Reliability Engineering and System Safety 277 (2027) 113180

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Integrating Naive Bayesian network and evidential reasoning for risk assessment in general cargo ship pre–stowage planning

Daihui Zhang [a], Zheng Qiao [a], Wenxin Wang [a], Ge Zhang [a], Zhuohua Qu [b], Zaili Yang [c,*]

[a] College of Navigation, Dalian Maritime University, Dalian 116026, China
[b] Liverpool Business School, Liverpool John Moorse University, L3 5UG, UK
[c] Liverpool Logistics, Offshore and Marine Reserach Institute, Liverpool John Moores University, L3 3AF, UK

ARTICLE INFO

Keywords:
General cargo ship pre–stowage plan
Naïve bayesian network
Evidential reasoning
Maritime safety
Navigation environment

ABSTRACT

In the current General Cargo Ship Pre–Stowage Planning (GCSPP), efficiency and economy are the two main dominant influential factors. Given the increased operational uncertainty and safety priority, a new solution that can incorporate the safety factor into the planning becomes highly demanded with urgency. Such a new solution is critical for reinforcing compliance with safety standards and regulations, ultimately enhancing the competitiveness of the general cargo shipping industry. To bridge this gap, we have developed an innovative data–driven quantitative GCSPP assessment framework that combines Naïve Bayesian network (NBN) modeling with evidence reasoning (ER) techniques. This framework focuses on quantitative analysis, supplemented by standardized qualitative input conversion, to incorporate safety, economy, and efficiency into the evaluation and selection of GCSPPs. The framework incorporates dynamic risks associated with ships navigating diverse environments into the planning process. Fourteen Risk influencing Factors (RIFs) related to GCSPPs were identified from general cargo ship accident reports from 2007 to 2021 and incorporated into a NBN model. Using forthcoming voyage information, navigation scenarios are constructed to determine the safety requirements for pre–stowage planning. These requirements are translated into weighting factors and, together with efficiency and economy, integrated into the ER method to support multi–criteria decision–making for GCSPPs. This enables the identification of the optimal pre–stowage plan for a specific voyage by considering safety, efficiency, and economy simultaneously. The findings show that, for the same ship carrying the same cargo, the most suitable pre–stowage plan varies with voyage–specific safety requirements. The study further contributes a robust solution to complex decision–making problems involving multiple dynamic attributes through the integrated application of NBN and ER techniques.

1. Introduction

Maritime shipping handles approximately 90% of global trade, making it a key force in world economic development due to its high transport capacity and cost–effectiveness [1,2]. Over the past decade, with the process of globalization, the shipping industry has been expanding, leading to increased maritime traffic density and heightened tensions. Various influential factors, such as complex maritime environments and weather conditions, often exacerbate the severity and frequency of maritime accidents that can cause significant economic losses and casualties, from which recovery is challenging [3,4]. Therefore, ensuring safety remains a fundamental priority in the shipping industry. Within this context, the movement and management of cargo onboard require further investigation and enhancement, particularly through the development of scientific decision–making methods that surpass traditional experience–based practices. Such methods should be capable of incorporating multiple attributes, e.g., cost, safety and efficiency, accounting for their dynamic features and associated uncertainties. For instance, the safety of cargo stowage shall be evaluated and systematically incorporated into the final choice of an appropriate pre–stowage plan. However, the task remains highly challenging, as the safety of cargo stowage is inherently complex and dynamic, influenced by the specific navigation environment of the ship's forthcoming voyage.

Typically, a proper and efficient pre–stowage plan is essential for managing cargo on board and is represented in the form of a cargo

* Corresponding author.
E-mail address: z.yang@ljmu.ac.uk (Z. Yang).

https://doi.org/10.1016/j.ress.2026.113180
Received 3 January 2026; Received in revised form 15 July 2026; Accepted 16 July 2026
Available online 18 July 2026

stowage chart. In practice, before loading begins, the first mate drafts the stowage chart on the basis of: 1) the nature of the cargo (e.g., the type, quantity, volume, weight), 2) the sequence of port calls, and 3) the ship's own conditions (e.g., the ship's deadweight, draught restrictions, stability requirements and capacity). However, there are significant uncertainties in the GCSPP as a result of the unpredictable and dynamic navigational risk associated with the voyage, which is a major challenge for scientifically and reasonably evaluating GCSPP. The majority of navigation–related uncertainty is due to the extremely complex and dynamic nature of the marine navigation environment, which consists of many interconnected, interdependent factors. First of all, marine meteorological states show great randomness such as typhoons, gales, thick fogs, and storm surges, which could take place during the voyage and real – time alterations in wind direction, wind strength, and wave height directly affect the vessel's stability and navigational performance thus bringing about unforeseen safety hazards to the GCSPP. Second, the maritime traffic settings are by nature complex and changeable with substantial differences in traffic density among various waterways (e.g., port channels, coastal waters, and the open seas). The nautical actions of other ships, the arrangement of nautical aids, and the appearance of uncharted obstacles add to operational uncertainties requiring more flexibility from cargo stowage regarding dynamic ship maneuvers. These dynamic environmental elements make traditional static GCSPP assessment insufficient for precisely forecasting real–time safety needs restricting their adaptability to the dynamic risks specific to each voyage. This shortcoming leads to the combination of NBN and ER techniques in this research. The developed hybrid assessment framework enables the dynamic quantification of navigational risks and incorporates them into the Multiple Criteria Decision Making (MCDM) process for GCSPP, thus improving the scientific precision and adaptability of stowage plan evaluations in response to nautical uncertainties.

The current assessment of GCSPPs remains overly rigid and compliance–oriented: it only verifies whether key safety indicators meet the minimum thresholds specified in conventions or rules (e.g., requiring the initial metacentric height GM $\geq$ 0.15 m, as mandated by the International Code on Intact stability,2008(2008 IS Code)). Once these basic static compliance requirements are satisfied, safety considerations are completely excluded from subsequent GCSPP decision–making. This static evaluation method cannot fully consider the dynamic security risks brought by the real navigation environment. For example, a loading plan that meets the minimum GM requirement (e.g., GM=0.15 m) may still face significant overturning risks when passing through typhoon prone waters (with strong wind–induced roll moments) or congested port waterways (sudden maneuvers amplify stability requirements). Therefore, there is a pressing need to explore a new dynamic safety assessment approach that reflects the specific risks of the ship's forthcoming voyage. The core research problem addressed in this study is that the current GCSPP fails to dynamically integrate specific navigation risks (e.g., typhoon prone waters, congested waterways) into GCSPP an MCDM problem involving (safety, economy, efficiency). It can easily lead to compromised navigation safety in high–risk environments or sacrifice efficiency/cost–effectiveness in low–risk environments.

Specifically, one significant research gap in cargo management is the limited integration of ship risk analysis into cargo stowage practices. Few studies incorporate dynamic risk analysis results into the evaluation system for GCSPP due to the high uncertainty in the relevant decision–making process. Meanwhile, various modeling approaches for uncertainty have been widely applied in risk assessment of other shipping activities (e.g., ship navigation). These approaches help determine and assess the RIFs like ship's own conditions, navigation environment, crew operational errors and management level [5]. For instance, using 161 accident reports from 2012–2017, Fan et al. [6] constructed a NBN structure for maritime accident analysis. Li et al. [7] established a Tree Augmented Naive Bayesian Network (TAN) based on accident records from 2017–2021 and their findings allowed for the development of appropriate preventive measures. Wang and Yang [8] constructed an Augmented Naive Bayesian Network (ANB) structure to analyze marine accident severity, drawing on 229 accident investigation reports collected from China's coastal and inland maritime routes via the Maritime Safety Administration (MSA) These studies clearly demonstrate that NBN is a robust probabilistic analysis tool that has been extensively developed in maritime accident analysis and risk modelling. However, it remains underutilized in more specific operational contexts such as the assessment of GCSPPs.

As an important segment of shipping practice, the assessment of cargo pre–stowage plans still subjectively relies on crew experience, lacking a systematic and rational framework which incorporates practical operational demands with risk awarenesses. To bridge the aforementioned gaps, this research proposes a novel GCSPP assessment framework that combines ER with NBN methods. By using NBN to obtain the dynamic importance of safety attributes under different navigation environments, safety weights are obtained to prioritize different GCSPPs involving multiple attributes (e.g. cost and efficiency). Based on NBN's quantitative analysis capabilities, ship navigation risk data is allowed to be effectively integrated into the ER approach as input information. Concentrating on practical stowage problems for general cargo ships and taking Tianjin Port waters as a case study, this study is among the first to incorporate quantitative and dynamic safety attributes into the evaluation of feasible GCSPPs across different navigation environments.

The primary outputs of the proposed GCSPP assessment framework are: 1) Dynamic safety weights quantification by NBN, reflecting real–time voyage risk levels derived from all the modelled RIFs; 2) A new ranking score for each candidate GCSPP by considering safety (stability, strength, ship overload, heeling angle), economy (cargo damage, broken space, trim), and efficiency (cargo location at ports) dimensions holistically; 3) Optimal GCSPP selection adaptive to specific navigation environments, balancing safety priority in high–risk navigation environment and economy/efficiency priority in low–risk navigation environment."

The rest sections are organized in the following section: Section 2 reviews and summarize the methodological applications of NBN and ER within the realm of risk analysis. Section 3 presents a hybrid methodology for GCSPP assessment. Section 4 exhibits and discusses the research findings in detail. Section 5 concludes the key insights of this study.

2. Literature review

2.1. Evidential reasoning

Evidential reasoning (ER) is a crucial method for dealing with uncertainty in risk evaluation and forms a key part of the hybrid framework in this research; this part will give an overall view of its underlying principles, clarify its connection to Dempster – Shafer (D–S) theory and how it is applied in maritime risk analysis, then point out its shortcomings when it comes to meeting the dynamic decision – making needs of the GCSPP, which makes it necessary to combine ER with the NBN so as to reduce its problems in modeling and adjusting to real–time navigation risks.

ER is a method with strong robustness, which can integrate information and deal with uncertainty. It has been used to support MCDM and risk assessment studies. This method was first proposed by Dempster in the 1960s and then completed by Shafer in 1976 [9]. It is based on evidence theory, using the basic probability assignment method, and can quantify the uncertainty in the process of performance evaluation [10–13]. It addresses the challenges facing the assessment of GCSPP where objective data and subjective judgment co–exist.

ER technique has also been widely used in maritime research. Specifically, Yang et al. [14] proposed a new framework for analyzing and synthesizing the risk of offshore engineering systems based on a generic fuzzy evidential reasoning (FER) approach. The method aimed to simplify the reasoning process and overcome the problems of traditional fuzzy rule–based risk analysis and decision–making methods Wu et al. [15] proposed an improved method for handling ships without command that combined ER and the Technique of Similar Preference Ordering for Ideal Solutions (TOPSIS) in emergency response decision making. The core of the method was to use ER to integrate the influencing factors indicated by incomplete information of each participating organization, and use TOPSIS method to make the final decision under the asymmetric distribution of information among participating organizations. Yu et al. [12] proposed a probabilistic framework for ship risk assessment based on a mixed–method and multiple data sources. In this study, data were aggregated through NBN and ER methods to assess the overall risk of ships in coastal waters. Dong et al. [16] developed a three–tiered hierarchical assessment framework considering ship conditions, traffic environment, distance from sensitive areas and natural environment. The extended IF–THEN rule was constructed to synthesize multiple influencing factors and derive specific shipwreck risk values through ER. This framework was then applied to the risk assessment of navigation barriers on the Yangtze River.

The above literature reveals that as a powerful uncertainty reasoning and information fusion method, ER has the advantages of being able to overcome the dependence of traditional Bayesian theory on conditional probabilities, using fewer constraints and having a powerful ability to combine evidence. Its ability to transform various forms of inputs into a unified confidence framework and to combine evidence gives ER the capability in dealing with both quantitative information and qualitative knowledge. However, different from NBN, ER lacks the capacity to dynamically capture changes in decision–making attributes in real time. There has long been a substantial research deficiency regarding the evaluation of the GCSPP, especially when it comes to integrating dynamic risk into the ER assessment procedure, and up to now, no research has dealt with the fundamental limitations of ER by including dynamic risk elements, which has led to a crucial disconnection between the GCSPP and the real – time risk changes in maritime navigation, so the current GCSPP evaluation approach is unable to sufficiently adjust to the varying safety requirements brought about by the continuously changing navigational environment.

To address this limitation, a key novelty of this research resides in the incorporation of NBN and ER. Specifically, NBN is employed to determine the weights of the safety attribute under different navigation environments, and these dynamically derived weights are then incorporated into the ER–based evaluation process to support the selection of GCSPPs.

2.2. BN in maritime risk analysis

Bayesian Network (BN) is a strong probabilistic modeling method useful for quantifying dynamic risks and form a basic part of the GCSPP assessment system created in this research; this part describes how BN is being applied at present and introduces some improved versions of it in maritime risk analysis, showing that they work well in assessing dynamic risks while also talking about why these models aren't fully utilized in GCSPP and finding out what causes this problem so as to set a clear path for developing BN models to help evaluate the safety of GCSPP in later parts.

BN, alternatively termed a belief network, is a directed acyclic graph representing probabilistic relationships between variables. Given its strengths in intuitively modelling these relationships, BN is widely utilized for maritime risk assessment, and often combined with other theoretical approaches to offer insights from probabilistic and quantitative perspectives.

For instance, Wu et al. [17] identified key factors relevant to offshore pipeline projects, utilizing Structural Modelling (ISM) to systematically obtain hierarchical networks and assist in constructing BN structures. Yang et al. [18] combined BN with game theory to develop a risky game model, which identified the optimal inspection strategies for port regulatory authorities across diverse scenarios subsequent to the new industrial revolution. Yu et al. [19]used ER to synthesize subjective information from experts and then applied if–then rules to transform it into a conditional probability table (CPT) to build the BN model to analyze the risk of collision between ships and offshore wind turbines. Xu et al. [20] integrated domain experts' specialized knowledge with actual voyage data to construct a Bayesian network, which was specifically designed to predict the probability of ships being beset by ice in the context of Arctic escort operations. Furthermore, Cao et al. [21] analyzed the RIFs of maritime accident severity based on a data–driven TAN–BN model, where accident type and ship–related factors were identified as key contributors to severe accidents.

Uflaz et al. [22] established a perfect framework for accurately grasping the occurrence, severity and detection rate of cyber risks in the ship bridge navigation system, and took AIS spoofing and GPS interference as examples to analyze the impact on maritime transportation. The research method integrates Failure Mode and Effects and Criticality Analysis (FMECA), ER and rule–based Bayesian networks (RBN), and focuses on the occurrence rate of failure, the severity of consequences and the ease of detection, and uses RPN to evaluate the risk priority." Li et al. [23] proposed a Fault Tree Analysis–Fuzzy Bayesian Network model (FTA–FBN) integrated model for quantitative risk evaluation of Maritime Autonomous Surface Ships (MASS) collisions, leveraging BN to forecast probabilities and identify core RIFs. Fan et al. [24] developed a new resilience framework to analyze the resilience performance of the straits and canals under uncertain risk coupling scenarios using a rule–based BN model The proposed framework employed a fuzzy logic reasoning and the BN model Aydin et al. [25] used a fuzzy Bayesian network (FBN) to analyze fire and explosion risks in bulk carriers, focusing on cargo properties, ignition sources, ventilation, crew operations and fire detection reliability as key risk parameters. Li et al. [26] proposed an innovative, data–driven Bayesian network framework for long–term prediction of Arctic shipping risks, which can provide a scientific basis for Arctic shipping policy formulation, emergency response planning, and risk assessment for insurance companies Fan et al. [27] proposed a novel data–driven BN framework that introduces Fisher's optimal segmentation method to address outliers and missing data, improving the accuracy of the model. Wang et al. [28] developed a BN model using AcciMap and ADASYN to handle maritime data issues, identifying "Standardized Operations" as the key risk factor for the first time. Aydin et al. [29] developed a new model that combines Z–numbers and Bayesian networks to assess the risk of autonomous vessels being subjected to cyber–attacks. By quantifying the probability of attacks, key risk factors such as insufficient software updates have been identified. This provides a proactive tool for maritime cybersecurity management. Jovanovic et al. [30] analyzed human factors in serious cargo ship accidents using a combined Fault Tree and Bayesian Network approach. Their study identified personnel management and crew resource management as primary contributors, offering a framework for enhancing maritime safety.

BN is widely used in maritime research for analyzing and predicting ship – related risks, but there exists a substantial research gap as the application of BN to the specific evaluation of GCSPP. It has not been explored yet due to the complexity and uncertainty of GCSPP variables such as cargo features, hull structural traits, and navigational environment. Therefore, applying BN to GCSPP assessment brings a new risk–based approach.

According to the description in the previous two sections, ER and BN exhibit obvious complementary characteristics in application methods and scenarios: ER is capable of combining subjective expert experience with objective quantitative data and can flexibly handle uncertainty information. However, its weight allocation is often static, making it unable to reflect real–time changes in decision attributes such as navigation environment. In contrast, BN can achieve dynamic risk quantification and real–time probability updates by modeling and analyzing causal relationships between variables, accurately describing the

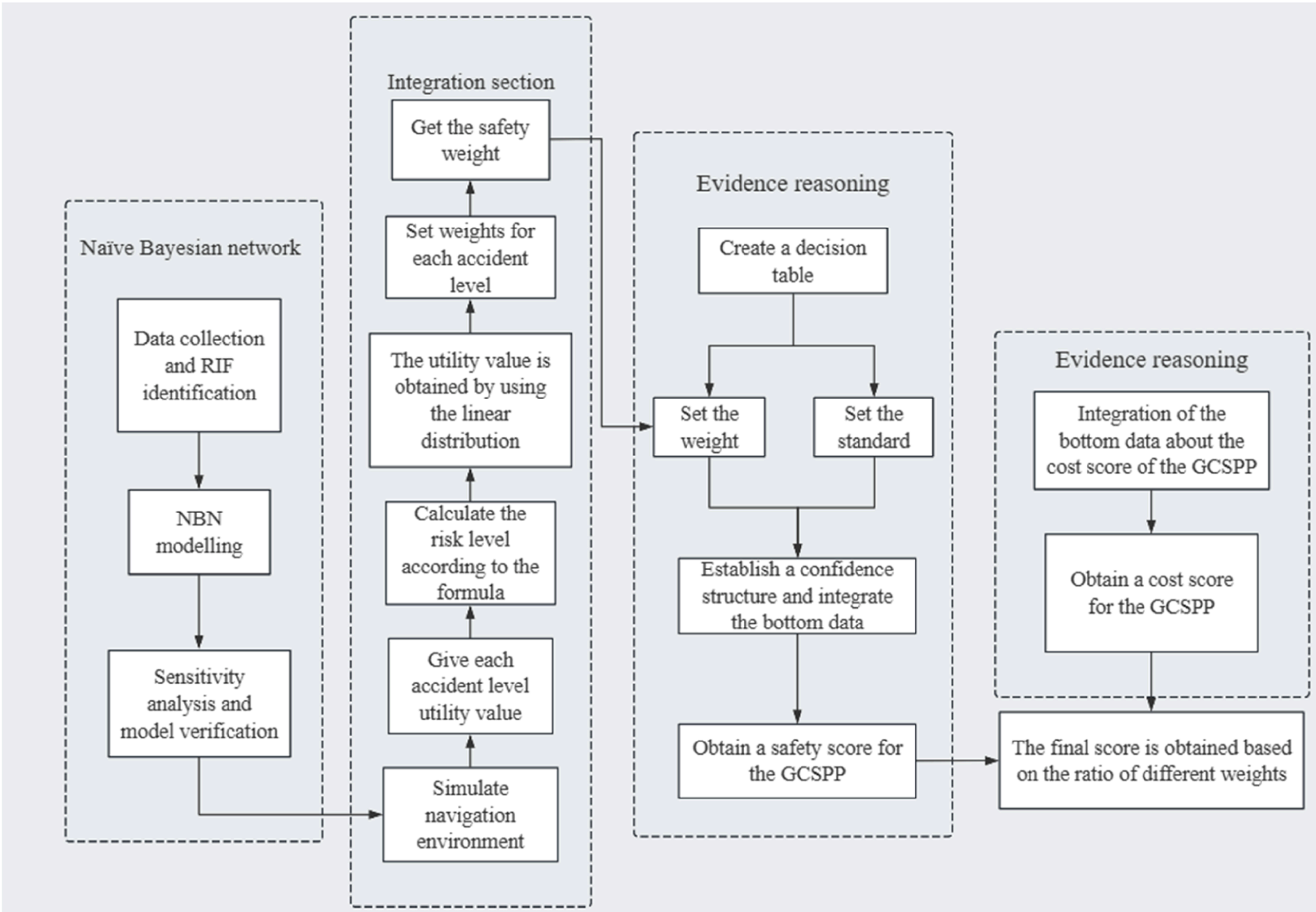

Fig. 1. The methodological flow of this study.

volatility risk of navigation environment. However, BN lacks the ability to integrate qualitative information and subjective experience. This inherent characteristic difference constitutes the core theoretical basis for the integration of the two methods in this study. It uses BN to address the shortcomings of ER in dynamic adaptability, while employing ER to enhance BN in uncertain information integration, thereby jointly meeting the multi criteria evaluation requirements of GCSPP in this study.

BN alleviates the inherent limitations of conventional ER by creating safety weights which fit the real – time navigational environment through the causal modeling of RIFs, offer a dynamic evaluation of how important different GCSPP decision criteria are depending on the specific navigational environment during a voyage, convert these evaluated importance values into weights for the criteria to back up ER – based MCDM, and make sure that the decision – making result more precisely represents the current navigational safety state thus getting over the problems of static weight assignment methods which don't consider the changing safety states of a voyage.

2.3. Study on cargo pre–stowage plan

Cargo ship pre–stowage plan is a basic part of maritime shipping operations and was the main focus of this study; this part gave a full review of the existing literature about it, especially emphasizing the differences in research methods and results between container ships and general cargo ships, also systematically classified the common evaluation methods and risk – reduction strategies found in current studies; the review ended by pointing out the main limitations in the current research on GCSPP, above all the lack of integrating dynamic navigation risks which set the primary reason for the present study.

To date, most studies of ship pre–stowage plan are positioned in the container ship context. Shen et al. [31] proposed a Deep Q–Learning Network (DQN) to solve the container pre–stowage planning problem on board. DQN performed a lot of computation and training in the pre–training phase, while in the application phase it could complete the pre–stowage plan in a few seconds. The results showed that DQN has good effectiveness in solving the ship pre–stowage planning problem. Cho and Ku [32] presented another optimized container pre–stowage planning study by using reinforcement learning. A two–stage pre–stowage planning system was designed, with the first stage using the Proximal Policy Optimization (PPO) algorithm for slot selection and the second stage determining the row and tier placement of containers. The results showed that the procedure enhanced the operational safety and stability of the vessel by generating loading plans more efficiently, reducing the number of reprocessing, and maintaining the uniformity of the weight distribution of the vessel compared to conventional methods. Rodero and Marrero [33] designed a cargo pre–stowage planning tool for roll on/roll off ships, which uses risk–scoring calculations to optimize cargo distribution in order to reduce fire risks. Zhou et al. [34] develop a meta–heuristic algorithm for pre–stowage planning on large containerships to reduce container shifts and demonstrated its effectiveness by completing real pre–stowage plans for large container ships. Sivertsen et al. [35] introduced a novel model for the Container Pre–Stowage Planning Problem (RCSPP), and demonstrated through experiments that omitting the constraints imposed by actual operations can lead to inaccurate estimates of twenty–foot equivalent units (TEU). Pham and Nguyen [36] applied a Random Forest model to accurately predict ship berthing time, which considers cargo volume as the most

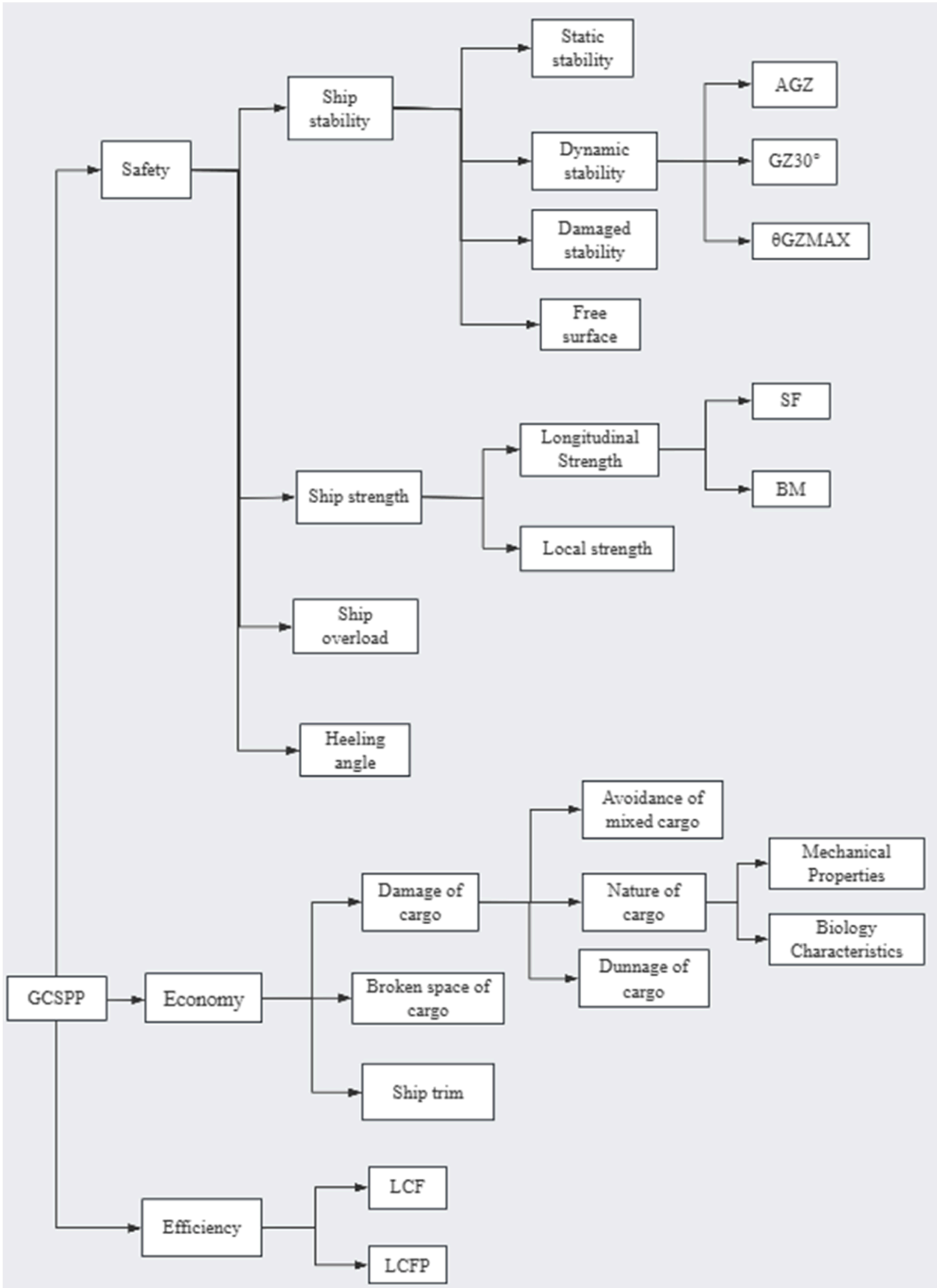

Fig. 2. System for evaluating GCSPP. AGZ: Area under GZ curve; GZ30°: The righting arm at a 30° transverse heel; θGZ_{MAX}: Angle at which the righting arm peaks; SF: Shear force; BM: Bending moment; LCF: The location of the first arrival port cargo; LCFP: The location of the final arrival port cargo.

significant influencing factor and provides an effective method for improving port loading and unloading efficiency.

