

Article

Not peer-reviewed version

Probabilistic Modeling of Maritime Accident Scenarios Leveraging Bayesian Network Techniques

[Shiguan Liao](#) , [Jinxian Weng](#) ^{*} , [Zhaomin Zhang](#) ^{*} , Zhuang Li , Fang Li

Posted Date: 22 June 2023

doi: 10.20944/preprints202306.1597.v1

Keywords: maritime traffic safety; maritime accident; Bayesian network (BN); accident scenario analysis; Netica



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Probabilistic Modeling of Maritime Accident Scenarios Leveraging Bayesian Network Techniques

Shiguan Liao ¹, Jinxian Weng ^{2,*}, Zhaomin Zhang ^{1,*}, Zhuang Li ³ and Fang Li ²

¹ School of Management, Shenzhen Polytechnic, Shenzhen 518055, China

² College of Transport and Communications, Shanghai Maritime University, Shanghai 201306, China

³ Naval Architecture and Shipping College, Guangdong Ocean University, Zhanjiang 524088, China

* Correspondence: jxweng@shmtu.edu.cn (J.W.); zhzhmin@szpt.edu.cn (Z.Z.)

Abstract: This paper presents a scenario evolution model for maritime accidents using Bayesian networks (BN) to predict the most likely causes of specific types of maritime incidents. The BN nodes encompass accident type, life loss contingency, accident severity, the quarter and time period of the accident, and the type and gross tonnage of the ships involved. We analyzed 5,660 global maritime accidents from 2005 to 2020. Using Netica software, we constructed a Tree Augmented Network (TAN) model, accounting for interdependencies among risk influencing factors. We validated the results through sensitivity analysis and historical accident records. Forward causal inference and reverse diagnostic inference were then performed on each node variable to investigate the accident development trend and evolution process under predetermined conditions. The findings indicate that the model can effectively predict the likelihood of various accident scenarios under specific conditions, as well as the extrapolation of accident consequences. Forward causal reasoning reveals that general cargo ships with a gross tonnage of 1-18,500 t are most likely to experience collision, grounding, and stranding accidents in the first quarter. Reverse diagnostic reasoning indicates that during early morning hours, container ships, general cargo ships, and chemical ships with a tonnage of 1-18,500 t are less likely to involve life loss in the event of collision accidents.

Keywords: maritime traffic safety; maritime accident; Bayesian network (BN); accident scenario analysis; Netica

1. Introduction

The intricately variable and multifaceted climatic conditions, endemic to the expansive marine environment, have perpetually underscored maritime transport as a vocation of considerable risk. Incidents disrupting maritime transit, encompassing a spectrum of occurrences from vessel collisions to groundings, and from onboard fires to devastating explosions, bear the hallmark of low frequency yet are marked by their profoundly destructive aftermath [1]. Upon the unfortunate manifestation of a maritime accident, a tsunami of undesirable outcomes typically ensues, notably substantial financial loss, a tragic toll of human casualties, or even the insidious onset of extensive environmental pollution [2]. As an imperative and fundamental step towards mitigating the inherent perils of maritime navigation, a comprehensive exploration into the multifactorial etiology of these traffic mishaps proves indispensable [3–5]. Moreover, it becomes increasingly pivotal to meticulously undertake a dynamic risk evaluation, focusing on the myriad facets of maritime operations [6–8]. Complementing this, the development and implementation of robust predictive models, which can potentially forecast the calamitous conjunction of human fatality and its probability in the event of an accident, can contribute significantly towards minimizing future maritime disasters [9–12].

Over the years, an impressive corpus of scholarly efforts has been devoted to enhancing our understanding of maritime traffic safety, including explorations into accident causation analysis [3,4,5,13], accident consequence assessment [14–16], and accident loss computation [17–18]. These

endeavors have given rise to an array of innovative evaluation methodologies. In one notable study, Hu et al. [3] skillfully harnessed the capabilities of the Human Failure Analysis and Classification System (HFACS) in tandem with Structural Equation Modeling (SEM) to disentangle the intricate web of causal factors underpinning Marine Traffic Accidents (MTAs). Chou et al. [4], in a synergistic integration of technologies, amalgamated the Automatic Identification System, Geographic Information System, and an electronic chart (e-chart) to scrutinize the interplay between environmental factors, geographical locations, and common causes of marine mishaps. By overlaying vessel traffic flows, accident sites, and environmental data on a shared e-chart, their research unfurled valuable insights for port authorities in streamlining ship traffic flow and curtailing the prevalence of marine accidents in the vicinity of ports. Meanwhile, Xue et al. [5] proffered a comprehensive analytical framework for investigating the peculiarities and causative factors of ship accidents, utilizing a decade's worth of historical data harvested from the capriciously fluctuating backwater expanse of the Three Gorges Reservoir region. Their extensive work yielded a thorough summary and visualization of vessel accident categories and severity, involved vessel types, spatial-temporal distribution characteristics, vessel accident losses, along with the underlying causes and lessons gleaned from pertinent accidents, achieved through a rigorous statistical and comparative analysis of historical data. Elsewhere, Fu et al. [13] engineered a bivariate probit model to delve into an array of 311 Arctic ship accidents spanning from 1998 to 2017. Their study brought to the fore influential factors such as gross tonnage, ship type, ship age, accident type, accident year, accident location, wind, and sea ice as primary contributors to accident severity. Simultaneously, their research unveiled an intriguing negative correlation between serious accidents and those resulting in pollution. As research on maritime traffic accidents has illuminated a gamut of potential causative factors, an increased granularity of available accident data has spurred a growing number of scholars to concentrate on the ramifications of these mishaps, specifically the evaluation of accident consequences and loss computation. Such undertakings have risen to prominence, particularly in the eyes of managers concerned with incidents that yield significant economic damage and human casualties. For instance, Chen et al. [15] presented an evidence-based Fuzzy Bayesian Network methodology to erect probabilistic models of marine accidents, thereby enabling the appraisal of accidents likely to spawn severe consequences. In a similar vein, Ventikos and Giannopoulos [16] introduced a criterion for assessing risk and repercussions within the maritime transport sector from the societal perspective, thereby formulating a novel framework for marine risk assessment that facilitates the comparison of disparate accident scales and characteristics, while accurately mirroring the risk threshold society is prepared to tolerate. Chen et al [17] pioneered an enhanced entropy weight-TOPSIS model to furnish a holistic analysis and appraisal of marine total loss incidents, encompassing a global scope from 1998 to 2018. These studies, though highly impactful, predominantly undertake analyses either from the standpoint of accident causation or evaluation of accident consequences. Rarely do these scholarly pursuits straddle both domains in a bidirectional inquiry.

In the realm of accident scenario analysis, methodological constructs like event tree analysis and accident tree analysis find frequent utilization in the assembly of traffic accident scenario evolution models [19–23]. However, the breadth of most accident-cause analyses often overshadows their specificity, impeding their ability to yield targeted recommendations for forestalling analogous events [24]. To bridge this gap, scholars may employ a Bayesian network-based maritime accident scenario modeling approach. Bayesian networks stand as a form of probabilistic graphical model, deftly equipped to encapsulate and deliberate over uncertain knowledge and nebulous relationships among variables. This versatile modeling approach, designed to embrace the labyrinthine and dynamic character of maritime activities, excels at discerning the contributory factors precipitating maritime accidents [3,5]. Employing a synergistic blend of historical data and expert acumen, this model can approximate both the likelihood of an accident's occurrence and the potential fallout arising from a range of accident scenarios [14,16]. Bayesian networks (BN) find broad application in confronting uncertain multi-factor causality inference, accident causation analysis, and scenario prediction, making them invaluable tools in road and waterway transportation sectors [25–32].

Various scholars have employed these tools in diverse studies: Zou and Yue [33] melded probabilistic risk analysis with BN theory to explore the origins of road traffic accidents; Yuan et al. [34] constructed a scenario-derived prediction model for the repercussions of fire accidents in oil and gas storage and transportation emergency processes, leveraging a defuzzification method and a dynamic BN model. Other researchers, such as Zhao et al. [35], used the ISM-BN model to assess the impact of varying factors on maritime safety, successfully pinpointing the critical risk components for different accident types. Afenyo et al. [36] utilized a BN model to sketch an Arctic shipping accident scenario and illuminate the crucial causative elements of a potential accident scenario. Similarly, Jiang et al. [37] proposed a Bayesian network-based risk analysis strategy for evaluating maritime accidents along the 21st century Maritime Silk Road (MSR), identifying the principal influencing factors to bolster accident prevention measures and ensure maritime transportation's safety and sustainability. In a more focused study, Si et al. [38] employed a BN structure learning algorithm that paired kernel density estimation with a model weighted average strategy to dissect the causative elements of container ship collisions, basing their analysis on a limited set of container ship collision sample data. Other studies like Fan et al. [39] and Hänninen et al. [40] proposed similar Bayesian network-based risk analysis approaches for understanding the contributing factors to maritime transport accidents, with the latter focusing more on maritime safety management and its relationship with maritime traffic safety. Despite these successes, these aforementioned studies suffer from a triad of limitations: 1) a paucity of sample data from maritime accidents, 2) a labor-intensive and time-consuming data collection process, and 3) the inherent difficulty in obtaining accident loss records. Summarily, while waterway transportation research has honed its focus on accident causality reasoning and accident causation analysis, there remains a conspicuous void in the research landscape pertaining to accident scenario modeling.

In light of this, the paper aims to build a BN model for the evolution of maritime accident scenarios using global maritime accident data. The data derive from the Global Integrated Shipping Information System (GISIS) established by the International Maritime Organization (IMO) and have been widely used by scholars in maritime accident studies [41–47]. A Bayesian network-based maritime traffic accident scenario modeling approach is proposed to analyze the causes of maritime traffic accidents, perform dynamic risk assessment of shipping activities, and predict the probability of occurrence and consequences of accidents. By identifying the influencing factors and simulating various accident scenarios, the proposed approach can help maritime stakeholders to implement appropriate preventive measures and improve the safety of maritime transportation. The proposed model can predict the most likely causal factors leading to specific accident consequences, thereby providing technical support for the practical direction of shipping safety management strategies.

The rest of this article is structured as follows. Section 2 briefly introduces the structure and construction method of Bayesian network, and further introduces the method tree augmented network (TAN) driven by Bayesian data. Section 3 builds the TAN model based on the data of 5660 maritime accidents, and carries out sensitivity analysis and simulation verification on the built model. Section 4 uses the two-way reasoning ability of TAN model to predict the accident chain and analyze the accident causes. Finally, the fifth part summarizes the full text.

2. BN Structure Learning—TAN

BN is a Directed Acyclic Graph (DAG) composed of nodes and directed edges, which is widely employed to illustrate the interdependence and strength of associations between variables. The network represents the interrelationships between variables through vectorial arcs, with the intensity of each association specified by a table of conditional probabilities.

There are two primary approaches to generating BN structures: 1) the expert knowledge method, and 2) the data-driven method. In the expert knowledge method, the BN structure is built by subjectively evaluating the causal relationships between variables. Conversely, the data-driven method is employed to uncover the interdependence between variables, based on the learning algorithm of the BN model and data correlations. In this study, since sufficient sample data were collected, the data-driven method was used to construct the BN structure.

Data-driven Bayesian approaches can be classified into three main categories: 1) the Naive Bayesian Network (NBN); 2) the Augmented Naive Bayesian Network (ABN); and 3) the Tree Augmented Network (TAN). Among these, TAN learning effectively combines the simplicity and robustness of NBN computation with the ability to characterize interaction dependencies among variables, thus providing insights into the key factors leading to the outcomes of specific accidents [48]. Therefore, this paper employs the data-driven TAN approach to construct the BN structure.

BN encodes the joint probability distribution over a set of random variables U . Let $U = \{A_1, \dots, A_n, C\}$, where n denotes the number of influencing factors, A_1, \dots, A_n represent the influencing factors, and C is a class variable (accident type). It is established that the set of parent nodes of C in U is empty, meaning $\Pi C = \emptyset$. Moreover, ΠA_i has at most one other node besides C that can have an associated edge pointing to it. The joint probability density distribution adheres to the following equation:

$$P(A_1, \dots, A_n, C) = P(C) \cdot \prod_{i=1}^n P(A_i|C), \quad (1)$$

In the process of learning the TAN structure, Chow and Liu [49] proposed an approach to optimize and construct the BN structure using the conditional mutual information of each attribute pair. The function is defined as:

$$I_p(A_i, A_j | C) = \sum_{a_{ii}, a_{ji}, c_i} P(a_{ii}, a_{ji}, c_i) \log \frac{P(a_{ii}, a_{ji} | c_i)}{P(a_{ii} | c_i)P(a_{ji} | c_i)}, \quad (2)$$

Where I_p denotes the conditional mutual information; and a_{ii} is the i -th state of the influencing factor A_i ; and a_{ji} is the i -th state of the influencing factor A_j .

3. Global Maritime Accident TAN Model

This section can be divided using subheadings to provide clear organization. It should offer a concise and accurate description of the experimental results, their interpretation, and the conclusions that can be drawn from the experiments.

3.1. Data Collection

This paper utilizes the Marine Casualties and Incidents (MCI) database in GISIS, which is managed by IMO [50]. GISIS is a comprehensive, global maritime information system. In accordance with IMO regulations, every country with sovereignty over its territorial sea is required to report maritime accidents occurring within its waters to the IMO. The MCI database contains two types of information related to global maritime accidents: first, factual data gathered from various sources; and second, detailed data obtained from casualty investigation reports submitted to the IMO.

The MCI database houses global maritime accident data dating back to 1973. Between 1973 and 2000, the annual number of recorded maritime accidents was quite limited. From 2001 onwards, the number of accidents documented in the MCI database has been more consistent. However, the accident timestamps during 2001-2004 are only accurate to the day, which is not sufficient for studying the specific time periods in which the accidents occurred. Consequently, we excluded the low-quality data from the early years and utilized a total of 5,660 maritime accident records from 2005 to 2020 to construct the BN model.

3.2. Node Variable Definitions

Based on the literature's studies on maritime accident factor analysis [1,43,51,52], there are 16 primary factors contributing to maritime accidents, including ship type, hull type, ship age, ship length, ship gross tonnage, ship operation, voyage segment, ship speed, ship condition, ship equipment or device condition, ship design, interaction information, weather condition, ocean condition, time period, and channel traffic condition. Combining these factors with the information available in the MCI database, we selected seven node variables for the BN model: accident quarter, accident period, accident type, ship type involved, total tonnage of the ship involved, life loss contingency, and accident severity.

Since BN nodes require discrete variables, it is necessary to discretize continuous variables in the accident statistics. We divided accident occurrence quarters into first quarter (January, February, March), second quarter (April, May, June), third quarter (July, August, September), and fourth quarter (October, November, December). Accident periods were categorized as dawn (0:00-5:59), early morning (6:00-8:59), morning (9:00-11:59), noon (12:00-13:59), afternoon (14:00-16:59), early evening (17:00-19:59), and evening (20:00-23:59). To discretize the gross tonnage of the ships involved, we used the collected data and the Centroid Clustering (CC) algorithm for classification. The CC algorithm employs the minimization error sum of squares as the objective function and terminates when the number of iterations reaches a preset maximum of 5,000 iterations. The optimal classification results yielded four groups based on the gross tonnage of the ships involved: (1-18,500 t), (18,501-57,500 t), (57,501-120,000 t), and (120,001-403,342 t). Among them, 403,342 t represents the maximum total tonnage of the ships involved in the collected data.

Furthermore, this paper classifies non-routine accidents, such as missing ships, life-saving equipment accidents, and numerous accident types with irregular or rare records, accounting for no more than 5%, as "other" [53]. Multipurpose ships, tugboats, supply and offshore vessels, unspecified ship types, and other ship types, representing no more than 10%, are categorized as "Other" [53]. Table 1 presents the names, classifications, frequency of occurrence, and percentages of each discrete variable category.

Table 1. Variables for building BN.

Variable Name	Classification	Frequency	Percentage/%	Variable Name	Classification	Frequency	Percentage/%
Accident quarter	a (the first quarter)	1539	27.19	Ship type	a (general cargo ship)	989	17.47
	b (the second quarter)	1353	23.90		b (bulk carrier)	255	4.50
	c (the third quarter)	1406	24.84		c (container ship)	370	6.54
	d (the fourth quarter)	1362	24.06		d (chemical tanker/oil tanker)	537	9.49
Accident time	a (dawn 0-5 a.m.)	1954	34.52	Vessel gross tonnage	e (passenger ship)	453	8.00
	b (early morning 5-8 a.m.)	562	9.93		f (fishing vessel)	634	11.20
	c (morning 8-11p.m.)	693	12.24		g (other)	2422	42.79
	d (noon 11-13 p.m.)	427	7.54		a (gross tonnage [1,18500])	4011	70.87
	e (afternoon 13-16 p.m.)	647	11.43		b (gross tonnage [18501,57500])	1219	21.54

Accident type	f (early evening 16-19 p.m.)	540	9.01	Life loss contingency	c (gross tonnage [57501,120000])	340	6.00
	g (evening 19-24 p.m.)	837	14.79		d (gross tonnage [120001,403342])	90	1.59
	a (collision)	1016	17.95		a (life loss)	1651	29.17
	b (reefed and stranded)	823	14.54		b (no life loss)	4009	70.83
	c (fire and explosion)	754	13.32		a (particularly serious accidents)	2837	50.12
	d (capsize)	365	6.45		b (serious accidents)	2034	35.94
	e (ship's machinery damage)	287	5.07		c (general accident)	622	10.99
	f (poor communication)	281	4.96		d (unspecified accident)	167	2.95
	g (other)	2134	37.70				

3.3. TAN Modeling

Based on the data processing results, we examined the relationship between six influencing factors and accident consequences. We used Netica software with a "learning network" function to develop a TAN model based on Equation (2), ensuring that all connections between nodes are meaningful. The initial structure of the TAN, depicted in Figure 1, is grounded in data-driven TAN training results that demonstrate realistic correlations between variables.

Utilizing the TAN model, Netica software employs basis functions to create a structure learning module and a parameter learning module that automatically learn the Conditional Probability Table (CPT) parameters from the sample dataset. By constructing the TAN and obtaining the CPT, we can calculate the posterior probability of each variable. The statistical results of the probabilistic variables facilitate the analysis of maritime safety considerations and aid in accident prevention. Figure 2 presents the TAN results for the random variables of interest.

