Xin, X, Yang, Z, Liu, K, Zhang, J and Wu, X

Multi-stage and multi-topology analysis of ship traffic complexity for probabilistic collision detection

Multi-stage and multi-topology analysis of ship traffic complexity for probabilistic collision detection

Xuri Xin[a,b], Zaili Yang[b], Kezhong Liu[a,c,d*], Jinfen Zhang[d,f], Xiaolie Wu[a]

[a] *School of Navigation, Wuhan University of Technology, Wuhan, Hubei, China*

[b] *Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool, UK*

[c] *Hubei Key Laboratory of Inland Shipping Technology, Wuhan, Hubei, China*

[d] *National Engineering Research Center for Water Transport Safety, Wuhan University of Technology, Wuhan, Hubei, China*

[f] *Intelligent Transportation System Center (ITSC), Wuhan University of Technology, Wuhan, Hubei, China*

* Corresponding author.

E-mail address: kzliu@whut.edu.cn (K. Liu), xrxin@whut.edu.cn (X. Xin), Z.Yang@ljmu.ac.uk (Z. Yang), jinfen.zhang@whut.edu.cn (J. Zhang), 657410102@qq.com (X. Wu)

Abstract: Maritime traffic situational awareness plays a vital role in the development of intelligent transportation-support systems. This state-of-the-art study focuses on near-miss collision risk between/among ships but reveals challenges in estimating large-scale traffic situations associated with dynamic and uncertain ship motions at a regional level. This study develops a systematic methodology to evaluate ship traffic complexity to comprehend the traffic situation in complex waters. In the new methodology, the topological and evolutionary characteristics of ship traffic networks and the uncertainty in ship movements are considered simultaneously to realise probabilistic collision detection. The methodology, through the effective integration of probabilistic conflict estimation and traffic complexity modelling and assessment, enables the evaluation of traffic complexity in a fine-grained hierarchical manner. With the AIS-based trajectory data collected from the world's largest port (i.e. Ningbo-Zhoushan Port), a thorough validation of the evaluation performance is conducted and demonstrated through scenario analysis and model robustness. Moreover, some critical research results are obtained in terms of traffic network heterogeneity analysis; statistics including occurrence frequency, temporal distribution, life cycle, and transition probability of traffic complexity patterns; and correlation examination between the number of ships and traffic complexity patterns. These findings offer new insights into improving maritime traffic awareness capabilities and promoting maritime traffic safety management.
Keywords: Maritime risk, Ship traffic complexity, Spatial-temporal traffic dynamics, Network topology, Situation evolution, AIS data.

1 Introduction

Maritime traffic safety has always been one of the most crucial research areas because failure to address it will cause catastrophic maritime accidents involving intolerable consequences such as life loss, property damage, and/or ocean pollution. Economic globalisation over the past decades has led to the rapid growth of shipping traffic volumes, making maritime traffic heavier, riskier, and more sophisticated (H. Yu, Fang, Murray, & Peng, 2019; H. Yu et al., 2021; W. Zhang, Feng, Goerlandt, & Liu, 2020), particularly in hotspots such as restricted waters. It is evident by the recent Suez Canal blockage (Lebedev, Lebedeva, & Butsanets, 2021). This has put tremendous pressure on practical monitoring tasks for maritime surveillance operators, particularly those who supervise ship traffic in waterways and port areas due to the high volume of traffic and complicated maritime transportation dynamics they normally encounter. In practice, surveillance operators need to ensure ship traffic safety in a navigating situation, generate and send an earlier warning alert when detecting collision risk, and assist navigators in path re-planning for conflict resolution. Although there is a significant appeal in assisting operators and controllers to monitor and regulate maritime traffic dynamics, existing technologies and systems, such as but not limited to vessel traffic service (VTS) systems, maritime traffic service networks, land radar, and intelligent traffic signalling systems (ITSSs), still have some limitations in their practical applications due to the increasing ship spatial-temporal movement uncertainty over the last decade. Advanced techniques that enable controllers to better comprehend traffic situations and strengthen maritime traffic situational awareness are therefore in high demand. This means there are increasing requirements for new modelling capability for perceiving forthcoming traffic situations and making decisions to proactively manage ship collision risks.

It is essential to conduct the quantitative analysis of a maritime traffic situation for proactive maritime traffic awareness and hence maritime traffic safety (Bukhari, Tusseyeva, & Kim, 2013; Chen, Huang, Mou, & van Gelder, 2019; M. Zhang, Kujala, & Hirdaris, 2022). The rapid development of the Automatic Identification System (AIS) has contributed to the accessibility of rich information on ship trajectories for guiding and improving vessel traffic situation assessments (R. W. Liu, Liang, Nie, Lim, et al., 2022; R. W. Liu, Liang, Nie, Yuan, et al., 2022). With improvements in data acquisition, storage, and processing, an increasing number of practical and advanced applications of AIS data have been developed to tackle various issues, such as maritime anomaly detection (Iphar, Ray, & Napoli, 2020; Rong, Teixeira, & Soares, 2019, 2022), marine traffic movement characterisation (Xin, Liu, Yang, Yuan, & Zhang, 2019; Q. Yu, Liu, Teixeira, & Soares, 2020), spatial-temporal ship traffic statistical analysis (Rong, Teixeira, & Soares, 2021; L. Zhang, Meng, & Fwa, 2019), and trajectory behaviour modelling and recognition (H. Li et al., 2022, 2020; Liang et al., 2021). Using AIS data-based maritime traffic analytics, it is possible to capture the characteristics of ship traffic complexity. Relevant studies have recently become popular (e.g., Sui et al., 2020; Y. Wen et al., 2015; M. Zhang, Zhang, Fu, Kujala, & Hirdaris, 2022). Ship traffic complexity analysis serves as a tool for traffic controllers to comprehend maritime traffic situations and gain traffic situation awareness, offering new insights on promoting maritime traffic safety management and supporting traffic situation warnings. However, it remains challenging to effectively integrate AIS-based trajectory information into real-time ship traffic situation estimation and evaluation. Specifically, an advanced model targeting a full understanding of the traffic situation requires the holistic integrity of multiple analyses of the spatial-temporal dynamics and uncertainty in ship movements; the collision risk-related interactions among ships; and the structural and evolutionary characteristics associated with the spatial-temporal proximity relationships among ships. Therefore, establishing a reliable model for ship traffic situation assessment and management is challenging. However, its success, defined as the condition for developing an intelligent transportation support system for autonomous ships, needs to be addressed urgently.

In this study, a systematic and holistic ship traffic complexity evaluation methodology that is adaptive to the complex waters involving heavy and dynamic traffic with high uncertainty is developed to realise real-time maritime traffic situation comprehension and advanced traffic awareness. Its success will aid in enhancing maritime traffic monitoring capabilities and assist in decision-making for surveillance controllers by accounting for traffic behaviour characteristics, including the stochastic and variant conflict relations among ships, traffic congestion, and the difficulties brought about by traffic structure for conflict resolution. To achieve this, a new framework is developed by

1) A probabilistic conflict detection (CD) technique incorporating the effects of dynamics and uncertainty inherent to spatial-temporal vessel motions to quantify the conflict criticality among ships.
2) A ship traffic complexity model that extends the application of complex network theory to describe the topological and evolutionary characteristics of various inter-ship influence relations.
3) A fuzzy clustering iterative (FCI)-based complexity evaluation approach to support a fine-grained and hierarchical complexity pattern assessment.
4) A node importance identification method to determine the key risk ships.

Its new feature lies in the development of a holistic and systematic framework for ship traffic complexity evaluation by incorporating both the multi-stage and multi-topology characteristics of ship traffic networks and ship spatial-temporal movement uncertainty simultaneously. The main contributions of this study are summarised as follows.

1) The proposed probabilistic CD scheme characterises the conflict probability accurately and reliably by considering both the dynamic and uncertain characteristics of spatial-temporal vessel movements. Compared with traditional ship collision analysis approaches, the proposed scheme can handle various encounter scenarios in complex maritime traffic waters, such as busy ports, thereby facilitating vessel drivers and maritime safety authorities to undertake real-time and effective collision risk management.
2) A new multi-stage and multi-topology-based ship traffic complexity model is established to provide a full description of a traffic situation by quantifying the structural characteristics of multi-ship interactions and incorporating the evolutionary trend of their conflict relations. A spatial-temporal state matrix accounts for the conflict-based spatial structural measurement and dynamic evolution of conflict criticality over time,

making it possible to analyse the traffic situation accurately.

3) To evaluate the traffic complexity matrix, an FCI-based assessment approach is proposed to facilitate surveillance controllers in identifying the traffic situation. This approach is practical and configurable because of its capability to provide a clear hierarchical description of traffic complexity patterns and its feasibility in achieving the fusion of subjective and objective knowledge without being dependent on any pre-defined assessment standards.

The remainder of this study is organised as follows. Section 2 presents an overview of state-of-the-art research related to ship collision risk, traffic complexity, and multiple-index evaluation. In Section 3, the proposed ship traffic complexity methodology is introduced in detail, including the probabilistic CD, traffic complexity matrix construction, FCI-based complexity evaluation, and node importance identification. The application results and performance tests using real AIS data-based experiments are presented in Section 4. Conclusions and potential future explorations are provided in Section 5.

2 Literature review

This section presents a critical analysis of ship collision risk detection and estimation. A systematic overview of existing traffic complexity models used to quantify and estimate regional traffic situations is then presented to contextualise the primary research of this study. Finally, the merits and weaknesses of various multiple-index assessment methods that can be adopted for ship traffic complexity evaluation are discussed.

2.1 Maritime collision risk detection and estimation

Maritime collision risk is a popular research topic in the maritime sector. Many researchers have pioneered efforts to analyse and quantify ship collision probabilities and severity using different tools, definitions, and methodologies. For example, IALA Waterway Risk Assessment Program (IWRAP) is a useful tool recommended by International Association of Marine Aids to Navigation and Lighthouse Authorities (IALA) and is commonly used to calculate the collision probabilities of ship traffic in specific geographical waters (Kim, Park, & Park, 2011). In these studies, collision risk detection and estimation are among the hottest topics because they serve as prerequisites for triggering an event that assists in noticing potential collisions and performing conflict resolution. Meanwhile, various non-accident criticality measurement concepts, such as traffic conflict (Lei, 2020), near-miss (W. Zhang, Goerlandt, Kujala, & Wang, 2016; W. Zhang, Goerlandt, Montewka, & Kujala, 2015), and collision candidates (Chen, Huang, Mou, & van Gelder, 2018) have been proposed to detect and characterise potential dangerous encounter events. Relevant works in this field are categorised into three groups:

ship-domain-based methods, synthetic index methods, and dangerous region-based methods.

The ship domain refers to the safety zone around a vessel within which all other vessels remain clear unless authorised. It is employed to estimate collision risks and detect potential conflicts in terms of invasion or overlap of the safety zones of encountering ships. In recent years, advances in diversified intelligent technologies and increasing AIS data sources have contributed to the development of various ship domain models with different shapes (e.g. circular and elliptical (Fujii & Tanaka, 1971; Szlapczynski & Szlapczynska, 2016), polygonal (Y. Wang & Chin, 2016), quaternion (K. Liu, Yuan, Xin, Zhang, & Wang, 2021; N. Wang, 2013), and risk-based (L. Zhang & Meng, 2019) domains), methodologies (e.g. empirical, knowledge-based, and analytical domains (Szlapczynski & Szlapczynska, 2017; L. Zhang & Meng, 2019)), and factors considered (e.g. vessel attributes, vessel manoeuvrability, knowledge and condition of navigators, and environmental conditions (J. Liu, Feng, Li, Wang, & Wen, 2016; Y. Wang & Chin, 2016)). These models and methodologies work well in quantitatively examining candidates with collision potential and undesired consequences; identifying collision risk hotspots; and deriving relations between collision candidates and extra information (e.g., historical accident databases and accident investigation reports). Thus, in this study, a conflict criticality metric is designed based on ship domain models to characterise the conflicts among encountering ships. Nevertheless, it is not feasible to purely use a ship domain model to perform real-time CD in highly dynamic waters. Hence, a ship motion prediction model is integrated with it for a proper solution to address this limitation.

Synthetic index methods usually integrate the factors that characterise the spatial-temporal proximity of encountering ships for collision risk index (CRI) calculations by employing mathematical expressions or black-box models. The two most common indices, namely the Distance to Closest Point of Approach (DCPA) and Time to Closest Point of Approach (TCPA), were initially synthesised using techniques such as a binary state (i.e., safe/dangerous) (Q. Liu, Pedersen, Okazaki, Fukuto, & Tanaka, 2006), linear regression (Chin & Debnath, 2009), and fuzzy theory (H.-J. Lee & Rhee, 2001) to measure the CRI. In this context, previous researchers have also improved and extended the synthesis by considering more proximity factors (e.g., relative bearing, relative speed, ship manoeuvrability, ship motion patterns, and stability conditions (Fang, Yu, Ke, Shaw, & Peng, 2018; Gil, 2021; Gil, Montewka, Krata, Hinz, & Hirdaris, 2020; Öztürk, Boz, & Balcisoy, 2021; M. Zhang, Montewka, Manderbacka, Kujala, & Hirdaris, 2021; M. Zhang, Zhang, Fu, Yan, & Goncharov, 2017)), adopting advanced fusion techniques (e.g., Analytical Hierarchical Process (Y. Zhao, Li, & Shi, 2016), Multilayer Perceptron (Ahn, Rhee, & You, 2012), and Dempster-Shafer evidence theory (B. Li & Pang, 2013)), and ensuring their applicability for various encountering scenarios (Goerlandt, Montewka, Kuzmin, & Kujala, 2015). In particular, two other parameters, Bow Crossing Range (BCR) and the Time to Bow Crossing Range, have gained much attention owing to their effectiveness in supporting collision evaluations in crossing scenarios (e.g. Goerlandt et al., 2015; W. Zhang et al., 2015). For example, Gil, Kozioł, Wróbel, & Montewka (2022) have revealed that the combination of the BCR and CPA can cope with various encounter situations to achieve effective early collision warnings. Despite that, the critical analysis of these studies reveals that most of them adhere to a fundamental hypothesis that the ships will maintain a constant speed during the look-head period (Huang, Chen, Chen, Negenborn, & van Gelder, 2020; Xiao, Fu, Zhang, & Goh, 2019). Such a fundamental assumption may lead to inaccurate estimations of collision risk. More specifically, when ships perform manoeuvres (e.g., course or speed change) during the encountering process, these methods will offer unreliable estimation results. In addition, it is very challenging for them to provide an explicit explanation of the CRI results due to the interaction effects among different risk factors. Improvements are possible by exploiting the spatial-temporal dynamic features of ship movements that correlate with actual collisions in complicated encountering scenarios.