Compared to pre–stowage planning in container ships, GCSPP is less explored due to its complexity despite the significance of general cargo ships in maritime transport and the high number of incidents caused by the inappropriate stowage in general cargo ships. Pérez–Canosa et al. [37] discussed how the navigation parameters of ships can be optimized to improve the stowage and fastening criteria of non–standard cargoes on board. The study analyzed the International Maritime Organization (IMO) and Classification Society standards, established a novel mathematical framework to calculate the lateral acceleration of a vessel under different sea states and navigation parameters, and presented 3D surface maps to help shipmasters quickly determine the optimal speed and heading. Ding et al. [38] conducted an empirical risk analysis of operational challenges faced by general cargo ship operators using a risk evaluation procedure based on Analytic Hierarchy Process (AHP). By identifying the risk factors and using expert questionnaires to construct a risk matrix, the study revealed four risk factors in the high–risk area. The risk management strategy based on expert recommendations was proposed to provide shipowners and charterers with a reference to the risk terms when signing chartering contracts. Zhang et al. [39] invited 18 domain experts from 15 global shipping enterprises across 5 countries to assign the weights to various safety criteria in GCSPP using the subjective AHP method. There has long been a considerable research deficiency in the area of GCSPP as dynamic navigation risk data is conspicuously left out of the existing multi – criteria evaluation frameworks, at present the methods in use show an excessive dependence on static safety compliance limits and the arbitrary allocation of weights for assessment which restricts the accuracy of these models. Therefore, a review of these limited studies further highlights another key contribution of this study, that is, incorporating the ship safety dynamics across different navigation environments into a systematic assessment framework of GCSPP.

3. Methodology

The framework began with multi–source data collection and risk factor identification. For data collection, we used accident reports of general cargo ships from 2007 to 2021, which were acquired from the China Maritime Safety Administration as the primary data source. The data set was supplemented with voyage data, including vessel characteristics, cargo list, and real–time navigation environment parameters. Through the analysis of accident reports and the comprehensive review of the relevant literatures, 14 RIFs were identified.

An NBN model was subsequently constructed to achieve safety quantification. The model took the 14 RIFs as input nodes and accident severity as the output one to calculate the navigation risk level under specific navigation scenarios. The risk level was further converted into changeable safety weights that reflect the real–time navigation scenarios and could be automatically calculated by the NBN model.

In order to effectively evaluate the GCSPP, the three–dimensional evaluation framework was built to ensure that the evaluation could be carried out in the aspects of safety, cost and efficiency. At the same time, ten experts from the shipping industry were employed to evaluate the cost and efficiency dimensions. Each of them had >15 years of work experience and had different but highly relevant backgrounds such as ship operators and captains.

We first calculated changeable safety weights using the NBN model. Afterwards, the ER method was employed to integrate all evaluation data covering safety, cost and efficiency dimensions. Using the integrated datasets, we worked out the final score for each candidate GCSPP. These scores served as the main reference for selecting the optimal GCSPP. (Fig. 1)

3.1. Establishment of the GCSPP decision–making system

This study developed two different GCSPPs for a same cargo list to make it easier to conduct a comparative analysis, adopted an evaluation framework consisting of a three–dimensional hierarchy of safety, economy, and efficiency from the validated methodology of Zhang et al. [39] (Fig. 2) which offered a strong basis for multi–criteria assessment, and since the original method had limitations such as its static weight allocation being separate from navigational risk and relying too much on subjective expert judgment this research made important improvements to the hybrid framework which were aimed at making it much more adaptable and objective as will be explained below.

Firstly, dynamic safety weight replaces static subjective weights. Former research, like that of Zhang et al. [39], depended on expert evaluation and the AHP to assign fixed safety weights which did not take into account the dynamic characteristics of real–time navigational environmental risks. In sharp contrast, this study utilized a data – driven NBN model to quantify safety weights in a dynamic manner and this method determined weights according to specific navigation environment thus guaranteeing that the safety of GCSPP was in line with actual operational demands.

Secondly, the safety criteria of the framework are linked to RIFs in particular operational situations and Zhang et al. [39] set up safety sub–criteria, that is to say, stability, strength, overload and heeling angle, in accordance with international regulations like the 2008 IS code while the original framework only checked whether the minimum requirements were met (e.g., GM $\geq$ 0.15 m), but in high–risk navigation environment, shipping firms usually adopt stricter safety standards than these minimum regulatory ones to improve navigation safety so this research builds on the original framework by adding RIFs, thus making it possible to have adaptable safety standards while meeting safety thresholds instead of just following traditional thresholds.

Thirdly, in order to improve the economic and efficiency evaluation of GCSPPs, a cost and efficiency – scoring method was developed which had been validated via expert assessment .Considering the complexity of this assessment, a group of ten experts who had a great deal of experience in general cargo shipping was formed including four certified general cargo ship captains, three people with doctoral degrees, four workers from the operation department of general cargo shipping company, five academic researchers and two officials from Maritime Safety Administration and each expert was given an equal weight of 0.1 to assess the costs and efficiency related to different GCSPPs in various navigation environment paying special attention to costs like dunnage costs and broken space charges which are specific to general cargo and qualitative expert opinions on economic criteria were changed into quantitative values by ER which reduced subjective bias in economic evaluation and improved the overall dependability of the suggested framework.

The standard grading of quantitative and qualitative criteria in the framework is consistent with the validation framework of Zhang et al. [39]. On the basis of the dynamic weight, hierarchical classification, and expert assessment of cost and efficiency scoring mentioned above, a belief structure was established based on the subordination relationship between criteria. Then, the framework combined basic data (such as stability parameters, cargo loading details) and dynamic safety weights derived from NBN to calculate the safety score of each GCSPP. At the same time, the expert cost and efficiency scores are integrated into the structure through an intelligent decision–making system (IDS), which is a specialized decision–support software tool used to implement ER for multi–attribute evaluation and comprehensive scoring, thereby generating a cost and efficiency score that supports comprehensive GCSPP ranking [40]. It is worth noting that the efficiency dimension quantified by criteria such as the cargo location of the first and final goods arrival ports has also been incorporated into the belief structure. Its score comes from the actual operability of cargo loading and unloading, directly reflecting the impact of GCSPP on port operation efficiency. The three–dimensional scoring system (safety, cost, efficiency) ensures the comprehensiveness of the assessment of GCSPP.

3.2. NBN model for safety evaluation against different navigation scenarios

3.2.1. Data collection

Accident records of maritime incidents between 2007 and 2021 were obtained from China's Maritime Safety Administration (MSA) and assessed systematically. Those accidents occurring in Tianjin Port, which is the largest general cargo ports in China and the leading one in the world, are selected to serve as the basic data population. Data screening is performed to retain only information related to the research scope (e.g. general cargo ship accidents). Following the exclusion of non–relevant accidents, 160 general cargo ship accident reports are retained as the research's input data.

Table 1
Definitions and classification of the employed RIFs.

Categories	Notation	Description	Values of state in NBN	Reference
Crew condition	R_C	Crew with valid certifications (sufficient quantity); Insufficient crew or invalid certifications	1,2	[41] [42] [43]
Gross tonnage	R_{GT}	500t or less; 500t to 3,000t; greater than 3,000t	1,2,3	[44] [45] [46]
Human factors	R_{HF}	No human factors involved; human factors involved	1,2	[47] [48] [49]
Ship length	R_{SL}	100 m or less; >100 m	1,2	[6] [7]
Fairway traffic	R_{FT}	Good, poor	1,2	[6]
Loading	R_L	Full loaded; ballast; overloaded	1,2,3	[50] [51]
Location	R_{Lo}	Quay; port channel; anchorage; water area adjacent to the port; water far from the port	1,2,3,4,5	[42] [52] [53]
Ship age	R_{SA}	0–5years; 6–10years; 11–15years; 16–20years; >20years	1,2,3,4,5	[44] [17] [54]
Ship flag	R_{SF}	China; the flag of convenience (FOC); other	1,2,3	[8]
Season	R_S	From April 16 to October 31; From November 1 to the following April 15	1,2	[46] [54]
Visibility	R_V	3nm or less; 3–5nm; greater than 5nm	1,2,3	[44] [54]
Equipment/ Device	R_E	Properly functioning onboard equipment; Underutilized or malfunctioning equipment	1,2	[6] [7]
Cargo stowage	R_{CS}	The location or nature of the cargo does not affect the navigation safety; the location or nature of the cargo is closely related to the accident (e.g. cargo combustion, cargo fluidization)	1,2	[39]
Weather condition	R_{WC}	Good (without rain, wind, or fog); poor (with rain, wind, or fog)	1,2	[6] [7]
Accident severity	S	Minor; major; critical; catastrophic	1,2,3,4	[8]

Table 2
Explanation of each level of accident severity (Source: [8]).

Severity of the accident	Specific explanation
Catastrophic	It refers to accidents that cause >30 deaths (including missing), or >100 serious injuries, or more than $13,738,400 of direct economic loss.
Critical	It refers to accidents that cause >10 and <30 deaths (including missing), or >50 and <100 serious injuries, or more than $6869,200 and less than $13,738,400 in direct economic loss.
Major	It refers to accidents that cause >3 and <10 deaths (including missing), or >10 and <50 serious injuries, or more than $1373,840 and less than $6869,200 in direct economic loss.
Minor	It refers to accidents that cause more than one person and less than three deaths (including missing), or more than one person and <10 serious injuries, or less than $1373,840 of direct economic loss.

3.2.2. RIF identification

Drawing from the 160 accident reports, a deeper exploration of the RIFs is conducted. Utilizing keywords such as "Bayesian network," "maritime risk," and "maritime accident," relevant literature is searched from the Web of Science. After meticulous screening of the abstracts and contents of the retrieved papers, studies presented in Table 1 are selected for in–depth analysis. These studies have identified a range of RIFs, including Gross tonnage, Crew, Ship age, Weather condition, Fairway traffic, Equipment/device, Season, Ship length, Visibility, and Human factors, etc. Considering the diversity in accident locations and ship nationalities within the general cargo ships under examination, another two novel RIFs, i.e., Location and Ship flag, are further incorporated into this study. Furthermore, given the study's exclusive focus on general cargo ships, Loading and Cargo stowage are also added as new RIFs in the NBN model–building process. Table 1 provides an overview of all 14 RIFs utilized in this study.

In this study, accident severity is considered as the dependent variable for NBN modelling, which is classified into four levels, i.e., catastrophic, critical, major and minor [8]. Table 2 provides the definitions of each level.

3.2.3. NBN modelling

In the BN model, each node represents a variable and links between the nodes indicate the conditional relationships between the variables. The basic principle is referred to the Bayes' theorem as follows [55]:

$$P(A_i|B) = P(B|A_i) \times P(A_i) \Big/ \sum_j P(B|A_j) \times P(A_j) \tag{1}$$

where $A_1 \ldots A_j$ is the complete event group, $P(A)$ is the prior probability of A and $P(B)$ is the prior probability of B. $P(A|B)$ is the conditional probability of A after B is known to occur, and $P(B|A)$ is the conditional probability of B given A is known to occur.

Modelling the identified RIFs in Section 3.2.2 as nodes, this study applies a data–driven approach to generating the NBN structure. Considering the trade–off between accuracy and computational complexity, the NBN structure is adopted in this study, where the accident severity is selected as the parent node (i.e., the dependent variable) and directly connects to other independent variables. Despite the limitations of the conditional independence assumption in NBN, this study still regards it as the core modeling method. The reason is that NBN has sufficient prediction accuracy and stability (verified by 95.83% test accuracy and 93.13% 10–fold cross validation accuracy in Section 4.1.4), making it more suitable for real–time GCSPP assessment in actual shipping operations.

3.2.4. Mutual information

Mutual information is used to describe the degree of interdependence between two variables. The larger the mutual information, the stronger the link between RIF and accident severity. In this study, the calculation of mutual information can be used to filter important variables among RIFs, which is referred to Eq. (2) as follows [56]:

$$I(X, Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log_2 \frac{p(x, y)}{p(x)p(y)} \tag{2}$$

Let the joint distribution of two random variables (X, Y) be $p(x,y)$, the edge distributions are $p(x)$ and $p(y)$, respectively. The mutual information $I(X, Y)$ is the relative entropy of the joint distribution $p(x,y)$and the marginal distribution $p(x)p(y)$. In this study, Y represents the accident node severity, X represents a RIF, and x represents a state in the RIF. RIFs can be ranked by their impact on accident severity through mutual information quantification.

3.2.5. Sensitivity analysis

By screening important RIFs based on mutual information, a scenario simulation approach is further taken to analyze the impact of each variable. In this study, the method proposed by Alyami et al. [57] is adopted. For a certain RIF, Highest Risk Impact (HRI) is generated by fixing the highest–impact state on accident severity to 100% probability, while Lowest Risk Impact (LRI) is obtained by setting the lowest–impact state to 100% probability. Total Risk Impact (TRI) is defined as the mean value of HRI and LRI, as shown below:

$$TRI = \frac{HRI + LRI}{2} \tag{3}$$

In other words, under specific accident severity conditions, the magnitude of a RIF's impact on accident severity is positively correlated with its *TRI*.