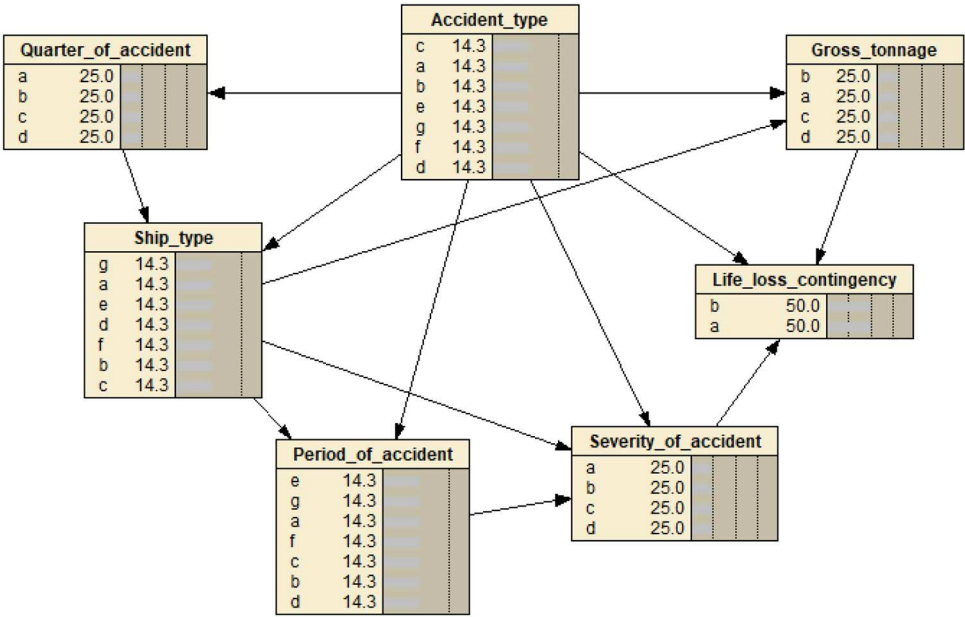


Figure 1. Initial structure of TAN model.

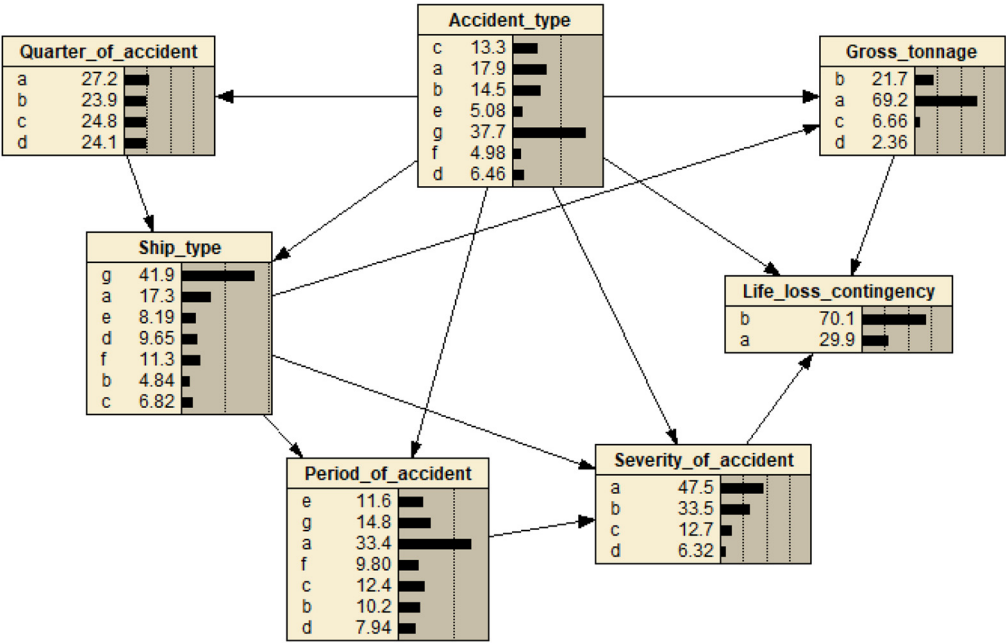


Figure 2. TAN of vessel navigation risk.

3.4. Sensitivity Analysis and Model Validation

3.4.1. Sensitivity Analysis

In the Netica software, we selected the accident type as the target node and conducted a sensitivity analysis on this node to identify the factors with the greatest influence on the target node within the TAN model.

The mutual information value represents the sensitivity level between two random variables; a higher value indicates greater sensitivity of the influencing factor to the target node and, conversely, lower sensitivity. We used the sensitivity analysis function in Netica software to calculate the mutual information value, percentage, and variance for each influencing factor and accident type, as displayed in Table 2. According to Table 2, the accident consequence and accident severity are the

factors most sensitive to accident type performance, with mutual information values of 0.14246 and 0.14033, respectively, which are notably higher than those of the other four factors. The results reveal that accident consequence and accident severity are the two most intuitive factors for determining the type of accident, followed by ship type, gross tonnage of the ship, time period, and quarter.

Table 2. Mutual information shared with "accident type".

Nodes	Mutual Information Value	Percentage /%	Variance
Life loss contingency	0.14246	5.800	0.0176774
Accident severity	0.14033	5.710	0.0088289
Ship type	0.04235	1.720	0.0013155
Vessel gross tonnage	0.02096	0.853	0.0004918
Time period	0.02006	0.817	0.0012170
Quarter	0.00421	0.171	0.0000869

3.4.2. Model Validation

To validate the TAN model's effectiveness, we randomly selected three offshore accident cases from 2021, each with varying accident consequences and severities, labeled as events 1, 2, and 3. We input the case data into the model for scenario analysis, and Table 3 presents the relevant data information for these accident cases.

Table 3. State values of real event factor variables.

Variables	Event Number		
	1	2	3
Quarterly	c	a	b
Time period	e	b	g
Ship type	g	a	g
Life loss contingency	a	b	b
Accident severity	a	b	c
Vessel gross tonnage	b	a	b
Accident type	g	b	a
Accident probability	75.1%	38.0%	44.4%

Based on the data from three randomly selected events, we set the probability of known nodes such as quarter, time period, vessel type, accident consequence, accident severity, and gross tonnage of the vessel to 100%. We then observed the types and probabilities of the predicted accidents. As illustrated in Table 3, the probability of other accident types occurring in event 1 was 75.1%; the probability of reefing and grounding in event 2 was 38.0%; and the probability of collision in event 3 was 44.4%. When compared to the original data's accident types, the predicted accident types for the three events matched, indicating that the model's predictions are accurate to some extent. Since the occurrence probability of other accident types in the original data is significantly higher than that of collision, reefing, and grounding, the data-driven TAN model's simulation results demonstrate better performance in predicting the occurrence probability of other accident types (e.g., Figure 3a) and average results in predicting collision and reefing and grounding accidents (e.g., Figure 3b and Figure 3c).