Dangerous region-based methods are dependent on the collection of the sets of ship's course or speed that will result in potential collision danger with other ships. A collision warning is then issued if the dynamic attributes of the ship fall into these sets. Classical solution approaches for dangerous region identification include Velocity Obstacle (VO) (Huang, van Gelder, & Wen, 2018), Projected Obstacle Area (Larson, Bruch, & Ebken, 2006), and Collision Threat Parameter Area (Lenart, 2015). Among them, VO has gained much popularity because of its simple operation and excellent performance in the search for collision-free schemes. Examples include integrating the ship domain with VO (Szlapczynski & Szlapczynska, 2015), developing Probabilistic VO and Generalised VO algorithms by loosening the linear ship motion assumption and taking ship manoeuvrability into consideration (Huang, Chen, & van Gelder, 2019; Huang et al., 2018), and extensions that consider the entire encountering process when detecting collision candidates (Chen et al., 2018). These studies demonstrate their strengths in detecting collision dangers in dynamic traffic situations, and can be applied in generalised scenarios. However, there are still unresolved issues from this perspective. Owing to the non-negligible computational burdens of mapping from the spatial-temporal proximity of the ship pairs to their velocity space, it is difficult for them to integrate the complex prohibit region models holistically. As a result, they are generally combined with simple risk measurement models, such as constant minimum safe distance and circular ship domain, for real-time collision detection. This has been deemed the main weakness of unveiling collision risk differences under different ship encounter situations.

In general, maritime collision risk estimation remains an active research field while simultaneously evolving towards a new stage where new advanced technologies aided by big AIS data can help improve its accuracy. The development of new technologies is justified by two reasons. First, few studies can simultaneously cope with the dynamic and uncertain characteristics of vessel motion. Most of them assumed that the encountering ships maintained constant velocity or that the ship's future trajectories could be accurately obtained when undertaking CD. However, it is challenging to precisely estimate a ship's future trajectories given the uncertainty caused by environmental disturbances, navigation plans or intentions, and physical and human factors. Consequently, these methods are not applicable to complex waters associated with high traffic density, variant manoeuvring behaviour, and unique geographical characteristics. The influence of trajectory uncertainty on collision risk estimation has attracted the attention of scholars in the field of air traffic by developing probabilistic models for CD (Hao, Zhang, Cheng, Liu, & Xing, 2018; C. H. J. Wang, Tan, & Low, 2019). Second, real-time collision risk assessment based on global/regional maritime traffic in busy water areas has rarely been investigated to date, thereby offering little insight into ship collision risk from a high-level traffic network perspective. It requires the incorporation of ship spatial-temporal dynamics, uncertainty, and collision risk interactions among regional maritime traffic.

2.2 Traffic complexity modelling and evaluation

Owing to increasing ship traffic movements and the emergence of large and fast vessels, maritime traffic has become more complex, resulting in growing concerns about global maritime traffic collision risk (Z. Liu, Wu, & Zheng, 2019; Zhen, Riveiro, & Jin, 2017). On this basis, some researchers started to attach importance to a new concept, called "ship traffic complexity" (Sui et al., 2020; Y. Wen et al., 2015; M. Zhang, Zhang, et al., 2022). It is a relatively new research topic in the maritime domain, with the work of (Y. Wen et al., 2015) who pioneered the quantitative assessment of traffic situations. However, in the aviation research field, air traffic complexity has been long-standing and applied for diverse purposes, including measuring the workload of traffic controllers (Cao et al., 2018), enhancing airspace capacity (Flener et al., 2007), assisting controllers in decision-making and conflict resolution (Radanovic, Eroles, Koca, & Gonzalez, 2018), and implementing traffic situation assessment (H. Wang, Song, Wen, & Zhao, 2016).

At an early stage, traffic density was treated as the basic feature for characterising traffic complexity

(Sridhar, Sheth, & Grabbe, 1998). Evidently, it is insufficient to capture the complete picture of complexity because many factors, including traffic flow characteristics and space structure associated with the overall traffic complexity (Cao et al., 2018) are not considered. For example, Fig. 1 illustrates two ship traffic scenarios with the same traffic density. The traffic situation in Fig. 1 (a) encounters more difficulties in controlling potential conflicts because many ship pairs will converge soon. In contrast, most ship pairs in Fig. 1 (b) show a dispersed trend and thus correspond to relatively low traffic complexity. Consequently, many types of complexity metrics, such as Interval Complexity (IC) (Flener et al., 2007) and Input–Output (IO) approach (Pallottino, Feron, & Bicchi, 2002), were formulated for traffic complexity assessment. Their critical design entails rationally screening and aggregating the complexity factors through a variety of combination means (Prandini, Piroddi, Puechmorel, & Brázdilová, 2011). To reflect between-aircraft influence relationships, some research communities have intensively studied complexity from the perspective of complex systems. They mathematically described the between-aircraft proximity (e.g. approaching effect and potential conflict) in terms of the traffic tracking information (e.g., location and velocity) to reveal the irregularity and disorder of the entire traffic situation using fractal dimension, topological entropy, and Lyapunov exponent (Delahaye & Puechmorel, 2010; Delahaye, Puechmorel, Hansman, & Histon, 2004; K. Lee, Feron, & Pritchett, 2009). These studies provided insights into traffic complexity but ignored the structural differences in interactions among aircraft.

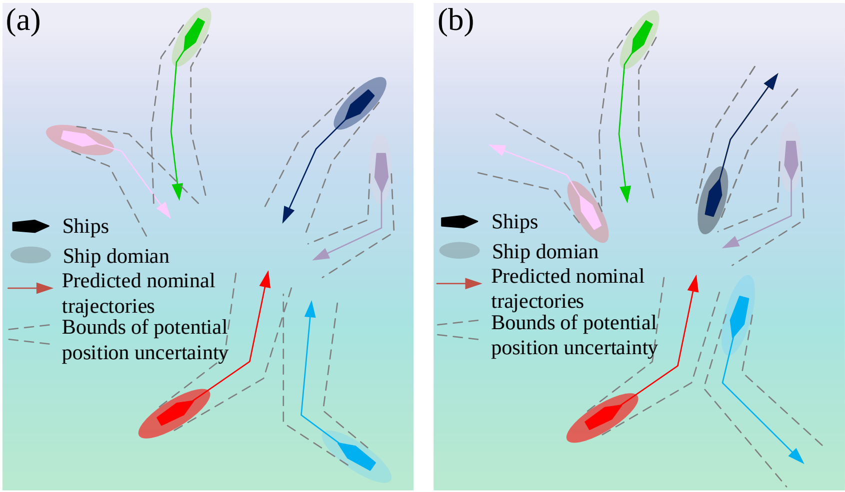

Fig. 1. Illustration of ship traffic encounter situations in complex waters.

Complex network theory is an important theoretical framework for characterising complex systems and revealing the topological properties of system structures. It is a useful tool for investigating the relations between different parts in a system to help understand the pattern and behaviour characteristics created by interacting elements. Many systems allow the deployment of complex network theory by abstracting component units as interactions between units (Barabási & Albert, 1999; Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006). With the rapid development of complexity science, complex network theory has become more prevalent across many traffic research topics, including vulnerability and resilience analysis of air transportation networks (Wong, Tan, Chandramouleeswaran, & Tran, 2020), dependence relation recognition between air traffic network structure and safety events (Carro, Valdés, García, & Comendador, 2019), difficulty measurement that controllers encounter in different traffic situations (H. Wang et al., 2016), and quantitative assessment of marine traffic systems (Sui et al., 2020), from a macroscopic or microscopic perspective. As air and maritime traffic have all the characteristics of time-variable complex systems, there has been a growing trend towards microscopic traffic network complexity modelling using complex network theory. They (Sui et al., 2020; H. Wang et al., 2016) characterised air/maritime traffic as a complex network by denoting aircraft/ships as nodes and between-aircraft/ship-pair conflict relations as edges, and then used topological metrics to highlight the overall network performance to enhance operators' perception capabilities for traffic operation situations.

Despite their popularity, the practical applications of previous models have been limited, including

- First, most studies ignored the fact that the traffic situation dynamically evolves with time. Obviously, it is inadequate to assess a traffic situation by only considering its structural characteristics because capturing its traffic dynamic evolving trend is equally important, as it can help analyse the urgency of different traffic cases that need immediate solutions to conflicts.
- Second, the interactions or conflict relations between aircraft-pairs/ship-pairs are formulated based on ideal hypotheses, such as regular or constant traffic moving speeds and no environmental disturbance on traffic motion. However, the dynamics and uncertainty in traffic movements are detrimental to exploiting real traffic conflict patterns, especially for complex traffic scenarios.
- Third, most of the studies merely concentrated on the overall complexity of the traffic situation but overlooked the investigation of how much influence each node has on the traffic network. A thorough complexity evaluation that enables the identification of key influential ships is of significant value for facilitating conflict resolution.
- Fourth, compared with traditional methods, which are more dependent on assessment standards and unsuitable for processing high-dimensional data, advanced new traffic complexity evaluation techniques can accommodate big traffic data to better identify traffic complexity patterns and support traffic alerts.

These challenges are yet to be addressed in the current literature. An overall solution will, therefore, no doubt require a new holistic framework involving the advanced models performed jointly, revealing the high complexity of the method in this study. However, it is, in return, very beneficial to develop a feasible and reliable ship traffic complexity evaluation model from both theoretical and practical perspectives. This will shift the paradigm of the current ship collision risk avoidance practise by moving from a local (ship to ship in the same water) to a regional traffic network (ships across different neighbouring waters but with potential encounters with time) level analysis.

2.3 Multiple index evaluation

Typically, traffic complexity evaluation is by nature a high-dimensional, complex, and multiple-index assessment issue. It involves many topological metrics in complex network theory to describe traffic complexity characteristics. Conventional mathematical solutions to index integration include the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) (K. Liu et al., 2016), fuzzy comprehensive evaluation (D. Liu et al., 2019), entropy weight method (Bian & Deng, 2018), grey relational analysis (J. Li et al., 2015), and support vector machine (SVM) (Yitian Xu, Xi, Lv, & Guo, 2012), among others. These approaches provide diverse perspectives for the interpretation of index interactions and help study complex systems. However, they have been criticised for their inherent weaknesses, such as subjectivity and difficulty in determining index weights, inability to process high-dimensional data, low resolution, and the requirement of a large number of training samples. In comparison, the FCI approach has recently been favoured by scholars because of its merits in handling high-dimensional index vectors and demonstrating quality clustering effects for multivariate data. The essence of this approach is to search for the optimal memberships of the assessment samples against different performance standards (e.g., risk levels and complexity patterns) based on the fuzzy set theory and an optimization algorithm (He, Zhou, Kou, Lu, & Zou, 2011). Compared with the above-named traditional methods, it is independent of the evaluation criteria by concentrating on the dataset's characteristics and can figure out the randomness, uncertainty, and fuzziness of the dataset in the evaluation process (Lu, Shang, & Li, 2017; Lyu, Shen, Zhou, & Zhou, 2019; D. Wu et al., 2013). In addition, the performance of the FCI model has been verified by solving typical benchmark test problems and actual multi-index assessment cases (He & Wan, 2020; He et al., 2011). Despite this, current applications of FCI are mainly constrained in the field of complex engineering (Q. Wang & Yang, 2020), and there is little research on applying FCI to evaluate ship traffic risk. Therefore, this study pioneers the application of FCI to perform a complexity assessment task for global and regional maritime traffic. In addition, some improvements to the original FCI model are made by combining it with an intelligent optimisation technique to enhance its searching ability for a global optimal solution and by designing a special optimisation objective function to achieve the integration of subjective and objective knowledge. In this way, it will produce better and more effective complexity evaluation results.

3 Methodology: ship traffic complexity modelling and evaluation

In this section, we provide a systematic description of the proposed ship traffic-complexity methodology. The methodology is outlined in four interactive steps, each of which is associated with methods of high dependency (see Fig. 2). Specifically, the conflict criticality between each pair of ships is first qualified. This step is achieved using a probabilistic CD framework, which can accurately estimate the conflict criticality in the presence of ship motion dynamics and uncertainty. Second, we construct a ship traffic complexity matrix based on the conflict interactions among ships to characterise the topological and evolutionary characteristics of ship traffic at a global level. After these two steps, the complexity matrix is evaluated to determine the complexity classes of the regional traffic situation. Finally, the key influential ships in the regional traffic situation are identified to assist in relieving traffic complexity. The critical supporting techniques in each step are elaborated in the subsequent subsections to demonstrate the logical flow of the methodology.

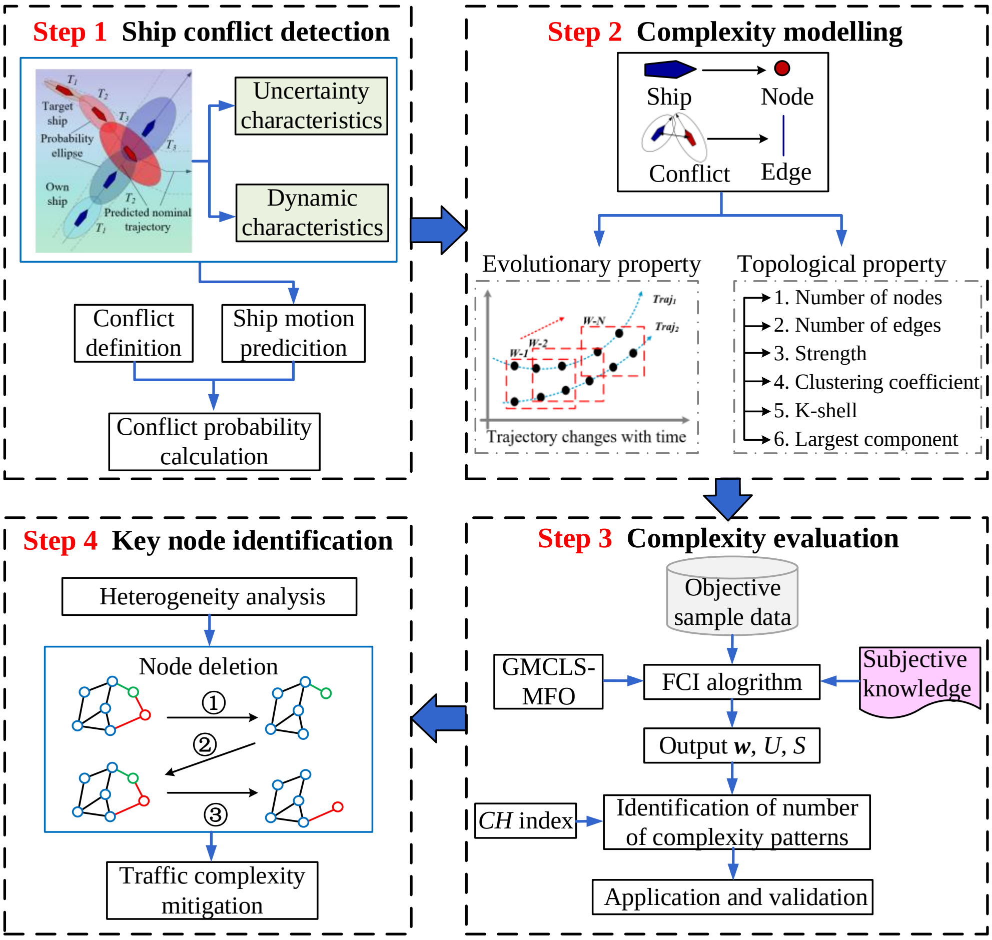

Fig. 2. Framework for ship traffic complexity modelling and evaluation.

