Journal of *Marine Science and Engineering*

MDPI

Article

Research on the Pattern and Evolution Characteristics of Global Dry Bulk Shipping Network Driven by Big Data

Haijiang Li [1,2,*], Xin Zhang [1,2], Peng Jia [1,2] and Qianqi Ma [1,2]

[1] School of Maritime Economics and Management, Dalian Maritime University, Dalian 116026, China; zhangxin922@dlmu.edu.cn (X.Z.); jiapeng@dlmu.edu.cn (P.J.); maqianqi@dlmu.edu.cn (Q.M.)
[2] Collaborative Innovation Center for Transport Studies, