on passage (44.1%); 4): > 20 (32.4%); 5): coastal waters (22.8%); 6): good (68.6%); 7): good (66.4%)	1): collision and grounding (29.6%); 2): container ship (22.2%); 3): on passage (50.2%); 4): > 20 (43.0%); 5): coastal waters (25.7%);	1): collision (67.4%); 2): cargo ship (35.3%); 3): on passage (36.7%); 4): [16,20] (42.3%); 5): coastal waters (36.9%); 6): good (95.4%); 7): good (66.7%)

(continued on next page)

Table 14 (*continued*)

RIFs		2017–2021 24 RIFs in total, of which **8** are important RIFs	2018–2021 24 RIFs in total, of which **6** are important RIFs	2019–2021 24 RIFs in total, of which **7** are important RIFs	2020–2021 24 RIFs in total, of which **7** are important RIFs	2021 24 RIFs in total, of which **10** are important RIFs
					6): good (86.3%); 7): good (65.1%)	
	Very serious (100%)	1): occupational accident (26.4%); 2): cargo ship (19.3%); 3): on passage (46.0%); 4): > 20 (28.1%); 5): open sea (31.8%); 6): good (66.6%); 7): good (63.5%)	1): occupational accident (29.4%); 2): bulk carrier (19.9%); 3): on passage (43.6%); 4): > 20 (27.9%); 5): open sea (33.6%); 6): good (70.9%); 7): good (67.4%)	1): occupational accident (33.9%); 2): bulk carrier (21.1%); 3): on passage (45.3%); 4): > 20 (26.7%); 5): open sea (36.7%); 6): good (67.2%); 7): good (61.7%)	1): occupational accident (20.3%); 2): bulk carrier (22.9%); 3): on passage (50.7%); 4): [6,10] (24.7%); 5): open sea (41.8%); 6): good (69.1%); 7): good (59.0%)	1): collision and sinking (27.6%); 2): fishing vessel (35.0%); 3): on passage (46.1%); 4): > 20 (35.4%); 5): open sea (45.8%); 6): good (55.7%); 7): bad (54.1%)

7. Conclusion

This study addresses significant research gaps in maritime accident casualty analysis, including a deficiency in comprehensive analysis, substantial datasets, and a lack of detailed model comparison and evolution analysis. A new global maritime casualty database is developed in this study, and then a data-driven BN model is constructed. The database, containing 402 records of global maritime accidents, provides a more comprehensive dataset for casualty analysis than previous research from a global perspective. This study details 24 RIFs with specific state divisions, enabling a comprehensive examination of the various factors influencing maritime accident casualties and enhancing the applicability of the findings to various regions and scenarios. Additionally, five models derived from maritime accident temporal trends are compared, validating the effectiveness of the proposed BN model and the accuracy of the selected 24 RIFs. In the meantime, it also reveals the evolution of maritime casualties against different timeframes. The evolution analysis of these five models reveals valuable findings and implications. Unlike expert-driven models, the proposed data-driven TAN model uncovers more intricate relationships within the 24 RIFs, providing more objective and predictive outcomes. The study delivers important findings that contribute to the reduction of casualties from maritime accidents. The new findings of this paper suggest that:

(1) Exploring the combined effect of significant RIFs on accident severity, it is found that lighter, smaller, wooden, single-hull fishing vessels are more likely to suffer ‘very serious’ accident severity when carrying out loading/unloading operations. In addition, various accident types have varying impacts on accident severity. For example, an occupational accident on a fishing vessel would probably result in a ‘very serious’ accident. However, in terms of sinking or grounding, the accident’s severity would be significantly reduced and tend to be ‘less serious’.
(2) The influence of environment-related and ship-related factors on the severity of accidents is also significant. Under daylight, good sea and weather conditions, ships are safer sailing in ‘canal’ with a low probability of serious accident consequences. Small, lightweight, single-hull, wooden ships are prone to cause ‘very serious’ accident severity. On the contrary, larger and double-bottom boats have a low probability of causing serious accident consequences.
(3) According to the BN model’s backward risk diagnosis analysis, mitigating the severity of accidents requires consideration of various factors such as the ‘time of day’, ‘voyage segment’, ‘weather conditions’, ‘sea state’, and ‘vessel condition’. Appropriate states of RIFs can significantly reduce the severity of accidents. Particular attention should be paid to occupation accidents, and relevant departments should formulate targeted policies to avoid them to reduce casualties and material losses.
(4) Different scenarios highlight specific causes of maritime accidents, such as contact or crush incidents, grounding, collision, and occupational hazards. These scenarios emphasise the need for enhanced navigation, collision avoidance, route planning, situational awareness, safety training, and safety checks.
(5) Environmental conditions are often favourable, but improvements in human factors and operational practices are crucial for reducing accident casualties.

