Fan, S, Yang, Z, Wang, J and Marsland, J

Shipping accident analysis in restricted waters: Lesson from the Suez Canal blockage in 2021

http://researchonline.ljmu.ac.uk/id/eprint/18430/

Article

Shipping accident analysis in restricted waters: Lesson from the Suez Canal blockage in 2021

Shiqi Fan [a,b], Zaili Yang [a,b,*], Jin Wang [a,b], John Marsland [b]

[a] *Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool, L3 3AF, UK*
[b] *School of Engineering, Liverpool John Moores University, L3 3AF, UK*

A B S T R A C T

Global shipping flows through strategically important restricted waters such as Panama and Suez Canals, and hence accidents occurring in the Canals will cause serious disruptions to global supply chains. In this paper, a new data-driven Bayesian network (BN) based risk model is developed to investigate how risk factors jointly generate impact on different types of maritime accidents within restricted waters. A new risk database involving 25 factors has been developed by manual analysis of all the recorded accidents from 2005 to 2021 that occurred in the world's important restricted waters including key maritime canals, channels and straits. A data-driven BN model is constructed to analyse the key risk influential factors (RIFs) contributing to such accidents. The model is verified by sensitivity analysis and real accident cases. Further, it is tested by the already known information of the 2021 Suez Canal blockage case to generate useful insights and draw the lessons to learn. In a retrospective analysis using the currently available limited information on the Suez Canal case, the implication of the case study shows a plausible explanation for the observed findings by scenario analysis. The findings can be applied to backward risk cause diagnosis for accident investigation and forward risk prediction for accident prevention in restricted waters to avoid the reoccurrence of similar accident to the Suze Canal blockage.

1. Introduction

Although growth in international maritime trade stalled in 2019 with volumes expanding at a marginal 0.5% owing to the slowdown in the world economy and trade (UNCTAD, 2020), it does not deny the fact that shipping plays a crucial role in the global economy. As maritime accidents expose high risks of casualties, economic loss, and severe maritime pollution, maritime safety represents the priority for the international shipping industry (Fan et al., 2020c). The investigation into maritime accidents helps understand the causes leading to the failures, provides recommendations on countermeasures taken for accident prevention, and guides ship owners and maritime authorities to manage risks rationally.

The occurrence of maritime accidents varies with voyage segments and ship operations. The spatial distributions of global maritime accidents reveal the value of investigating accidents in maritime-accident-prone regions or restricted waters. For example, 12% of global shipping uses the Suez Canal, which can easily cause supply chain disruptions, fuelling shortages and hike prices if there are any delays by accidents (Topham, 2021). On the March 23, 2021, one of the world's

largest container ships (*i.e.* Ever Given) ran aground on the banks of the Suez Canal, which caused an approximate £7bn loss a day in trade due to the blockage of the Canal and up to £10.9m incoming loss each day for the Canal (Michaelson and Safi, 2021). Because of the unique characteristics of the navigational environments in restricted waters (incl. canals, channels and straits), the findings from previous studies relating to general maritime accidents/risk analysis only provide limited insights on both theoretical and practical implications within the context of restricted waters. Thus, it is essential and beneficial to initiate a new study on maritime accident analysis in restricted waters to bridge the research gap.

This paper aims to develop a new data-driven Bayesian network (BN) based risk model to analyse risk factors contributing to maritime accidents in restricted waters and use the Suez Canal blockage case to generate insights for accident prevention. It extracted the data from three public databases to train prior probabilities in the risk-based BN through the historical data-driven model. From the newly developed database, it is found that the accidents in restricted waters share many commonalities which are different from those in other waters/regions. It further justifies the necessity of this investigation, as well as the

---

rationale of using such a developed database to analyse the accidents. The findings can benefit backward risk diagnosis for accident investigation and forward risk prediction for accident prevention in restricted waters. Specifically, it firstly investigates maritime accident reports from the Global Integrated Shipping Information System (GISIS), as well as accident reports within transit voyages from the Marine Accident Investigation Branch (MAIB) and the Transportation Safety Board of Canada (TSB). Secondly, it uses classical statistical analysis to identify the most frequently appearing risk influential factors (RIFs). Thirdly, to analyse the impact of RIFs, a Tree Augmented Network (TAN) is generated to aid the construction of a data-driven model between different nodes. In addition, the conditional causal relationship among the RIFs is purely configured by historical data. By doing so, this research makes new contributions as follows.

- Developing a new data-driven risk model that enables the systematic analysis of the risk factors of maritime accidents in restricted waters.
- Revealing the dependencies between the RIFs for maritime accidents in restricted waters through a BN model, which is constructed and trained by a TAN learning method.
- Conducting the forward risk prediction for accident analysis and prevention, which generates insights for maritime authorities to reduce the shipping risk in restricted waters.
- Providing useful insights to guide the Suze Canal blockage accident investigation on its risk root causes through backward risk cause diagnosis.

The rest of the paper is organised as follows. Section 2 summarises and critically reviews the literature. Section 3 describes the methodology development for data collection, modelling and validation. In Section 4, the case analysis results show the robustness and reliability of the risk model and reveal the importance of each RIF in the model. Then the accident of the Suez Canal blockage is analysed to generate the implications for accident prevention. Section 5 concludes the paper.

2. Literature review

2.1. Shipping accident analysis

Maritime accidents and catastrophes raise public concerns in terms of casualties, economic loss, and maritime pollution. Owing to the severe-consequence and low-frequency of maritime accidents, there are qualitative and quantitative methods to analyse accidents with limited historical data. Some qualitative studies were proposed to analyse the maritime accidents from systematic perspectives (Kim et al., 2016; Puisa et al., 2018; Uddin and Ibn Awal, 2020). In addition, the Functional Resonance Analysis Model (FRAM) was applied to analyse how shipboard operations cause accidents (Salihoglu and Besikci, 2021). Due to unavailable or nonrepresentative data, Martins et al. (2020) proposed a methodology to assess and quantify the probabilities of occurrence of undesired events based on expert opinions combined with fuzzy analysis. In the light of Maritime Autonomous Surface Ships (MASS), several risk factors have been identified using qualitative methods (Chang et al., 2021; Fan et al., 2020a; Liu et al., 2022). A framework for MASS operating at the third degree of autonomy (remotely controlled ship without seafarers on board) was proposed using literature review and expert knowledge, including 23 human factors, 12 ship factors, 8 environment factors, and 12 technology factors (Fan et al., 2020a). The complex mechanism of system failures was illustrated by these methods to analyse the occurrence of the relevant accidents. However, such qualitative analysis was largely based on expert judgement, causing concerns over subjective bias of the findings.

To further illustrate the causal analysis, historical data from maritime accidents is integrated with expert knowledge to conduct the (semi-) quantitative analysis of maritime accidents. Statistical analysis, Multiple Correspondence Analysis (MCA), and hierarchical clustering were

utilised to explain the causal factors contributing to maritime accidents from a statistical perspective (Chauvin et al., 2013; Ugurlu and Cicek, 2022; Wang et al., 2021). In addition, a statistical analysis was integrated with the implementation scenarios for autonomous ships to quantify the potential reduction in loss of life given autonomous shipping (De Vos et al., 2021). Moreover, the BN method has been applied to predict the occurrence of shipping accidents (Khan et al., 2020; Ung, 2021). Fan et al. (2020b) proposed a data-driven BN model to reflect the interdependencies among human factors, which generated rational scenarios for accident prevention. A dynamic fuzzy BN was developed to list unsafe preconditions and unsafe supervision as the top two considerations for human factors analysis in maritime accidents, especially for supervision failures of shipping companies and ship owners (Qiao et al., 2020). Furthermore, a Human Factor Analysis and Classification System (HFACS) framework was proposed to identify human factors using multiple linear regression (Hasanspahic et al., 2021) and fault tree model (Zhang et al., 2019c). It can also be integrated with BN to reveal accident formation patterns and show that the inland and aged vessels were important factors in sinking and grounding incidents (Ugurlu et al., 2020). In addition, the quantitative research shows that multiple levels of risk factors interacted with each other affect the event chain in maritime accidents, including environmental factors, vessel factors, human and organisational factors, and management factors (Fan et al., 2020c). Among them, human errors and human factors are significant risk factors in shipping accidents (Weng et al., 2019; Chen et al., 2020).

To conduct quantitative analysis of accidents, the database is often used as one of the most available sources to obtain the primary data, including the GISIS (Pristrom et al., 2016; Wang et al., 2022), automatic identification system (AIS) data (Zhang et al., 2019a), and the historical accident data collected from national/regional maritime administrations (Liu et al., 2021; Xu and Hu, 2019; Coraddu et al., 2020). However, such databases reveal different formats to present the results of accident analysis and have no uniform criteria to assess risks. From this perspective, maritime accident reports are commonly used sources with detailed information for the analysis such as: navigational environment, operational process, direct or indirect causes of the accidents, and the actions taken after each event failure (Wang et al., 2021). Also, the potential hazards and causal analysis for various factors, which are not stated in the given public database, are demonstrated in the accident reports with detailed information. However, there are rare studies in which primary data is extracted from maritime accident reports due to the time-consuming process of collecting data from the reports and limited records within public sources (Chauvin et al., 2013; Yildiz et al., 2021; Zhang et al., 2019b). With more regional accident records, accident reports from regional maritime administration were utilised to generate vital risk factors influencing the severity of accidents (Wang and Yang, 2018). To analyse the spatial patterns of global maritime accidents, density analysis and clustering analysis were utilised to find that approximately 60% of serious and very serious accidents happened within 30 nm to the coastline (Wang et al., 2022). In addition, it was found that the small general cargo ships are the riskiest in the coastal waters of China through the analysis of public and national databases (Liu et al., 2021). Therefore, even studies utilising limited content of the data sources revealed rare implications on how to use outcomes of the model to predict the scenarios of maritime accidents.