3.2.6. Model validation

In this study, to validate the established model, two approaches are used to assess how multiple RIFs jointly impact accident severity. Firstly, two axioms are adopted in this study [58]:

Axiom 1: A slight increase or decrease in the prior probability of each test node should contribute to a corresponding increase or decrease in the posterior probability of the target node.

Axiom 2: The practical effect of combined probabilistic changes in evidence should not be less than that of an evidence subset.

Further to this, the data obtained from the accident reports is also divided into two parts, with 85% data inputted as the training dataset and 15% data is applied as testing datasets to examine the predictive performance of the model.

For enhanced model reliability validation, K–fold cross–validation is integrated into the training dataset's validation procedure [59]. The specific implementation process is as follows: the 85% training dataset is randomly divided into K mutually exclusive subsets of approximately equal size (this study adopts K = 10, i.e., 10–fold cross–validation), and the remaining 15% is used as an independent test set; In each round of validation, one subset is selected as the "validation set" for model performance assessment, and the remaining K–1 subsets are combined as the "training subset" for model training; This process is repeated K times so that each subset is used as a validation set exactly once; Finally, the reliability of the model is determined based on the evaluation results. Among them, area under the receiver operating characteristic (ROC) curve (AUC) and F1–score are considered two key evaluation indicators.

AUC: It can represent the model's ability to distinguish different categories through numerical values, with a range of 0 to 1. The closer the value is to 1, the better the model's ability to distinguish categories.

F_1–score: It is the harmonic mean of precision and recall, with a value range of 0 to 1. The closer the value is to 1, the better the overall classification performance of the model. The formula is as follows:

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

Where '*Precision*' refers to the proportion of results predicted as positive class by the model that are truly positive. '*Recall*' refers to the proportion predicted by the model results among all samples that are truly positive.

3.3. Hybrid ER and NBN

3.3.1. Risk–based gcspp through NBN modelling

As reviewed in Section 2.3, among the limited studies assessing GCSPP, a common evaluation approach relies primarily on subjective inputs, such as questionnaires or interviews with experts and stakeholders [39]. While this approach provides a basis for modelling the relative importance of different safety criteria for GCSPP, especially in the absence of real–world data, it also has inherent limitations. Overreliance on subjective preferences and experience may compromise objectivity in weight allocation, ultimately impacting the final decision–making process.

Moreover, in real–world shipping operations, the risk level associated with general cargo ship navigation varies across different scenarios. However, in earlier studies, risk weights are typically assigned in a relatively fixed manner, limiting their applicability. To overcome these shortcomings, this work pioneers the integration of the data–driven NBN risk assessment and ER into the GCSPP framework. This approach enables a more robust determination of safety weights across different navigation environments, boosting the reliability of the risk–based GCSPP evaluation system to support effective cargo management aboard ships.

3.3.2. Assessing the safety of voyages across different navigation environments

For better handling of the extensive qualitative and quantitative information in GCSPP risk assessment, a hierarchical structure was established using the following relevant terms to achieve the final evaluation. Here are interpretations of some relevant terms in this study:

Utility Value: In the ER method, it can transform qualitative evaluation results into computable values, providing more reliable basis for final decisions.

Safety Weight: By using the risk analysis results of NBN, its value can accurately reflect the safety status of GCSPP under different navigation environment. The weights of safety (ω_s), economy/cost (ω_c) and efficiency (ω_e) satisfy that the sum of the three is equal to 1 (100%), as shown in Eq. (5).

$$\omega_s + \omega_c + \omega_e = 1 \tag{5}$$

Safety Score: By constructing a framework for evidence–based reasoning through specific utility values and safety weights, the calculated results represent the safety score of GCSPP, with higher scores indicating higher levels of safety.

Cost and Efficiency Score: Experts evaluate the economic and efficiency performance of GCSPP by analyzing various operational factors, and the score reflects the advantages and disadvantages of GCSPP in relation to cost and efficiency.

Final Score: Calculate the final score by combining the weights of safety, cost, and efficiency.

Firstly, different navigation environments will result in different navigation risks. When a ship faces higher navigation risks, the safety in GCSPP will have to carry more weight. In contrast, when a ship faces lower navigation risks, the economy and efficiency of a GCSPP will be given higher priorities. Therefore, safety weights in the evaluation system of the GCSPPs should be increased or decreased accordingly with changes of navigation risks. To determine the safety weights, the results from the risk analysis conducted through the NBN structure are transformed into ER to obtain the quantitative safety weights. These safety weights are then used to calculate the safety scores of different GCSPPs. By incorporating these safety scores into the overall evaluation, the selection of GCSPPs can be better aligned with the specific navigation risks encountered.

Based on this goal, four accident severity levels in the NBN model are quantified by assigning a set of utility values. Specifically, the severity gradually increases from minor, major, critical to catastrophic. The utility values of 2, 4, 6 and 8 are assigned to each to represent their corresponding severity (S), respectively [60]. Subsequently, the probabilities (*P*) assigned to each grade of accident severity (*S*) can be used to calculate the risk level (*R*) a ship would have to face in a particular navigation environment, as shown in Eq. (6).

$$R = \sum P \times S \tag{6}$$

After obtaining the risk values, the distribution interval of each value can be determined, and then the actual safety weight can be calculated by ratio analysis of the risk value relating to two endpoints of the

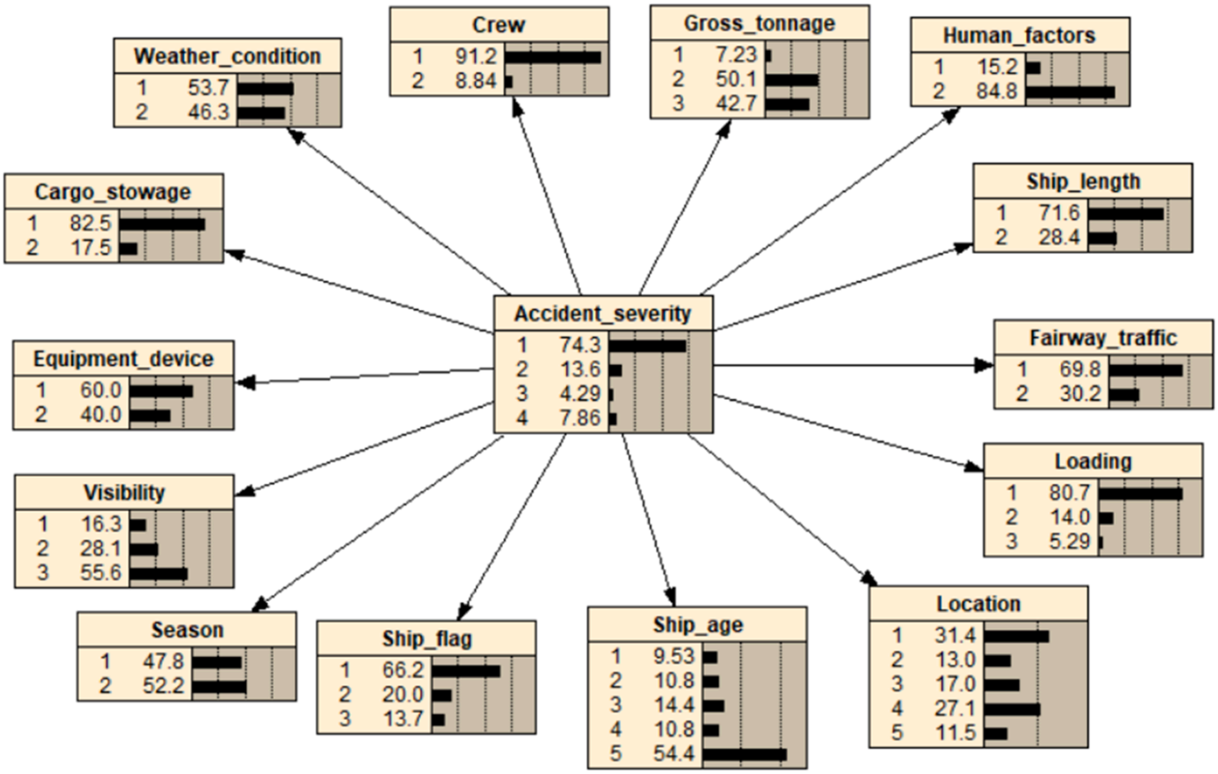

Fig. 3. The established NBN structure.

Table 3
Mutual information.

Node	Mutual information	Percentage	Variance of beliefs
Accident severity	1.1927	100	0.2565
Location	0.0951	7.97	0.0149
Cargo stowage	0.0930	7.8	0.0115
Ship flag	0.0787	6.6	0.0135
Equipment/device	0.0532	4.46	0.0082
Fairway traffic	0.0493	4.14	0.0056
Visibility	0.0448	3.76	0.0039
Weather condition	0.0352	2.95	0.0004
Season	0.0186	1.56	0.0011
Loading	0.0173	1.45	0.0029
Crew	0.0170	1.43	0.0021
Gross tonnage	0.0168	1.41	0.0004
Ship length	0.0160	1.34	0.0017
Ship age	0.0073	0.609	0.0009
Human factors	0.0060	0.499	0.0001

interval. For instance, when the probability of risk is 50% minor risk and 50% catastrophic risk, the risk value at this time is $2 \times 50\% + 8 \times 50\% = 5$.

Here, for the lowest accident severity, “minor”, the safety weight is set at 50%. Given the rising risk levels of “minor”, “major”, “critical” and “catastrophic”, the corresponding safety weights also rise. Thus, 60%, 70% and 80% are assigned to the safety weights for “major”, “critical” and “catastrophic” to reflect how safety weightage increases or decreases with risks. Since the critical safety weights corresponding to major and critical are set to 60% and 70%, the actual safety weights are $60\%*50\% + 70\%*50\% = 65\%$.

Furthermore, this study aims to link the results of the risk analysis with the GCSPPs with consideration of different risk situations. As a result, dynamic safety weights are set to correspond to real–time navigation environments. The general cargo ships navigating on different routes face different degrees of risk, and the differences in risk on different routes lead to changes in safety weights. Therefore, crews and shipping companies can dynamically adjust the safety weights in the GCSPP according to the risk characteristics of specific routes to ensure safe navigation.

4. Results and discussion

4.1. NBN modelling

The NBN model was first specified by defining the variables and states from the case data, ensuring consistent discretization for categorical learning. Next, the structure was fixed to a naïve star–shaped topology. This followed the class–centric design where the target node is connected to all features known as child nodes. With the topology defined, parameter learning was performed by importing the case dataset and letting the Netica software estimate the conditional probability tables (CPTs) via Bayesian updating. The Netica’s interface then allowed immediate belief updating and selection of value ranges as evidence.

In this regard, based on the accident reports, the identified RIFs, and the model instructions presented in Section 3.2, the proposed NBN structure was developed and illustrated in Fig. 3, providing a graphical representation of the relationships among the RIFs and accident severity. The network adopts a naïve topology with accident severity as the target node and all other identified RIFs modeled as direct parents. This design assumes the parents are conditionally independent given accident severity, which keeps inference transparent, reduces parameter burden, and performs robustly with limited case counts.

Conceptually, the incoming arcs represent complementary dimensions of exposure (weather, visibility, season, fairway traffic, location), susceptibility (loading condition, cargo stowage integrity, equipment/device condition, human and crew factors), and inherent vessel attributes (gross tonnage, length, age, flag). Based on the obtained result, the minor accident has the highest probability of 74.3%, followed by major at 13.6%. Among ship factors, 50.1% of ships have a GT between 500 and 3000 tons, while 42.7% of ships have a tonnage of >3000

Table 4
Individual RIF states set to 100%.

RIFs	Minor	Major	Critical	Catastrophic
Initial	74.3	13.6	4.29	7.86
Location				
1	**87.6**	9.39	**1.36**	**1.67**
2	84.8	**4.55**	6.61	4.04
3	80.9	10.4	2.52	6.16
4	60.9	19.6	**7.91**	11.6
5	**47.8**	**25.7**	3.73	**22.8**
Cargo stowage				
1	**80.6**	**13.2**	**2.23**	**3.97**
2	**44.4**	**15.5**	**14**	**26.1**
Ship flag				
1	**83.6**	**7.81**	4.04	**4.56**
2	63.0	22.6	**5.35**	9.06
3	**45.9**	**28.2**	**3.9**	**22.0**
Equipment/device				
1	**82.6**	**7.92**	**5.11**	**4.37**
2	**61.8**	**22.0**	**3.06**	**13.1**
Fairway traffic				
1	**80.0**	**13.6**	**2.63**	**3.75**
2	**61.0**	**13.5**	**8.12**	**17.4**
Visibility				
1	**81.8**	**7.94**	**6.58**	**3.71**
2	**64.7**	**25.3**	5.17	4.29
3	76.9	9.30	**2.89**	**10.9**

Table 5
TRI for each RIF.

RIFs	Minor	Major	Critical	Catastrophic	Average
Location	19.90	10.58	3.28	10.57	11.08
Cargo Stowage	18.10	1.15	5.89	11.07	9.50
Ship flag	18.85	10.20	0.73	8.72	9.62
Equipment/device	10.40	7.04	1.03	4.37	5.71
Fairway traffic	9.50	0.05	2.75	6.82	4.78
Visibility	8.55	8.68	1.85	3.60	5.67

Table 6
Influence of changes of target nodes on different accident severity levels.

Changes		5%	5%	5%	5%	5%	5%
Fairway traffic			5%	5%	5%	5%	5%
Equipment/ device				5%	5%	5%	5%
Ship flag					5%	5%	5%
Cargo Stowage						5%	5%
Location							5%
1	74.30	75.10	76.10	77.10	78.90	80.50	82.20
2	13.60	14.40	14.40	15.20	16.30	16.40	17.60
3	4.29	4.47	4.76	4.87	4.95	5.64	6.07
4	7.86	8.22	8.92	9.41	10.40	11.70	13.10

tons. Ships with the length of <100 m and those with the length of >100 m account for 71.6% and 28.4% respectively. Ships with the age of 20 years or more account for 54.4%. In addition, most accidents happen in the waters near piers and harbors, accounting for 31.4% and 27.1% respectively. Moreover, 84.8% of accidents are related to human factors.