The implications for different stakeholders are listed below.

(1) Maritime authorities and regulators should consider implementing targeted safety measures and regulations for smaller vessels to address their changing role in maritime incidents. Shipowners and operators of smaller ships should prioritise safety protocols and training.
(2) Safety measures and regulations should be tailored to address the increasing risk associated with smaller vessels. Shipowners and operators of larger vessels should continue to adhere to safety protocols, as accidents can still occur.
(3) Stakeholders, including ship operators and navigators, should be cautious about the dangers of high speed in navigation practices, especially in high-traffic areas or adverse conditions. Enhanced training and safety measures should be considered.
(4) Improved weather forecasting, better preparation, and training for adverse weather conditions should be prioritised by maritime authorities and ship operators. Ship design should also be enhanced to withstand harsh weather.
(5) Continuous improvements in the sea-readiness of vessels and crew training should be maintained and further enhanced through ongoing efforts.
(6) While there are improvements in port safety and operations, stakeholders should continue their efforts to maintain and enhance safety protocols, given the complexity of port operations.
(7) Stakeholders should focus on operational safety, particularly during ‘manoeuvring’ activities, as they still represent a relatively high level of risk. Continuous attention to safety practices is crucial.
(8) Maritime authorities, shipowners, operators, and crew members should tailor their safety strategies and training programs to address the specific causes of accidents highlighted in different scenarios. Safety vigilance is essential, even in favourable environmental conditions.
(9) Maritime industry stakeholders should prioritise continuous improvement in human factors, training, and operational practices to enhance safety and reduce accident casualties, regardless of environmental conditions.

The findings and implications from the data-driven BN model for the maritime sector and stakeholders are groundbreaking, which can also aid them in developing risk prevention policies and reducing accident losses. Nevertheless, there are several limitations in this study. For instance, human factors are not further classified into multiple states, and the impact of human factors on maritime accident severity could therefore not be explored in detail. In future studies, data collection and analysis in this area should be strengthened to make the BN model more comprehensive in performance and more widely applicable.

CRediT authorship contribution statement

Kaiwen Zhou: Writing – original draft, Visualization, Validation, Software, Resources, Investigation, Formal analysis, Data curation. **Wenbin Xing:** Writing – original draft, Visualization, Validation, Software, Resources, Investigation, Formal analysis, Data curation, Conceptualization. **Jingbo Wang:** Writing – original draft, Visualization, Validation, Software, Resources, Investigation, Formal analysis, Data curation. **Huanhuan Li:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Zaili Yang:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Zaili Yang reports financial support was provided by European Research Council.

Data availability

Data will be made available on request.

Acknowledgements

This work is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 864724).

Appendix I

The confusion matrices of four models are presented in Tables 15–18.

Table 15
Confusion matrix of the one-year model.

Predicted	serious	very serious	Actual total	Accuracy rate(100%)
serious	1	0	1	100
very serious	0	2	2	100
Predicted total	1	2	3	100

Table 16
Confusion matrix of the two-year model.

Predicted	less serious	serious	very serious	Actual total	Accuracy rate(100%)
less serious	1	0	0	1	100
serious	0	5	1	6	83.33
very serious	0	0	7	7	100
Predicted total	1	5	8	14	92.86

Table 17
Confusion matrix of the three-year model.

Predicted	less serious	serious	very serious	Actual total	Accuracy rate(100%)
less serious	3	0	0	3	100
serious	0	10	0	10	100
very serious	0	2	13	15	86.66
Predicted total	3	12	13	28	92.86

Table 18
Confusion matrix of the four-year model.

Predicted	less serious	serious	very serious	Actual total	Accuracy rate(100%)
less serious	3	0	1	4	75
serious	1	11	1	13	84.62
very serious	0	2	33	35	94.29
Predicted total	4	13	35	52	90.38

The constructed models based on the four datasets are visualised in Figs. 10–13.