3.1 Probabilistic conflict detection

A CD approach is first developed from a probabilistic risk viewpoint, in which the influence of uncertainty related to various sources on the potential conflict is considered. It is designed to quantify the conflict criticality between ship pairs and provide essential functionality for establishing a ship traffic network. The probabilistic CD approach is characterised by a conflict criticality measure and conflict probability estimation.

3.1.1 Conflict criticality measure

In general, a potential conflict exists if the predicted positions between ship pairs violate the safe separation criteria during the look-ahead period (Hernandez-Romero, Valenzuela, & Rivas, 2019; Mitici & Blom, 2018). This study characterises vessel conflict using a commonly used ship-domain model (Fujii & Tanaka, 1971). It is adaptive to restricted and high-density traffic areas but has low model complexity for practical applications. Fig. 3 shows an example of conflict identification. In the figure, a conflict between ships *A* and *B* is defined when the following inequality is satisfied during the look-head period.

$Dist_{AB}(t) \leq SD_A(t) + SD_B(t)$ (1)

where $Dist_{AB}$ is the distance between the two ships, and SD_A and SD_B represent the distances between the ship centres and their domain boundaries, respectively. Because we focus on identifying ship conflicts with ship motion uncertainty, inequality (1) is a probabilistic issue.

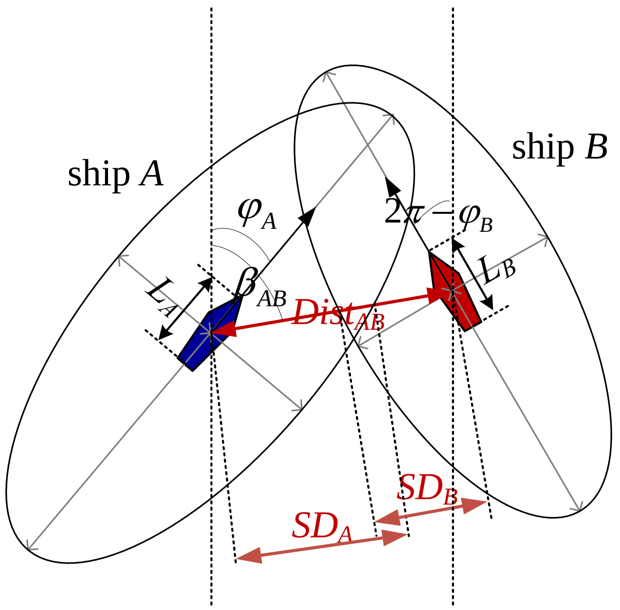

Fig. 3. Ship conflict illustration.

Given that the probability density function (PDF) for the unsafe separation (i.e. $Dist_{AB}(t) \leq SD_A(t)+SD_B(t)$)

between the two ships at time t is denoted as $f_{L(t)}$, the instantaneous occurrence probability of a conflict is expressed as follows:

$PC(t) = \Pr[L(t) \le 0] = \int_{-\infty}^{0} f_{L(t)}(\rho)d\rho$ (2)

where $L(t)$ equals $Dist(t)$-$SD_A(t)$-$SD_B(t)$.

To further quantify the criticality of a conflict, an appropriate supporting metric in terms of the maximum conflict probability over a finite time horizon is designed to characterise the conflict as follows:

$C(\gamma) \coloneqq \max_{t\in[0,T]} PC(t)$ (3)

where T represents the predicted time range.

3.1.2 Conflict probability estimation

The probabilistic CD is implemented in terms of the predicted ship trajectories with uncertainty. Hence, we extend CD into a probabilistic framework by developing a ship motion model for future uncertain trajectory prediction and a two-stage Monte Carlo (TSMC) simulation algorithm for precise conflict probability calculation.

According to the survey in (Huang et al., 2020), interaction-aware prediction is the most accurate approach. This is because this category of methods assumes that ships estimate each other's trajectories by exchanging their navigation plans or intentions through communications, whereas each ship has more relevant information related to its future trajectory. Consequently, a ship motion model is built by assuming that each ship obtains other ships' navigation plans or planned trajectory information through interaction.

Typically, a ship's navigation plan is associated with a series of waypoints, denoted as $WP_{i=1,2,\ldots,n+1}$, where the lines between consecutive waypoints constitute the movement route. This study considers the uncertainties inherent to the CD issue and models ship motion as a deterministic motion related to the navigation plan and an uncertain motion expressed by the probability distribution functions. Therefore, the ship motion model comprises three main modules: 1) a continuous dynamic model revealing the physical motion of the ship; 2) a discrete dynamic model given by the navigation plan; and 3) an uncertainty module coinciding with diverse disturbances such as physical, environmental, and human factors. Specifically, the predicted position of ship A at the future moment T is computed using the following equation:

$\overrightarrow{S_A}(t_c + T) = \overrightarrow{S_A}(t_c) + \int_{t_c}^{t_c+T} \overrightarrow{V_A}(t)dt + R(\vartheta_A^T) \cdot \overrightarrow{Q_A}(T)$ (4)

where t_c denotes the current moment, $\overrightarrow{S_A}(t_c)$ denotes the ship A's initial position, $\overrightarrow{V_A}(t)$ is the ship A's nominal speed and is described as a piecewise constant function related to the navigation plan, $R(\vartheta_A^T)$ represents the rotation matrix related to the ship's nominal heading course ϑ_A^T at time moment T, and $\overrightarrow{Q_A}(T)$ represents the time-dependent uncertain components associated with the ship's predicted position in the heading and lateral directions. Upon obtaining each ship's uncertain trajectory distributions over the prediction horizon, whether inequality (1) holds can be further identified. Because of the uncertainty of the future trajectories of encountering ships, the existence of conflicts is also uncertain.

The addition of the ship motion uncertainty module increases the computational cost of the conflict probability estimation. Therefore, a TSMC algorithm that enables fast computation of the conflict probability and provides the quantitative bounds of the approximation error is designed. This algorithm is developed in terms of the fact that only the maximum conflict probability during the prediction horizon needs to be precisely identified. Therefore, the time points with high conflict probabilities are first roughly extracted before conducting a larger number of Monte Carlo iterations. Thus, the algorithm can efficiently and accurately estimate the conflict criticality. A complete description of the probabilistic CD process can be found in (Xin, Liu, Yang, Zhang, & Wu, 2021).

3.2 Ship traffic network complexity modelling

As mentioned above, this study extends the application of complex network theory to ship traffic complexity modelling. Graph theory is the essence of complex networks. It considers the units in the network as nodes and the interrelations between the units as edges, thereby constructing a graph structure. This study identifies whether an edge exists between a ship pair based on the conflict criticality between them. If the conflict criticality between a ship pair exceeds a pre-defined safety threshold, they are connected by an edge, and the weight of the edge is measured by its actual conflict criticality value. In this manner, a ship traffic network can be built to lay the foundation for traffic network analysis. Such new quantitative and qualitative features allow for simultaneous analysis of the structural and evolutionary characteristics of the traffic network. The relevant work is elaborated as follows:

3.2.1 Topological metrics of traffic network

The complex network theory provides various lists of network indices to reveal the structural properties of a network. To evaluate a ship traffic network, we need to select indices that not only adapt to the basic features of ship traffic but also have high representativeness for network performance. Based on these two criteria, six indicators are chosen for the traffic network structure characterisation.

(1) Number of nodes

The number of nodes (N), which is known as the network scale, is a basic characteristic of a network. The structure of a network strongly depends on its number. A larger number of ships in a ship traffic network is associated with a busier and more complex traffic environment.

(2) Number of edges

The number of edges (E) represents the number of connections (i.e., potential conflicts) between all ship pairs. Surveillance operators encounter tremendous regulatory pressure when many ships are in potential conflict.

(3) Strength

Node strength refers to the sum of the weights of the edges connected to one node. This study adopts the sum of all node strengths to reveal the total conflict intensity of a traffic situation, which is formulated using the following equation:

$$S = \sum_{i=1}^{N} \sum_{j=1}^{d_i} w_{ij}/2 \quad (5)$$

where N is the number of nodes, d_i is the number of nodes connected to node i, and w_{ij} is the edge weight between nodes i and j. The higher the value of S, the more dangerous the associated traffic situation.

(4) Clustering coefficient

The clustering coefficient (CC) describes and measures the clustering/proximity of nodes in a network. The local CC of node i is defined as the ratio of "the actual number of edges ($N_\Delta(i)$) between the nodes connecting with node i" to "the theoretical maximum number of edges between these nodes" as follows:

$$C_i = \frac{N_\Delta(i)}{d_i(d_i-1)/2} \quad (6)$$

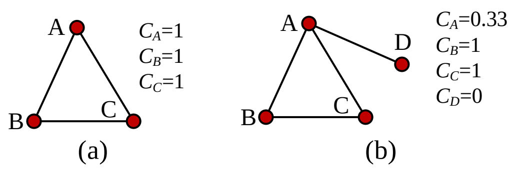

Fig. 4. Example of clustering coefficient calculation.

Although this indicator reveals the proximity of one node to its nearby nodes, it cannot satisfy the general rules of ship traffic network complexity modelling, that is, adding a new ship or conflict edge should not result in a decrease in traffic complexity. This can be proven by observing Fig. 4. It is found that the clustering coefficient of node A decreased after node D and a new edge was added to the network. Therefore, some adjustments are made based on the original metric, and the sum of C_i' is used to represent the clustering coefficient of the entire traffic network, which is expressed as follows:

$C = \sum_{i=1}^{N} C_i' = \sum_{i=1}^{N} N_\Delta(i)$ (7)

The improved indicator can effectively quantify the degree of cross-conflict among ships while satisfying traffic network modelling principles. When the traffic network structure corresponds to a large value of *C*, this implies that conflicts among ships become challenging to resolve. This is because there is a high probability that the action taken by one ship for conflict resolution with another ship results in a higher collision risk with nearby ships. Hence, a large value of *C* is associated with a complex traffic situation in which conflict-resolution strategies are difficult to implement.

(5) K-shell decomposition

K-shell decomposition is a widely adopted technique in graph theory that divides the network structure into different layers/shells in terms of the position of the nodes in the network. It assigns an identical positive index (ks_i) to the nodes in the same layer to help identify the central nodes and rank their global influential power (see Fig. 5). Therefore, the sum of ks_i is adopted to reflect how the nodes are grouped from a global perspective, as follows:

$KS = \sum_{i=1}^{N} ks_i$ (8)

A traffic network structure with a higher *KS* value indicates that many ships are close together with complex conflict relations; thus, more attention should be received from surveillance operators.

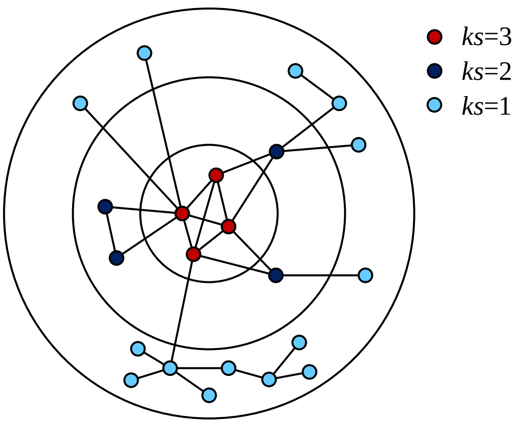

Fig. 5. A schematic diagram of k-shell decomposition.

(6) Largest component

The largest component (*LC*) is a useful measure of the local cohesiveness of a network. It is the largest subgraph in a network, where there is at least one path between all node pairs. The size of the *LC* is expressed as

$\eta = |S|$ (9)

where $|S|$ represents the number of nodes in the LC. LC reveals the network's local structural nature, and the larger the size of the LC, the more critical and complex the traffic network.

3.2.2 Multi-stage and multi-topology traffic complexity matrix

The ship traffic network structure is highly dependent on the conflict relations between the ship pairs, which further determines the outputs of the chosen topological indices. In this study, the criticality of a conflict is measured in terms of the maximum conflict probability over a finite prediction time horizon. Hence, the traffic network structure may differ significantly when different prediction time ranges are adopted. For a given ship traffic situation, if the detection time range is too short, there will be few connections between ship pairs. In contrast, the edges between ship pairs could be too large. Additionally, the shorter the conflict detection periods, the more urgent the appropriate collision avoidance manoeuvres to conflict (Goerlandt et al., 2015; Hu et al., 2019; Zhen et al., 2017). Thus, this study defines short prediction periods as associated with urgent traffic networks, and long periods as related to relatively non-urgent traffic networks.

To offer a complete description of the structural and evolutionary characteristics of traffic situations, a multi-stage and multi-topology-based ship traffic network complexity model is developed to quantify the complexity of traffic situations. We establish a two-dimensional spatial-temporal state matrix *X* using the row dimension to describe the evolutionary trend of a traffic situation and the column dimension to quantify the corresponding structural characteristics, as follows:

$$X = (x_{ts})_{o\times p} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{o1} & x_{o2} & \cdots & x_{op} \end{bmatrix} \quad (10)$$

where o denotes the number of prediction time range thresholds for CD and p represents the number of chosen topological indices. The element x_{ij} in the state matrix represents the output of the jth topological metric corresponding to the network structure constructed based on the ith selected conflict-detection time threshold. Because ships could be in conflict at some time points (Chai, Weng, & Xiong, 2017; Weng, Meng, & Qu, 2012; Weng & Shan, 2015; X. Wu, Mehta, Zaloom, & Craig, 2016), especially for high traffic waters, the first-layer traffic network structure is constructed based on real-time conflict relations between ship pairs. Furthermore, 5, 10, and 15 min are adopted as the detection time thresholds for network construction of the 2nd, 3rd and 4th layers, respectively. This is because we focus on the CD in the medium-term time range, that is, the order of tens of minutes, whereas the long-term traffic motion and conflict prediction, that is, over a time horizon of several hours, are generally highly inaccurate and, on many occasions, not essential (Jilkov, Ledet, & Li, 2018; Prandini, Hu, Lygeros, & Sastry, 2000; Xiao et al., 2019). Consequently, a four-layer and six-topology network model is built to fully characterise current and near-future traffic situations. It is noteworthy that this model can be extended by setting different detection time thresholds or incorporating more useful network indices according to the traffic features of the regulatory waters and surveillance operators' demands.

3.2.3 Similarity measurement between different complexity matrices

Based on the established traffic state matrix, a distance metric must be designed to calibrate the degree of closeness between the different traffic state matrices. The traditional Euclidean distance is commonly used to measure the similarity degree but ignores the influence differences of the elements in the state matrix. It is evident that each topological indicator contributes differently to the complexity of the traffic situation. In addition, the rows associated with urgent traffic network structures should be given higher weights because small detection time thresholds correspond to urgent conflict relations to be resolved, thereby having a more substantial influence on traffic complexity. Consequently, a weighted Euclidean distance metric is introduced to rationally define the distance between different traffic state matrices through the combination of a time-weighted matrix and a space-weighted matrix.

The forms of time-weighted and space-weighted matrices are expressed as follows:

$$W_t = \begin{pmatrix} w_{t,1} & & & \\ & w_{t,2} & & \\ & & \ddots & \\ & & & w_{t,p} \end{pmatrix}$$

$$W_s = \begin{pmatrix} w_{s,1} & & & \\ & w_{s,2} & & \\ & & \ddots & \\ & & & w_{s,p} \end{pmatrix}$$

where $w_{t,i} = f^i / \sum_{i=1}^{p} f^i$, f denotes a fading time factor that satisfies $0<f<1$, $0 \le w_{s,i} \le 1$ and $\sum_{i=1}^{o} w_{s,i} = 1$.

Before computing the distance between two different state matrices, each element in matrix X should be normalised to eliminate the influence of the data dimensions using Eq. (11).

$$v_{ts} = (x_{ts} - x_{\min,ts})/(x_{\max,ts} - x_{\min,ts}) \quad (11)$$

where $x_{\max,ts}$ and $x_{\min,ts}$ are the maximum and minimum values in the tth row and the sth column, respectively. Consequently, matrix X can be transformed into matrix V (i.e. a normalised matrix), as follows:

$$V = (v_{ts})_{o\times p} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1p} \\ v_{21} & v_{22} & \cdots & v_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ v_{o1} & v_{o2} & \cdots & v_{op} \end{bmatrix} \quad (12)$$

The distance between matrices V_a and V_b is further computed by

$$D_{ab} = \|W_t V_a W_s - W_t V_b W_s\| \quad (13)$$

Note that the complexity state matrices and weight matrices must be transformed into vector forms to make them adaptive to the complexity assessment approach in Section 3.3. The distance metric is fine-tuned as follows:

$$D_{ab} = \|\boldsymbol{w}(\boldsymbol{r_a} - \boldsymbol{r_b})\| = \sqrt{\sum_{i=1}^{m}[w_i(r_{ia} - r_{ib})]^2} \quad (14)$$

where $\boldsymbol{w} = [w_{t,1}\cdot w_{s,1}, w_{t,2}\cdot w_{s,1}, \ldots, w_{t,p}\cdot w_{s,o}] = [w_1, w_2, \ldots, w_m]$, $\boldsymbol{r_a} = [v_{a,11}, v_{a,21}, \ldots, v_{a,po}] = [r_{1a}, r_{2a}, \ldots, r_{ma}]$ and $m = o\times p$.

3.3 Traffic complexity evaluation

One of the significant obstacles to using the traffic complexity state matrix for maritime traffic situation comprehension is the assessment of complexity classes or patterns (e.g., low-complexity, medium-complexity, and high-complexity). This is because the traffic complexity state matrix involves many evaluation indices, and there are nonlinear, strong coupling, and uncertain relationships among these indices. This study adopts the FCI approach as a solution. The direct application of the original FCI to ship traffic complexity evaluation has several technical deficiencies. Hence, further improvements and extensions are essential to produce reliable and practical evaluation results.

3.3.1 Fuzzy clustering iterative approach

A typical FCI approach uses an iterative computation procedure among the membership matrix, clustering centre matrix, and index weight vector to constantly optimise the target function until it finds the optimal U and S (Q. Wang & Yang, 2020). The operation details of the FCI approach are presented in the following section.

Assume that there are n samples, each of which has m assessment indices, $A_j = (a_{1j}, a_{2j}, \ldots, a_{mj})$. The sample dataset can be described as an $m\times n$ matrix, as follows:

$$A = (a_{ij})_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (15)$$

where a_{ij} is the eigenvalue of index i for sample j; $i = 1, 2, \ldots, m$; and $j = 1, 2, \ldots, n$.

To make different indices scale-invariant, dataset A is transformed into a normalised eigenvalue matrix $R = (r_{ij})_{m\times n}$ using Eq. (11). Meanwhile, we denote the fuzzy membership matrix as $U = (u_{kj})_{c\times n}$, and the fuzzy class centre matrix as $S = (s_{ik})_{m\times c}$, subject to constraints $0 \le u_{kj} \le 1$, $\sum_{k=1}^{c} u_{kj} = 1$ and $0 \le s_{ik} \le 1$, where c is the number of classes in which n samples are clustered, u_{kj} is the relative membership degree of sample j belonging to class k, and s_{ik} is the class centre of index i in class k.

To obtain the optimal U and S, the objective function can be established by minimising the quadratic sum of the weighted Euclidean distance from all samples to different fuzzy class centres, which is given by the following expression:

$$\min\left[F\left(w_i, u_{kj}, s_{ik}\right)\right] = \min\left\{\sum_{j=1}^{n}\sum_{k=1}^{c}\left({u_{kj}}^2 \sum_{i=1}^{m}(w_i(r_{ij} - s_{ik}))^2\right)\right\} \quad (16)$$

where w_i is the weight of the ith index, satisfying $0 \le w_i \le 1$ and $\sum_{i=1}^{m} w_i = 1$.

Based on the Lagrange multiplier method (He et al., 2011; Lyu et al., 2019), the optimal U and S can be obtained through an iterative way using Eq. (17)-(18).

$$s_{ik} = \sum_{j=1}^{n} {u_{kj}}^2 {w_i}^2 r_{ij} \Big/ \sum_{j=1}^{n} {u_{kj}}^2 {w_i}^2 \quad (17)$$

$$u_{kj} = \left[\sum_{h=1}^{c} \frac{\sum_{i=1}^{m}[w_i(r_{ij} - s_{ik})]^2}{\sum_{i=1}^{m}[w_i(r_{ij} - s_{ih})]^2}\right]^{-1} \quad (18)$$

until the following termination conditions are met.

$$\max_{kj} |u_{kj}^{l+1} - u_{kj}^{l}| \le \varepsilon_1 \text{ and } \max_{ik} |s_{ik}^{l+1} - s_{ik}^{l}| \le \varepsilon_2$$

where l represents the iteration counter and ε_1 and ε_2 are the precisions for calculating u_{kj} and s_{ik}, respectively.

Although the FCI approach performs well in coping with complex nonlinear and high-dimensional problems, several limitations can be found through its application: (1) it is highly sample-dependent; (2) it is easily trapped in local convergence; and (3) it is strongly dependent on the users to identify the optimal number of classes c beforehand.

With respect to the first issue, a constraint is added to the original target function to integrate the decision maker's subjective cognition and objective information derived from the sample dataset. For ship traffic complexity evaluation, the number of nodes/ships is regarded as the basic characteristic of regional ship traffic, which generally has a more significant contribution to traffic complexity than other indices from the perspective of subjective cognition (Y. Wen et al., 2015). Therefore, the objective function of the FCI in this study can be redefined using Eq. (19).

$$\min f = \min\left[F\left(w_i, u_{kj}, s_{ik}\right) + F(w_i)\right]$$
$$= \min\left[F\left(w_i, u_{kj}, s_{ik}\right) + M \cdot \max\left(\max_{2 \ll i \ll m} w_i - w_1, 0\right)^2\right] \quad (19)$$

where M is a penalty factor and w_1 denotes the weight of the index of the number of nodes. By doing so, a more practical evaluation that combines subjective knowledge with objective information can be achieved. The form of Eq. (19) can be conveniently adjusted when more expert knowledge is available.

With regard to the second issue, intelligent optimization techniques are often combined with the FCI to increase the possibility of finding the globally optimal solution (He et al., 2011; Zou, Liao, Ding, & Qin, 2019). Thus, an extended moth-flame optimisation algorithm which is combined with Gaussian mutation and chaotic local search, called GMCLS-MFO, is adopted to enhance the global optimisation performance of the FCI model. The relevant details of the extended MFO operation can be found in (Yueting Xu et al., 2019).

Regarding the third issue, the Calinski–Harabasz (CH) index is used to identify the optimal number of classes. It is computed based on the ratio of the inter-cluster means and the intra-cluster sum of squares, as follows:

$$CH = \frac{\left\{\sum_{k=1}^{c} \sum_{j=1}^{n} {u_{kj}}^2 \sum_{i=1}^{m} (w_i(s_{ik} - z_i))^2\right\}/(c-1)}{\left\{\sum_{k=1}^{c} \sum_{j=1}^{n} {u_{kj}}^2 \sum_{i=1}^{m} (w_i(r_{ij} - s_{ik}))^2\right\}/(n-c)} \quad (20)$$

where $z = (z_1, z_2, \ldots, z_m)$ is the centre of the entire dataset. The larger the CH index, the better the classification performance.

3.3.2 Extending FCI-based traffic complexity evaluation

In this study, the enhanced FCI-based traffic complexity assessment approach is among the most important methodological contributions. It aims to provide a feasible and reliable evaluation associated with a hierarchical description of the traffic complexity patterns. The implementation of the approach comprises the following steps:

Step 1: Sample extraction, pre-processing, and normalization

Extract the sample dataset of ship traffic complexity state matrices in terms of real AIS-based trajectory data in the investigated waters. Subsequently, transform the complexity state matrices into vector forms to construct the sample dataset matrix. Subsequently, standardise the sample matrix to the normalised form for performing the extended FCI.

Step 2: Parameter initialisation

Initialise the parameters, including precisions ε_1 and ε_2, variable dimensionality to be optimised d, and number of complexity patterns c.

Step 3: Implement GMCLS-MFO

Use the GMCLS-MFO to search for the optimal index weight vector.

Step 4: Implement FCI

Return the optimal fuzzy membership matrix U and the optimal clustering centre matrix S using the FCI.

Step 5: Identify the optimal number of complexity patterns

Determine the search range for the number of classes c (e.g., c = 2-10). Then, run step 2 to step 4 repeatedly for each number of classes and compute the corresponding CH index. Furthermore, identify the optimal c based on the CH index.

Step 6: Output the evaluation results

Use the following two indicators to evaluate each traffic complexity sample.

$$H_j^1 = \arg\max_k u_{kj} \quad (21)$$

$$H_j^2 = R_j \cdot w \quad (22)$$

where H_j^1 represents the complexity pattern/grade of the jth sample and H_j^2 denotes its complexity value. The first indicator is determined based on the maximum membership of each sample, whereas the second indicator is identified in terms of the weighted sum of the normalised network complexity indices.

Step 7: Test the validity

Validate the effectiveness and reliability of the approach based on sensitivity analysis. It is performed in terms of the following two axioms (Fan, Blanco-Davis, Yang, Zhang, & Yan, 2020; Q. Yu, Liu, Chang, & Yang, 2020): (1) a slight increase/decrease in a single factor should lead to a corresponding increase/decrease in the evaluation results; (2) the influence of multiple factor changes should be no less than that of a single factor change. Here, the effects of factor change on the assessment results are examined based on the degree of complexity pattern change and quantitative utility values using the following equations:

$$CD = \frac{n_{pc}}{n} \quad (23)$$

$$I = \sum_{k=1}^{c} n_k \cdot k \quad (24)$$

where n represents the total number of traffic complexity samples, n_{pc} is the number of samples with changed complexity patterns after the values of factors/indices are adjusted, and n_k represents the number of samples belonging to complexity pattern k.

3.4 Key node identification in the traffic network

Surveillance operators are not only required to proactively monitor a dynamic maritime traffic situation but also to assist navigators in resolving potential conflicts. Therefore, identifying the key influential ships is crucial because the guidance for these key ships from a global perspective is helpful in relieving regional traffic complexity. It also provides a reference for operators at a local level to reduce the difficulty in controlling and mitigating collision risk. A node deletion method is the most typical system analysis method, which is adopted to determine the key influential ships in a traffic network. It has already been successfully applied to identify the key ports and airports in the traffic domain (e.g., H. Liu, Tian, Huang, & Yang, 2018; X. Wen, Tu, & Wu, 2018). The main idea of this method is to calculate the changes in network performance after removing different nodes from the network. The more significant the drop in network performance, the more influential the node is. This method follows the procedure described in.

- First, the complexity value of the entire ship traffic network with n_t nodes at a given time moment is calculated, denoted by H_t^2.
- Second, the ship traffic complexity network is reconstructed after deleting the qth (q = 1, 2, …, n_t) node and the corresponding complexity with the remaining n_t-1 nodes is calculated, denoted by H_{tq}^2.
- Third, the margin of H_{tq}^2 against H_t^2 is figured out, defined by $\Delta H_{tq}^2 = H_t^2 - H_{tq}^2$.
- Finally, the key nodes are identified based on the values of ΔH_{tq}^2, a higher value of ΔH_{tq}^2 indicates that the node has a more significant impact on the overall traffic complexity.

4 Applications and case study results

In this section, we apply and evaluate the performance of the proposed traffic complexity evaluation method. It is organised as follows: Section 4.1 illustrates the case study and relevant AIS data. Section 4.2 presents the training results for the extended FCI-based evaluation model. Section 4.3 demonstrates how the proposed methodology improves regional maritime traffic interpretation and helps in relieving traffic complexity. Section 4.4 performs a few statistical analyses of the traffic complexity patterns in the investigated waters, while Section 4.5 conducts a model validation and comparison analysis.

4.1 Study area and data description

The Ningbo-Zhoushan port is selected as the experimental research area to test the proposed ship traffic complexity methodology. According to cargo throughput, it is the largest port globally, exposing high traffic density, restricted navigable waters, and sophisticated time-varying uncertainty in vessel movement. Therefore, the port is representative of the complex waters. AIS data records from the port are collected in November. 1, 2018 to November. 30, 2018. The region for maritime traffic complexity evaluation is bounded between longitudes 122°06E-122°16E and latitudes 29°49N-29°57N (see Fig. 6). Because vessels such as fishing ships, pilot boats, and tugboats often exist together for their missions and may not abide by conventional maritime safety rules, the AIS messages of general merchant ships, including cargo ships and tankers, are used for experimental analysis.

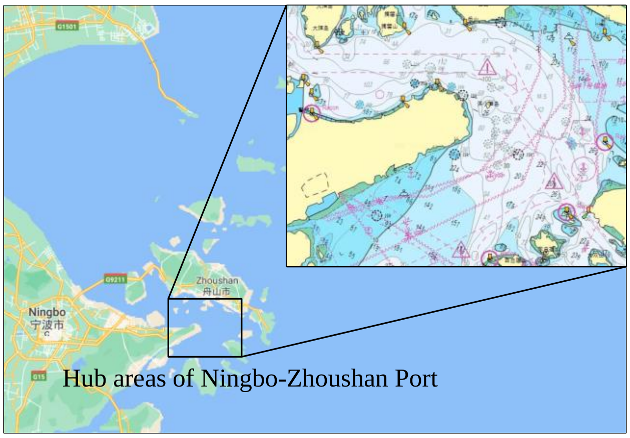

Fig. 6. Illustration of Ningbo-Zhoushan Port, China.

Although AIS provides a powerful and easy-to-access data source for maritime traffic research, the collected data often encounters issues such as missing data or position and speed information errors due to various technical failures. Therefore, the AIS data pre-processing module is essential for enhancing data quality. This study adopted a systematic pre-processing procedure to clean the data, including noise elimination of each vessel attribute (Kang, Meng, & Liu, 2018), trajectory extraction and separation, trajectory consistency confirmation (L. Zhao, Shi, & Yang, 2018), and trajectory interpolation (L. Zhang et al., 2019). In this manner, we can eliminate possible noise and reconstruct clean and reliable trajectory data for experimental analysis. In other words, this data cleansing process is a pre-condition for real-time analysis.

4.2 Model training results

This subsection presents the offline training results of the extended FCI-based complexity evaluation model in terms of the procedure described in Section 3.3.2. We also demonstrate how the proposed evaluation approach improves based on the three deficiencies of the original FCI model mentioned in Section 3.3.1.

Fig. 7 first shows the boxplot of *CH* index statistics with the different number of complexity classes. According to the assessment standard of the *CH* index, the number of classes $c = 2$ corresponds to the best classification performance. The problem is that one may not take advantage of the FCI model to support a hierarchical complexity pattern assessment when c is too small. This is because the small c is insufficient to aid maritime surveillance operators in identifying how complex the traffic situation is at a more fine-grained scale.

Therefore, by considering both the values of the *CH* index and the application effects of the FCI model, the traffic complexity is divided into six classes/patterns: very low (VL), low (L), slightly low (SL), slightly high (SH), high (H), and very high (VH) complexity. Additionally, it is shown that searching for the optimal number of classes within 2-10 is acceptable as the keen point occurs when $c = 6$ and the *CH* index then keeps falling.

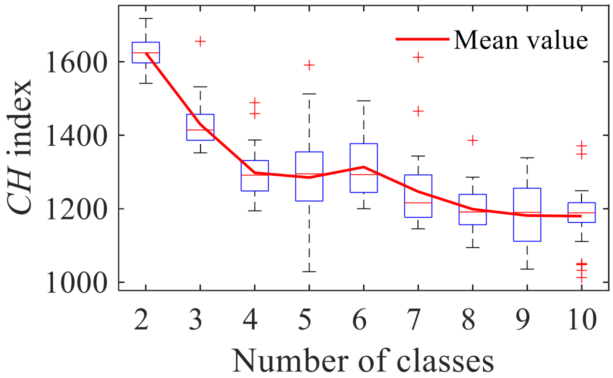

Fig. 7. Change in *CH* index with the increase in the number of complexity classes/patterns.

In addition to making use of the *CH* index to address the third issue in the original FCI model, improvements in terms of the first two issues in FCI are also analysed. Fig. 8 further illustrates the training results of the index weight *w* and clustering centre *S* when the number of complexity patterns is six. According to Fig. 8 (a), the first five indices have equal importance to traffic complexity, whereas *LC* is relatively less influential. However, if no subjective knowledge, that is, the index of the number of ships has the largest contribution to the traffic complexity, is incorporated into Eq. (16), the training results of each index weight are overdependent on the features of the sample dataset, which may not meet the general knowledge of maritime surveillance operators. It is also observed that the cluster centre distributions of the network indices at different traffic network layers depict similar modes (see Fig. 8 (b)). These training results assist in identifying real-time complexity pattern memberships of new traffic scenarios. In addition, the optimisation performance of the proposed GMCLS-MFO-FCI model is compared with that of the MFO-FCI model. Here, the original FCI is not adopted for comparison because its procedure is infeasible for the constrained objective function, Eq. (19). As shown in Fig. 9, the proposed model converges faster and exhibits better stability, thereby confirming its effectiveness and superiority.

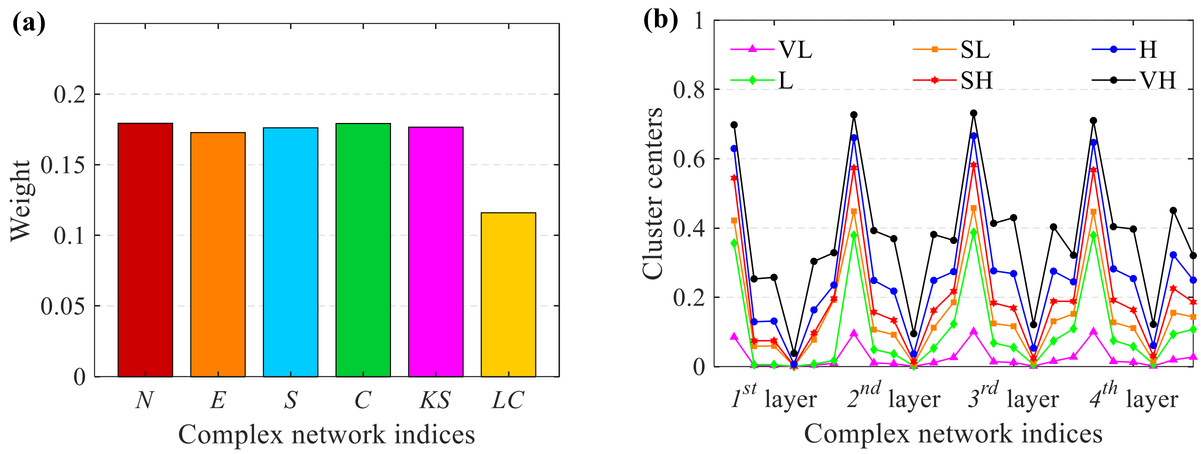

Fig. 8. (a) Weights of different network indices; (b) clustering centres of different traffic complexity patterns.

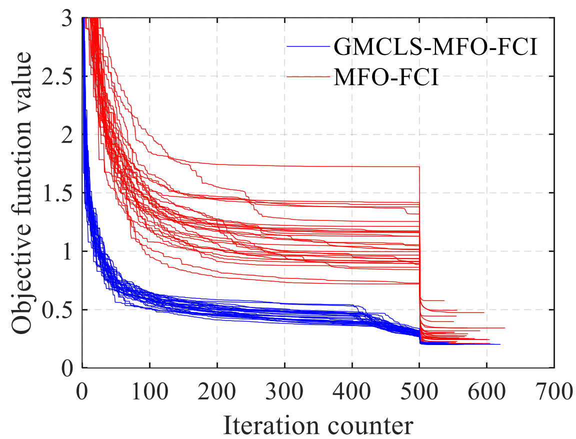

Fig. 9. Iteration optimisation processes of GMCLS-MFO-FCI and MFO-FCI model.

4.3 Application case analysis

In this subsection, the performance of the proposed traffic complexity approach is demonstrated using application case analysis. It starts by illustrating its online application in traffic complexity recognition using a case example, which is related to methodological steps 1-3 in Section 3. Subsequently, the feasibility and effectiveness of the approach for traffic complexity mitigation are shown via traffic network heterogeneity analysis and key influential ship identification, which are associated with methodological step 4.

4.3.1 Real-time regional traffic complexity evaluation

An example of the ship traffic complexity evolution over three hours within the research area is shown in Fig. 10. In Fig. 10 (a), the memberships provide a fine-grained hierarchical description of the traffic situations. It can be clearly seen that the membership distributions of different complexity patterns change over time. Moreover, it is relatively easy to find that the periods from 1.2-1.4 and 1.65-2.05 hours are in high traffic complexity situations because the membership of very high (VH) and high (H) complexity patterns occupies a large proportion. This finding can support and assist surveillance operators in monitoring and offering timely hazard warnings. Fig. 10 (b) shows the changes in the two comprehensive evaluation indices (Eq. (21) and (22)) over time: These two indices exhibit consistent trends, which are beneficial for comprehending traffic complexity and enhancing the perceptual abilities of traffic controllers.

To better understand the proposed complexity evaluation approach, Fig. 10 (c) illustrates the process of obtaining complexity evaluation results at a random moment. In the figure, steps I–II show the conflict probability computation for ship pairs based on the probabilistic CD approach. More information on the performance of the probabilistic CD model can be found in the previous work (Xin et al., 2021). Steps III-IV demonstrate the process of constructing the multi-stage and multi-topology-based traffic network complexity model, which depends on the conflict interactions among ships. Step V presents the memberships of the traffic complexity patterns and corresponding evaluation results based on the improved FCI approach.

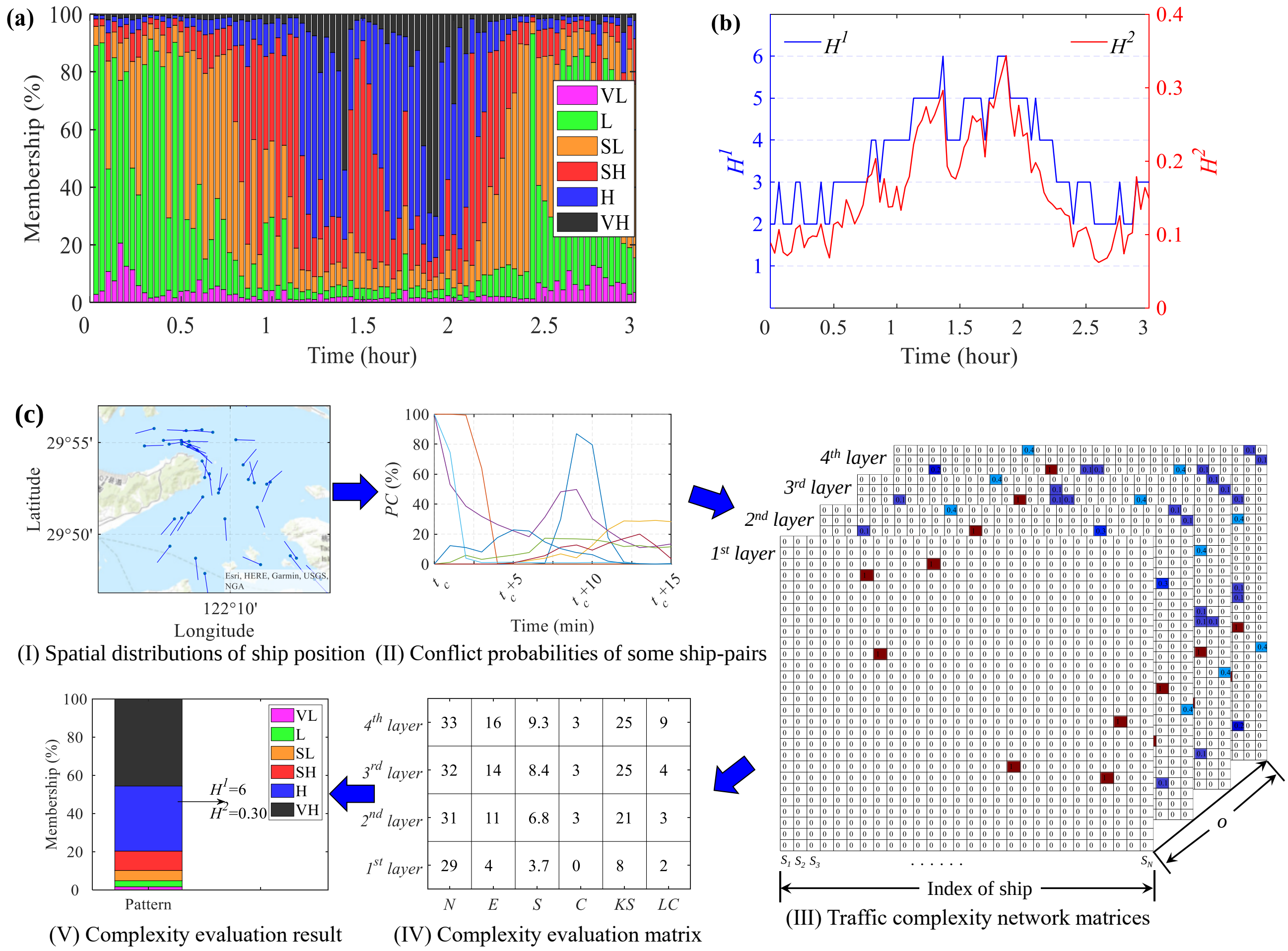

Fig. 10. Ship traffic situation evolution. (a) Membership histogram of ship traffic complexity patterns over time; (b) curves of traffic complexity assessment indicators over time; (c) illustration of traffic complexity

computation at time t = 110 min.

4.3.2 Traffic network heterogeneity analysis and key risk ship identification

To investigate the features of the traffic network topology structure, the variance of the node degrees (i.e., the number of edges to one node) is applied to measure the heterogeneity of the node degree distributions under different traffic situations. According to Fig. 11, a high-complexity pattern/value typically indicates a large degree of variance. These results suggest that when the traffic complexity is higher, the distributions of node degrees are more unbalanced. There is a high probability that several key influential ships occupy the majority of conflict edges over the entire traffic network. By contrast, the remaining ships are associated with very few edges. Therefore, the identification of key risk ships, especially in high ship traffic complexity situations, would be of great practical significance for conflict resolution and traffic complexity reduction.

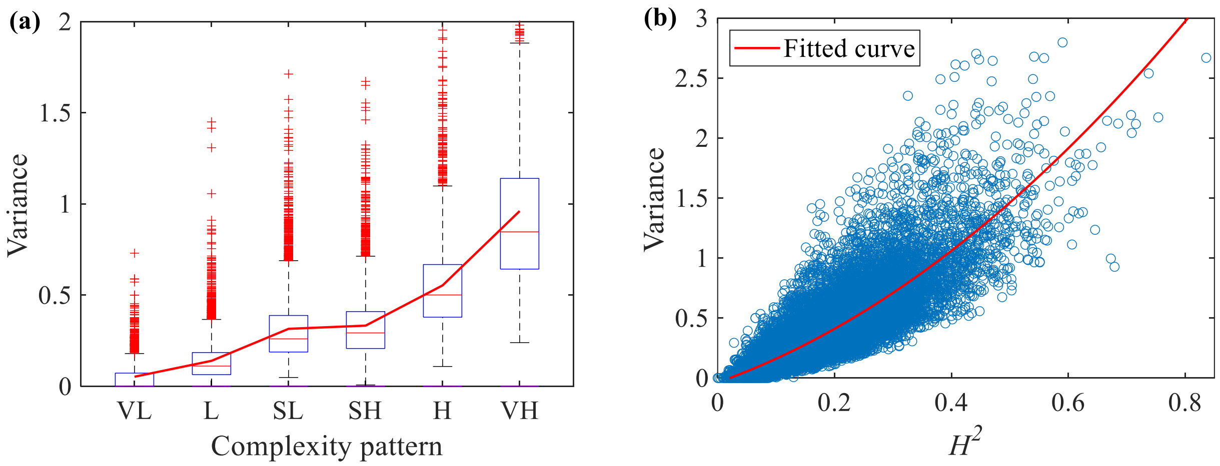

Fig. 11. (a) Boxplot of variance of degree distributions with respect to different traffic complexity patterns; (b) relation between variance of degree distributions and H^2.