4.1.1. Mutual information

Accordingly, the mutual information between each RIF and accident severity is quantified. As presented in Table 3, accident location has the largest mutual information value, followed by cargo stowage, ship flag, equipment/device, fairway traffic, and visibility. The mean value of mutual information is 0.03916, and the first six RIFs on the table have mutual information greater than this mean value. Thus, these six RIFs are recognized as having a higher impact and analyzed further below.

Table 7
Predictive performance testing.

Testing data No.	Prediction results	Testing data No.	Prediction results
1	97.30, 1.10, 0.62, 0.94	13	93.60, 6.22, 0.10,0.04
2	94.2, 4.49, 0.75, 0.53	14	97.50, 1.97, 0.28, 0.23
3	92.8, 2.62, 2.05, 2.49	15	96.00, 3.88, 0.06, 0.09
4	95.4, 2.88, 0.26, 1.41	16	98.30, 1.59, 0.08, 0.03
5	98.00, 1.82, 0.14, 0.08	17	91.20, 4.20, 3.50, 1.06
6	25.8, 7.51, 0.22, 66.40	18	96.00, 3.88, 0.06, 0.09
7	97.20, 2.17, 0.09, 0.57	19	98.30, 1.59, 0.08, 0.03
8	98.20, 0.91, 0.71, 0.19	20	97.20, 1.80, 0.17, 0.84
9	95.80, 1.15, 2.31, 0.70	21	97.20, 2.17, 0.09, 0.57
10	97.80, 1.35, 0.19, 0.63	22	97.30, 2.21, 0.12, 0.38
11	95.10, 0.85, 3.59, 0.51	23	96.60, 0.99, 1.99, 0.38
12	94.10, 2.84, 2.74, 0.31	24	96.10, 3.48, 0.12, 0.27

4.1.2. Sensitivity analysis

For these six important RIFs screened on the basis of mutual information, this study set each of their states to a 100% probability individually, then the probability result of four severity levels is summarized in Table 4. Bold text indicates the states in each variable that have the largest and smallest impact on the accident severity. For example, for the RIF of location, state 1 has the greatest impact on minor accidents, with the probability of 87.6% and 13.3% (i.e., the HRI) higher than the initial value of 74.3%. While state 5 has the least impact on minor accidents, with the probability of 47.8% and 26.5% (i.e., the LRI) lower than the initial value. The average of these two values, 19.9%, is the TRI. The same approach is taken for the other RIFs and their individual states. Table 5 presents the TRI values for the six selected key RIFs, and the influence levels of these six most significant factors are outlined below:

Location>Ship flag> Cargo Stowage>Equipment/device>Visibility>Fairway traffic

4.1.3. Model validation

For enhanced model accuracy validation, the six identified RIFs were subjected to the two axioms from Section 3.2.6, with the variation set to 5%, and changes in accident severity were observed and recorded in Table 6. For example, this study increased the probability by 5% of the RIF state that most strongly impacts minor accidents, and decreased the probability by 5% of the state with the weakest impacted on these accidents, and then applied the same approach to all RIFs in turn. Similarly, the same approach was taken for the other three accident severity levels. From Table 6, it could be seen that as the RIFs were adjusted, the patterns were in accordance with Axiom 1 and Axiom 2. Therefore, the effectiveness of the NBN structure could be verified.

4.1.4. Predictive performance testing

Furthermore, to deliver predictive performance testing, 136 accident records are imported into Netica for data learning to generate NBN, and then 24 accident records are utilized as the testing dataset. If the probability of the actual situation being predicted is greater than 90%, then the prediction is considered to be accurate. As illustrated in Table 7, the prediction outcomes corresponding to the four severity levels are listed from left to right within each row, where 23 out of 24 accident records

Table 8
10–fold cross validation statistical summary.

Assessment Metric	Mean (10–Fold)	Std. Dev. (10–Fold)
Mean Log–Likelihood (Mean LL)	−0.80	0.03
Accuracy	0.93	0.04
Weighted F1–score	0.92	0.03
Macro–AUC	0.94	0.02

Table 9
Ship Information.

Vessel detail	Numerical value
Tonnage	2500t
Ship's breadth	21.2m
Ship flag	China
Ship length	90m
Ship age	21 years

Table 10
Cargo List.

NO.	Name of cargos	Package	Weight (tons)	Stowage volume(m^3)	Unloading port
1	Milk powder	Drums	204	224.4	Yantai port
2	Camphor	Cases	51	178.5	Yantai port
3	Teas	Cases	42	92.4	Yantai port
4	Willow products	Cartons	34	95.2	Yantai port
5	Steel plates	Drums	203	152.25	Yantai port
6	Chromium nitrate	Bags	34	44.2	Tianjin port
7	Lithopone	Bags	118	141.6	Tianjin port
8	Aluminum ore	In bulk	507	608.4	Tianjin port
9	Honey	Drums	91	91.0	Tianjin port
10	Cotton fabrics	Cases	51	127.5	Tianjin port

Table 11
Status setting of each RIF.

RIFs	Navigation environment I	Navigation environment II
Crew	2	1
Gross tonnage	2	2
Human factors	2	2
Ship length	1	1
Fairway traffic	2	1
Loading	2	2
Location	5	1
Ship age	5	5
Ship flag	1	1
Season	2	2
Visibility	1	1
Equipment/device	2	1
Cargo stowage	2	2
Weather condition	2	1

are correctly predicted, with an accuracy rate of 95.83%.

To further confirm the model's robustness, a 10–fold cross–validation was carried out. The outcomes indicated that the accuracy of the NBN in predicting the severity of accidents reached 93.13%, and the standard deviation of all indicators was below 0.04, indicating its high stability. The standard deviation of all multiples is only 0.038. This indicates that the prediction accuracy of the model is high, and its score in multi–class AUC is 94.1%, with a weighted F1 score of 92.3%, indicating that the model also has strong class discrimination ability. The main data is shown in Table 8 below.

4.2. Risk analysis–based gcspp assessment in different navigation environments

4.2.1. Scenario setting

For the purpose of comparative evaluation, two representative GCSPPs, namely GCSPP I and GCSPP II, were designed for the selected ship and cargo list. GCSPP I was formulated with a focus on safety, while GCSPP II paid more attention to operational cost and loading–unloading efficiency. These two GCSPPs are used in the following scenario analysis to verify the effectiveness of the proposed integrated framework

To further develop the GCSPP assessment based on the risk analysis, a general cargo ship was adopted as the analysis subject. The specific information about the selected ship is shown in Table 9. The cargoes carried by this ship include milk powder, camphor, teas, willow products, steel plates, chromium nitrate, lithopone, aluminum ore, honey and cotton fabrics, where the cargo manifest is shown in Table 10. The vessel is equipped with 5 holds, each of which was partitioned into a bilge and an upper deck. When designing GCSPP I and GCSPP II, the stowage volume of the cargos is estimated based on the stowage factor and the weight of the cargos provided by the cargo owner, and it is necessary to ensure that the stowage volume of the cargos was smaller than the volume of each compartment. The general cargo ship will depart from Qingdao port, pass through Yantai port (indicated as B) and finally arrive at Tianjin port (indicated as C).

Moreover, to measure the effect of distinct safety weight allocations on the final assessment of various GCSPPs, two sailing environments are simulated in this study, as shown in Table 11. Navigation environment I is set as an extremely high–risk navigation environment, with all RIFs pointing towards unfavorable conditions that increase the ship accident risk. For navigation parameters, the ship sails in international waters (Location, state5) the waterway is very crowded (Fairway traffic, state2), with visibility is only 2 nautical miles (Visibility, state1) and it will encounter typhoons (weather condition, state 2). In terms of ship operation and status, the crew lacks qualifications (Crew, state2) and the equipment is not functioning properly (Equipment/device, state2), and there are violations of operating regulations (Human factors, state 2). In terms of inherent attributes of ships, it belongs to small and medium––sized general cargo ships (Gross tonnage, state2; Ship length, state1) and has an age of 21 years (Ship age, state 5). In terms of cargo loading, the vessel is not fully loaded (Loading, state2), and there are safety hazards in the position of special cargo (Cargo stowage, state2).

Navigation environment II is a simulation of typical navigation environment, where the state of RIFs varies and is considered a simulation of relatively low–risk situations.

Based on different environmental conditions and according to the cargo list, two different GCSPPs (i.e., GCSPP I and GCSPP II) are developed for a detailed analysis. Specifically, GCSPP I is developed by the first mate and focuses more on the safe navigation of ships, while GCSPP II is developed by the port loading manager and focuses more on the cost of ships. Fig. 4 and Fig. 5 present the detailed configurations of the two GCSPPs.

4.2.2. Risk analysis–based gcspp assessment

According to the lowest level data formulated by D. Zhang et al. [39], these two loading plans were transformed into the IDS. After inputting the bottom–level criteria data (Table 12) into the IDS software, subsequent analysis will focus on determining the safety weights. In this section, a detailed calculation of these safety weights will be conducted using the combined framework. This approach ensures that the safety weights are derived from a robust, data–driven analysis, enhancing the objectivity and accuracy of the assessment.

As introduced in Section 3.1, ten experts with over 15 years of experience scored the economic and efficiency dimensions (Tables 13–16), focusing on core cost criteria (e.g., cargo damage, broken space, ship trim, and dunnage costs) and efficiency criteria (e.g., LCF, LCFP). Through the ER method, these scores reduced subjective

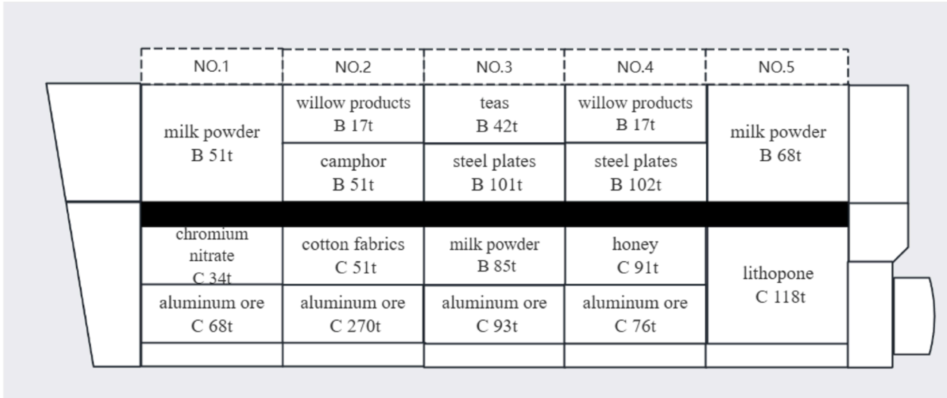

Fig. 4. GCSPP I.

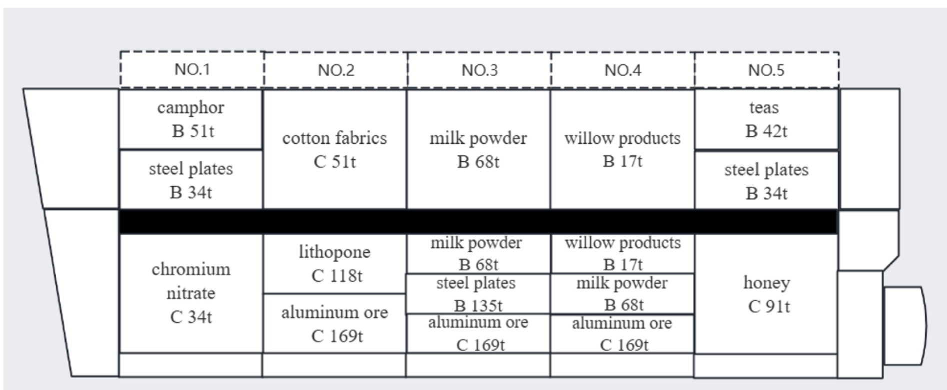

Fig. 5. GCSPP II.

bias and provide a quantitative basis for GCSPP feasibility. As shown in Fig. 6, GCSPP II achieves higher cost and efficiency scores in both navigation environments, benefiting from optimized cargo arrangement and reduced space waste. Table 15; Table 14.

By inputting the corresponding RIFs information into the NBN structure, the risk probabilities of each severity level were calculated in Fig. 7 and Fig. 8.

By applying the integration approach to both environments, the overall risk value corresponding to the first navigation environment is:

$$P_1 = 1.96\% \times 2 + 2\% \times 4 + 6.73\% \times 6 + 89.3\% \times 8 = 7.67$$

Similarly, the overall risk value of the second navigation environment can be obtained as:

$$P_2 = 94.4\% \times 2 + 4.05\% \times 4 + 1.07\% \times 6 + 0.49\% \times 8 = 2.15$$

As a result, the risk value for the first navigation environment of 7.67 lies in the range of 6–8, namely between critical and catastrophic. The value of risk for the second navigation environment is 2.15, which lies in the range of 2–4, between minor and major. The ratio analysis is further applied for further analysis. By calculation, the first navigation environment has a catastrophic risk of 83.35% and a critical risk of 16.65%. The second navigation environment therefore has a major risk of 7.67% and a minor risk of 92.33%.