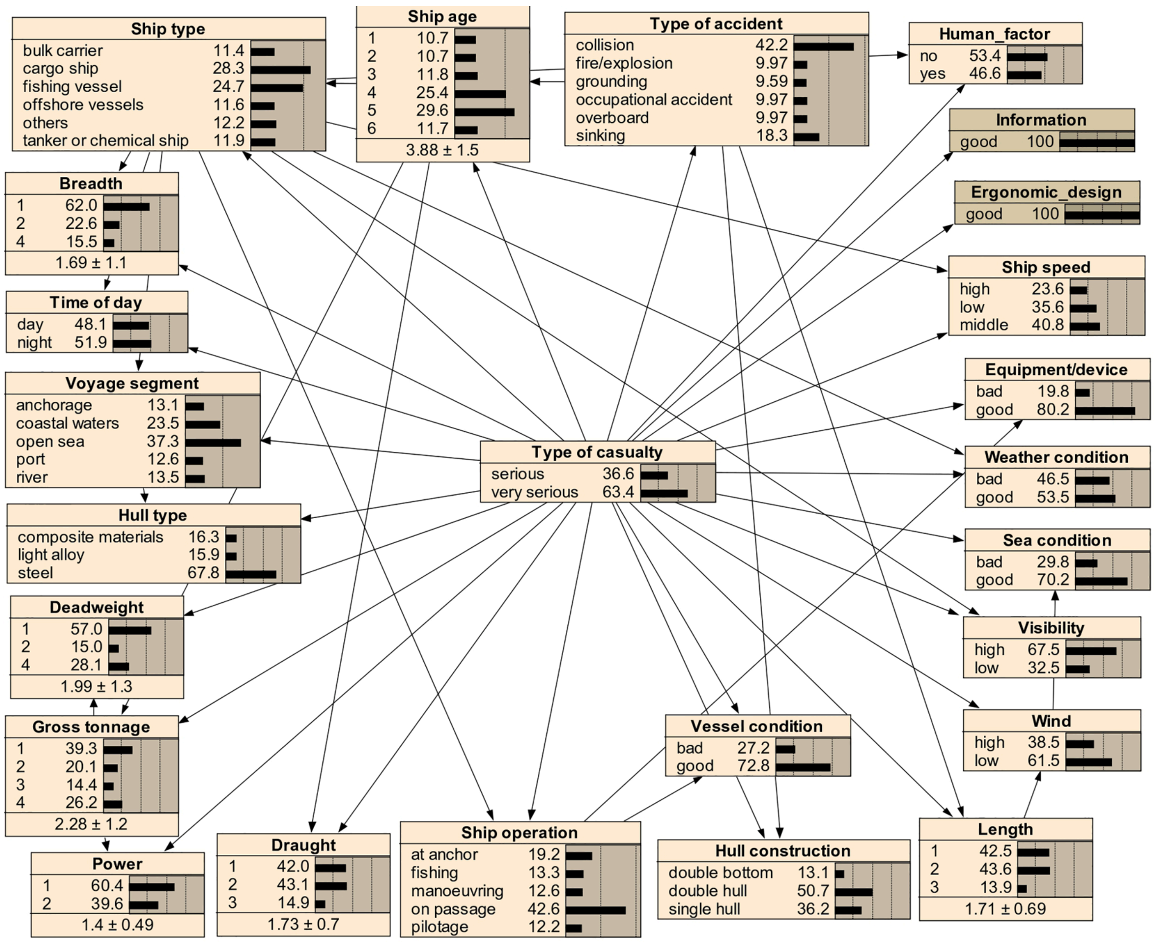

Fig. 10. The constructed BN model based on the data from 2021.