Fig. 12 presents an example of the key node identification based on the node deletion method in Section 3.4. In Fig. 12 (a), it is easy to see which ships have a significant influence on the entire traffic network based on the colour index, assisting operators in resolving potential conflicts. Fig. 12 (b) further illustrates the decline in traffic complexity after deleting each ship from the traffic network. From the figure, 5 ships are more critical to the traffic situation, while the rest are less influential. This result implies that the investigated ship traffic network is heterogeneous. In addition, the dynamic changes of the traffic complexity against the ratio of removed ships based on deliberate deletion (i.e., deleting the ships based on the descending rank of their importance) and random deletion are analysed. According to Fig. 12 (c), the deliberate deletion contributes to a significant reduction in traffic complexity. In particular, the complexity index significantly decreases from 0.26

to 0.15 after the first two crucial ships are separated from the traffic network. In contrast, the random deletion leads to a relatively small decline. This phenomenon means that the identification of key influential ships as well as the risk management for these ships are essential and beneficial to relieve traffic complexity, especially from a global traffic surveillance management perspective.

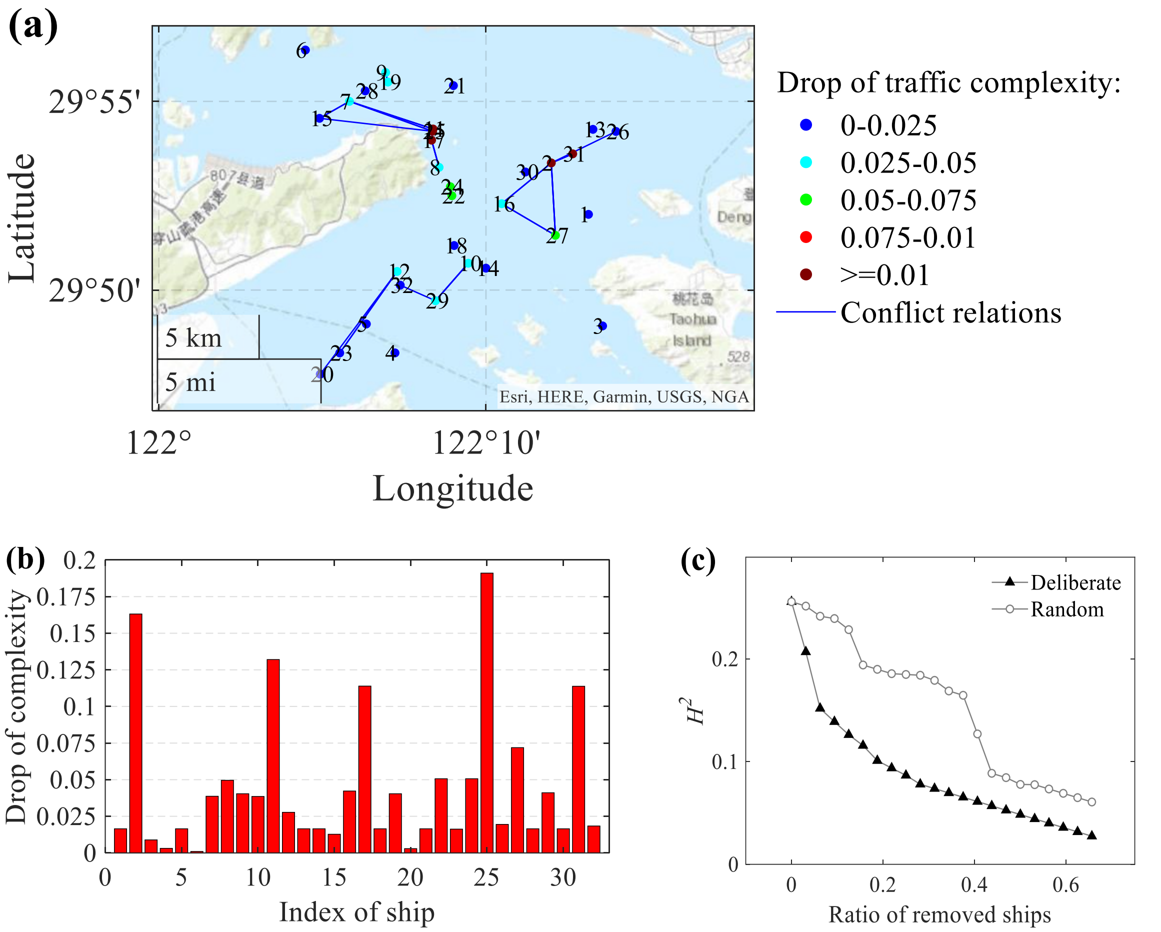

Fig. 12. (a) Ship traffic network topology; (b) traffic complexity margins of separated nodes; (c) dynamics of traffic complexity against the ratio of removed ships.

4.4 Statistics analysis of traffic complexity characteristics

Statistical analyses of real AIS data are conducted to reveal the characteristics of ship traffic complexity patterns. Fig. 13 shows the probability distributions of the different traffic complexity patterns within the study area. It is noticeable that the very low (VL) and very high (VH) complexity patterns jointly account for a minor occurrence frequency, whereas the remaining patterns occupy a large proportion. This indicates that ship traffic situations primarily exhibit relatively medium-level complexity patterns. In addition, the relatively small percentage of VL and VH complexity patterns also means that the proposed complexity pattern classification approach would neither result in traffic controllers being subject to extremely severe safety pressure nor make them inefficient in supervising maritime traffic.

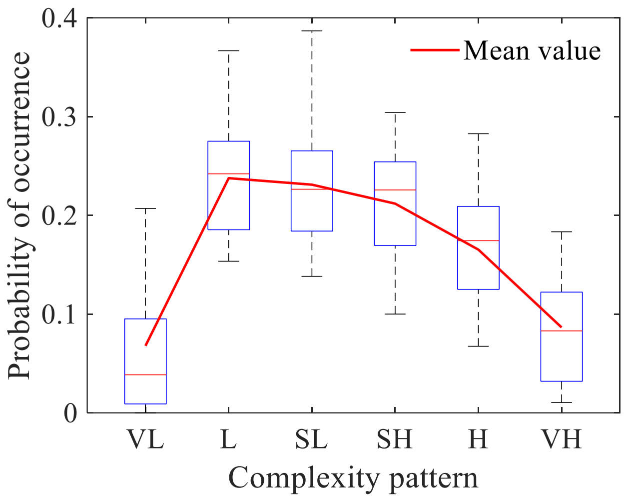

Fig. 13. Occurrence probability of different traffic complexity patterns.

Fig. 14 illustrates the temporal distribution of traffic complexity patterns for every hour of the day. In the figure, the peak periods associated with higher traffic complexity patterns can be observed during 10:00-16:00, which matches the water traffic situation where high traffic intensity usually occurs during the daytime. Safety-related navigation departments can implement appropriate risk-control strategies based on statistical analysis to ensure the safety of traffic management. The analysis results also serve as a practical reference for balancing the operators' workloads because the monitoring pressure within different periods varies significantly.

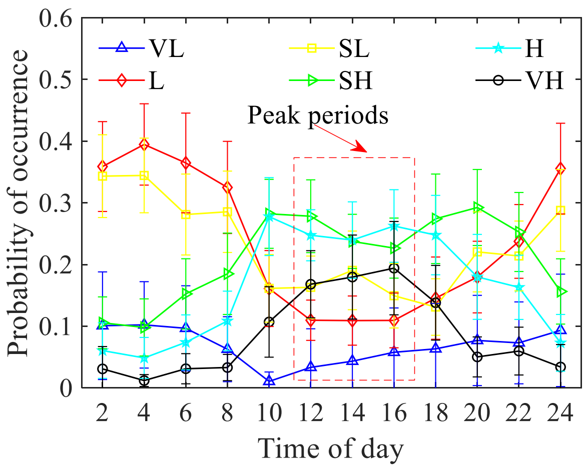

Fig. 14. Probabilities of different traffic complexity patterns against the time of day. The horizontal axis with label 2 represents 00:00-01:59.

To identify the dynamic evolutionary characteristics of the traffic complexity patterns, the life cycle of the complexity patterns is investigated. This is defined as the existing duration of a traffic complexity pattern from occurrence to transfer to another pattern. According to Fig. 15, a high traffic complexity pattern is associated with a short lifecycle. This is primarily attributed to the fact that the higher the traffic complexity pattern, the more frequent the sailing activities, and the lower the predictability of the traffic situation. Thus, this analysis reveals the relationship between traffic complexity patterns and the life cycle and may facilitate the implementation of risk management measures.

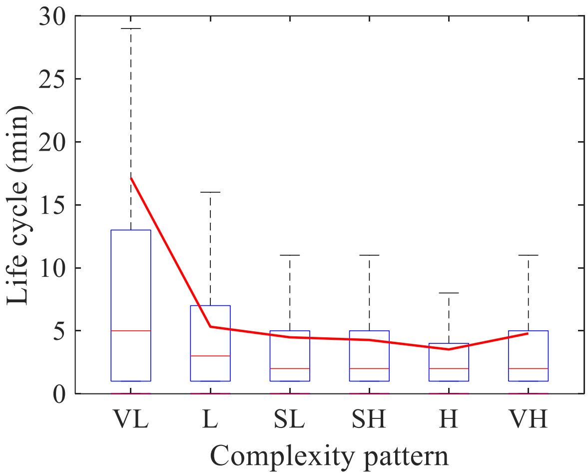

Fig. 15. Life cycle of different complexity patterns.

Furthermore, the transition probability of traffic complexity patterns is explored to fully understand the evolving mechanism in traffic situations. As listed in Table 1, two exciting phenomena are observed. First, one pattern has the highest proportion of keeping the same complexity pattern in the next moment, followed by the transfer to its nearest patterns, whereas jump transitions are less likely to occur. This implies that the ship traffic situation will evolve steadily in most cases but may occasionally show dramatic variation. Second, a higher-complexity pattern corresponds to a larger probability of transferring to other complexity patterns. This may be explained by the fact that maritime operators generally take measures to mitigate potential collision risk when facing complex traffic situations. Hence, high-complexity patterns will not last too long. These investigations provide a prerequisite for future traffic complexity pattern prediction and will enhance proactive traffic management.

Table 1. Transition probabilities between different complexity patterns.

Pattern	Very Low	Low	Slight low	Slight high	High	Very high
Very Low	94.2%	5.2%	0.4%	0.1%	0.1%	0.1%
Low	1.5%	81.2%	14.2%	3.0%	0.0%	0.0%
Slight low	0.1%	15.3%	77.7%	6.1%	0.7%	0.1%
Slight high	0.0%	2.7%	7.3%	76.6%	13.1%	0.2%
High	0.0%	0.0%	1.0%	16.9%	71.6%	10.4%
Very high	0.0%	0.0%	0.4%	0.4%	20.1%	79.1%

In addition, Fig. 16 shows the statistical relationships between the number of ships and traffic complexity. Both the complexity patterns and complexity values are significantly positively correlated with the number of ships. However, high traffic complexity patterns do not necessarily mean high traffic density. This can be proved by observing Fig. 16 (b), where the number of ships is in the range of 15-42 when the traffic situations have very high (VH) complexity patterns. In other words, using one metric, such as the number of ships, is insufficient to capture the entire picture of traffic complexity. In addition, it is found that an indicator H^2 greater than 0.27 usually implies a very high complexity pattern. Therefore, it can be set as the threshold to declare high traffic complexity and assist operators in triggering early warnings.

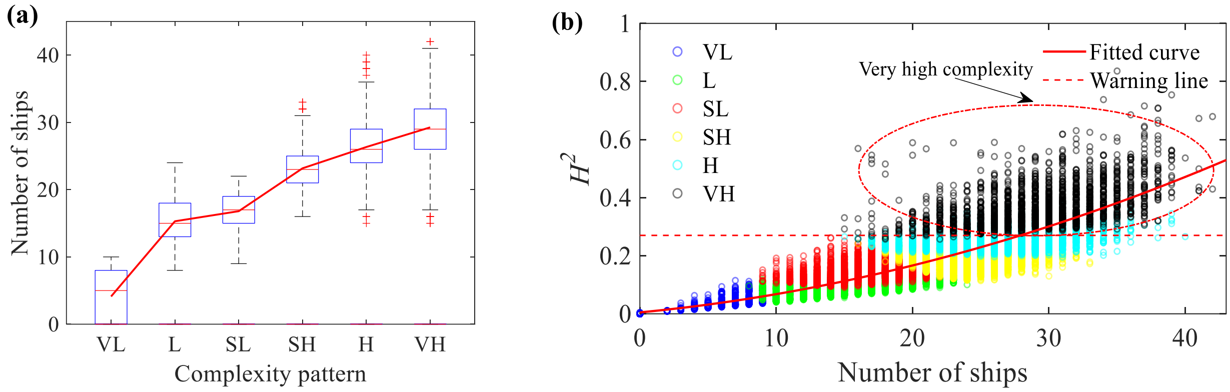

Fig. 16. (a) Boxplot of number of ships to different traffic complexity patterns; (b) relationship between the number of ships and H^2.

4.5 Model validation and comparison analysis

4.5.1 Sensitivity analysis results

Model validation is an essential step to ensure the effectiveness of traffic complexity evaluation results. According to step 7 in Section 3.3.2, this study applies a sensitivity analysis to check the robustness of the new methodology. Table 2 summarises the single-factor change analysis by considering factors including the number of nodes and the number of edges associated with the traffic network structure in the first stage as examples. It can be seen that the increases/decreases of a single factor contribute to the changes in *CD* and *I*. A more considerable change in a single factor leads to a more significant variation in the assessment results. In other words, the model coincides with the first axiom.

Table 2. Model validity test (1).

Change rate of N^1	*CD*	*I*	Change rate of *I*	Change rate of E^1	*CD*	*I*	Change rate of *I*
Prior value	0	74231	-	Prior value	0	74231	-
+10%	+11.86%	77134	+3.91%	+10%	+0.88%	74424	+0.26%
+20%	+22.59%	79882	+7.61%	+20%	+1.76%	74615	+0.52%
-10%	+12.32%	71299	-3.95%	-10%	+0.87%	74041	-0.26%
-20%	+25.55%	68076	-8.29%	-20%	+1.84%	73829	-0.54%

Table 3. Model validity test (2).

Change rate of N^l	/	+10%	+10%	+10%	+10%	+10%
Change rate of E^l	/	/	+10%	+10%	+10%	+10%
Change rate of S^l	/	/	/	+10%	+10%	+10%
Change rate of C^l	/	/	/	/	+10%	+10%
Change rate of KS^l	/	/	/	/	/	+10%
CD	0	+11.86%	+12.72%	+13.66%	+13.68%	+15.00%
I	74231	77134	77326	77536	77540	77836
Change rate of *I*	-	+3.91%	+4.17%	+4.45%	+4.46%	+4.86%

Table 3 further illustrates the multiple factor change analysis. It is found that with the increased number of changed factors, the *CD* continually grows from 0 to 15.00%. Similarly, certain responses related to the index *I* can also be observed. Therefore, the model is in line with the second axiom. These analyses partially validate the feasibility and performance of the proposed methodology.