To bridge the gap, one novelty of this study lies in that it proposes a generic model by extracting all key RIFs from maritime accident reports from a comprehensive perspective, overcoming the drawbacks on risk analysis and prediction with insufficient data sources.

2.2. Risk assessment of maritime accidents in restricted waters

Shipping in restricted waters faces more significant challenges in navigational safety compared to that in open waters, such as hydrodynamic and bank effects. Previous studies show that 46% of collisions occurred in restricted waters (rivers or fairways), and most of those were

under the conduct of a pilot (Chauvin et al., 2013). Regarding the accidents in restricted waters, communication problems on board and bridge resource management (BRM) deficiencies are frequently highlighted issues for collisions in restricted waters. Also, in confined waters such as the narrow channel, ship handling is significantly affected by moments acting between the ship and bank, hydrodynamic forces, and sidewalls of the channel. Among them, the hydrodynamic force is closely linked to the water depth, the ship's speed and longitudinal and lateral distance. This inevitable force has effects on the ship manoeuvring. Therefore, when a ship is approaching the narrow or restricted waters, it may encounter a high risk of collision, grounding, or contact due to the combined effect of various factors.

To assess the risks of shipping in restricted waters, the hydrodynamic effects on vessels from an individual perspective have been investigated, such as hydrodynamic interaction effects between two large vessels in narrow waterways, the minimum safe distance, and appropriate safe speed required to avoid the accident (Lee et al., 2016). With regard to ship handling, Maimun et al. (2013) investigated the manoeuvring performance of an LNG tanker considering the ship bank interaction effects, showing that the interaction effects, fitting fins and enlarging rudder size greatly influenced the ship handling in restricted waters. To investigate the emergency response in confined waters, a three-stage decision-making framework was developed to select the best risk control options in inland waterways: proposing options in the first stage; selecting the most feasible options by comparing cases in the second stage; and making decisions using BN in the third stage (Wu et al., 2017). Previous studies focusing on ship manoeuvring in restricted waters provided advice such as safe distance and safe velocity for ships. However, it cannot generate causal analysis and risk assessment in view of factor interactions in maritime accidents.

In addition, there are few studies on risk assessment of maritime accidents with a focus on restricted waters. A BN model was constructed by domain experts and the Conditional Probability Tables (CPTs) were developed from historical data, to model the collision consequences in downstream of the Yangtze River (Wu et al., 2020). The risk causal model in traffic-intensive waters was constructed to reveal the key risk causal transmission process (Chen et al., 2019). A fuzzy DEMATEL (Decision Making Trial and Evaluation Laboratory) method was used to emphasise that human errors and weaknesses in organisational factors are primarily responsible for accidents in enclosed spaces (Soner, 2021). The significance of investigating human and organisational factors with regard to maritime accidents in restricted waters was illustrated. In addition, BN was utilised to estimate the occurrence likelihood of grounding accidents in the fluctuating backwater zone, which found that the fluctuating backwater zone, the month, and water level were essential factors for grounding accidents in the Three Gorges Reservoir (Jiang et al., 2021).

Moreover, the risk analysis can be found in maritime accident reports. M/V NEW KATERINA grounding in the Suez Canal bank was entirely due to the human error factor, such as lack of personal capability and lack of knowledge (Authority, 2016). The same blame for human factors in the accidents was found in the collision of M/V CHUANHE in the channel of Xiamen (Authority, 2013). Because the master of the ship was unable to keep proper lookout considering it was in the fairway. In addition, there were environmental factors and vessel factors in restricted waters, including drift, wind, high superstructure, hydrodynamic effects, and limited space. It was reported that M/V COMMANDER drifted on to the Round Reef in the channel due to loss of steering and pushed by the wind (Authority, 2014). In the collision on the Kiel Canal between HANSE VISION and BIRKA EXPRESS (Seeunfalluntersuchung, 2010), the high superstructure of the BIRKA EXPRESS made her particularly susceptible to wind. Then an unexpectedly strong turning motion occurred to her due to the caused hydrodynamic effects, which was worsened by the wind. With regard to accident features in narrow waterways, the attempt of ships to take effective actions in limited space with hydrodynamic effects was risky, which enabled the

ship to run aground, even blocking the channel or canal.

Previous maritime accident reports focused on the identification of the risk factors in restricted waters. The thorough literature analysis reveals that there is yet to be a study to investigate how all RIFs influence each other and how RIFs work co-ordinately to contribute to an accident in restricted waters. The Suez Canal blockage accident leading to catastrophic consequences and significant effect, causes a growing concern and public safety awareness on the effectiveness of the current maritime risk analysis methods within the context of restrict waters, as it shows a typical low frequency but high consequence risk feature. In the meantime, they share many common navigational and geographical characteristics that are different with the ones in other regions/waters (e.g. open sea). From this perspective, the presented research makes new contributions by developing a new data-driven BN based risk model, with a special focus on the systematic analysis of the risk factors of maritime accidents in restricted waters. To reveal the dependencies between RIFs for maritime accidents in restricted waters, a BN model is constructed and trained by using the TAN learning method. It conducts backward risk cause diagnosis (after the occurrence of an accident) and forward risk prediction (when the navigational environment triggers a high risk) for accident prevention, which generates insight for maritime authorities to reduce risks. When applied in the recent accidents (e.g., Suze Canal blockage), the model makes a significant contribution to accident investigation, which aids the finding of root causes leading to the occurrence from an applied research perspective.

3. Methodology

3.1. Data collection

To identify risk factors from maritime accidents in restricted waters, data was obtained from the following sources: 1) database established from previous studies (Fan et al., 2020c, 2020d) between January 2012 to December 2017 from MAIB and TSB, and 2) accident reports in English available from the Marine Casualties and Incidents (MCI) module of the GISIS dating from January 1, 2005 to April 9, 2021.

To generate the RIFs in restricted waters, the procedure consists of reports screening, refining, and RIF selection. All available accident reports within the canal or the 'transit' voyage segment are selected for analysis. In this regard, the study generates a database with 61 vessels, including 24 vessels from the GISIS, and 37 vessels from the MAIB and TSB. Referring to the framework by Fan et al. (2020b), records from the GISIS have been refined and analysed manually in the context of the given risk factors.

The risk factors extracted from maritime accident reports are presented in Table 1. However, due to the size of the database (i.e., 61 vessels), the risk factors utilised for the model need to be filtered. The occurrence frequencies of each factor are calculated to rank them. Based on the distribution curve of the ranked factors (from the maximum value to the minimum value), several values (i.e., 0.3443, 0.1475 and 0.0820) leading to a shape change as the turning points are observed and used for the threshold selection. Secondly, after building different models using the turning point values, sensitivity analysis is carried out to determine the appropriate threshold value. When a too-low threshold value is chosen, many RIFs retain in the model, and the model is not sensitive to minor input changes. When the value is set too high, only a few number of nodes are taken into account, leading to the result not being logical and accurate. The threshold value of 0.1475 was selected through many trials and case-by-case comparisons. The model based on this value is proved to be optimal in terms of both the delivery of accurate results and model representation by the retention of key RIFs.

Simultaneously, domain experts (a scientist in maritime risk assessment and an expert in marine engineering with industry experience) were then invited to verify the definitions of RIFs purely based on the accident records. In this process, 'situation awareness' is excluded because it relates to different vague descriptions of the events in

Table 1

Risk factors extracted from maritime accident reports.

Risk factor	Description	Frequency
Weather condition	wind, fog or poor visibility	0.3934
Sea condition	tide, current and waves contribute to the accidents	0.5082
Fairway condition	complex geographic environment, dense traffic, bank effects or hydrodynamic effects that cause a sheer to port side	0.3443
Ship speed	too fast	0.3607
Vessel condition	poor condition of vessels, or increasing complexity of propulsion arrangements, or modification made to vessels and size contributes to the accidents	0.2295
Equipment/device	devices and equipment not fully utilised or operated correctly (*e.g.*, BNWAS (Bridge Navigational Watch & Alarm System) switched off, alarm system not in the recommended position or not noticed)	0.2623
Ergonomic design	ergonomic impact of innovative bridge design (*e.g.*, visual blind sector ahead, motion illusion)	0.2459
Information	insufficient or lack of updated information (*e.g.*, poor quality of equipment data, falsified records of information, reliance on a single piece of navigational equipment, without working indicators or light for necessary observing)	0.5902
Communication	poor communication and coordination with team	0.4754
Supervision	ineffective supervision and supports (lone watchkeeper or working isolated, improper supervision of loading operation)	0.3443
Passage plan	no detailed passage plan or revised passage plan being unsafe	0.1639
Over-reliance on device	over-reliance on devices (*e.g.* AIS and GPS) but omit others, or poor lookout	0.1639
Clear order	no clear order (not accurately interpreting and applying the requirements of a safe manning document)	0.3115
Limited time	limited time to respond	0.1803
Situation awareness	lack of situation awareness	0.1475
Fatigue	fatigue/asleep/tiredness and desire to rest	0.0656
knowledge	unfamiliar with/lack of equipment knowledge, inexperienced, ill-prepared	0.3443
Complacent	complacent about the duties or underestimation of the severity of the condition (low state of alertness)	0.1475
Drugs and alcohol	recreation drugs, alcohol	0.0492
Mental workload	cognitively overloading	0.0820
Distracted	distracted/insufficient attention	0.1639
Regulation	inappropriate or ambiguous code, endorsement, regulations, procedure, instructions, formally published guidance, operation manual, mandatory requirement	0.1639
Risk assessment	lack of risk assessment	0.2787
Management	dysfunctional management system (shore management, maintenance management, bridge source management, bridge team management, on board management, safety management systems, port service, qualification examination, inadequate training, practice, emergency drill)	0.4262
Safety culture	lack of safety culture, precautionary thought	0.2459

restricted waters from accident reports. Furthermore, 'ergonomic design' is excluded due to its uneven distribution in different databases. For instance, it occurs only once in the GISIS but is frequently mentioned in the MAIB database.