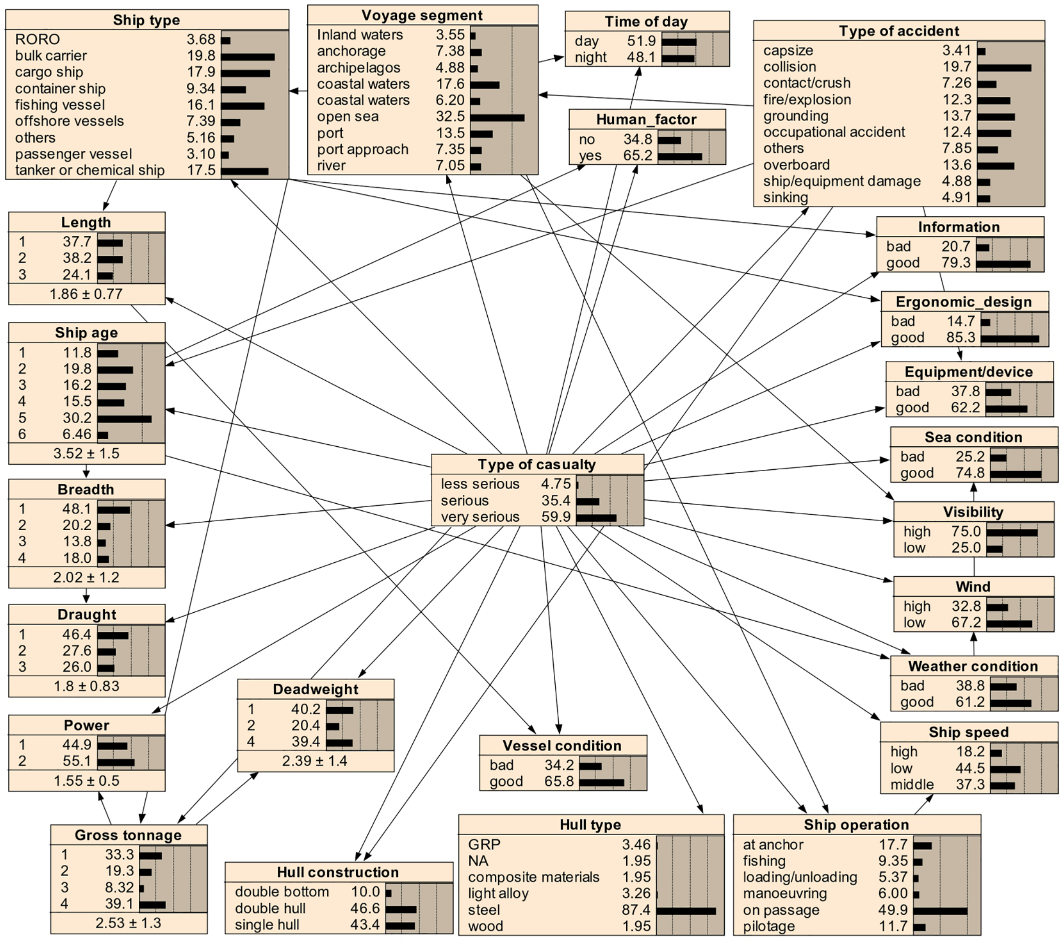

Fig. 11. The constructed BN model based on the data from 2020 to 2021.

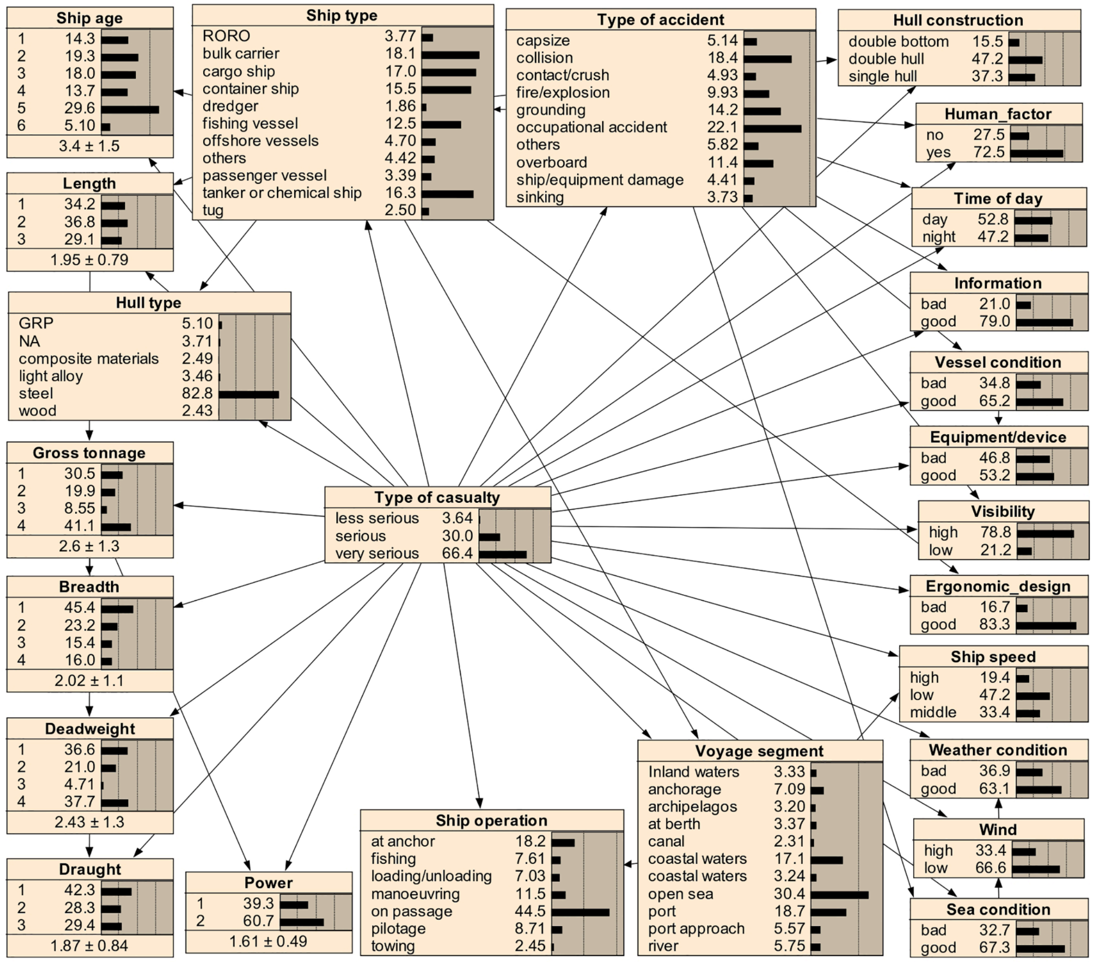

Fig. 12. The constructed BN model based on the data from 2019 to 2021.