Since the criticality safety weights corresponding to catastrophic, critical, major and minor accidents are 80%, 70%, 60% and 50% respectively. Substituting the utility values corresponding to the risk levels involved in the two navigation environments, the actual safety weight corresponding to the first navigation environment is: $83.35\% \times 80\% + 16.65\% \times 70\% = 78.34\%$, and the actual safety weight corresponding to the second navigation environment is:$7.67\% \times 60\% + 92.33\% \times 50\% = 50.77\%$. Therefore, for three defined dimensions, i.e., safety, economy and efficiency, the respective weights in the first navigation environment are 78.34%, 14.44%, and 7.22%. The respective weights in the second navigation environment are 50.77%, 32.82%, and 16.41%. These weight values reflect the risk levels that a ship has to face when navigating through a special voyage. They naturally formulate the actual decision–making logic and priority preferences of shipping companies when selecting a GCSPP. In high–risk navigation environment, shipping companies will prioritize more on navigation safety over cost and efficiency in order to avoid serious accidents, casualties, and huge economic losses; therefore, a significantly higher safety weight is applied. In low–risk navigation environment, given that basic safety requirement is satisfied, shipping companies will pay more attention on cost and efficiency to reduce operating costs and enhance port handling efficiency. Consequently, the final score calculated using these weights can better guide shipping companies to choose the most suitable GCSPP that matches their real operational priorities under different navigation risk scenarios. After inputting the above weights, the safety scores of the two GCSPPs are determined, as shown in Fig. 9 and Fig. 10.

From Fig. 9, it can be seen that in extremely high–risk environments, the safety score of GCSPP I is significantly higher than that of GCSPP II. GCSPP I with better safety performance performs better in high–risk environments. From Fig. 10, it can be seen that the score of GCSPP II exceeds that of GCSPP I . This indicates that when general cargo ships are in low–risk navigation environments, they can prioritize the "efficiency oriented" GCSPP while balancing loading and unloading efficiency and economic costs.

The score reversal in Figs. 9 and 10 demonstrates the significant impact of dynamic environment on GCSPP selection. If GCSPP is

Table 12
Bottom data.

The bottom–level criteria	Assessment of the GCSPP I	Assessment of the GCSPP II
Static stability	0% bad, 0% average, 100% good	0% bad, 0% average, 100% good
AGZ	0% bad, 40% average, 60% good	0% bad, 20% average, 80% good
GZ30°	81.25% 0.2 m,18.75% 0.4 m,0% 0.6m	0% 0.2 m, 77.5% 0.4 m, 22.5% 0.6m
θGZMAX	0% 25°, 30% 35°, 70% 45°	55.79% 25°,44.21% 35°, 0% 45°
Damaged stability	0% 1.2 m, 50% 2.7 m, 50% 3.2m	0% 1.2 m, 80% 2.7 m,20% 3.2m
Free–surface effect	0% 1.1 m, 48% 0.6 m, 52% 0.1m	0% 1.1 m, 48% 0.6 m, 52% 0.1m
SF	15% very big, 35% big, 35% moderate, 15% small, 0% very small	20% very big,35% big,35% moderate,10% small, 0% very small
BM	20% very big, 30% big, 40% moderate, 10% small, 0% very small	25% very big, 30% big,40% moderate, 5% small, 0% very small
Local strength	0% very weak, 0% weak, 50% moderate, 30% strong, 10% very strong, 10% unknown	0% very weak, 0% weak, 35% moderate, 45% strong, 10% very strong, 10% unknown
Ship overload	0% bad, 0% average, 100% good	0% bad, 0% average, 100% good
Heeling angle	0% 0.3°, 0% 0.1°, 100% 0°	0% 0.3°, 0% 0.1°, 100% 0°
Avoidance of mixed cargo	10% very bad, 20% bad, 30%average, 0% good, 0% very good, 40% unknown	0% very bad, 10% bad, 60% average, 20% good, 0% very good, 10% unknown
Mechanical properties	0% fragile, 30% firm, 40% very strong, 30% unknown	10% fragile, 30% firm, 30% very strong, 30% unknown
Dunnage of cargo	0% bad, 60% average, 10% good, 30% unknown	15% bad, 45% average, 0% good, 40% unknown
Broken space of cargo	20% very big, 35% big, 20% average, 0% small, 0% very small, 25% unknown	0% very big, 20% big, 30% average, 20% small, 10% very small, 20% unknown
Ship trim	0% 0 m, 37.5% 0.4 m, 62.5% 0.8 m, 0% 1.2 m, 0% 1.6m	0% 0 m,0% 0.4 m,45% 0.8 m,55% 1.2 m,0% 1.6m
Locations of cargos in the first arrival port	0% bad, 20% average, 80% good	0% bad, 15% average, 85% good
Locations of cargos in the final arrival port	0% bad, 20% average, 80% good	0% bad, 15% average, 85% good

Table 13
Cost and efficiency scoring of GCSPP I in the navigation environment I.

Cost/Efficiency	Very high	High	Average	Low	Very low
Expert1	0/0	0.3/0.1	0.2/0.2	0.3/0.3	0.2/0.4
Expert2	0.1/0.0	0.2/0.1	0.3/0.2	0.3/0.3	0.1/0.4
Expert3	0.0/0.0	0.2/0.1	0.4/0.2	0.3/0.3	0.1/0.4
Expert4	0.1/0.0	0.3/0.1	0.2/0.2	0.3/0.3	0.1/0.4
Expert5	0.0/0.0	0.2/0.1	0.3/0.2	0.4/0.3	0.1/0.4
Expert6	0.1/0.0	0.1/0.1	0.3/0.2	0.3/0.3	0.2/0.4
Expert7	0.0/0.0	0.3/0.1	0.2/0.2	0.3/0.3	0.2/0.4
Expert8	0.1/0.0	0.2/0.1	0.3/0.2	0.2/0.3	0.2/0.4
Expert9	0.0/0.0	0.2/0.1	0.4/0.2	0.3/0.3	0.1/0.4
Expert10	0.1/0.0	0.1/0.1	0.2/0.2	0.4/0.3	0.2/0.4

selected solely based on experience, it is easy to encounter safety hazards in high–risk environments or waste efficiency and cost in low–risk environments.

Finally, the safety, cost and efficiency scores are integrated via the following equation.

$$FS = \omega_s \times F_s + \omega_c \times F_c + \omega_e \times F_e \quad (7)$$

where FS is the final score of GCSPP; ω_s, ω_c and ω_e are the weights of safety, cost and efficiency, respectively; F_s, F_c and F_e are the corresponding safety score, cost score and efficiency score of the GCSPP.

Table 14
Cost and efficiency scoring of GCSPP II in the navigation environment I.

Cost/Efficiency	Very high	High	Average	Low	Very low
Expert1	0.0/0.0	0.2/0.2	0.4/0.3	0.2/0.3	0.2/0.2
Expert2	0.0/0.0	0.1/0.1	0.2/0.2	0.4/0.3	0.3/0.4
Expert3	0.0/0.0	0.1/0.1	0.3/0.2	0.4/0.3	0.2/0.4
Expert4	0.0/0.0	0.2/0.2	0.3/0.3	0.4/0.3	0.1/0.2
Expert5	0.0/0.0	0.1/0.1	0.4/0.2	0.3/0.3	0.2/0.4
Expert6	0.0/0.0	0.1/0.1	0.2/0.2	0.5/0.3	0.2/0.4
Expert7	0.0/0.0	0.2/0.2	0.3/0.3	0.3/0.3	0.2/0.2
Expert8	0.0/0.0	0.1/0.1	0.2/0.2	0.5/0.3	0.2/0.4
Expert9	0.0/0.0	0.2/0.2	0.3/0.3	0.4/0.3	0.1/0.2
Expert10	0.0/0.0	0.1/0.1	0.3/0.2	0.4/0.3	0.2/0.4

Table 15
Cost and efficiency scoring of GCSPP I in the navigation environment II.

Cost/Efficiency	Very high	High	Average	Low	Very low
Expert1	0.1/0.0	0.1/0.1	0.2/0.2	0.3/0.3	0.3/0.4
Expert2	0.0/0.0	0.1/0.1	0.3/0.2	0.3/0.3	0.3/0.4
Expert3	0.0/0.0	0.1/0.1	0.2/0.2	0.4/0.3	0.3/0.4
Expert4	0.0/0.0	0.2/0.1	0.3/0.2	0.3/0.3	0.2/0.4
Expert5	0.1/0.0	0.1/0.1	0.2/0.2	0.4/0.3	0.2/0.4
Expert6	0.0/0.0	0.1/0.1	0.2/0.2	0.4/0.3	0.3/0.4
Expert7	0.1/0.0	0.1/0.1	0.2/0.2	0.3/0.3	0.3/0.4
Expert8	0.0/0.0	0.1/0.1	0.3/0.2	0.4/0.3	0.2/0.4
Expert9	0.0/0.0	0.2/0.1	0.2/0.2	0.4/0.3	0.2/0.4
Expert10	0.1/0.0	0.1/0.1	0.2/0.2	0.3/0.3	0.3/0.4

Table 16
Cost and efficiency scoring of GCSPP II in the navigation environment II.

Cost/Efficiency	Very high	High	Average	Low	Very low
Expert1	0.0/0.0	0.1/0.2	0.4/0.3	0.3/0.3	0.2/0.2
Expert2	0.0/0.0	0.2/0.2	0.2/0.2	0.3/0.3	0.3/0.4
Expert3	0.0/0.0	0.1/0.1	0.3/0.2	0.3/0.3	0.3/0.4
Expert4	0.1/0.0	0.1/0.1	0.3/0.2	0.3/0.3	0.2/0.4
Expert5	0.0/0.0	0.1/0.1	0.3/0.2	0.3/0.3	0.3/0.4
Expert6	0.1/0.0	0.1/0.1	0.3/0.2	0.3/0.3	0.2/0.4
Expert7	0.0/0.0	0.1/0.2	0.4/0.3	0.3/0.3	0.2/0.2
Expert8	0.0/0.0	0.2/0.2	0.2/0.2	0.3/0.3	0.3/0.4
Expert9	0.0/0.0	0.1/0.1	0.4/0.2	0.3/0.3	0.2/0.4
Expert10	0.1/0.0	0.1/0.1	0.3/0.2	0.3/0.3	0.2/0.4

To investigate the impact of three weight parameters on the evaluation results, a comparison analysis is delivered by setting $(\omega_c + \omega_e)/\omega_s$ from 0.1 to 10 and satisfying $\omega_s + \omega_c + \omega_e = 1$. The results are plotted as a line graph as shown in Fig. 11 and Fig. 12. The horizontal coordinate is the ratio of combined cost and efficiency weights to safety weights, and the vertical coordinate is the final score of the GCSPP.

Fig. 11 shows the performance of two GCSPPs in high–risk navigation environment I. When $(\omega_c + \omega_e)/\omega_s \leq 1.4$, it represents that shipping companies prioritize safety over cost and efficiency in high–risk navigation environments, which is a common practice in shipping decision–making. At this point, the score of GCSPP I is higher. When $(\omega_c + \omega_e)/\omega_s > 1.4$, it represents that shipping companies place extreme emphasis on cost and efficiency while neglect safety in high–risk navigation environments, which does not exist in real decision–making. Only in such extreme cases will the GCSPP II final score exceeds GCSPP I .

Fig. 12 corresponds to the performance of two GCSPPs in low–risk navigation environment II, showing completely different characteristics from high–risk navigation environment. When$(\omega_c + \omega_e)/\omega_s \leq 1.1$, it represents that shipping companies attach equal importance to safety versus cost and efficiency in low–risk navigation environment, and GCSPP II leads GCSPP I by a slight advantage. When $(\omega_c + \omega_e)/\omega_s > 1.1$, this represents the increasing importance that shipping companies attach to cost and efficiency in low–risk navigation environments,

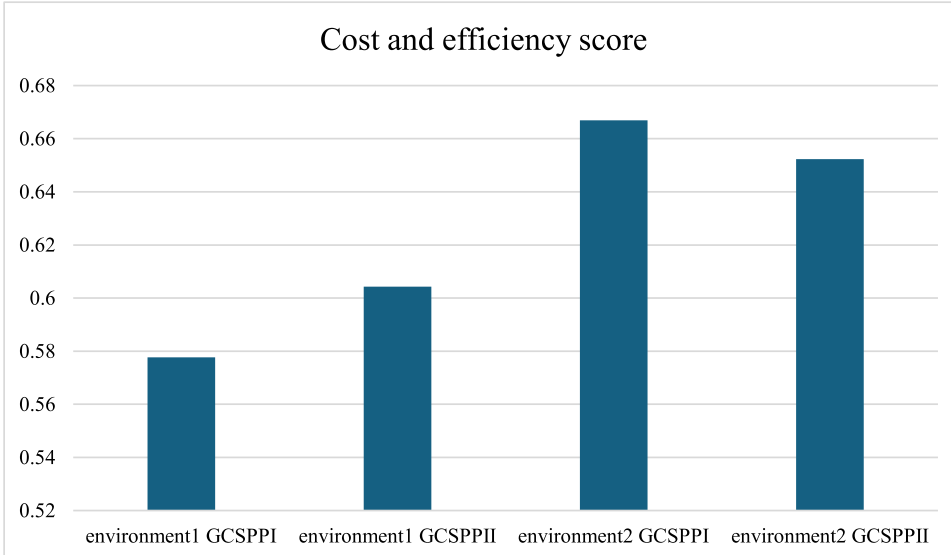

Fig. 6. Cost and efficiency scores for both GCSPPs in different navigation environments.