3.2. Tree Augmented Network (TAN) modelling

To reveal the dependencies between RIFs for maritime accidents in restricted waters, a BN model is constructed and trained by a learning algorithm. The historical data collected from maritime accident reports

is treated as the training data used to calculate the CPTs in the BN model. In this process, the conditional causal relationship among the RIFs is purely configured by historical data without subject input from domain experts. The data-driven BN model is constructed in two steps: the first step is to generate the structure between different nodes in the network by using the TAN learning method and the second step is to calculate the CPTs by the Gradient method. A BN illustrates a joint probability distribution over a number of random variables, which is an annotated directed acyclic graph (DAG). The definition and applications of BN in accident analysis can be found in the literature (Fan et al., 2020c; Yang et al., 2018a; Wang and Yang, 2018). TAN learning is an optimisation method by training the data collected from accident reports, which defines the optimised tree structure using conditional mutual information between attributes. The details of finding a tree structure and main steps for TAN learning were demonstrated in Yang et al. (2018b) and Fan et al. (2020b), followed by CPTs calculation.

There are several algorithms that can be used to learn CPTs. The features, advantages, and disadvantages of these algorithms are shown in Table 2. Because the proposed model considers missing data within a number of nodes in the BN, Gradient decent is selected as the appropriate learning algorithm to calculate CPTs.

3.3. Model validation

The model validation is conducted using methods including D-separation, expert knowledge, sensitivity analysis, and case study. The expert knowledge and D-separation concept are firstly used to investigate the relationship between nodes. Then, the sensitivity analysis of the BN model and a real case study are conducted to validate the model quantitatively.

3.3.1. Expert knowledge and D-separation

The BN-based risk model requires validation to determine the model's robustness and reliability. For instance, using expert knowledge continues to be a reasonable way to validate when there is limited historical data in the quantitative analysis. Experts are required to provide judgements on whether the BN structure and variables selected for the nodes in the network are appropriate. The BN model can be validated through a panel of experts to utilise their knowledge and experience to ensure consistency (Yu et al., 2020).

The links of nodes in the networks represent the propagation of information between events rather than the causal relationships. It

Table 2

Typical Bayesian learning algorithm.

Name	Features	Advantage	Disadvantage
Counting-learning algorithm (Fan et al., 2020c)	Learn parameters of CPTs from a file of cases	The simplest and fastest algorithm	No missing data or uncertain findings for the learning nodes or their parents
Expectation-maximization (EM) algorithm (Neapolitan, 2004; Lauritzen, 1995)	Compute the expected value of missing data, then search over Bayes net CPTs to maximise the probability of the data given the Bayes net	Learn the latent variables and is more robust	More time-consuming than Gradient
Gradient descent (Yang et al., 2018a)	An iterative process to search the space of Bayes net parameters by using the negative log likelihood as an objective function it is trying to minimise	Converges faster than EM learning	May be more susceptible to local maxima

explains why the anti-causal links existing in the risk based BNs might be valid (Jensen, 1996). From this point of view, the D-separation concept is introduced (Jensen, 2001), which represents conditional independence in the Bayesian probability theory and can be applied to check and modify the BN structure (Yang et al., 2010).

3.3.2. Sensitivity analysis

Mutual information is derived from entropy theory, which describes the uncertainty of the dataset and interprets the entropy reduction. Specifically, it represents the dependence between two variables in the probabilistic theory (Yang et al., 2018b). In this study, mutual information shows how strong the connection between the RIF and the node 'accident type', which can be defined as follows:

$$I(s, \alpha_i) = - \sum_{s,i} P(s, \alpha_{ij}) \log_b \frac{P(s, \alpha_{ij})}{P(s)P(\alpha_{ij})} \qquad (1)$$

where s is 'accident type', α_i represents the ith RIF, α_{ij} represents the jth state of the ith RIF, $P(a_{ij})$ is the probability of the jth state of the ith RIF, $P(s)$ is the probability of s, $P(s, a_{ij})$ represents the joint probability of α_{ij} and s, $I(s, \alpha_i)$ is the mutual information between 'accident type' and the ith RIF in restricted waters. In this way, the value of mutual information works in filtering out the RIFs that show less relevance to the node 'accident type'. By doing this, the remaining RIFs are important variables regarding the node 'accident type' in the BN.

Then, scenario simulation is conducted to determine the effects of different RIFs in a combined way as another form of sensitivity analysis. The traditional method of upgrading the states of one node with the other nodes locked is applicable for two-state variables but does not fit variables with more than two states (Yang et al., 2018b). To overcome this disadvantage, a method proposed by Alyami et al. (2019) calculates the True Risk Influence (TRI) for the multi-state variable against a type of accident (*e.g.* grounding). For instance, to calculate the TRI value between one RIF and grounding, it firstly calculates the High Risk Influence (HRI) by increasing the probability of the state producing the strongest influence on grounding to 100%. Secondly, the Low Risk Inference (LRI) of grounding is calculated by increasing the probability of the state generating the lowest influence on the grounding to 100%. Next, the average value of HRI and LRI is the TRI of each RIF in the 'grounding' accident type. Then, a similar procedure is applied to other accident types respectively. Therefore, the TRI values of variables in different accident types are obtained, which illustrates the RIFs' influences on accident types. The average TRI values representing RIFs' effects on 'accident type' rank the variables' effects on the 'accident type'.

Apart from the above, there are two axioms to be satisfied in sensitivity analysis (Fan et al., 2020c; Yang et al., 2009; Zhang et al., 2013).

Axiom 1. A slight increase or decrease in the prior probabilities of each RIF, should contribute to the corresponding increase or decrease in the posterior probability of the target node.

Axiom 2. The total influence of the integration of the probability variations of x parameters should be not smaller than the one from the set of y (y∈x) RIFs.

To meet the axioms, minor changes of variables are updated in the scenario simulation. Specifically, the state of the first node generating the highest changed value of the first state of 'accident type' is increased by 10%, while the state of the first node generating the lowest changed value of the first state in 'accident type' is decreased by 10%, referring to Yang et al. (2018b). Then, the same approach is applied to the second node of RIFs, and the value is updated. Next, the third node is also included in the same process. In this way, the updated values of the first state of 'accident type' are gradually increasing or decreasing when more RIFs are included. Subsequently, similar procedures are applied to the second, third and fourth states of the node 'accident type'. If the

updated values of 'accident type' are gradually increasing or decreasing along with the continuously changing RIFs, two axioms are satisfied.

3.3.3. Scenarios setting and real case testing

The model validation is also conducted by simulating past maritime accidents. Given the observed states of several nodes, how the target node and other nodes reflect in the model implies whether the model is consistent with reality. By simulating the past maritime accidents with the associated parameter settings, the configuration can be tested when some states of nodes are given. In addition, the real case testing also provides a plausible explanation to the other nodes for the observed findings, which generates insights for maritime accident investigation.

4. Results

4.1. Factor selection

From the established database extracted from the GISIS, MAIB and TSB, there are in total 61 vessels for maritime accident analysis. 'Accident type' is set as the root node in the model, and the state of the root node 'accident type' is shown in Table 3.

With regard to the accidents occurring in canals or in transit voyage segments, the most common types are collision (44.3%) and grounding (21.3%) due to the limited space for ships to take effective manoeuvring and the hydrodynamic effects of the waterways. Besides, contact or crush on the bank also occurs in the canal or channel. Other accidents, including capsize, sinking and falling overboard are not commonly observed in restricted waters. In terms of the location of shipping accidents, the details are shown in Table 4.

Although the database is developed based on 61 vessels, the relevant RIFs are very similar in each investigated accident and hence representative. The RIFs used in this study are filtered based on the frequency of occurrence in Section 3.1. Regarding the factor selection, the details of each RIF are presented in Table 5.

The states of the RIFs are defined according to the literature and accident reports. For example, 'ship age', 'length', 'gross tonnage' and 'time of day' were graded according to Wang and Yang (2018) and Fan et al. (2020b). 'Ship type' is defined according to the records in accident reports. In addition, most two-state RIFs are graded based on whether they are blamed for the failures in accidents. Specifically, some RIFs are similar to risk factors in previous maritime accident analyses but reveal their characteristics within restricted waters. The detailed explanations of such RIFs are illustrated below:

(1) Weather condition

The weather condition of maritime accidents refers to wind, fog, or visibility. Good weather is described in the reports as clear sky, light/low intensity wind, fresh breeze, overcast and good visibility. However, poor weather refers to rainy, drizzling, heavy fog, and poor visibility (300–400m or less than 100m). Specifically, the contexts highlight that wind pushes the vessel towards one side, or the vessel is susceptible to wind, are also explained as poor weather conditions. For example, the BIRKA EXPRESS in Section 2.2 is particularly susceptible to the wind because of the high superstructure around her forecastle.