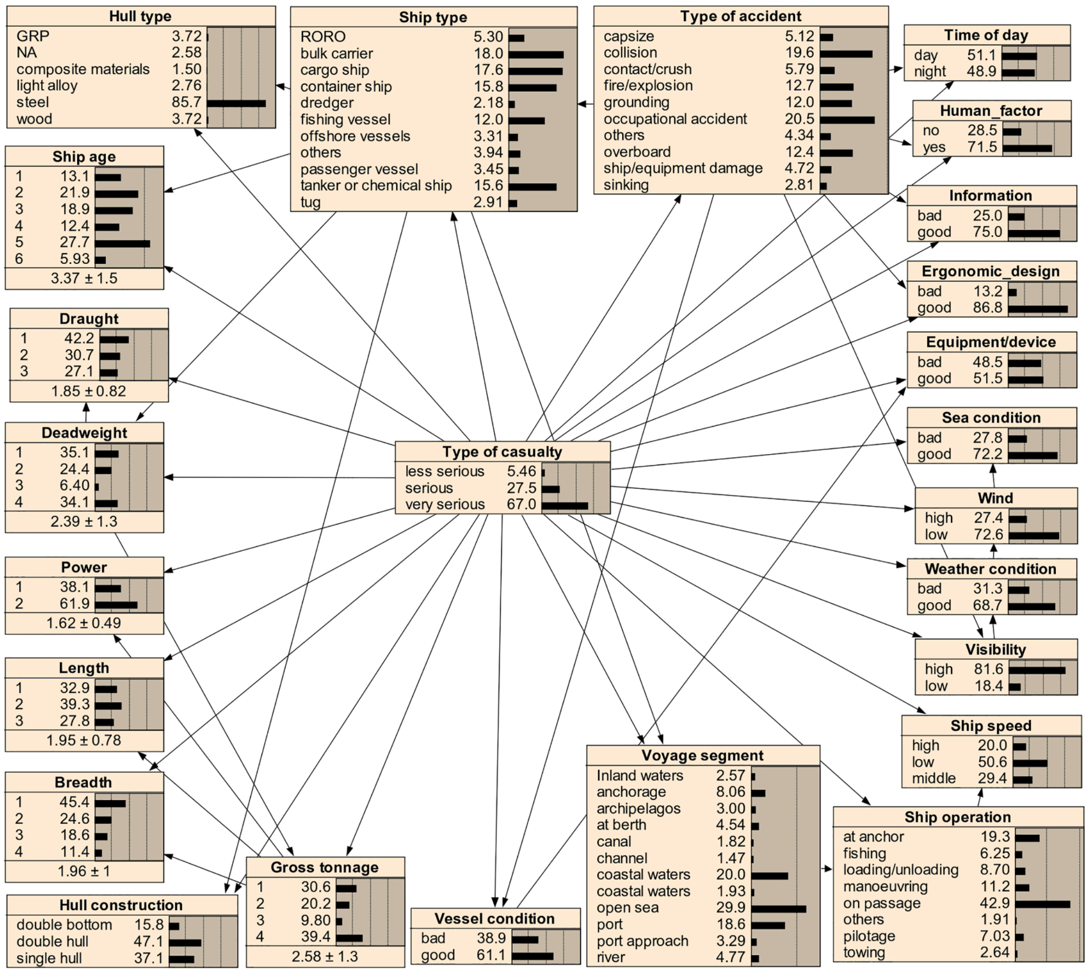

Fig. 13. The constructed BN model based on the data from 2018 to 2021.

References

[1] Li H, Jiao H, Yang Z. AIS data-driven ship trajectory prediction modelling and analysis based on machine learning and deep learning methods. Transp Res Part E Logist Transp Rev 2023;175:103152. https://doi.org/10.1016/j.tre.2023.103152.

[2] Li H, Xing W, Jiao H, Yang Z, Li Y. Deep bi-directional information-empowered ship trajectory prediction for maritime autonomous surface ships. Transp Res Part E: Logist Transp Rev 2024;181:103367. https://doi.org/10.1016/j.tre.2023.103367.

[3] Puisa R, Montewka J, Krata P. A framework estimating the minimum sample size and margin of error for maritime quantitative risk analysis. Reliab Eng Syst Saf 2023;235:109221. https://doi.org/10.1016/j.ress.2023.109221.

[4] Xin X, Liu K, Loughney S, Wang J, Li H, Yang Z. Graph-based ship traffic partitioning for intelligent maritime surveillance in complex port waters. Expert Syst Appl 2023;231:120825. https://doi.org/10.1016/j.eswa.2023.120825.