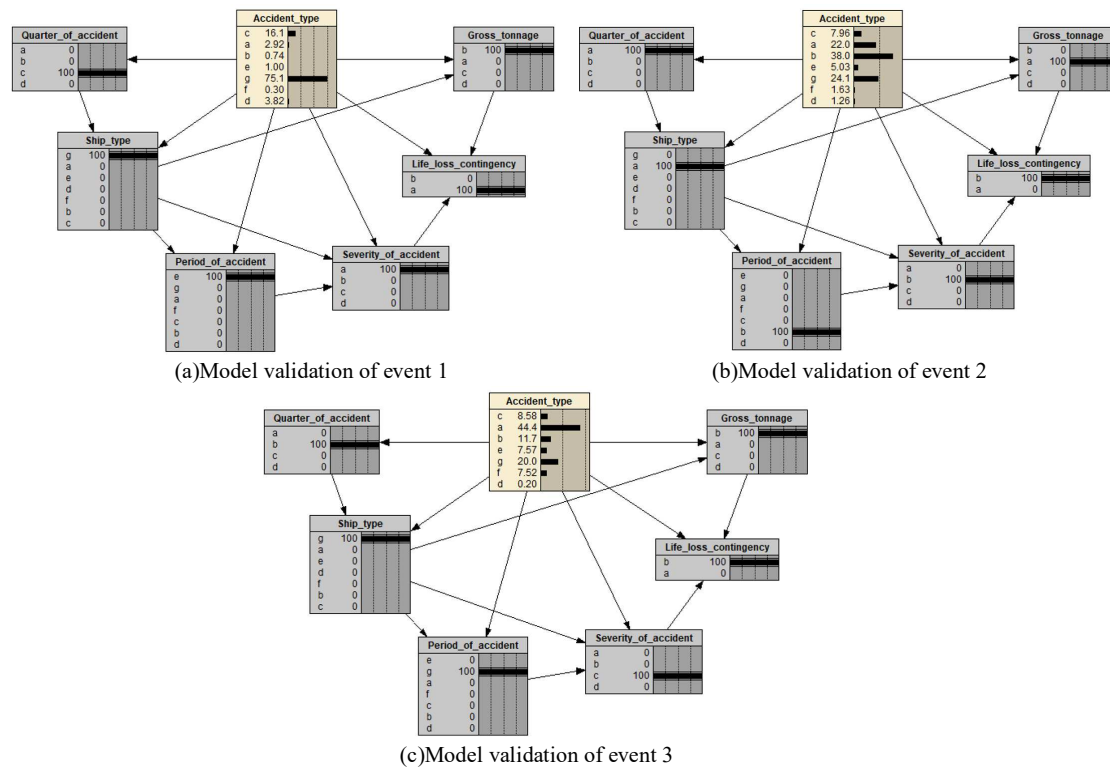


Figure 3. Event model validation.

4. Model Reasoning

The TAN model has the ability to reason bidirectionally and helps explain the most likely scenarios associated with a specific accident type. The data-driven TAN-based model examines the correlations between various influencing factors of maritime accidents and accident types, as well as accident consequences. This analysis enables the prediction of the likelihood of various accident scenarios and the extrapolation of accident consequences under specific conditions.

4.1. Accident Chain Forecast

After using the Netica software to manipulate the TAN model, we obtained the relationships between the influencing factors and the accident type, life loss contingency, and probability of each node. By adjusting the placement bar of a single node or multiple nodes, we observed the target node's probability trend, forming a judgment of the potential trend and consequences of the accident.

We first simulated the parameters of the conditions for maritime accidents by changing a single node and observing the target node's changes. When changing the ship type, more significant changes occur in the probability of each accident type. For example, when the ship type is set to a chemical ship, the probability of fire and explosion accident type increases significantly. When the ship type is a bulk carrier, the probability of collision accident type increases notably. The study shows that different ship types can lead to significant differences in the occurrence of accident types. Additionally, the ship's gross tonnage and the accident's quarter and time also impact the accident type.

Since the accident type is influenced by the joint decision of several nodes, the influence of a single node on the accident type is more one-sided. Therefore, we set the accident quarter to the first quarter (with the variable node's confidence bar set to 100%), the ship type to general cargo ship, and the gross tonnage to 1-18,500 t, as shown in Figure 4. At this time, the change in accident type and accident severity node probability from early morning to evening is shown in Figure 5. As seen in Figure 5, among the types of maritime accidents throughout the day, the probability of fire and explosion on ships is low, except for the afternoon time period, which is 18.2%; the probabilities of ship capsizing, machinery damage, and poor communication are also low, at below 10%. Among

other accidents, the probabilities of ship collision and reefing and grounding accidents are significantly higher, around 20%. Additionally, it is observed that the occurrence probability of reefing and grounding accidents is significantly higher during dawn and evening than in other periods.

In summary, the highest probabilities of collision and grounding occur at dawn, with the accident severity being particularly serious. Collision is most likely to occur at noon, with high accident severity. Particularly severe collision and grounding accidents are more likely to occur in the evening. It is worth mentioning that although the probabilities of collision and grounding and stranding of ships are higher in this scenario, the probability of life loss is relatively low, and the accident consequences are less affected by the time of the accidents.