(2) Sea condition

Table 3
The states of node 'accident type'.

State	Accident type
1	Collision
2	Grounding
3	Contact
4	Others

Table 4
The location of shipping accidents.

No	Location	No	Location
1	Strait of Istanbul	32	The end of South Breakwater
2	Suez Canal, Egypt (Between 59 and 60 km)	33	Entering Gijon, Spain
3	Houston Ship Channel	34	Kings Reach, River Thames, London
4	St Croix Channel	35	Kings Reach, River Thames, London
5	Suez Canal Southbound	36	The entrance to Poole Harbour, UK
6	Channel Of Xiamen, China	37	Grand écueil d'Olmeto shoal
7	Channel Of Xiamen, China	38	Lerwick Harbour, Shetland Islands
8	Port Said, entrance of the Suez Canal	39	Tor Bay
9	Outside of Suez Canal (South Side)	40	English Channel
10	St. Lawrence Seaway	41	Off the coast of north-west Wales
11	Kiel Canal	42	Belfast Lough
12	Kiel Canal	43	Belfast Lough
13	Kiel Canal, canal kilometre (ckm) 95.5	44	6 nm south of Dungeness, UK
14	Kiel Canal, ckm 95.5	45	6 nm south of Dungeness, UK
15	Chile/Canal Sarmiento	46	Transiting the Fraser River, British Columbia
16	Sacramento River Deep Water Ship Channel	47	St. Lawrence River, Montréal, Quebec
17	Kiel Canal	48	St. Lawrence River, Montréal, Quebec
18	Kiel Canal	49	St. Lawrence River, Montréal, Quebec
19	Kiel Canal, canal kilometre (ckm) 6.1	50	St. Lawrence River, Montréal, Quebec
20	Kiel Canal, ckm 6.1	51	Milligan's Wharf, Prince Edward Island
21	Rusterbergen Pilot Station (Kiel)	52	North of Merry Island, British Columbia
22	Rusterbergen Pilot Station (Kiel)	53	Burnside, Newfoundland and Labrador
23	Mongla/Zulfiquar Channel, Bangladesh	54	Port of Saint John, New Brunswick
24	Coast Of Falconara (Italy)	55	Port of Saint John, New Brunswick
25	Varne Bank in the English Channel	56	Port of Montréal, Quebec
26	Lochmaddy, North Uist, Scotland	57	Port of Montréal, Quebec
27	River Humber, UK	58	St. Lawrence Seaway near Iroquois, Ontario
28	River Humber, UK	59	Roberts Bank, British Columbia
29	Bramble Bank, The Solent, UK	60	Gulf of St. Lawrence near Sept-lles, Quebec
30	Jebel Ali, United Arab Emirates	61	Gulf of St. Lawrence near Sept-lles, Quebec
31	Jebel Ali, United Arab Emirates		

Table 5
The details of each RIF.

RIF	Node in the BN	State/Description	State
Ship type	Ship_type	Passenger vessel, tug/barge, container ship, bulk carrier, RORO, tanker or chemical ship, cargo ship, others	1,2,3,4,5,6,7,8
Ship age (years)	Ship_age	0 to 5, 6 to 10, 11 to 15, 16 to 20, more than 20, NA	1, 2, 3, 4, 5, 6
Length (metres)	Length	100 or less, more than 100	1, 2
Gross tonnage (GT)	Gross_tonnage	300 or less, 300 to 10000, greater than 10000	1, 2, 3
Weather condition	Weather_condition	Good, poor (wind, fog, visibility)	1 (good), 2 (poor)
Sea condition	Sea_condition	Good, poor (falling tide, current, waves)	1, 2
Time of day	Time_of_day	07:00 to 19:00, other	1, 2
Fairway condition	Faiway_condition	Good, poor (dense traffic, a bank suction effect or hydrodynamic effects)	1, 2
Ship speed	Ship_speed	Normal, fast	1, 2
Vessel condition	Vessel_condition	Good condition of vessels, or the condition of vessel has nothing to do with the accidents; Poor condition of vessels, or increasing complexity of propulsion arrangements, or modification made to vessels and size contributes to the accidents	1, 2
Equipment/device	Equipment_device	Devices and equipment on board operate correctly; Devices and equipment not fully utilised or operated correctly (*e.g.*, BNWAS switched off, alarm system not in the recommended position or not noticed)	1, 2
Information	Information	Effective and updated information provided; insufficient or lack of updated information (*e.g.*, poor quality of equipment data, falsified records of information, reliance on a single piece of navigational equipment, without working indicators or light for necessary observation)	1, 2
Communication	Communication	Poor communication and coordination with team	1, 2
Supervision	Supervision	Ineffective supervision and supports (lone watchkeeper or working isolated, improper supervision of loading operation)	1, 2
Passage plan	Passage_plan	No detailed passage plan or revised	1, 2

(continued on next page)

The sea condition of maritime accidents refers to the tides, currents, and waves. Good sea condition is stated as slight seas and swell, calm, nil swell, flat sea, or weak currents, while the poor condition is swelling with significant waves, neap tide or strong currents.

(3) Fairway condition

The fairway condition factor represents the density of traffic for general maritime accidents. With regard to the characteristics of transport in restricted waters, the fairway condition is closely linked with bank effects or hydrodynamic effects that cause a sheer to the port side and drift induced by the current.

(4) Ship speed

Each ship has a certain speed that it may attain in canals, but cannot exceed, which is due to the large physical displacement of water when the vessel passes through the canal compared with the canal's width and depth. For example, the permissible velocity through water for non-

Table 5 (*continued*)

RIF	Node in the BN	State/Description	State
		passage plan was unsafe	
Over-reliance on device	OR_device	Over-reliance on devices (AIS, GPS …) but omit others, including poor lookout	1, 2
Clear order	Clear_order	No clear order (not accurately interpret and apply the requirements of a safe manning document)	1, 2
Limited time	Limited_time	Limited time to respond	1, 2
Knowledge	Knowledge	Unfamiliar with/lack of equipment knowledge, inexperienced, ill-prepared	1, 2
Complacent	Complacent	Complacent about the duties or underestimation of the severity of the condition (low state of alertness)	1, 2
Distracted	Distracted	Distracted/insufficient attention	1, 2
Regulation	Regulation	Inappropriate or ambiguous code, endorsement, regulations, procedure, instructions, formal published guidance; operation manual, requirement	1, 2
Risk assessment	Risk_assessment	Lack of risk assessment	1, 2
Management	Management	Dysfunctional management system (shore management, maintenance management, bridge source management, bridge team management, onboard management, safety management systems, port service, qualification examination, inadequate training, practice, emergency drill)	1, 2
Safety culture	Safety_culture	Lack of safety culture, precautionary thought	1, 2

tanker vessels is 16 km/h (approximately 8.5 knots) and 14 km/h for tanker vessels, referring to the New Suez Canal Regulations. Due to narrow waterways for the channel, the effect of shallow waters and high tide also accelerates the vessel's speed.

4.2. Data-driven modelling for maritime accidents

After collecting the maritime accident records, 25 RIFs are analysed to illustrate interdependencies in the BN model. In this way, the structure of BN is trained and optimised by calculating the conditional mutual information in Section 3.2. Then, it is carefully checked by domain experts and D-separation to ensure that all the links between the variables are meaningful.

Once the TAN structure is generated, the parameter learning of CPTs from the cases is conducted by using 'Learn using Gradient' (Yang et al., 2018a). The arrows from one RIF to the others represent their causal relationship. Such a relationship is learnt from the historical accident data. It means that the relationship is developed based on the statistical correlation in terms of the RIFs and their contributions to the occurrence of accidents in restricted waters. The magnitude of such relationship is modelled by the CPTs of each pair of related RIFs. After CPTs are obtained, the posterior probabilities of each node can be calculated, and the results are shown in Fig. 1. The states and explanations of each node (i.e. RIF) are explained in Table 5.

The statistical analysis of the probabilities of nodes represents initial findings regarding various RIFs. Among the accidents, collision is the most frequent accident type, accounting for 40.7%. As for the accidents in restricted waters, most of these ships are in relatively new age, from 0 to 5 years, accounting for 32.4%. Approximately 36.6% of the accidents in canals occur under poor weather conditions, which is predominantly affected by wind or limited visibility. In addition, 47.4% occur with poor sea conditions, which is closely linked to currents, tides and waves. Referring to the features of fairway in restricted waters, 36.7% of vessels involving in the accidents are affected by hydrodynamic effects.

4.3. TAN model validation

The qualitative BN model is first validated by face validation using domain experts. 17 world-leading scholars in the maritime risk area (in terms of the number of SCI-cited core journal publications) were invited via emails[1]. 13 of them provided their feedback, among which 9 agreed with the TAN trained network and 4 suggested the removal of some links. With regard to such links, an extra test was conducted on their impact to the model result accuracy. New scenario tests and analyses were conducted using real cases (20% of the sample size) to compare their prediction performance with the original BN. The results showed that the original model in Fig. 1 (with a 100% accuracy for all cases) outperforms the modified BN (with an average 98% accuracy in 13 real cases and 27.4% accuracy in 1 real case). As a result, the BN structure remains at this stage of the model validation.