4.5.2 Comparison analysis with other approaches

Currently, ship traffic complexity evaluation is an emerging topic under development in the field of maritime traffic with increasing traffic in narrow waters, such as ports, and the development of autonomous ships. To demonstrate the superiority and practical usability of the proposed complexity methodology, it is compared with recently developed models related to ship traffic complexity or regional collision risk. Because the parameters in these existing models are difficult to determine and are inapplicable to the selected complex waters, we focus on a theoretical comparison analysis instead of a quantitative performance comparison. According to Table 4, all of these existing models are specific to open waters. Regional traffic complexity analysis that is adaptive to dynamic and uncertain ship traffic in complex and restricted waters has never been approached. Specifically, they estimate the converging/diverging trends or potential conflicts among ships based on the assumption that the ships maintain their course and speed during the look-ahead period, which hinders their application in nonlinear ship motion cases. Furthermore, one model (Sui et al., 2020) used complex network theory to describe the topological characteristics of ship traffic networks but failed to account for traffic's dynamic evolving properties. Hence, it is difficult to reveal how urgent different traffic scenarios require immediate action for collision avoidance. However, the proposed methodology fully considers the dynamic and uncertain ship motion behaviours and the structural and evolutionary features related to the conflict interactions among ships, which therefore adapts to more general traffic scenarios and provides a better understanding than other relevant methods.

Table 4. Comparison with other ship traffic complexity or regional collision risk models.

Methods	Research focus	Model handling dynamic and uncertain	Model adaptive to water areas	Traffic information dimension	Model considering influencing factors

		traffic			
Wen et al. (2015)	Marine traffic complexity	No	Open sea	One (space)	Traffic density, ship dynamic attributes
W. Zhang et al. (2019)	Regional vessel collision risk	No	Open sea	One (space)	Traffic density, ship dynamic attributes and ship domain
Sui et al. (2020)	Marine traffic complexity	No	Open sea	One (space)	Ship dynamic attributes, traffic topological features
Z. Liu et al., (2019)	Regional vessel collision risk	No	Open sea	Two (time and space)	Ship domain, collision avoidance manoeuvre
van Westrenen & Ellerbroek (2015)	Single ship complexity in multi-ship situations	No	Open sea	Two (time and space)	Ship domain, conflict resolution space
Proposed approach	Marine traffic complexity	Yes	Any	Two (time and space)	Ship domain, traffic topological and evolutionary features

5 Conclusion and future research

This study introduces a systematic ship traffic complexity evaluation approach to comprehend the maritime traffic situation and enhance maritime traffic situational awareness. It integrates probabilistic conflict detection, traffic complexity modelling, and evaluation to produce reliable perception outputs, which effectively justifies the methodological complexity caused by the involvement of multiple methods. If applied, the results could contribute to significant changes in the current ship anti-collision best practices. More specifically, the potential encounters among ships across different neighbouring waters in the same region can be considered as early as possible from a traffic network perspective to minimise the possible conflict probability of ships when they encounter the same water locally. It could even present a higher value when the entire shipping industry moves away from manned to autonomous ships. The developed methodology is embedded with the following new features: 1) the dynamics and uncertainty in ship movements are considered to ensure their applicability to various complex traffic scenarios; 2) a multi-stage and multi-topology dynamic traffic network model is developed by extending the application of complex network theory, which fully characterises the structural and dynamic evolving properties of traffic situations; and 3) the employed FCI-based evaluation approach is practical, reliable, and has the ability to present a fine-grained hierarchical description of the traffic complexity patterns rather than being solely based on a single assessment value.

Real data-based experiments are conducted to demonstrate and test the performance of the proposed methodology. The results show that the proposed method can effectively detect ship collision scenarios and allow the traffic complexity patterns to be evaluated hierarchically, and hence could be used in practise to help improve traffic situation awareness. In addition, the complexity method can identify the key influential ships in a complex traffic network, which could facilitate the resolution of traffic conflicts from a strategic level instead of being dependent on the navigators to tackle them locally. In addition, some critical research findings and possible implications are discussed in terms of traffic network heterogeneity analysis; statistics including occurrence frequency, temporal distribution, life cycle, and transition probability of different traffic complexity patterns; and the correlation between the number of ships and traffic complexity levels. The validity and superiority of the new method are further investigated through sensitivity analysis and comparative analysis with existing relevant research. The existing techniques could result in inaccurate estimations of uncertain traffic movements and have difficulty revealing dynamic traffic evolutionary characteristics. In contrast, the developed complexity methodology estimates traffic complexity better by simultaneously incorporating the topological and evolutionary characteristics of traffic networks and ship movement uncertainty. Its methodological contributions could be further demonstrated by its generalisation and possible tailoring for application in all types of traffic scenarios.

Current ship collision avoidance still focuses on ship pair analysis at the local level. However, in high-traffic waters, the countermeasures used by ship *A* to avoid conflict with ship *B* could increase the risk of conflict with the other ships (i.e., *C* and *D*) in the same water. This will force ships to keep manipulating their courses and speeds, which is very costly for both human-crewed and autonomous ships, or mixed traffic. This study shifts the ship conflict risk analysis from the ship pair local to the global/regional level, where the conflict risk of multiple ship encounters can be better managed. Therefore, implementing the proposed approach in an intelligent traffic service system can support maritime traffic situational awareness and traffic safety management in complex waters.

Future research will focus on the following aspects: First, the performance of the proposed model requires further testing before it can be practically applied in complex port waters. This study only checked the traffic complexity evaluation models in traffic scenarios generated using historical AIS data. In the future, an effectiveness evaluation can be conducted by incorporating the proposed models into the current commercial systems for maritime transportation surveillance. Second, the influence of the risk perception difference on probabilistic collision detection, especially for the interaction between large and small ships, could be investigated. When encountering ships that are spatially close to each other, large ships are often subject to high collision avoidance pressure and require an earlier alert. Hence, further efforts will be made to integrate the risk perception difference into the traffic complexity evaluation by constructing directed ship traffic networks. Third, ship manoeuvrability needs to be further factored into traffic complexity estimation in that ships with good manoeuvrability usually have higher capabilities to reduce collision risk. Fourth, the traffic complexity model could be expanded to forecast future traffic complexity peaks using advanced machine learning techniques, such as deep learning and support vector machines. This will facilitate the perception of forthcoming traffic complexity in advance and enhance proactive traffic surveillance capacity. Fifth, the aggregation effect of multi-ship control on traffic complexity mitigation requires further attention and investigation. For example, we assume that the importance of ships A and B in a regional traffic situation is calculated as E_a and E_b, respectively. When ships A and B are simultaneously deleted from the traffic network, the traffic complexity reduction may hold $E_{a,b} > E_a + E_b$. This means that the best solution for conflict risk reduction could be derived from the joint guidance of multiple ships in a traffic network. Therefore, how to efficiently examine the importance of multiple ships in a regional traffic situation and then design appropriate conflict resolution strategies can be researched in the future. Finally, we consider developing an improved model to achieve an adaptive switch between regional and local traffic complexity measurements so that surveillance operators can gain maritime traffic situational awareness at different scales based on their real-time demands.

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.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (Grant No. 52031009), and China Scholarship Council (No. 202006950039). This work is also financially supported by the EU H2020 ERC Consolidator Grant program (TRUST Grant No. 864724).

References

Ahn, J.-H., Rhee, K.-P., & You, Y.-J. (2012). A study on the collision avoidance of a ship using neural networks and fuzzy logic. *Applied Ocean Research*, *37*, 162–173.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509–512.

Bian, T., & Deng, Y. (2018). Identifying influential nodes in complex networks: A node information

dimension approach. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *28*(4), 43109.
Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, *424*(4–5), 175–308.
Bukhari, A. C., Tusseyeva, I., & Kim, Y.-G. (2013). An intelligent real-time multi-vessel collision risk assessment system from VTS view point based on fuzzy inference system. *Expert Systems with Applications*, *40*(4), 1220–1230.
Cao, X., Zhu, X., Tian, Z., Chen, J., Wu, D., & Du, W. (2018). A knowledge-transfer-based learning framework for airspace operation complexity evaluation. *Transportation Research Part C: Emerging Technologies*, *95*, 61–81.
Carro, I. G.-O., Valdés, R. M. A., García, J. M. C., & Comendador, F. G. (2019). The influence of the air traffic network structure on the occurrence of safety events: A data-driven approach. *Safety Science*, *113*, 161–170.
Chai, T., Weng, J., & Xiong, D. Q. (2017). Development of a quantitative risk assessment model for ship collisions in fairways. *Safety Science*, *91*, 71–83.
Chen, P., Huang, Y., Mou, J., & van Gelder, P. (2018). Ship collision candidate detection method: A velocity obstacle approach. *Ocean Engineering*, *170*, 186–198.
Chen, P., Huang, Y., Mou, J., & van Gelder, P. (2019). Probabilistic risk analysis for ship-ship collision: state-of-the-art. *Safety Science*, *117*, 108–122.
Chin, H. C., & Debnath, A. K. (2009). Modeling perceived collision risk in port water navigation. *Safety Science*, *47*(10), 1410–1416.
Delahaye, D., & Puechmorel, S. (2010). Air traffic complexity based on dynamical systems. *49th IEEE Conference on Decision and Control (CDC)*, 2069–2074. IEEE.
Delahaye, D., Puechmorel, S., Hansman, J., & Histon, J. (2004). Air traffic complexity map based on non linear dynamical systems. *Air Traffic Control Quarterly*, *12*(4), 367–388.
Fan, S., Blanco-Davis, E., Yang, Z., Zhang, J., & Yan, X. (2020). Incorporation of human factors into maritime accident analysis using a data-driven Bayesian network. *Reliability Engineering & System Safety*, 107070.
Fang, Z., Yu, H., Ke, R., Shaw, S.-L., & Peng, G. (2018). Automatic identification system-based approach for assessing the near-miss collision risk dynamics of ships in ports. *IEEE Transactions on Intelligent Transportation Systems*, *20*(2), 534–543.
Flener, P., Pearson, J., Ågren, M., Garcia-Avello, C., Celiktin, M., & Dissing, S. (2007). Air-traffic complexity resolution in multi-sector planning. *Journal of Air Transport Management*, *13*(6), 323–328.
Fujii, Y., & Tanaka, K. (1971). Traffic Capacity. *Journal of Navigation*, *24*(04), 543–552. doi: 10.1017/S0373463300022384
Gil, M. (2021). A concept of critical safety area applicable for an obstacle-avoidance process for manned and autonomous ships. *Reliability Engineering & System Safety*, *214*, 107806.
Gil, M., Kozioł, P., Wróbel, K., & Montewka, J. (2022). Know your safety indicator–A determination of merchant vessels Bow Crossing Range based on big data analytics. *Reliability Engineering & System Safety*, *220*, 108311.
Gil, M., Montewka, J., Krata, P., Hinz, T., & Hirdaris, S. (2020). Determination of the dynamic critical maneuvering area in an encounter between two vessels: Operation with negligible environmental disruption. *Ocean Engineering*, *213*, 107709.
Goerlandt, F., Montewka, J., Kuzmin, V., & Kujala, P. (2015). A risk-informed ship collision alert system: framework and application. *Safety Science*, *77*, 182–204.

Hao, S., Zhang, Y., Cheng, S., Liu, R., & Xing, Z. (2018). Probabilistic multi-aircraft conflict detection approach for trajectory-based operation. *Transportation Research Part C: Emerging Technologies*, *95*, 698–712.

He, Y., & Wan, J. (2020). Flood hazard assessment based on fuzzy clustering iterative model and chaotic particle swarm optimization. *Journal of Ambient Intelligence and Humanized Computing*, 1–10.

He, Y., Zhou, J., Kou, P., Lu, N., & Zou, Q. (2011). A fuzzy clustering iterative model using chaotic differential evolution algorithm for evaluating flood disaster. *Expert Systems with Applications*, *38*(8), 10060–10065.

Hernandez-Romero, E., Valenzuela, A., & Rivas, D. (2019). A probabilistic approach to measure aircraft conflict severity considering wind forecast uncertainty. *Aerospace Science and Technology*, *86*, 401–414.

Hu, L., Naeem, W., Rajabally, E., Watson, G., Mills, T., Bhuiyan, Z., … Pekcan, C. (2019). A multiobjective optimization approach for COLREGs-compliant path planning of autonomous surface vehicles verified on networked bridge simulators. *IEEE Transactions on Intelligent Transportation Systems*, *21*(3), 1167–1179.

Huang, Y., Chen, L., Chen, P., Negenborn, R. R., & van Gelder, P. (2020). Ship collision avoidance methods: State-of-the-art. *Safety Science*, *121*, 451–473.

Huang, Y., Chen, L., & van Gelder, P. (2019). Generalized velocity obstacle algorithm for preventing ship collisions at sea. *Ocean Engineering*, *173*, 142–156.

Huang, Y., van Gelder, P., & Wen, Y. (2018). Velocity obstacle algorithms for collision prevention at sea. *Ocean Engineering*, *151*, 308–321.

Iphar, C., Ray, C., & Napoli, A. (2020). Data integrity assessment for maritime anomaly detection. *Expert Systems with Applications*, *147*, 113219.

Jilkov, V. P., Ledet, J. H., & Li, X. R. (2018). Multiple model method for aircraft conflict detection and resolution in intent and weather uncertainty. *IEEE Transactions on Aerospace and Electronic Systems*, *55*(2), 1004–1020.

Kang, L., Meng, Q., & Liu, Q. (2018). Fundamental diagram of ship traffic in the Singapore Strait. *Ocean Engineering*, *147*, 340–354.

Kim, D.-W., Park, J.-S., & Park, Y.-S. (2011). Comparison Analysis between the IWRAP and the ES Model in Ulsan Waterway. *Journal of Navigation and Port Research*, *35*(4), 281–287.

Larson, J., Bruch, M., & Ebken, J. (2006). Autonomous navigation and obstacle avoidance for unmanned surface vehicles. *Unmanned Systems Technology VIII*, *6230*, 623007. International Society for Optics and Photonics.

Lebedev, A. O., Lebedeva, M. P., & Butsanets, A. A. (2021). Could the accident of “Ever Given” have been avoided in the Suez Canal? *Journal of Physics: Conference Series*, *2061*(1), 12127. IOP Publishing.

Lee, H.-J., & Rhee, K. P. (2001). Development of collision avoidance system by using expert system and search algorithm. *International Shipbuilding Progress*, *48*(3), 197–212.

Lee, K., Feron, E., & Pritchett, A. (2009). Describing airspace complexity: Airspace response to disturbances. *Journal of Guidance, Control, and Dynamics*, *32*(1), 210–222.