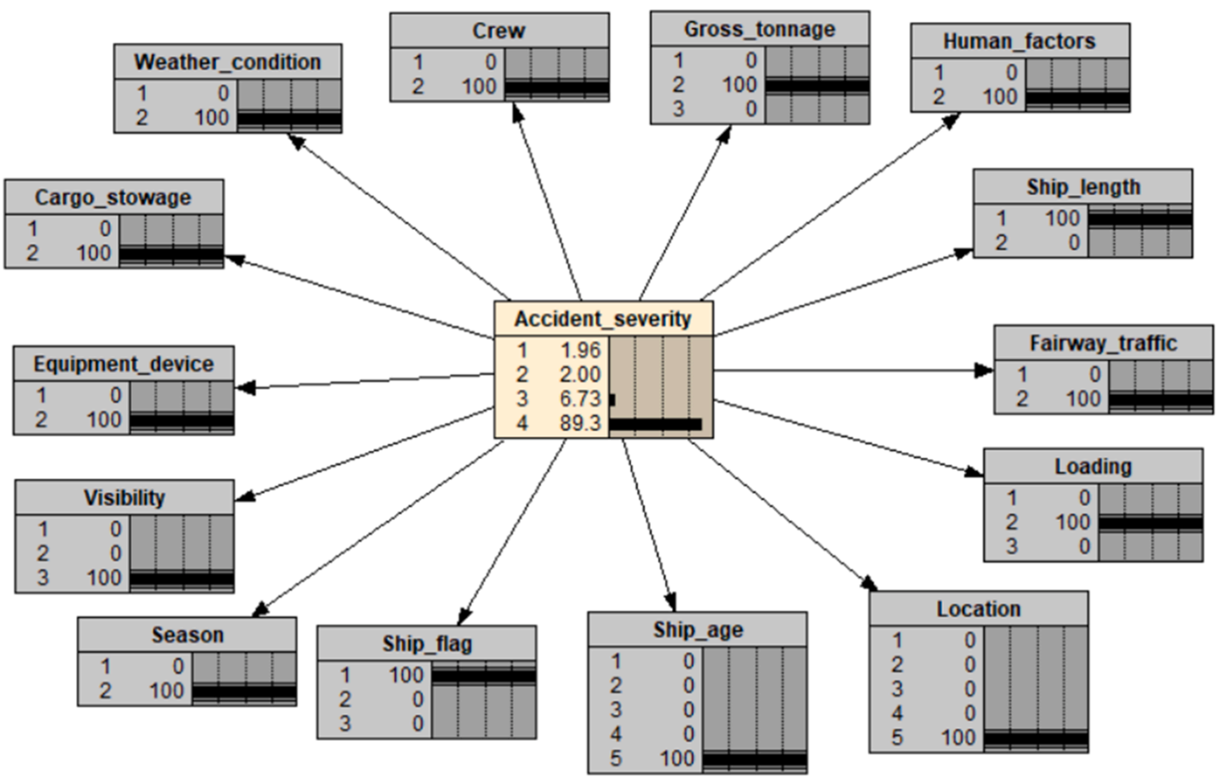

Fig. 7. Navigation environment I.

resulting in a GCSPP I final score exceeding GCSPP II and the leading advantage continuing to expand.

The completely different features in the two figures demonstrate that the optimal results of GCSPP have significant 'weight dependence'. In high–risk environments, higher safety weights result in GCSPP I being the choice of shipping companies. In low–risk environments, the more suitable GCSPP should be dynamically selected based on the actual weight allocation. Based on this conclusion, shipping companies can set the weight allocation range of safety, cost and efficiency according to the specific risk level of the navigation environment, and obtain GCSPP with better performance.

5. Conclusion

Existing research on maritime risks mostly focuses on "macro accident analysis" of specific risk types, but has long overlooked the important role of dynamic navigation environment and cargo stowage factors in GCSPP risk assessment. This study fills this gap and makes the research more closely related to the practical needs of general cargo ship operations.

In this study, the integration of NBN and ER techniques is utilized to evaluate the GCSPP on board. By meticulously analyzing the accident reports of general cargo ships in Tianjin waters from 2007 to 2021, a

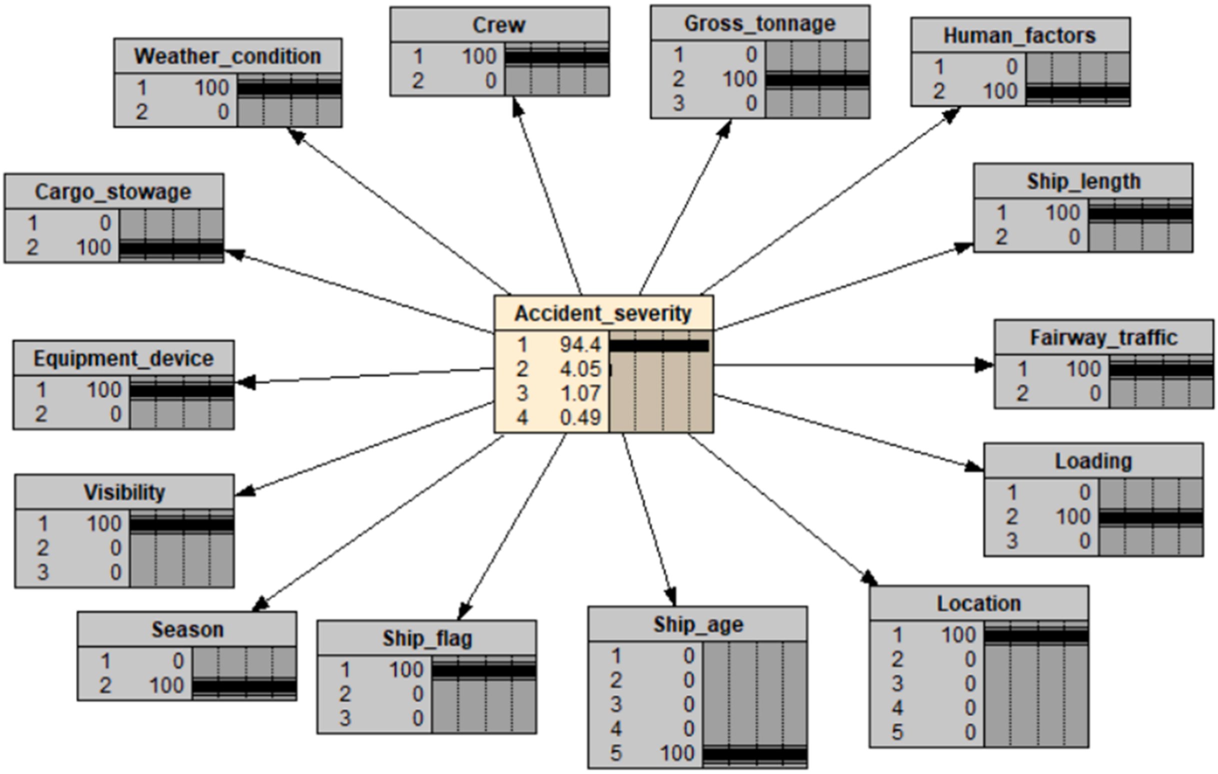

Fig. 8. Navigation environment II.

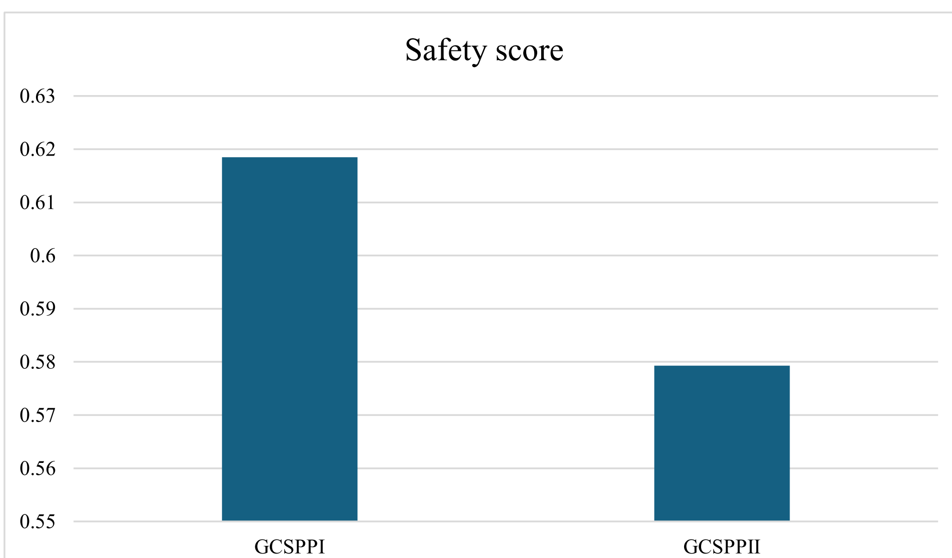

Fig. 9. Evaluation of two GCSPPs in the high–risk navigation environment.

comprehensive database that has allowed us to pinpoint 14 significant RIFs is established. Using this data, an NBN structure has been developed to conduct an in–depth risk analysis of general cargo ships navigating in Tianjin waters. Subsequently, by integrating the ER method, the combined framework achieves comprehensive assessment of GCSPPs.

Unlike prior studies in the field, this study brings the risk analysis of marine traffic accidents involving general cargo ships into cargo stowage planning. Novel factors such as Loading and Cargo stowage are incorporated, which are unique to general cargo ships and essential for GCSPP evaluation. The risk values obtained from the NBN structure are fed into the evaluation system, allowing us to quantitatively determine relatively reliable GCSPPs in different navigation environments. By further applying the ER method to integrate the expert assessments of GCSPP costs and efficiency, the safety scores, cost and efficiency scores of the GCSPP can be integrated into a final score. The findings reveal that when analyzed from safety, cost and efficiency perspectives, the prioritization of different GCSPPs may vary for the same ship under different navigation conditions.

Given this, the real–world implications of this research are far–reaching. For shipping operators, the proposed framework offers a robust decision–making tool that can optimize cargo pre–stowage

Fig. 10. Evaluation of two GCSPPs in the low–risk navigation environment.

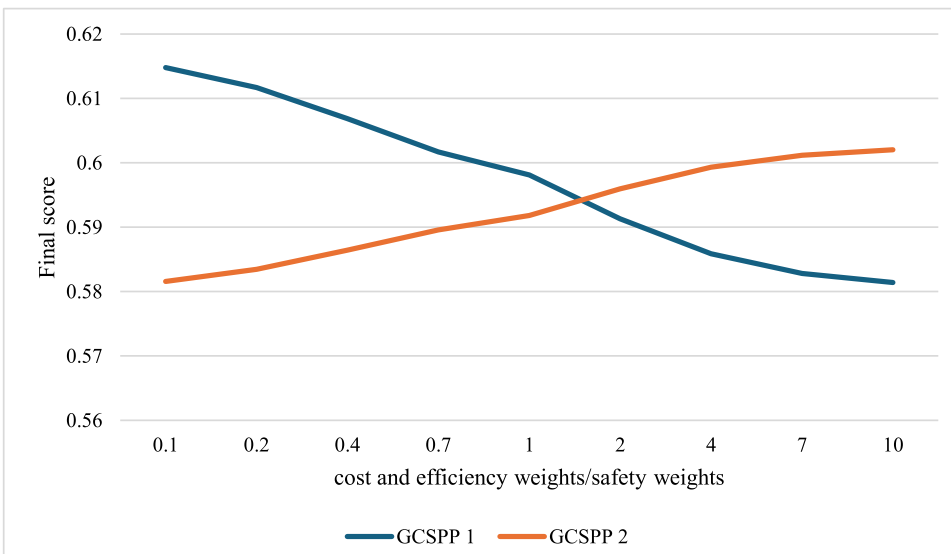

Fig. 11. Comparison of two GCSPPs in the navigation environment I.

planning, thereby reducing the risk of accidents and associated costs. By incorporating safety, cost and efficiency considerations into the GCSPP evaluation process, shipping companies can enhance the overall efficiency and competitiveness of their operations. In addition, maritime safety regulators can also benefit from this study. The framework provides a systematic approach to monitoring and improving the safety of general cargo ship operations. By adopting this framework, regulators can establish more effective safety standards and guidelines for cargo pre–stowage planning. Technically, the integration of NBN and ER techniques in this study showcases a novel methodological approach that can be applied to other complex decision–making problems in the maritime industry and beyond. This study serves as a valuable example of how advanced computational techniques can be applied to enhance safety and operational efficiency in practice. The study results and performance comparison are shown in Table 17 below.

Looking to the future, there are several directions to enhance the current study. To strengthen the benchmark for assessing GCSPP, more cutting–edge data can be applied to provide fresh insights from the industry. Additionally, employing more data–driven techniques such as AI and Digital Twin can further enhance the objectivity and universal acceptance, potentially improving the safety and cargo management on board general cargo ships.

CRediT authorship contribution statement

Daihui Zhang: Writing – review & editing, Writing – original draft, Data curation, Conceptualization. **Zheng Qiao:** Writing – original draft, Resources, Methodology, Data curation, Conceptualization. **Wenxin Wang:** Validation, Resources, Investigation, Data curation. **Ge Zhang:** Methodology, Conceptualization. **Zhuohua Qu:** Writing – review & editing. **Zaili Yang:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

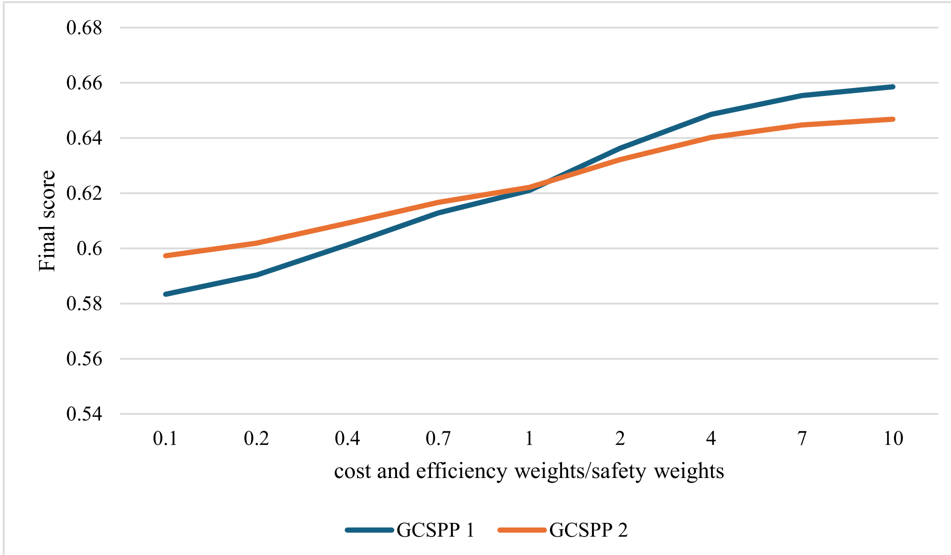

Fig. 12. Comparison of two GCSPPs in the navigation environment II.