Furthermore, more model validation is conducted using other methods, including D-separation, sensitivity analysis, and case study. The validation for the rationality and consistence of the BN is conducted. The D-separation concept is firstly used to investigate the relationship between nodes 'length' and 'weather condition'. With the evidence of node 'accident type', the connection between nodes 'length' and 'weather condition' is independent. Therefore, they are d-separated (conditionally independent) and suit the concept of d-separation with sound links and directions. Then, similar investigations are conducted in other nodes and links, providing confidence that the BN structure is rational.

Sensitivity analysis of the BN model is also conducted. The mutual information between RIFs and accident types is illustrated in Table 6. Referring to Eq.(1), higher $I(s, \alpha_i)$ reveals essential impacts of the RIF on 'accident type'. From this point of view, 'ship type' with corresponding mutual information value of 0.5458 has the most critical impact on the accident type. Variables with higher mutual information are selected as essential RIFs. In order to set a threshold for the selection, the average value of all RIFs is calculated, 0.1051. The variables with values higher than the threshold are extracted for further discussion, including 'ship type', 'ship age', 'gross tonnage', 'passage plan', 'information', 'risk assessment', and 'weather condition'. They are calculated in terms of the quantitative extent to which one RIF influences another in the BN model.

In view of TRI calculation between RIFs and accident types, Table 7 presents the TRI value of 'ship type' against collision where TRI is equal to (HRI + LRI)/2. Table 8 shows all TRI values of RIFs for all accidents.
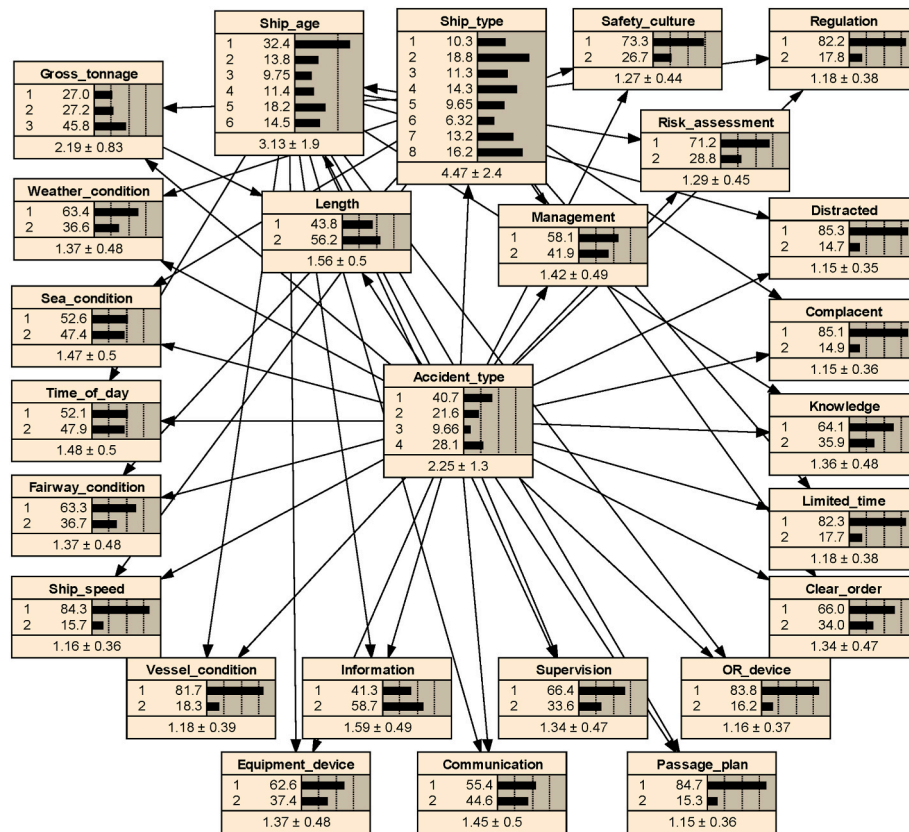
Fig. 1. BN modelling of maritime accidents.

Table 6
Mutual information between RIFs and accident types.

Node	Mutual Information
Ship_type	0.5458
Ship_age	0.2057
Gross_tonnage	0.1810
Passage_plan	0.1618
Information	0.1250
Risk_accessment	0.1174
Weather_condition	0.1171
Knowledge	0.1036
Vessel_condition	0.0942
Safety_culture	0.0882
Clear_order	0.0870
Supervision	0.0797
Length	0.0790
Fairway	0.0781
OR_device	0.0780
Distracted	0.0777
Sea_condition	0.0764
Limited_time	0.0593
Equipment_device	0.0546
Time_of_day	0.0514
Communication	0.0505
Management	0.0505
Ship_speed	0.0489
Complacent	0.0134
Regulation	0.0029

From the results of the updated value of the dependent node, this model is proved to be in line with Axiom 1 (Fan et al., 2020b).

The first row in Table 7 represents the base-case scenario, and the following rows show the different scenarios when each state of the variable reaches 100%. Then the TRI values of 'ship type' against collision are obtained. Through calculating the TRI values of RIFs

against every accident type, key factor identification against different accident types is illustrated. From Table 8, the most important factor for collision is 'ship type' with a TRI value of 39.25, and the most important factor for grounding is 'passage plan' with a TRI value of 28.20. By comparing the average TRI values, the most important variables for 'accident types' are as follows:

Ship type > Ship age > Passage plan > Gross tonnage > Weather condition > Risk assessment > Information

To further validate the BN model, it is examined by testing the combined effect of multiple RIFs on the accident types, referring to Section 3.3.2 written as '~10%' in Table 9. Specifically, the first column of Table 9 represents the original values of accident types in TAN, and the following columns show the updated results with changed values of RIFs. However, each state of 'accident type' is calculated separately from each other, i.e., each row representing each accident type is calculated through the updated change of states of RIFs. In view of the results, the updated values of the 'accident type' are gradually increasing along with the continuously changing RIFs, which is in line with Axiom 2.

Furthermore, a reported maritime accident (which has not been included in the database for the BN construction) is tested to validate the model. Based on the details of accident reports, particular states of the selected relevant variables are given with a probability of 100%. Then the probability of each state of the 'accident type' node is updated accordingly, reflecting the predictive accident type. In this case, it will validate the model if it is consistent with the actual accident type. In the case of grounding of the passenger vessel Royal Iris of the Mersey while manoeuvring toward Eastham Locks at the entrance to the Manchester Ship Canal on the July 10, 2016, the states of nodes in the BN model are assigned in Fig. 2, based on the details of accident report published by the MAIB (report number 11/2017).

Table 7
TRI calculation between ship type and collision.

1	2	3	4	5	6	7	8	Collision	HRI	LRI	TRI
/	/	/	/	/	/	/	/	40.7	53.5	25.0	39.25
100%	0	0	0	0	0	0	0	28.7			
0	100%	0	0	0	0	0	0	15.7			
0	0	100%	0	0	0	0	0	71.4			
0	0	0	100%	0	0	0	0	29.3			
0	0	0	0	100%	0	0	0	62.9			
0	0	0	0	0	100%	0	0	94.2			
0	0	0	0	0	0	100%	0	33.7			
0	0	0	0	0	0	0	100%	37.6			

Table 8
TRI of RIFs for all accident types.

Node	TRI				
	1 Collision	2 Grounding	3 Contact	4 Others	Average
Ship_type	39.25	24.90	18.54	40.93	30.90
Ship_age	19.65	16.85	13.94	23.09	18.38
Gross_tonnage	9.40	8.80	7.19	25.36	12.69
Passage_plan	13.90	28.20	1.86	16.18	15.04
Information	8.80	7.75	8.15	7.10	7.96
Risk_assessment	14.65	1.55	6.92	9.30	8.11
Weather_condition	18.00	3.25	7.46	7.30	9.00

(1) 'Passenger ferry' (state 1) for ship type, 'built in 1959' (state 5) for ship age, '46.45m' (state 2) for length, '464' (state 2) for gross tonnage.

(2) 'July 10, 2016 at 1254 (UTC+1)' (state 1) for time of day.

(3) 'Good visibility' (state 1) for the weather condition.

(4) 'The bridge team were navigating solely by eye and incorrectly assessed that the ferry was in safe water', which reflected the deficiencies for bridge team management and the improper use of the equipment and device onboard.

The result of the updated BN model illustrates the high probability of this accident type being in state 2, i.e., grounding, which is consistent with reality. Moreover, this model reflects more information associated with the accident investigation. According to the statement of the accident report, the information shown on the chart of the area in terms of the status of the mooring dolphins was inaccurate but did not contribute to the grounding. However, the risk factor 'information' with a high probability of being in state 2 is observed in the proposed model, as shown in Fig. 2, which supports the investigation result of the 'information' factor in the accident. It further demonstrates the reliability as well as the usefulness of the constructed BN model.

On the other hand, this model helps illustrate the possibility of the state of each single risk factor by setting parameters in the proposed BN, which benefits the ongoing/future investigation of maritime accidents in restricted waters. For example, one of the world's largest container ships (the Ever Given), an Ultra-Large Container Vessel (ULCV) capable of carrying over 20,000 shipping containers, ran aground at Suez Canal's east bank and caused the blockage, on the March 23, 2021.