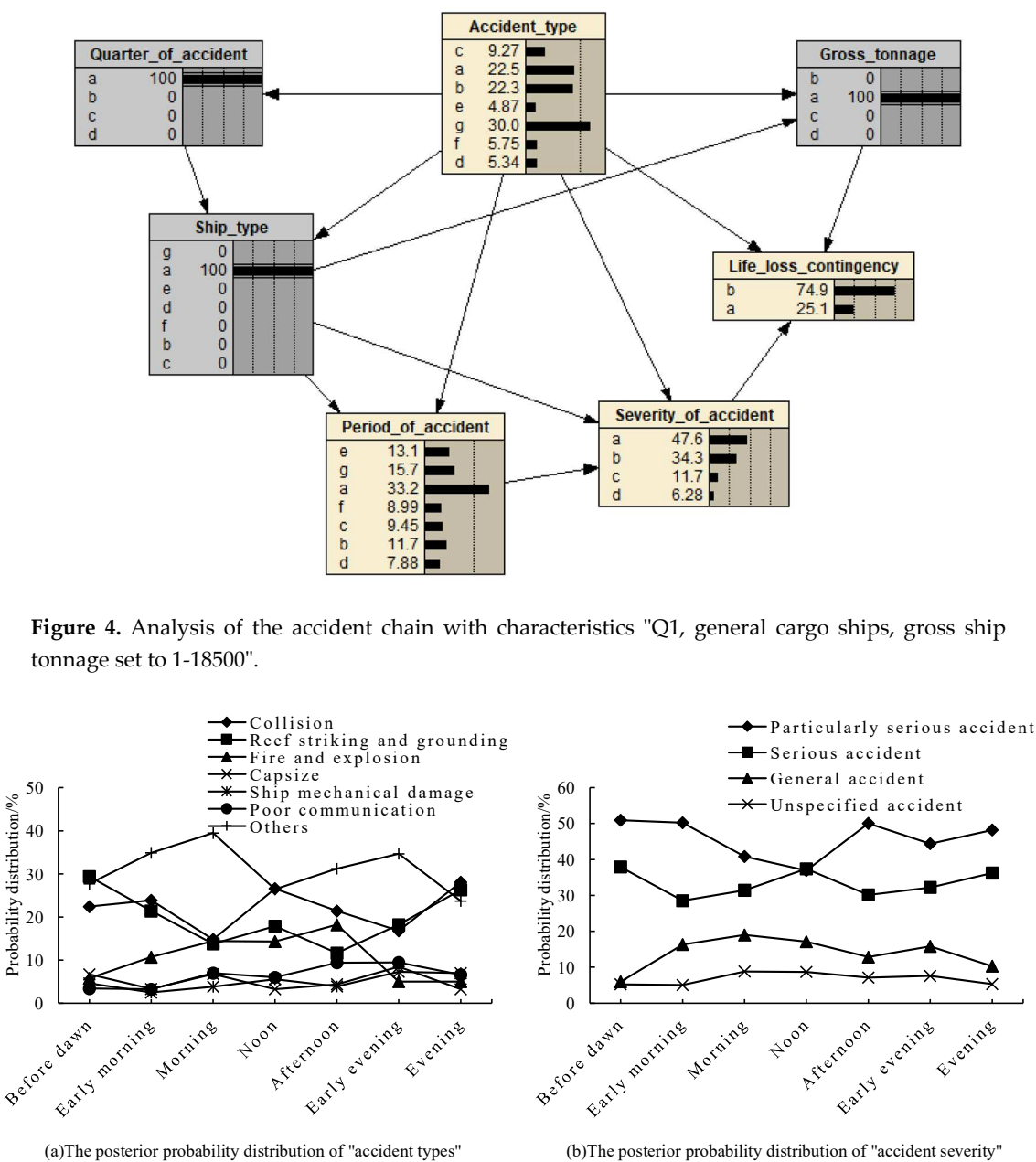


Figure 5. The posterior probability of "accident types and accident severity" in specific accident scenarios.

4.2. Accident Cause Analysis

The bidirectional reasoning of the TAN model refers to causal reasoning and diagnostic reasoning. Causal reasoning can be applied to accident chain prediction, while diagnostic reasoning can be applied to accident cause analysis. Diagnostic reasoning aims to determine the type, consequence, and severity of an accident to more intuitively understand the causes and mechanisms of offshore accident formation.

As shown in Figure 6, when the accident type is determined to be a collision accident, the accident severity is a general accident, and it does not involve life loss. The study shows that the probability of nodes for container ships, general cargo ships, and chemical ships among the ship types is significantly higher than for other ship types, and the probability of collision accidents for these three ship types is higher than for other ship types. Additionally, observation of accident time and ship tonnage shows that ships with a tonnage of 1-18,500 t are more prone to collisions during the dawn hours. Furthermore, the consequences of accidents involving container ships, general cargo ships, and chemical ships of such tonnage indicate a lower probability of causing life loss. This could be because it is easier for personnel to escape from small ships in distress, or because the rescue of small ships in distress is relatively easy and more successful.

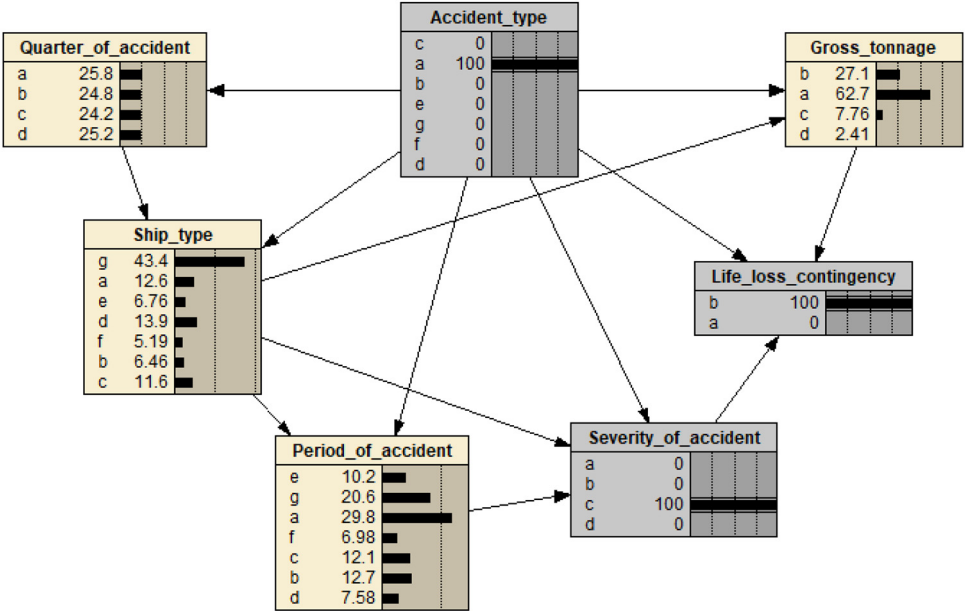


Figure 6. Analysis of accident scenarios with characteristics "collision, general accident, no loss of life involved".

5. Conclusions

Ship safety has always been a major concern in the maritime transportation industry. In this paper, a TAN model for maritime traffic accident risk assessment is constructed to analyze the relationship between the consequences of maritime accidents and various influencing factors, and to use model simulation to analyze how different risk factors affect different types of maritime accidents.

The TAN model was constructed based on data from a total of 5,660 maritime accidents from 2005 to 2020. In addition to other accident types, the accident type with the highest probability of occurrence among maritime traffic accidents is collision, followed by grounding and stranding, and then fire and explosion.

The sensitivity analysis and simulation validation of the constructed model showed that accident consequences and accident severity are the two most intuitive factors in determining the type of accident occurrence, followed by ship type, gross tonnage of the ship, time period, and season.