[5] Li H, Yang Z. Incorporation of AIS data-based machine learning into unsupervised route planning for maritime autonomous surface ships. Transp Res Part E: Logist Transp Rev 2023;176:103171. https://doi.org/10.1016/j.tre.2023.103171.

[6] Liang M, Li H, Liu RW, Lam JS, Yang Z. PiracyAnalyzer: Spatial Temporal Patterns Analysis of Global Piracy Incidents. Relia Eng Syst Saf 2024;243:109877.

[7] Li H, Lam JSL, Yang Z, Liu J, Liu RW, Liang M, et al. Unsupervised hierarchical methodology of maritime traffic pattern extraction for knowledge discovery. Transp Res Part C: Emerg Technol 2022;143:103856. https://doi.org/10.1016/j.trc.2022.103856.

[8] Guo Y, Jin Y, Hu S, Yang Z, Xi Y, Han B. Risk evolution analysis of ship pilotage operation by an integrated model of FRAM and DBN. Reliab Eng Syst Saf 2023;229:108850. https://doi.org/10.1016/j.ress.2022.108850.

[9] Chen J, Bian W, Wan Z, Yang Z, Zheng H, Wang P. Identifying factors influencing total-loss marine accidents in the world: analysis and evaluation based on ship types and sea regions. Ocean Eng 2019;191:106495. https://doi.org/10.1016/j.oceaneng.2019.106495.

[10] Fan S, Blanco-Davis E, Yang Z, Zhang J, Yan X. Incorporation of human factors into maritime accident analysis using a data-driven Bayesian network. Reliab Eng Syst Saf 2020;203:107070. https://doi.org/10.1016/j.ress.2020.107070.

[11] Yang Z, Yang Z, Yin J, Qu Z. A risk-based game model for rational inspections in port state control. Transp Res Part E Logist Transp Rev 2018;118:477–95. https://doi.org/10.1016/j.tre.2018.08.001.

[12] Fan S, Yang Z, Wang J, Marsland J. Shipping accident analysis in restricted waters: lesson from the Suez Canal blockage in 2021. Ocean Eng 2022;266:113119. https://doi.org/10.1016/j.oceaneng.2022.113119.

[13] Wan C, Yan X, Zhang D, Qu Z, Yang Z. An advanced fuzzy Bayesian-based FMEA approach for assessing maritime supply chain risks. Transp Res Part E-Logist Transp Rev 2019;125:222–40. https://doi.org/10.1016/j.tre.2019.03.011.

[14] Fan S, Yang Z. Accident data-driven human fatigue analysis in maritime transport using machine learning. Reliab Eng Syst Saf 2024;241:109675. https://doi.org/10.1016/j.ress.2023.109675.

[15] Xin X, Liu K, Loughney S, Wang J, Li H, Ekere N, et al. Multi-scale collision risk estimation for maritime traffic in complex port waters. Reliab Eng Syst Saf 2023;240:109554. https://doi.org/10.1016/j.ress.2023.109554.

[16] Li H, Yang Z. Towards safe navigation environment: the imminent role of spatio-temporal pattern mining in maritime piracy incidents analysis. Reliab Eng Syst Saf 2023;238:109422. https://doi.org/10.1016/j.ress.2023.109422.

[17] MdT Amin, Khan F, Imtiaz S. Dynamic availability assessment of safety critical systems using a dynamic Bayesian network. Reliab Eng Syst Saf 2018;178:108–17. https://doi.org/10.1016/j.ress.2018.05.017.

[18] Cao Y, Wang X, Wang Y, Fan S, Wang H, Yang Z, et al. Analysis of factors affecting the severity of marine accidents using a data-driven Bayesian network. Ocean Eng 2023;269:113563. https://doi.org/10.1016/j.oceaneng.2022.113563.

[19] Yip TL, Jin D, Talley WK. Determinants of injuries in passenger vessel accidents. Accid Anal Prev 2015;82:112–7. https://doi.org/10.1016/j.aap.2015.05.025.

[20] Wang J, Li M, Liu Y, Zhang H, Zou W, Cheng L. Safety assessment of shipping routes in the South China Sea based on the fuzzy analytic hierarchy process. Saf Sci 2014;62:46–57. https://doi.org/10.1016/j.ssci.2013.08.002.

[21] Wu B, Tian H, Yan X, Soares C. A probabilistic consequence estimation model for collision accidents in the downstream of Yangtze River using Bayesian Networks. Proc Inst Mech Eng Part O-J Risk Reliab 2020;234:422–36. https://doi.org/10.1177/1748006X19825706.