The Suez Canal is an artificial sea-level waterway with a length of 193 km, connecting the Mediterranean Sea and the Red Sea. It is the shortest maritime route between Europe and South Asia, which conveys 12% of global shipping. According to the available information currently released from the press, parameter settings for some variables in the proposed BN model can be obtained, as shown in Fig. 3.

(1) 'Container ship' for ship type, '3 years' for ship age (the year of build is 2018).

(2) Weather condition is flawed, because the ship 'was suspected of being hit by a sudden strong wind, causing the hull to deviate from the waterway and accidently hit bottom' (Maguire, 2021). Besides, the 'sail effect' worsens the situation when containers piled high on the top of the large vessel are more susceptible to strong winds (Michaelson and Safi, 2021).

(3) As for the fairway condition, it has been reported that the ship ran fifth in a northbound convoy and in the queue behind it sat fifteen vessels when it ran aground. For just six days, there was traffic jam in both directions of over two hundred vessels (International, 2021).

(4) '07:40' for the time of day.

Under these circumstances, shown in Fig. 3, it can provide a plausible configuration for the observed findings. The above information is used to assign CPTs of the corresponding nodes in the model. Specifically, it reveals a very high probability of 82.0% for the vessel to be involved in grounding, which further validates the proposed BN model, as shown in Fig. 3. Moreover, the ship length has a probability of 97.2% to be in state 2 (i.e., more than 100 m), and gross tonnage has a probability of 96.4% to be in state 3 (i.e., more than 10000 GT), further demonstrating the consistency with the reality ('220940 GT' for the gross tonnage, '399.94 m' for the length).

4.4. Risk prediction and implications from the Suez Canal blockage

This study explains differences among critical factors, contributing to different types of accidents in restricted waters, and provides the most probable scenarios with reference to specific conditions. The Suez Canal
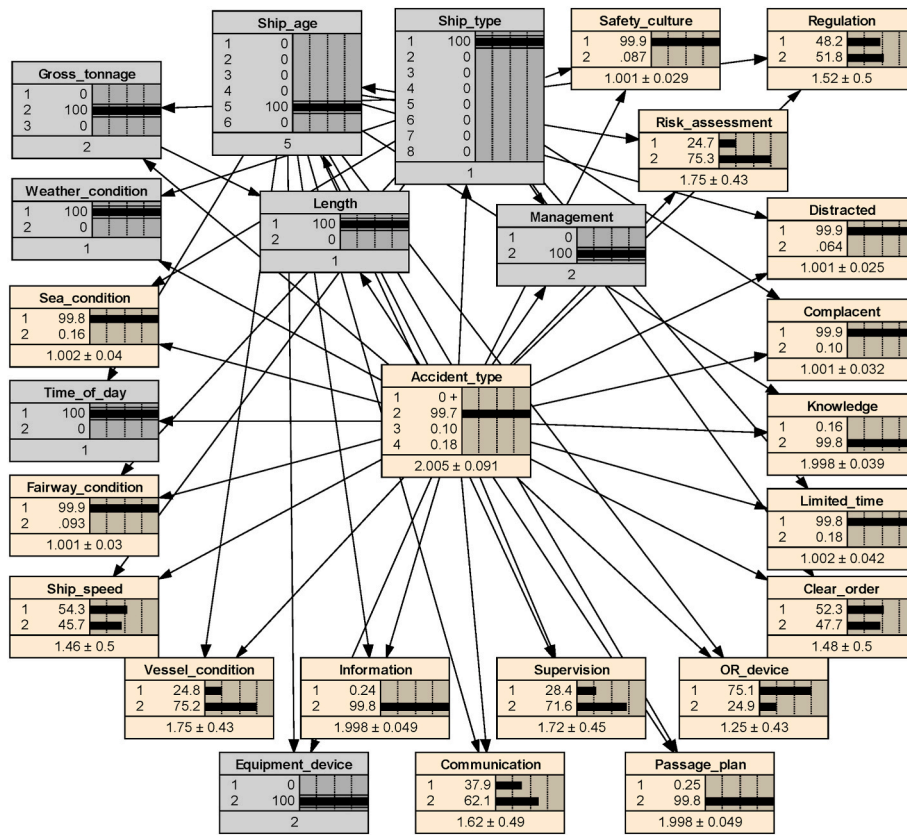
Table 9
Accident rate of the minor change in variables.

Ship type	/	~10%	~10%	~10%	~10%	~10%	~10%	~10%
Ship age	/	/	~10%	~10%	~10%	~10%	~10%	~10%
Gross tonnage	/	/	/	~10%	~10%	~10%	~10%	~10%
Passage plan	/	/	/	/	~10%	~10%	~10%	~10%
Information	/	/	/	/	/	~10%	~10%	~10%
Risk assessment	/	/	/	/	/	/	~10%	~10%
Weather condition	/	/	/	/	/	/	/	~10%
S1 Collision	40.7	41.4	42.0	42.5	45.0	45.7	47.9	49.9
S2 Grounding	21.6	22.4	22.8	23.3	28.2	28.9	28.9	29.5
S3 Contact	9.7	10.0	10.5	10.9	11.1	11.8	12.7	14.0
S4 Others	28.1	29.4	28.9	30.3	33.2	33.8	35.6	36.9

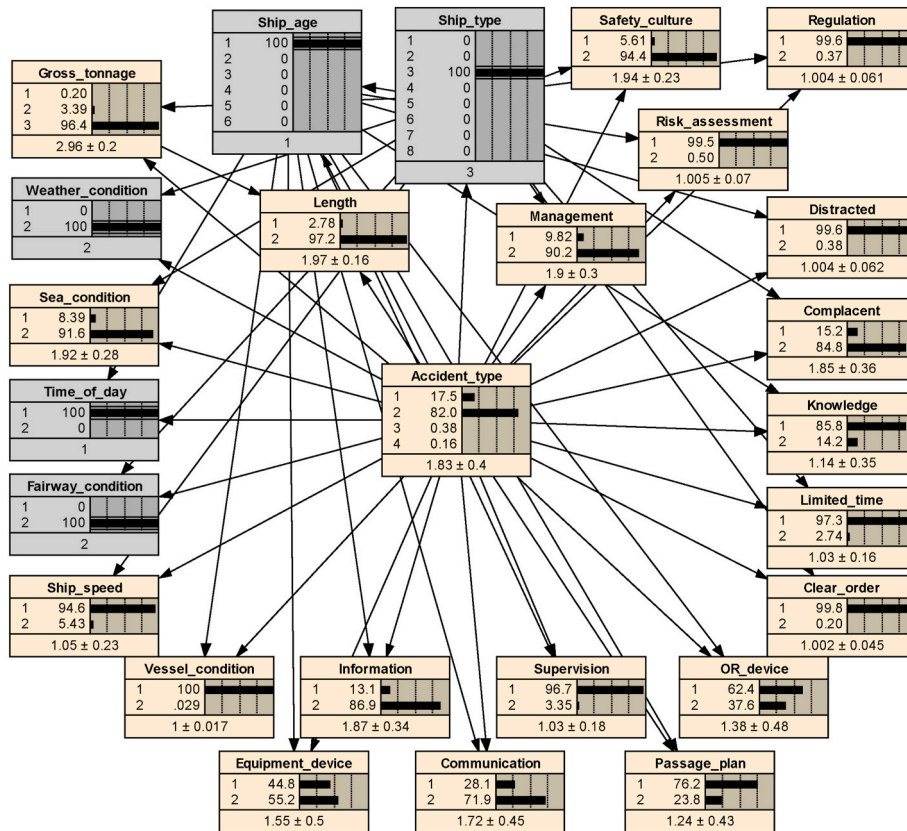**Fig. 2.** Model validation using the past accident.

Fig. 3. Results based on Suez Canal Blockage.

case study generates insights for accident analysis and accident prevention by explaining the most probable scenarios.

It has been addressed by Suez Canal Authorities in a press conference that weather conditions were not the main causes for the ship's grounding, and technical or human errors may exist. Although there is so far no public statement on contributing factors for the accident, this study provides a systematic perspective considering vessel factors, environmental factors, and human factors. Specifically, the states of nodes in Fig. 3 show implications for occurrence probabilities of identified contributing factors.

(1) There is a probability of 86.9% for the Ever Given having insufficient information.

Vessels need to obtain adequate and updated information in shipping, especially when in restricted waters. Insufficient information may relate to poor quality of equipment data, falsified records of information, reliance on a single piece of navigational equipment, no working indicators or light for necessary observation.

(2) There is a 71.9% probability for the Ever Given having poor communication.

Communication makes the teamwork onboard safe and effective, while poor communication not only affects the daily routine duties with high risks, but also deteriorates the situation in an emergency response.

(3) There is a probability of 84.8% for the Ever Given involving in the complacent issue.

The complacent issue means the situation when seafarers are complacent about the duties and underestimate the severity of the condition, which leads to actions contributing to subsystem failures. It emphasises the high risk of complacent issues contributing to the misconduct of seafarers and wrong decision-making.

(4) There is a probability of 90.2% for the Ever Given under dysfunctional management system.

The management system consists of shore management, maintenance management, bridge resource management, bridge team management, port service, safety management system, qualification examination, training, practice, and emergency drill. In restricted waters, dysfunctional management system contributes to technical or nontechnical errors due to human and organisational factors.

(5) There is a probability of 94.4% for the Ever Given with inadequate safety culture.