The constructed model can effectively predict the likelihood of various accident scenarios and accident consequence projections under specific conditions.

According to the causal reasoning analysis of the TAN model, under the condition of "first quarter, general cargo ship, and ship's gross tonnage of 1-18,500 t," the probability of ship collision and grounding and stranding accidents is higher, while the probability of life loss is relatively low, and the consequences of the accident are less affected by the time of the accident. According to the analysis of model diagnostic reasoning, in the general collision accident chain without loss of life, container ships, general cargo ships, and chemical ships are the main types of ships involved in such accidents. Ships with a tonnage of 1-18,500 t are more likely to have such accidents in the dawn hours, but the probability of causing loss of life is lower.

Author Contributions: Conceptualization, methodology, software, writing—original draft, S.L. (Shiguan Liao); supervision, J.W. (Jinxian Weng); writing—review and editing, Z.Z. (Zhaomin Zhang); data curation, F.L. (Fang Li) and Z.L. (Zhuang Li); visualization, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China under Grants 52072237, the Guangdong Basic and Applied Basic Research Foundation of China under Grants 2022A1515110776 and the Project of Educational Commission of Guangdong Province of China under Grants 2021WTSCX222.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fan, S., Blanco-Davis, E., Yang, Z., Zhang, J., & Yan, X. (2020). Incorporation of human factors into maritime accident analysis using a data-driven Bayesian network. *Reliability Engineering & System Safety*, 203, 107070.
2. Zhang, G., & Thai, V. V. (2016). Expert elicitation and Bayesian Network modeling for shipping accidents: A literature review. *Safety science*, 87, 53-62.
3. Hu, S., Li, Z., Xi, Y., Gu, X., & Zhang, X. (2019). Path analysis of causal factors influencing marine traffic accident via structural equation numerical modeling. *Journal of marine science and engineering*, 7(4), 96.
4. Chou, C. C., Wang, C. N., Hsu, H. P., Ding, J. F., Tseng, W. J., & Yeh, C. Y. (2022). Integrating AIS, GIS and E-chart to analyze the shipping traffic and marine accidents at the Kaohsiung Port. *Journal of marine science and engineering*, 10(10), 1543.
5. Xue, J., Papadimitriou, E., Reniers, G., Wu, C., Jiang, D., & van Gelder, P. H. A. J. M. (2021). A comprehensive statistical investigation framework for characteristics and causes analysis of ship accidents: A case study in the fluctuating backwater area of Three Gorges Reservoir region. *Ocean Engineering*, 229, 108981.
6. Göksu, S., & Arslan, Ö. (2020). Quantitative analysis of dynamic risk factors for shipping operations. *Journal of ETA Maritime Science*, 8(2).
7. Li, Z., Yao, C., Zhu, X., Gao, G., & Hu, S. (2022). A decision support model for ship navigation in Arctic waters based on dynamic risk assessment. *Ocean Engineering*, 244, 110427.
8. Guo, Y., Jin, Y., Hu, S., Yang, Z., Xi, Y., & Han, B. (2023). Risk evolution analysis of ship pilotage operation by an integrated model of FRAM and DBN. *Reliability Engineering & System Safety*, 229, 108850.
9. Zhang, L., Wang, H., Meng, Q., & Xie, H. (2019). Ship accident consequences and contributing factors analyses using ship accident investigation reports. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of risk and reliability*, 233(1), 35-47.
10. Otay, E. N., & Özkan, S. (2003, September). Stochastic Prediction of Maritime Accidents in the strait of Istanbul. In *Proceedings of the 3rd International Conference on Oil Spills in the Mediterranean and Black Sea regions* (pp. 92-104).

11. Koromila, I., Nivolianitou, Z., & Giannakopoulos, T. (2014, May). Bayesian network to predict environmental risk of a possible ship accident. In Proceedings of the 7th International Conference on Pervasive Technologies Related to Assistive Environments (pp. 1-5).
12. Zhang, C., Zou, X., & Lin, C. (2022). Fusing XGBoost and SHAP models for maritime accident prediction and causality interpretability analysis. *Journal of Marine Science and Engineering*, 10(8), 1154.
13. Fu, S., Liu, Y., Xi, Y., Zhang, M., & Hu, S. (2019). Analysis of influencing factors on severity of ship accidents in ice-covered Arctic waters. *China Safety Science Journal*, 29(10), 84. (in Chinese)
14. Baksh, A. A., Abbassi, R., Garaniya, V., & Khan, F. (2018). Marine transportation risk assessment using Bayesian Network: Application to Arctic waters. *Ocean Engineering*, 159, 422-436.
15. Chen, P., Zhang, Z., Huang, Y., Dai, L., & Hu, H. (2022). Risk assessment of marine accidents with Fuzzy Bayesian Networks and causal analysis. *Ocean & Coastal Management*, 228, 106323.
16. Ventikos, N. P., & Giannopoulos, I. F. (2013). Assessing the consequences from marine accidents: Introduction to a risk acceptance criterion for Greece. *Human and Ecological Risk Assessment: An International Journal*, 19(3), 699-722.
17. Chen, J., Bian, W., Wan, Z., Yang, Z., Zheng, H., & Wang, P. (2019). Identifying factors influencing total-loss marine accidents in the world: Analysis and evaluation based on ship types and sea regions. *Ocean Engineering*, 191, 106495.
18. Chen, J., Zhang, F., Yang, C., Zhang, C., & Luo, L. (2017). Factor and trend analysis of total-loss marine casualty using a fuzzy matter element method. *International journal of disaster risk reduction*, 24, 383-390.
19. Nivolianitou, Z. S., Leopoulos, V. N., & Konstantinidou, M. (2004). Comparison of techniques for accident scenario analysis in hazardous systems. *Journal of Loss Prevention in the Process Industries*, 17(6), 467-475.
20. Wang, W., Jiang, X., Xia, S., & Cao, Q. (2010). Incident tree model and incident tree analysis method for quantified risk assessment: an in-depth accident study in traffic operation. *Safety Science*, 48(10), 1248-1262.
21. Liu, P., Yang, L., Gao, Z., Li, S., & Gao, Y. (2015). Fault tree analysis combined with quantitative analysis for high-speed railway accidents. *Safety science*, 79, 344-357.
22. Ung, S. T. (2019). Evaluation of human error contribution to oil tanker collision using fault tree analysis and modified fuzzy Bayesian Network based CREAM. *Ocean Engineering*, 179, 159-172.
23. Ahn, Y. J., Yu, Y. U., & Kim, J. K. (2021). Accident cause factor of fires and explosions in tankers using fault tree analysis. *Journal of Marine Science and Engineering*, 9(8), 844.
24. Kim, K. I., Jeong, J. S., & Lee, B. G. (2017). Study on the analysis of near-miss ship collisions using logistic regression. *Journal of advanced computational intelligence and intelligent informatics*, 21(3), 467-473.
25. Yu, H., Khan, F., & Veitch, B. (2017). A flexible hierarchical Bayesian modeling technique for risk analysis of major accidents. *Risk analysis*, 37(9), 1668-1682.
26. Hänninen, M. (2014). Bayesian networks for maritime traffic accident prevention: Benefits and challenges. *Accident Analysis & Prevention*, 73, 305-312.
27. Abistado, K. G., Arellano, C. N., & Maravillas, E. A. (2014). Weather forecasting using artificial neural network and Bayesian network. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 18(5), 812-817.
28. Wang, J.; Zhang, M.; Huang, X.; Chen, J. (2019). Scenario analysis of road transportation accidents of inflammable and explosive hazardous chemicals. *China Safety Science Journal*, 29(5), 171. (in Chinese)
29. Qiao, W., Liu, Y., Ma, X., & Liu, Y. (2020). Human factors analysis for maritime accidents based on a dynamic fuzzy Bayesian network. *Risk analysis*, 40(5), 957-980.
30. Zhang, G., Thai, V. V., Yuen, K. F., Loh, H. S., & Zhou, Q. (2018). Addressing the epistemic uncertainty in maritime accidents modelling using Bayesian network with interval probabilities. *Safety science*, 102, 211-225.
31. Li, K. X., Yin, J., Bang, H. S., Yang, Z., & Wang, J. (2014). Bayesian network with quantitative input for maritime risk analysis. *Transportmetrica A: Transport Science*, 10(2), 89-118.
32. Wang, L., & Yang, Z. (2018). Bayesian network modelling and analysis of accident severity in waterborne transportation: A case study in China. *Reliability Engineering & System Safety*, 180, 277-289.
33. Zou, X., & Yue, W. L. (2017). A bayesian network approach to causation analysis of road accidents using netica. *Journal of advanced transportation*, 2017.
34. Yuan, C., Hu, Y., Zhang, Y., Zuo, T., Wang, J., & Fan, S. (2021). Evaluation on consequences prediction of fire accident in emergency processes for oil-gas storage and transportation by scenario deduction. *Journal of Loss Prevention in the Process Industries*, 72, 104570.