Table 17
Study results and performance comparison.

literature	RIFs	Factor Nature	Methodology	Target Object	Approach Accuracy
Current paper	14	Objective + Subjective	NBN+ER	GCSPP assessment	Explicitly quantified
[61]	24	Objective	TAN	Global maritime accident casualty analysis	Explicitly quantified
[25]	Not specified	Subjective	FBN	Bulk carrier cargo hold/deck fire & explosion risks	Sensitivity analysis
[55]	20	Objective + Subjective	DEMATEL+ISM +FBN	China maritime dangerous goods transport accidents	Quantified
[62]	Not specified	Objective + Subjective	BN	Merchant ship piracy/robbery in Western Indian Ocean/East Africa	Quantified
[7]	23	Objective + Subjective	TAN	Global maritime accidents	Explicitly quantified
[27]	30	Objective + Subjective	TAN	Maritime accident severity	Explicitly quantified

Methodology: (NBN: Bayesian Network;TAN:Tree–Augmented Naive Bayes; FBN: Fuzzy Bayesian Network; NBN: Naive Bayesian Network; ISM: Interpretative Structural Modeling Method; DEMATEL: Decision–Making Trial and Evaluation Laboratory).

Data availability

Research Link Provided

References

[1] Jiang M, Lu J, Yang Z, Li J. Risk analysis of maritime accidents along the main route of the Maritime Silk Road: a Bayesian network approach. Marit Policy Manag 2020;47(6):815–32.
[2] Nguyen S, Shu–Ling Chen P, Du Y. Risk assessment of maritime container shipping blockchain–integrated systems: an analysis of multi–event scenarios. Transp Res E: Logist Transp Rev 2022;163(May):102764.
[3] Notteboom TE. Container shipping and ports: an overview. Rev Netw Econ 2009;3 (2):86–106.
[4] Zhang G, Thai VV. Expert elicitation and bayesian network modeling for shipping accidents: a literature review. Saf Sci 2016;87:53–62.
[5] Khan RU, Yin J, Mustafa FS, Liu H. Risk assessment and decision support for sustainable traffic safety in Hong Kong waters. IEEE Access 2020;8:72893–909.
[6] Fan S, Yang Z, Blanco-Davis E, Zhang J, Yan X. Analysis of maritime transport accidents using Bayesian networks. Proc Inst Mech Eng O: J Risk Reliab 2020;234 (3):439–54.
[7] Li H, Ren X, Yang Z. Data-driven Bayesian network for risk analysis of global maritime accidents. Reliab Eng Syst Saf 2023;230:108938.
[8] Wang L, Yang Z. Bayesian network modelling and analysis of accident severity in waterborne transportation: a case study in China. Reliab Eng Syst Saf 2018;180 (July):277–89.
[9] Shafer G. A mathematical theory of evidence. Princeton, NJ: Princeton University Press; 1976.
[10] Gao Q, Xu DL. An empirical study on the application of the Evidential Reasoning rule to decision making in financial investment. Knowl Based Syst 2019;164: 226–34.
[11] Sadeghi A, Farhad H, Mohammadzadeh Moghaddam A, Jalili Qazizadeh M. Identification of accident–prone sections in roadways with incomplete and uncertain inspection–based information: a distributed hazard index based on evidential reasoning approach. Reliab Eng Syst Saf 2018;178(November 2017): 278–89.
[12] Yu Q, Teixeira ÂP, Liu K, Rong H, Guedes Soares C. An integrated dynamic ship risk model based on Bayesian Networks and Evidential Reasoning. Reliab Eng Syst Saf 2021;216:106981.
[13] Zhang D, Yan X, Zhang J, Yang Z, Wang J. Use of fuzzy rule–based evidential reasoning approach in the navigational risk assessment of inland waterway transportation systems. Saf Sci 2016;82:352–60.
[14] Yang Z, Wang J. Use of fuzzy risk assessment in FMEA of offshore engineering systems. Ocean Eng 2015;95:195–204.
[15] Wu B, Zong L, Yan X, Guedes Soares C. Incorporating evidential reasoning and TOPSIS into group decision–making under uncertainty for handling ship without command. Ocean Eng 2018;164(April):590–603.
[16] Dong X, Zhao C, Wu B, Guedes Soares C. A fuzzy evidential reasoning–based approach for assessing the navigation obstruction risk of shipwrecks in the Yangtze river. Marit Policy Manag 2024;00(00):1–18.
[17] Wu B, Wang Y, Zhang J, Savan EE, Yan X. Effectiveness of maritime safety control in different navigation zones using a spatial sequential DEA model: yangtze River case. Accid Anal Prev 2015;81:232–42.
[18] Yang Z, Yang Z, Yin J. Realising advanced risk–based port state control inspection using data–driven bayesian networks. Transp Res A: Policy Pract 2018;110 (November 2017):38–56.
[19] Yu Q, Liu K, Yang Z, Wang H. Realising advanced risk assessment of vessel traffic flows near offshore wind farms. Reliab Eng Syst Saf 2020;203:106667.
[20] Xu S, Kim E, Haugen S, Zhang M. A Bayesian network risk model for predicting ship besetting in ice during convoy operations along the Northern Sea Route. Reliab Eng Syst Saf 2022;223(March):108475.
[21] Cao Y, Wang X, Wang Y, Fan S, Wang H, Yang Z, Liu Z, Wang J, Shi R. Analysis of factors affecting the severity of marine accidents using a data–driven bayesian network. Ocean Eng 2023;269(January):113563.
[22] Uflaz E, Sezer SI, Tunçel AL, Aydin M, Akyuz E, Arslan O. Quantifying potential cyber–attack risks in maritime transportation under Dempster–shafer theory FMECA and rule–based bayesian network modelling. Reliab Eng Syst Saf 2024;243 (September 2023):109825.
[23] Li P, Wang Y, Yang Z. Risk assessment of maritime autonomous surface ships collisions using an FTA–FBN model. Ocean Eng 2024;309(P2):118444.

[24] Fan H, Lyu J, Wang S, Gong X. Leverage rule–based Bayesian network on assessing straits/canals resilience performance on a risk coupling analysis perspective. Marit Policy Manag 2024;00(00):1–19.
[25] Aydin M, Akyuz E, Boustras G. A holistic safety assessment for cargo holds and decks fire & explosion risks under fuzzy bayesian network approach. Saf Sci 2024;176(May):106555.
[26] Li W, Henke M, Pundt R, Miller–Hooks E. A data–driven Bayesian network methodology for predicting future incident risk in Arctic maritime–based cargo transit. Ocean Eng 2025;320(September 2024):120299.
[27] Fan H, Wang J, Chang Z, Lyu J, Jia H. Embracing imperfect data: a novel data–driven bayesian network framework for maritime accidents severity risk assessment. Ocean Eng 2025;329(March):121212.
[28] Wang J, Fan H, Chang Z, Lyu J. Unleashing data power : driving maritime risk analysis with bayesian networks. Reliab Eng Syst Saf 2025;264(PA):111310.
[29] Aydin M, Ilke S, Akyuz E, Gardoni P. Improved Z–number based bayesian network modelling to predict cyber–attack risk for maritime autonomous surface ship (MASS). Appl Soft Comput 2025;180(March 2024):113416.
[30] Jovanović I, Perčić M, Vladimir N. Assessment of human contribution to cargo ship accidents using Fault Tree Analysis and Bayesian Network Analysis. Ocean Eng 2025;323:109275.
[31] Shen Y, Zhao N, Xia M, Du X. A deep Q–learning network for ship stowage planning problem. Pol Marit Res 2017;24(S3):102–9.
[32] Cho JH, Ku NK. Developing a container ship loading–planning program using reinforcement learning. J Mar Sci Eng 2024;12(10):1832.
[33] Rodero F, Marrero Á. Fire hazard reduction in ro–ro spaces by means of using a cargo distribution algorithm including fire risk management during the stowage process. Transp Res Procedia 2023;71:331–8.
[34] Zhou S, Guo Z, Chen J, Jiang G. Large containership stowage planning for maritime logistics: a novel meta–heuristic algorithm to reduce the number of shifts. Adv Eng Inform 2025;64(November 2024):102962.
[35] Sivertsen A, Reinhardt L, Jensen RM. A representative model and benchmark suite for the container stowage planning problem. Transp Res E: Logist Transp Rev 2025;201(November 2024):104224.
[36] Pham TY, Nguyen PN. A data–driven framework for predicting ship berthing time and optimizing port operations at Tan Cang Cat Lai Port. Vietnam Case Stud Transp Policy 2025;20(December 2024):101441.
[37] Pérez-Canosa JM, Orosa JA, de Osés FXM, Lama-Carballo FJ. Optimization of ship's navigational parameters to improve the stowage and securing criteria of non-standardized cargo in ships. J Mar Sci Eng 2023;11(9):1241.
[38] Ding JF, Tseng WJ, Sung YJ. An evaluation of operational risks for general cargo ship operators. Brodogradnja 2024;75(1):1–19.
[39] Zhang D, Qu Z, Wang W, Yu J, Yang Z. New uncertainty modelling for cargo stowage plans of general cargo ships. Transp Res E: Logist Transp Rev 2020;144 (November 2018):102151.
[40] Xu DL, Yang JB. Intelligent decision system based on the evidential reasoning approach and its applications. J Telecommun Inf Technol 2005;3:73–80.
[41] Hänninen M, Kujala P. Influences of variables on ship collision probability in a bayesian belief network model. Reliab Eng Syst Saf 2012;102:27–40.
[42] Mullai A, Paulsson U. A grounded theory model for analysis of marine accidents. Accid Anal Prev 2011;43(4):1590–603.
[43] Yang ZL, Bonsall S, Wall A, Wang J, Usman M. A modified CREAM to human reliability quantification in marine engineering. Ocean Eng 2013;58:293–303.
[44] Balmat JF, Lafont F, Maifret R, Pessel N. MAritime RISk Assessment (MARISA), a fuzzy approach to define an individual ship risk factor. Ocean Eng 2009;36(15–16):1278–86.
[45] Hansen HL, Jepsen JR, Hermansen K. Factors influencing survival in case of shipwreck and other maritime disasters in the Danish merchant fleet since 1970. Saf Sci 2012;50(7):1589–93.
[46] Knapp S, Bijwaard G, Heij C. Estimated incident cost savings in shipping due to inspections. Accid Anal Prev 2011;43(4):1532–9.
[47] Antão P, Guedes Soares C. Causal factors in accidents of high–speed craft and conventional ocean–going vessels. Reliab Eng Syst Saf 2008;93(9):1292–304.
[48] Chen ST, Wall A, Davies P, Yang Z, Wang J, Chou YH. A Human and Organisational factors (HOFs) analysis method for marine casualties using HFACS–Maritime Accidents (HFACS–MA). Saf Sci 2013;60:105–14.
[49] Sotiralis P, Ventikos NP, Hamann R, Golyshev P, Teixeira AP. Incorporation of human factors into ship collision risk models focusing on human centred design aspects. Reliab Eng Syst Saf 2016;156:210–27.
[50] Soares CG, Teixeira AP. Risk assessment in maritime transportation. Reliab Eng Syst Saf 2001;74(3):299–309.
[51] Akyuz E, Celik M. A hybrid decision–making approach to measure effectiveness of safety management system implementations on–board ships. Saf Sci 2014;68:169–79.
[52] Sun X, Yan X, Wu B, Song X. Analysis of the operational energy efficiency for inland river ships. Transp Res D: Transp Environ 2013;22:34–9.
[53] Weng J, Yang D. Investigation of shipping accident injury severity and mortality. Accid Anal Prev 2015;76:92–101.
[54] Zhang D, Yan XP, Yang ZL, Wall A, Wang J. Incorporation of formal safety assessment and Bayesian network in navigational risk estimation of the Yangtze River. Reliab Eng Syst Saf 2013;118:93–105.
[55] Ma L, Ma X, Lan H, Liu Y, Deng W. A methodology to assess the interrelationships between contributory factors to maritime transport accidents of dangerous goods in China. Ocean Eng 2022;266(P3):112769.
[56] Yang Z, Yang Z, Yin J, Qu Z. A risk–based game model for rational inspections in port state control. Transp Res E: Logist Transp Rev 2018;118(April):477–95.
[57] Alyami H, Yang Z, Riahi R, Bonsall S, Wang J. Advanced uncertainty modelling for container port risk analysis. Accid Anal Prev 2019;123:411–21.
[58] Yang Z, Bonsall S, Wang J. Fuzzy rule–based Bayesian reasoning approach for prioritization of failures in FMEA. IEEE Trans Reliab 2008;57(3):517–28.
[59] Zhang X, Liu CA. Model averaging prediction by K–fold cross–validation. J Econom 2023;235(1):280–301.
[60] Yu Q, Liu K, Chang CH, Yang Z. Realising advanced risk assessment of vessel traffic flows near offshore wind farms. Reliab Eng Syst Saf 2020;203(October 2019):107086.
[61] Zhou K, Xing W, Wang J, Li H, Yang Z. A data-driven risk model for maritime casualty analysis: A global perspective. Reliab Eng Syst Saf 2024;244:109925.
[62] Pristrom S, Yang Z, Wang J, Yan XP. A novel flexible model for piracy and robbery assessment of merchant ship operations. Reliab Eng Syst Saf 2016;155:196–211.