Safety culture defines how safety is managed onboard a vessel and can be illustrated as the way of doing things on board. It reflects the perceptions and values of the crew concerning safety, which may be influenced by management factors and commercial pressure. However, it can be difficult to measure or quantify.

The Suez Canal case study results show the probability of node states when the states of other nodes are observed, which provides a plausible explanation to the investigated nodes for the observed findings. Therefore, this model explains risk factors' states and reveals the interrelationships between risk factors, helping investigate the hidden causes of maritime accidents in restricted waters. By using scenario testing, the proposed model provides clues about the deficiency in the system and the probability of ineffective responses to the accidents. In addition, the proposed model helps explore the possibility of reducing risks. The states of other nodes also contribute to the value of grounding probability in the BN model. For example, by only assigning the state of 'communication' to be '1' (good communication), the grounding probability of such an accident decreases from 82% to 61.4%. By only assigning the state of 'complacent' to be '1' (good and proper perceptions of duties and situations), the grounding probability decreases to 0.037%. Thus, it reflects insights for applying this model to the past and ongoing investigation of maritime accidents in restricted waters, which benefits backward risk cause diagnosis (after the occurrence of an accident) and forward risk prediction (when RIFs trigger a high risk) for accident prevention.

5. Conclusion

This study proposes a data-driven TAN approach to investigate how different RIFs generate impact on maritime accident types in restricted waters. Maritime accident reports from the MAIB and TSB from 2012 to 2017 and the GISIS from 2005 to 2021, are reviewed and refined to develop the primary database to identify RIFs. Then, the BN model is constructed by the TAN approach to analyse risk factors in maritime accidents. The sensitivity analysis and case study are conducted to validate the model. Lastly, the Suez Canal blockage case study is analysed to provide insights for risk assessment and implications for accident investigation.

According to the mutual information and TRI calculations, important RIFs for shipping accidents in restricted waters are ranked against accident types, *i.e.*, 'ship type', 'ship age', 'passage plan', 'gross tonnage', 'weather condition, 'risk assessment', and 'information'. The model shows that ship length has less risk contribution in restricted waters than in ports. Because a ship needs to change its course in ports while maintaining a certain course following the geological characteristics of the passing canals and channels. It is difficult for a large ship to change course. In other words, it is easy to maintain the course of a large ship in canals/channels. The statistics also support that among all the investigated accidents, most ships involved in accidents are small to medium ships (200 m or less). Although the ship length as an individual factor has shown limited impact on the accidents in restricted waters, its combined effect with the other factors, such as 'gross tonnage', has shown its importance indirectly. Specifically, factor identification against different accident types is demonstrated according to the TRIs of RIFs against each accident type. Meanwhile, the case study shows implications from a plausible explanation for the observed findings by scenario analysis. There is a high probability for the accident of Ever Given with insufficient information, poor communication, a complacent issue, a dysfunctional management system, and inadequate safety culture.

Compared with the established general models in inland waterways (Zhang et al., 2013) and coastal waters (Wang and Yang, 2018), the proposed model shows very different results, reflecting the unique characteristics of restricted waters. The general models could not be used to analyse the shipping risk in restricted waters because they fail to incorporate a few key RIFs such as 'fairway condition'. Oppositely, some RIFs concerned in the general models such as 'season' are not applicable in restricted waters. Although different accident investigation organisations have a variety of methods or frameworks to conduct investigations, this model identifies contributing factors by predicting the probabilities of nodes' states in the model regarding human factors, environmental factors and vessel factors. Therefore, it helps identify the potential hazards by predicting contributing factors and effectively assists maritime authorities with accident investigation.

As for the lessons learned from maritime accident reports in restricted waters, it is evident that the manoeuvring of a vessel is heavily affected due to the effect of lateral banks. Effective recommendations are given to provide awareness training to the crews to have preparedness for manoeuvring in restricted waters. For example, the multipurpose cargo ship BBC STEINHOEFT (C0008890 -M11C0001, IMO number 9358046) in the South Shore Canal of the St. Lawrence Seaway was grounded on the March 31, 2011. As the vessel approached the entrance to the narrower area of the channel, it suffered the bank suction effect,

which caused a sheer to port. In addition, factors such as wind, current, and drift are also closely linked with grounding and collision of the vessels in restricted waters. The joint impact analysis using the proposed model in this paper can effectively prevent the occurrence of a similar accident in future. In some cases, these factors cannot be blamed for the leading causes of the accidents, as risk factors are interacted with each other, especially human factors. Therefore, lessons from marine casualty highlight the significance of human factors regarding negligence and good seamanship. For example, good communication will decrease the grounding probability of Suez Canal accident from 82% to 61.4%. The proper perceptions of duties and situations will significantly decrease the probability of grounding. The success of crews in intervening in minor mistakes or violations benefits the risk control of the system. From this perspective, risk assessment of maritime accidents in restricted waters provides insights for accident prevention strategies.

Generally, the results from the proposed approach present differentiations among the vital RIFs contributing to different types of accidents in restricted waters, which helps provide implications for accident investigation and prevention. As the wind is not the leading cause of the accidents, the case study of the Suez Canal blockage shows a plausible explanation of the scenario to find the potential hazards from a systematic perspective. It implies clues for the accident investigation. There is a recommendation for captains to take training courses aiming at safely navigating ships in restricted waters within wind and current effects for the individuals.

Furthermore, the proposed model is capable of simulating the interactions between risk factors and presenting the probability of states of investigated nodes, which helps provide guides for relevant investigation. For maritime authorities and ship owners, it is possible to obtain information from the proposed model to investigate accidents, manage risk levels of the voyage, and eliminate the reoccurrence of the similar accidents that cause enormous economic losses or casualties influencing the associated whole supply chain. Nevertheless, the limitation of this study lies in the comparatively small number of data records used in the model for a particular restricted water, meaning that the findings are only representative to the generic restrict waters. When more data becomes available with regards to a particular restrict water/region, the model in this study can be used as the basis to support further in-depth analysis (e.g. location-related ones) to generate the results of more specific implications for accident prevention. A thorough analysis shows that many accidents in narrow waters in specific and in other waters in general lacked detailed information recorded on the investigated RIFs. Therefore, a compromising solution is proposed to analyse all the maritime accident reports and derive the primary data ourselves. This work identifies and refines every piece of data with detailed attributes, which increases the quality of data. Collecting the primary data also benefits the feasibility of adjusting the RIFs (i.e., nodes of BN) to obtain an optimal solution between accuracy and easiness. Through this fine-tuning process, the model delivers robust results even for some recent accidents. It helps create a number of scenarios in the form of IF-THEN by locking a few nodes (including the occurrence of the accident), thus producing the possible causes (the change of root causes). This further verifies the model and hence generates useful guidance for accident investigation and prevention.

CRediT authorship contribution statement

Shiqi Fan: Methodology, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration. **Zaili Yang:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Jin Wang:** Validation, Formal analysis, Investigation, Resources, Writing – review & editing, Supervision. **John Marsland:** Resources, Conceptualization, Writing – review & editing, Supervision.

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.

References

Alyami, H., Yang, Z., Riahi, R., Bonsall, S., Wang, J., 2019. Advanced uncertainty modelling for container port risk analysis. Accid. Anal. Prev. 123, 411–421.

Authority, P.M., 2013. DIAM. M/V "CHUANHE" R- 062-2013.

Authority, P.M., 2014. M/V"COMMANDER"R-036 -2014. DIAM.

Authority, P.M., 2016. M/V"NEW KATERINA"R-018-2016. DIAM.

Chang, C.H., Kontovas, C., Yu, Q., Yang, Z.L., 2021. Risk Assessment of the Operations of Maritime Autonomous Surface Ships. Reliability Engineering & System Safety, p. 207.

Chauvin, C., Lardjane, S., Morel, G., Clostermann, J.P., Langard, B., 2013. Human and organisational factors in maritime accidents: analysis of collisions at sea using the HFACS. Accid. Anal. Prev. 59, 26–37.

Chen, D.J., Pei, Y.L., Xia, Q., 2020. Research on human factors cause chain of ship accidents based on multidimensional association rules. Ocean Eng. 218, 9.

Chen, Y.J., Liu, Q., Wan, C.P., 2019. Risk causal analysis of traffic-intensive waters based on infectious disease dynamics. J. Mar. Sci. Eng. 7, 19.

Coraddu, A., Oneto, L., De Maya, B.N., Kurt, R., 2020. Determining the most influential human factors in maritime accidents: a data-driven approach. Ocean Eng. 211.

De Vos, J., Hekkenberg, R.G., Banda, O.A.V., 2021. The Impact of Autonomous Ships on Safety at Sea - A Statistical Analysis. Reliability Engineering & System Safety, p. 210.

Fan, C.L., Wrobel, K., Montewka, J., Gil, M., Wan, C.P., Zhang, D., 2020a. A framework to identify factors influencing navigational risk for Maritime Autonomous Surface Ships. Ocean Eng. 202.

Fan, S., Blanco-Davis, E., Yang, Z., Zhang, J., Yan, X., 2020b. Incorporation of human factors into maritime accident analysis using a data-driven Bayesian network. Reliab. Eng. Syst. Saf. 203, 107070.

Fan, S., Yang, Z., Blanco-Davis, E., Zhang, J., Yan, X., 2020c. Analysis of maritime transport accidents using Bayesian networks. Proc. Inst. Mech. Eng. O J. Risk Reliab. 234, 439–454.