[22] Lan H, Ma X, Qiao W, Deng W. Determining the critical risk factors for predicting the severity of ship collision accidents using a data-driven approach. Reliab Eng Syst Saf 2023;230:108934. https://doi.org/10.1016/j.ress.2022.108934.

[23] Wang H, Liu Z, Wang X, Graham T, Wang J. An analysis of factors affecting the severity of marine accidents. Reliab Eng Syst Saf 2021;210:107513. https://doi.org/10.1016/j.ress.2021.107513.

[24] Wang H, Liu Z, Wang X, Huang D, Cao L, Wang J. Analysis of the injury-severity outcomes of maritime accidents using a zero-inflated ordered probit model. Ocean Eng 2022;258:111796. https://doi.org/10.1016/j.oceaneng.2022.111796.

[25] Cakir E, Sevgili C, Fiskin R. An analysis of severity of oil spill caused by vessel accidents. Transp Res Part -Transp Environ 2021;90:102662. https://doi.org/10.1016/j.trd.2020.102662.

[26] Wang H, Liu Z, Liu Z, Wang X, Wang J. GIS-based analysis on the spatial patterns of global maritime accidents. Ocean Eng 2022;245:110569. https://doi.org/10.1016/j.oceaneng.2022.110569.

[27] Weng J, Yang D. Investigation of shipping accident injury severity and mortality. Accid Anal Prev 2015;76:92–101. https://doi.org/10.1016/j.aap.2015.01.002.

[28] Montewka J, Manderbacka T, Ruponen P, Tompuri M, Gil M, Hirdaris S. Accident susceptibility index for a passenger ship-a framework and case study. Reliab Eng Syst Saf 2022;218:108145. https://doi.org/10.1016/j.ress.2021.108145.

[29] Talley WK, Jin D, Kite-Powell H. Determinants of the severity of passenger vessel accidents. Marit Policy Manag 2006;33:173–86. https://doi.org/10.1080/03088830600612971.

[30] Eliopoulou E, Papanikolaou A. Casualty analysis of large tankers. J Mar Sci Technol 2007;12:240–50. https://doi.org/10.1007/s00773-007-0255-8.

[31] Tan B, Otay EN. Modeling and analysis of vessel casualties resulting from tanker traffic through narrow waterways. Nav Res Logist NRL 1999;46:871–92. https://doi.org/10.1002/(SICI)1520-6750(199912)46:8<871::AID-NAV1>3.0.CO;2-I.

[32] Talley WK, Jin D, Kite-Powell H. Determinants of the damage cost and injury severity of ferry vessel accidents. WMU J Marit Aff 2008;7:175–88. https://doi.org/10.1007/BF03195130.

[33] Talley WK. The safety of ferries: an accident injury perspective. Marit Policy Manag 2002;29:331–8. https://doi.org/10.1080/03088830210132641.

[34] Nwaoha TC, Yang Z, Wang J, Bonsall S. Adoption of new advanced computational techniques to hazards ranking in LNG carrier operations. Ocean Eng 2013;72:31–44. https://doi.org/10.1016/j.oceaneng.2013.06.010.

[35] Khan RU, Yin J, Mustafa FS, Anning N. Risk assessment for berthing of hazardous cargo vessels using Bayesian networks. Ocean Coast Manag 2021;210:105673. https://doi.org/10.1016/j.ocecoaman.2021.105673.

[36] Jin D. The determinants of fishing vessel accident severity. Accid Anal Prev 2014;66:1–7. https://doi.org/10.1016/j.aap.2014.01.001.

[37] Lu C-S, Tsai C-L. The effects of safety climate on vessel accidents in the container shipping context. Accid Anal Prev 2008;40:594–601. https://doi.org/10.1016/j.aap.2007.08.015.

[38] Montewka J, Ehlers S, Goerlandt F, Hinz T, Tabri K, Kujala P. A framework for risk assessment for maritime transportation systems-A case study for open sea collisions involving RoPax vessels. Reliab Eng Syst Saf 2014;124:142–57. https://doi.org/10.1016/j.ress.2013.11.014.

[39] Dong Y, Frangopol DM. Probabilistic ship collision risk and sustainability assessment considering risk attitudes. Struct Saf 2015;53:75–84. https://doi.org/10.1016/j.strusafe.2014.10.004.

[40] Liu Y, Frangopol DM. Probabilistic risk, sustainability, and utility associated with ship grounding hazard. Ocean Eng 2018;154:311–21. https://doi.org/10.1016/j.oceaneng.2018.01.101.

[41] Qu X, Meng Q, Suyi L. Ship collision risk assessment for the Singapore Strait. Accid Anal Prev 2011;43:2030–6. https://doi.org/10.1016/j.aap.2011.05.022.

[42] Wu B, Zong L, Yip TL, Wang Y. A probabilistic model for fatality estimation of ship fire accidents. Ocean Eng 2018;170:266–75. https://doi.org/10.1016/j.oceaneng.2018.10.056.