Lei, P.-R. (2020). Mining maritime traffic conflict trajectories from a massive AIS data. *Knowledge and Information Systems*, *62*(1), 259–285.

Lenart, A. S. (2015). Analysis of collision threat parameters and criteria. *The Journal of Navigation*, *68*(5), 887–896.

Li, B., & Pang, F.-W. (2013). An approach of vessel collision risk assessment based on the D–S evidence

theory. *Ocean Engineering*, *74*, 16–21.
Li, H., Lam, J. S. L., Yang, Z., Liu, J., Liu, R. W., Liang, M., & Li, Y. (2022). Unsupervised hierarchical methodology of maritime traffic pattern extraction for knowledge discovery. *Transportation Research Part C: Emerging Technologies*, *143*, 103856.
Li, H., Liu, J., Yang, Z., Liu, R. W., Wu, K., & Wan, Y. (2020). Adaptively constrained dynamic time warping for time series classification and clustering. *Information Sciences*, *534*, 97–116.
Li, J., Li, X., Lv, N., Yang, Y., Xi, B., Li, M., … Liu, D. (2015). Quantitative assessment of groundwater pollution intensity on typical contaminated sites in China using grey relational analysis and numerical simulation. *Environmental Earth Sciences*, *74*(5), 3955–3968.
Liang, M., Liu, R. W., Li, S., Xiao, Z., Liu, X., & Lu, F. (2021). An unsupervised learning method with convolutional auto-encoder for vessel trajectory similarity computation. *Ocean Engineering*, *225*, 108803.
Liu, D., Zhang, G., Li, H., Fu, Q., Li, M., Faiz, M. A., … Khan, M. I. (2019). Projection pursuit evaluation model of a regional surface water environment based on an Ameliorative Moth-Flame Optimization algorithm. *Ecological Indicators*, *107*, 105674.
Liu, H., Tian, Z., Huang, A., & Yang, Z. (2018). Analysis of vulnerabilities in maritime supply chains. *Reliability Engineering & System Safety*, *169*, 475–484.
Liu, J., Feng, Z., Li, Z., Wang, M., & Wen, L. R. (2016). Dynamic Ship Domain Models for Capacity Analysis of Restricted Water Channels. *Journal of Navigation*, *69*(3), 481–503.
Liu, K., Yuan, Z., Xin, X., Zhang, J., & Wang, W. (2021). Conflict detection method based on dynamic ship domain model for visualization of collision risk Hot-Spots. *Ocean Engineering*, *242*, 110143.
Liu, K., Zhang, J., Yan, X., Liu, Y., Zhang, D., & Hu, W. (2016). Safety assessment for inland waterway transportation with an extended fuzzy TOPSIS. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, *230*(3), 323–333. doi: 10.1177/1748006X16631869
Liu, Q., Pedersen, E., Okazaki, T., Fukuto, J., & Tanaka, K. (2006). Direct perception interface for ship-ship collision avoidance. *2006 IEEE International Conference on Systems, Man and Cybernetics*, *1*, 808–813. IEEE.
Liu, R. W., Liang, M., Nie, J., Lim, W. Y. B., Zhang, Y., & Guizani, M. (2022). Deep learning-powered vessel trajectory prediction for improving smart traffic services in maritime Internet of Things. *IEEE Transactions on Network Science and Engineering*.
Liu, R. W., Liang, M., Nie, J., Yuan, Y., Xiong, Z., Yu, H., & Guizani, N. (2022). STMGCN: Mobile Edge Computing-Empowered Vessel Trajectory Prediction Using Spatio-Temporal Multi-Graph Convolutional Network. *IEEE Transactions on Industrial Informatics*.
Liu, Z., Wu, Z., & Zheng, Z. (2019). A cooperative game approach for assessing the collision risk in multi-vessel encountering. *Ocean Engineering*, *187*, 106175.
Lu, S., Shang, Y., & Li, Y. (2017). A research on the application of fuzzy iteration clustering in the water conservancy project. *Journal of Cleaner Production*, *151*, 356–360.
Lyu, H.-M., Shen, S.-L., Zhou, A., & Zhou, W.-H. (2019). Data in flood risk assessment of metro systems in a subsiding environment using the interval FAHP–FCA approach. *Data in Brief*, *26*, 104468.
Mitici, M., & Blom, H. A. P. (2018). Mathematical models for air traffic conflict and collision probability estimation. *IEEE Transactions on Intelligent Transportation Systems*, *20*(3), 1052–1068.
Öztürk, Ü., Boz, H. A., & Balcisoy, S. (2021). Visual analytic based ship collision probability modeling for ship navigation safety. *Expert Systems with Applications*, *175*, 114755.
Pallottino, L., Feron, E. M., & Bicchi, A. (2002). Conflict resolution problems for air traffic management

systems solved with mixed integer programming. *IEEE Transactions on Intelligent Transportation Systems*, *3*(1), 3–11.
Prandini, M., Hu, J., Lygeros, J., & Sastry, S. (2000). A probabilistic approach to aircraft conflict detection. *IEEE Transactions on Intelligent Transportation Systems*, *1*(4), 199–220.
Prandini, M., Piroddi, L., Puechmorel, S., & Brázdilová, S. L. (2011). Toward air traffic complexity assessment in new generation air traffic management systems. *IEEE Transactions on Intelligent Transportation Systems*, *12*(3), 809–818.
Radanovic, M., Eroles, M. A. P., Koca, T., & Gonzalez, J. J. R. (2018). Surrounding traffic complexity analysis for efficient and stable conflict resolution. *Transportation Research Part C: Emerging Technologies*, *95*, 105–124.
Rong, H., Teixeira, A. P., & Soares, C. G. (2019). Ship trajectory uncertainty prediction based on a Gaussian Process model. *Ocean Engineering*, *182*, 499–511.
Rong, H., Teixeira, A. P., & Soares, C. G. (2021). Spatial correlation analysis of near ship collision hotspots with local maritime traffic characteristics. *Reliability Engineering & System Safety*, *209*, 107463.
Rong, H., Teixeira, A. P., & Soares, C. G. (2022). Maritime traffic probabilistic prediction based on ship motion pattern extraction. *Reliability Engineering & System Safety*, *217*, 108061.
Sridhar, B., Sheth, K. S., & Grabbe, S. (1998). Airspace complexity and its application in air traffic management. *2nd USA/Europe Air Traffic Management R&D Seminar*, 1–6.
Sui, Z., Wen, Y., Huang, Y., Zhou, C., Xiao, C., & Chen, H. (2020). Empirical analysis of complex network for marine traffic situation. *Ocean Engineering*, *214*, 107848.
Szlapczynski, R., & Szlapczynska, J. (2015). A Target Information Display for Visualising Collision Avoidance Manoeuvres in Various Visibility Conditions. *Journal of Navigation*, *68*(06), 1041–1055. doi: 10.1017/S0373463315000296
Szlapczynski, R., & Szlapczynska, J. (2016). An analysis of domain-based ship collision risk parameters. *Ocean Engineering*, *126*, 47–56.
Szlapczynski, R., & Szlapczynska, J. (2017). Review of ship safety domains: Models and applications. *Ocean Engineering*, *145C*, 277–289.
van Westrenen, F., & Ellerbroek, J. (2015). The effect of traffic complexity on the development of near misses on the North Sea. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *47*(3), 432–440.
Wang, C. H. J., Tan, S. K., & Low, K. H. (2019). Collision risk management for non-cooperative UAS traffic in airport-restricted airspace with alert zones based on probabilistic conflict map. *Transportation Research Part C: Emerging Technologies*, *109*, 19–39.
Wang, H., Song, Z., Wen, R., & Zhao, Y. (2016). Study on evolution characteristics of air traffic situation complexity based on complex network theory. *Aerospace Science and Technology*, *58*, 518–528.
Wang, N. (2013). A novel analytical framework for dynamic quaternion ship domains. *Journal of Navigation*, *66*(02), 265–281. doi: 10.1017/S0373463312000483
Wang, Q., & Yang, X. (2020). Investigating the sustainability of renewable energy–An empirical analysis of European Union countries using a hybrid of projection pursuit fuzzy clustering model and accelerated genetic algorithm based on real coding. *Journal of Cleaner Production*, *268*, 121940.
Wang, Y., & Chin, H.-C. (2016). An Empirically-Calibrated Ship Domain as a Safety Criterion for Navigation in Confined Waters. *Journal of Navigation*, *69*(2), 257–276.
Wen, X., Tu, C., & Wu, M. (2018). Node importance evaluation in aviation network based on “No Return” node deletion method. *Physica A: Statistical Mechanics and Its Applications*, *503*, 546–559.
Wen, Y., Huang, Y., Zhou, C., Yang, J., Xiao, C., & Wu, X. (2015). Modelling of marine traffic flow

complexity. *Ocean Engineering*, *104*, 500–510.

Weng, J., Meng, Q., & Qu, X. (2012). Vessel Collision Frequency Estimation in the Singapore Strait. *Journal of Navigation*, *65*(2), 207–221.

Weng, J., & Shan, X. (2015). Ship Collision Frequency Estimation in Port Fairways: A Case Study. *Journal of Navigation*, *68*(3), 602–618.

Wong, A., Tan, S., Chandramouleeswaran, K. R., & Tran, H. T. (2020). Data-driven analysis of resilience in airline networks. *Transportation Research Part E: Logistics and Transportation Review*, *143*, 102068.

Wu, D., Yan, D.-H., Yang, G.-Y., Wang, X.-G., Xiao, W.-H., & Zhang, H.-T. (2013). Assessment on agricultural drought vulnerability in the Yellow River basin based on a fuzzy clustering iterative model. *Natural Hazards*, *67*(2), 919–936.

Wu, X., Mehta, A. L., Zaloom, V. A., & Craig, B. N. (2016). Analysis of waterway transportation in Southeast Texas waterway based on AIS data. *Ocean Engineering*, *121*, 196–209.

Xiao, Z., Fu, X., Zhang, L., & Goh, R. S. M. (2019). Traffic pattern mining and forecasting technologies in maritime traffic service networks: A comprehensive survey. *IEEE Transactions on Intelligent Transportation Systems*, *21*(5), 1796–1825.

Xin, X., Liu, K., Yang, X., Yuan, Z., & Zhang, J. (2019). A simulation model for ship navigation in the "Xiazhimen" waterway based on statistical analysis of AIS data. *Ocean Engineering*, *180*, 279–289.

Xin, X., Liu, K., Yang, Z., Zhang, J., & Wu, X. (2021). A probabilistic risk approach for the collision detection of multi-ships under spatiotemporal movement uncertainty. *Reliability Engineering & System Safety*, 107772.

Xu, Yitian, Xi, W., Lv, X., & Guo, R. (2012). An improved least squares twin support vector machine. *Journal of Information and Computational Science*, *9*(4), 1063–1071.

Xu, Yueting, Chen, H., Heidari, A. A., Luo, J., Zhang, Q., Zhao, X., & Li, C. (2019). An efficient chaotic mutative moth-flame-inspired optimizer for global optimization tasks. *Expert Systems with Applications*, *129*, 135–155.

Yu, H., Fang, Z., Murray, A. T., & Peng, G. (2019). A Direction-Constrained Space-Time Prism-Based Approach for Quantifying Possible Multi-Ship Collision Risks. *IEEE Transactions on Intelligent Transportation Systems*, *22*, 131–141.

Yu, H., Murray, A. T., Fang, Z., Liu, J., Peng, G., Solgi, M., & Zhang, W. (2021). Ship Path Optimization That Accounts for Geographical Traffic Characteristics to Increase Maritime Port Safety. *IEEE Transactions on Intelligent Transportation Systems*, *23*, 5765–5776.

Yu, Q., Liu, K., Chang, C.-H., & Yang, Z. (2020). Realising advanced risk assessment of vessel traffic flows near offshore wind farms. *Reliability Engineering & System Safety*, *203*, 107086.

Yu, Q., Liu, K., Teixeira, A. P., & Soares, C. G. (2020). Assessment of the Influence of Offshore Wind Farms on Ship Traffic Flow Based on AIS Data. *The Journal of Navigation*, *73*(1), 131–148.

Zhang, L., & Meng, Q. (2019). Probabilistic ship domain with applications to ship collision risk assessment. *Ocean Engineering*, *186*, 106130.

Zhang, L., Meng, Q., & Fwa, T. F. (2019). Big AIS data based spatial-temporal analyses of ship traffic in Singapore port waters. *Transportation Research Part E: Logistics and Transportation Review*, *129*, 287–304.

Zhang, M., Kujala, P., & Hirdaris, S. (2022). A machine learning method for the evaluation of ship grounding risk in real operational conditions. *Reliability Engineering & System Safety*, 108697.

Zhang, M., Montewka, J., Manderbacka, T., Kujala, P., & Hirdaris, S. (2021). A big data analytics method for the evaluation of ship-ship collision risk reflecting hydrometeorological conditions. *Reliability*

Engineering & System Safety, *213*, 107674.
Zhang, M., Zhang, D., Fu, S., Kujala, P., & Hirdaris, S. (2022). A Predictive Analytics Method for Maritime Traffic Flow Complexity Estimation in Inland Waterways. *Reliability Engineering & System Safety*, 108317.
Zhang, M., Zhang, D., Fu, S., Yan, X., & Goncharov, V. (2017). Safety distance modeling for ship escort operations in Arctic ice-covered waters. *Ocean Engineering*, *146*, 202–216.
Zhang, W., Feng, X., Goerlandt, F., & Liu, Q. (2020). Towards a Convolutional Neural Network model for classifying regional ship collision risk levels for waterway risk analysis. *Reliability Engineering & System Safety*, *204*, 107127.
Zhang, W., Feng, X., Qi, Y., Feng, S., Zhang, Y., & Wang, Y. (2019). Towards a model of regional vessel near-miss collision risk assessment for open waters based on AIS data. *The Journal of Navigation*, *72*(6), 1449–1468.
Zhang, W., Goerlandt, F., Kujala, P., & Wang, Y. (2016). An advanced method for detecting possible near miss ship collisions from AIS data. *Ocean Engineering*, *124*, 141–156.
Zhang, W., Goerlandt, F., Montewka, J., & Kujala, P. (2015). A method for detecting possible near miss ship collisions from AIS data. *Ocean Engineering*, *107*, 60–69.
Zhao, L., Shi, G., & Yang, J. (2018). Ship trajectories pre-processing based on AIS data. *The Journal of Navigation*, *71*(5), 1210–1230.
Zhao, Y., Li, W., & Shi, P. (2016). A real-time collision avoidance learning system for Unmanned Surface Vessels. *Neurocomputing*, *182*, 255–266.
Zhen, R., Riveiro, M., & Jin, Y. (2017). A novel analytic framework of real-time multi-vessel collision risk assessment for maritime traffic surveillance. *Ocean Engineering*, *145*, 492–501.
Zou, Q., Liao, L., Ding, Y., & Qin, H. (2019). Flood classification based on a fuzzy clustering iteration model with combined weight and an immune grey wolf optimizer algorithm. *Water*, *11*(1), 80.