Fan, S.Q., Zhang, J.F., Blanco-Davis, E., Yang, Z.L., Yan, X.P., 2020d. Maritime accident prevention strategy formulation from a human factor perspective using Bayesian Networks and TOPSIS. Ocean Eng. 210, 12.

Hasanspahic, N., Vujicic, S., Francic, V., Campara, L., 2021. The role of the human factor in marine accidents. J. Mar. Sci. Eng. 9.

International, U.P., 2021. *Grounded container ship blocks traffic on Suez Canal* [Online]. Available: https://gephardtdaily.com/national-international/grounded-container-ship-blocks-traffic-on-suez-canal/.

Jensen, F.V., 1996. An Introduction to Bayesian Networks. UCL press London.

Jensen, F.V., 2001. Causal and Bayesian Networks. *Bayesian Networks and Decision Graphs*. Springer.

Jiang, D., Wu, B., Cheng, Z.Y., Xue, J., Van Gelder, P., 2021. Towards a Probabilistic Model for Estimation of Grounding Accidents in Fluctuating Backwater Zone of the Three Gorges Reservoir. Reliability Engineering & System Safety, p. 205.

Khan, B., Khan, F., Veitch, B., 2020. A Dynamic Bayesian Network model for ship-ice collision risk in the Arctic waters. Saf. Sci. 130.

Kim, T.-E., Nazir, S., Øvergård, K.I., 2016. A STAMP-based causal analysis of the Korean Sewol ferry accident. Saf. Sci. 83, 93–101.

Lauritzen, S.L., 1995. The EM algorithm for graphical association models with missing data. Comput. Stat. Data Anal. 19, 191–201.

Lee, C.-K., Moon, S.-B., Jeong, T.-G., 2016. The investigation of ship maneuvering with hydrodynamic effects between ships in curved narrow channel. Int. J. Nav. Archit. Ocean Eng. 8, 102–109.

Liu, K.Z., Yu, Q., Yuan, Z.T., Yang, Z.S., Shu, Y.Q., 2021. A Systematic Analysis for Maritime Accidents Causation in Chinese Coastal Waters Using Machine Learning Approaches, vol. 213. Ocean & Coastal Management.

Liu, W., Liu, Y., Bucknall, R., 2022. Filtering based multi-sensor data fusion algorithm for a reliable unmanned surface vehicle navigation. J. Mar. Eng. Technol. 1–17.

Maguire, D., 2021. The Ship Blocking the Suez Canal Is Called Ever Given, Even Though 'Evergreen' Is Written across it in Huge Letters. ABC News.

Maimun, A., Priyanto, A., Rahimuddin, Siana, Y., Awal, Z.I., Celement, C.S., , NURCHOLIS, Waqiyuddin, M., 2013. A mathematical model on manoeuvrability of a LNG tanker in vicinity of bank in restricted water. Saf. Sci. 53, 34–44.

Martins, M.R., Pestana, M.A., Droguett, E.A.L., 2020. A methodology for assessing the probability of occurrence of undesired events in the Tiete-Parana inland waterway based on expert opinion. Risk Anal. 40, 1279–1301.

Michaelson, R., Safi, M., 2021. Who pays for Suez blockage? Ever Given grounding could spark years of litigation [Online]. Available: https://www.theguardian.com/busine ss/2021/mar/30/years-of-litigation-could-lie-ahead-as-insurers-eye-cost-of-ever-gi ven-grounding.

Neapolitan, R.E., 2004. Learning Bayesian Networks. Pearson Prentice Hall, Upper Saddle River, NJ.

Pristrom, S., Yang, Z., Wang, J., Yan, X., 2016. A novel flexible model for piracy and robbery assessment of merchant ship operations. Reliab. Eng. Syst. Saf. 155, 196–211.

Puisa, R., Lin, L., Bolbot, V., Vassalos, D., 2018. Unravelling causal factors of maritime incidents and accidents. Saf. Sci. 110, 124–141.

Qiao, W.L., Liu, Y., Ma, X.X., Liu, Y., 2020. Human factors analysis for maritime accidents based on a dynamic fuzzy Bayesian network. Risk Anal. 40, 957–980.

Salihoglu, E., Besikci, E.B., 2021. The use of Functional Resonance Analysis Method (FRAM) in a maritime accident: a case study of Prestige. Ocean Eng. 219.

Seeunfalluntersuchung, B.F., 2010. Collision on the Kiel Canal (NOK) between HANSE VISION and BIRKA EXPRESS.

Soner, O., 2021. Application of fuzzy DEMATEL method for analysing of accidents in enclosed spaces onboard ships. Ocean Eng. 220.

Topham, G., 2021. How the Suez Canal Blockage Can Seriously Dent World Trade [Online]. Available: https://www.theguardian.com/business/2021/mar/26/how-th e-suez-canal-blockage-can-seriously-dent-world-trade.

Uddin, M.I., Ibn Awal, Z., 2020. Systems-theoretic approach to safety of inland passenger ship operation in Bangladesh. Saf. Sci. 126, 12.

Ugurlu, H., Cicek, I., 2022. Analysis and assessment of ship collision accidents using fault tree and multiple correspondence analysis. Ocean Eng. 245, 17.

Ugurlu, O., Yildiz, S., Loughney, S., Wang, J., Kuntchulia, S., Sharabidze, I., 2020. Analyzing collision, grounding, and sinking accidents occurring in the Black sea utilizing HFACS and Bayesian networks. Risk Anal. 40, 2610–2638.

UNCTAD, 2020. Review of Maritime Transport 2020. United nations conference on trade and development.

Ung, S.T., 2021. Navigation Risk Estimation Using a Modified Bayesian Network Modeling-A Case Study in Taiwan. Reliability Engineering & System Safety, p. 213.

Wang, H.X., Liu, Z.J., Liu, Z.C., Wang, X.J., Wang, J., 2022. GIS-based analysis on the spatial patterns of global maritime accidents. Ocean Eng. 245.

Wang, H.X., Liu, Z.J., Wang, X.J., Graham, T., Wang, J., 2021. An analysis of factors affecting the severity of marine accidents. Reliab. Eng. Syst. Saf. 210, 12.

Wang, L., Yang, Z., 2018. Bayesian network modelling and analysis of accident severity in waterborne transportation: a case study in China. Reliab. Eng. Syst. Saf. 180, 277–289.

Weng, J.X., Yang, D., Chai, T., Fu, S.S., 2019. Investigation of occurrence likelihood of human errors in shipping operations. Ocean Eng. 182, 28–37.

Wu, B., Tian, H.B., Yan, X.P., Soares, C.G., 2020. A probabilistic consequence estimation model for collision accidents in the downstream of Yangtze River using Bayesian Networks. Proc. Inst. Mech. Eng. O J. Risk Reliab. 234, 422–436.

Wu, B., Yan, X.P., Wang, Y., Zhang, D., Soares, C.G., 2017. Three-stage decision-making model under restricted conditions for emergency response to ships not under control. Risk Anal. 37, 2455–2474.

Xu, S., Hu, H., 2019. Development of a maritime safety management database using relational database approach. Int. J. Shipp. Transp. Logist. (IJSTL) 11, 334–353.

Yang, Z., Bonsall, S., Wang, J., 2010. Facilitating uncertainty treatment in the risk assessment of container supply chains. J. Mar. Eng. Technol. 9, 23–36.

Yang, Z., Yang, Z., Yin, J., Qu, Z., 2018a. A risk-based game model for rational inspections in port state control. Transport. Res. E Logist. Transport. Rev. 118, 477–495.

Yang, Z.L., Wang, J., Bonsall, S., Fang, Q.G., 2009. Use of fuzzy Evidential Reasoning in maritime security assessment. Risk Anal. 29, 95–120.

Yang, Z.S., Yang, Z.L., Yin, J.B., 2018b. Realising advanced risk-based port state control inspection using data-driven Bayesian networks. Transport. Res. Pol. Pract. 110, 38–56.

Yildiz, S., Ugurlu, O., Wang, J., Loughney, S., 2021. Application of the HFACS-PV Approach for Identification of Human and Organizational Factors (HOFs) Influencing Marine Accidents. Reliability Engineering & System Safety, p. 208.

Yu, Q., Liu, K.Z., Chang, C.H., Yang, Z.L., 2020. Realising Advanced Risk Assessment of Vessel Traffic Flows Near Offshore Wind Farms. Reliability Engineering & System Safety, p. 203.

Zhang, D., Yan, X.P., Yang, Z.L., Wall, A., Wang, J., 2013. Incorporation of formal safety assessment and Bayesian network in navigational risk estimation of the Yangtze River. Reliab. Eng. Syst. Saf. 118, 93–105.

Zhang, L.Y., Meng, Q., Fwa, T.F., 2019a. Big AIS data based spatial-temporal analyses of ship traffic in Singapore port waters. Transport. Res. E Logist. Transport. Rev. 129, 287–304.

Zhang, L.Y., Wang, H., Meng, Q., Xie, H.B., 2019b. Ship accident consequences and contributing factors analyses using ship accident investigation reports. Proc. Inst. Mech. Eng. O J. Risk Reliab. 233, 35–47.

Zhang, M.Y., Zhang, D., Goerlandt, F., Yan, X.P., Kujala, P., 2019c. Use of HFACS and fault tree model for collision risk factors analysis of icebreaker assistance in ice-covered waters. Saf. Sci. 111, 128–143.