35. Zhao, J.; Xie, L.; Yang, Y.; Hu, X.; Ou, C.; Zeng, R. (2022). An ISM-BN based research on navigation risk factors of inland waterway vessels. *China Safety Science Journal*, 32(8), 37. (in Chinese)
36. Afenyo, M., Khan, F., Veitch, B., & Yang, M. (2017). Arctic shipping accident scenario analysis using Bayesian Network approach. *Ocean Engineering*, 133, 224-230.
37. Jiang, M., Lu, J., Yang, Z., & Li, J. (2020). Risk analysis of maritime accidents along the main route of the Maritime Silk Road: a Bayesian network approach. *Maritime Policy & Management*, 47(6), 815-832.
38. Si, D.; Zhang, J.; Lang, K. (2019). Causation analysis of container ship collision accidents based on improved BN. *China Safety Science Journal*, 29(10), 31. (in Chinese)
39. Fan, S., Yang, Z., Blanco-Davis, E., Zhang, J., & Yan, X. (2020). Analysis of maritime transport accidents using Bayesian networks. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 234(3), 439-454.
40. Hänninen, M., Banda, O. A. V., & Kujala, P. (2014). Bayesian network model of maritime safety management. *Expert Systems with Applications*, 41(17), 7837-7846.
41. Zhao, X., Yuan, H., & Yu, Q. (2021). Autonomous vessels in the Yangtze river: A study on the maritime accidents using data-driven bayesian networks. *Sustainability*, 13(17), 9985.
42. Acharya, T. D., Yoo, K. W., & Lee, D. H. (2017). GIS-based spatio-temporal analysis of marine accidents database in the coastal zone of Korea. *Journal of Coastal Research*, (79 (10079)), 114-118.
43. Uğurlu, Ö., Köse, E., Yıldırım, U., & Yükseskyıldız, E. (2015). Marine accident analysis for collision and grounding in oil tanker using FTA method. *Maritime Policy & Management*, 42(2), 163-185.
44. Antão, P., Teixeira, A. P., & Soares, C. G. (2021). Statistical characterization of risk influencing factors in ship collision accidents. In *Developments in Maritime Technology and Engineering* (pp. 221-229). CRC Press.
45. Magda, B. (2015). Fires as a cause of ship accidents-a statistical approach. *Safety & Fire Technology*, 37, 171-180.
46. Li, H., Ren, X., & Yang, Z. (2023). Data-driven Bayesian network for risk analysis of global maritime accidents. *Reliability Engineering & System Safety*, 230, 108938.
47. Huang, D. Z., Hu, H., & Li, Y. Z. (2013). Spatial analysis of maritime accidents using the geographic information system. *Transportation research record*, 2326(1), 39-44.
48. Friedman, N., Geiger, D., & Goldszmidt, M. (1997). Bayesian network classifiers. *Machine learning*, 29, 131-163.
49. Chow, C. K. C. N., & Liu, C. (1968). Approximating discrete probability distributions with dependence trees. *IEEE transactions on Information Theory*, 14(3), 462-467.
50. International Maritime Organization-Global Integrated Shipping Information System. Marine Casualties and Incidents, <https://gis.imo.org/Public/MCI/Default.aspx>. [accessed 21 June 2022].
51. Zhu, L., Lu, L., Zhang, W., Zhao, Y., & Song, M. (2019). Analysis of accident severity for curved roadways based on bayesian networks. *Sustainability*, 11(8), 2223.
52. Liu, L., Ye, X., Wang, T., Yan, X., Chen, J., & Ran, B. (2022). Key factors analysis of severity of automobile to two-wheeler traffic accidents based on Bayesian network. *International journal of environmental research and public health*, 19(10), 6013.
53. Zhang, Y., Sun, X., Chen, J., & Cheng, C. (2021). Spatial patterns and characteristics of global maritime accidents. *Reliability Engineering & System Safety*, 206, 107310.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.