[43] Pitman SJ, Wright M, Hocken R. An analysis of lifejacket wear, environmental factors, and casualty activity on marine accident fatality rates. Saf Sci 2019;111:234–42. https://doi.org/10.1016/j.ssci.2018.07.016.

[44] Domeh V, Obeng F, Khan F, Bose N, Sanli E. Risk analysis of man overboard scenario in a small fishing vessel. Ocean Eng 2021;229:108979. https://doi.org/10.1016/j.oceaneng.2021.108979.

[45] Aydogdu YV. A Comparison of Maritime Risk Perception and Accident Statistics in the Istanbul Straight. J Navig 2014;67:129–44. https://doi.org/10.1017/S0373463313000593.

[46] Sahin B, Kum S. Risk assessment of arctic navigation by using improved fuzzy-AHP approach. Int J Marit Eng 2015;157. https://doi.org/10.5750/ijme.v157iA4.965.

[47] Liu K, Yu Q, Yuan Z, Yang Z, Shu Y. A systematic analysis for maritime accidents causation in Chinese coastal waters using machine learning approaches. Ocean Coast Manag 2021;213:105859. https://doi.org/10.1016/j.ocecoaman.2021.105859.

[48] Zhang J, Teixeira AP, Guedes Soares C, Yan X, Liu K. Maritime transportation risk assessment of tianjin port with bayesian belief networks. Risk Anal 2016;36:1171–87. https://doi.org/10.1111/risa.12519.

[49] Zhang D, Yan XP, Yang ZL, Wall A, Wang J. Incorporation of formal safety assessment and Bayesian network in navigational risk estimation of the Yangtze River. Reliab Eng Syst Saf 2013;118:93–105. https://doi.org/10.1016/j.ress.2013.04.006.

[50] Zhang D, Yan X, Yang Z, Wang J. An accident data-based approach for congestion risk assessment of inland waterways: a Yangtze River case. Proc Inst Mech Eng Part O-J Risk Reliab 2014;228:176–88. https://doi.org/10.1177/1748006X13508107.

[51] Le Blanc LA, Rucks CT. A multiple discriminant analysis of vessel accidents. Accid Anal Prev 1996;28:501–10. https://doi.org/10.1016/0001-4575(96)00020-6.

[52] Banda OAV, Goerlandt F, Montewka J, Kujala P. A risk analysis of winter navigation in Finnish sea areas. Accid Anal Prev 2015;79:100–16. https://doi.org/10.1016/j.aap.2015.03.024.

[53] Yip TL. Port traffic risks - A study of accidents in Hong Kong waters. Transp Res Part E-Logist Transp Rev 2008;44:921–31. https://doi.org/10.1016/j.tre.2006.09.002.

[54] Li H, Ren X, Yang Z. Data-driven Bayesian network for risk analysis of global maritime accidents. Reliab Eng Syst Saf 2023;230:108938. https://doi.org/10.1016/j.ress.2022.108938.

[55] Wang L, Yang Z. Bayesian network modelling and analysis of accident severity in waterborne transportation: a case study in China. Reliab Eng Syst Saf 2018;180:277–89. https://doi.org/10.1016/j.ress.2018.07.021.

[56] Kamal B, Cakir E. Data-driven Bayes approach on marine accidents occurring in Istanbul strait. Appl Ocean Res 2022;123:103180. https://doi.org/10.1016/j.apor.2022.103180.

[57] Yang Z, Yang Z, Yin J. Realising advanced risk-based port state control inspection using data-driven Bayesian networks. Transp Res Part Policy Pract 2018;110:38–56. https://doi.org/10.1016/j.tra.2018.01.033.

[58] Jones B, Jenkinson I, Yang Z, Wang J. The use of Bayesian network modelling for maintenance planning in a manufacturing industry. Reliab Eng Syst Saf 2010;95:267–77. https://doi.org/10.1016/j.ress.2009.10.007.

[59] Yang ZL, Wang J, Bonsall S, Fang QG. Use of fuzzy evidential reasoning in maritime security assessment. Risk Anal 2009;29:95–120. https://doi.org/10.1111/j.1539-6924.2008.01158.x.

[60] Alyami H, Yang Z, Riahi R, Bonsall S, Wang J. Advanced uncertainty modelling for container port risk analysis. Accid Anal Prev 2019;123:411–21. https://doi.org/10.1016/j.aap.2016.08.007.

[61] Alyami H, Lee PT-W, Yang Z, Riahi R, Bonsall S, Wang J. An advanced risk analysis approach for container port safety evaluation. Marit Policy Manag 2014;41:634–50. https://doi.org/10.1080/03088839.2014.960498.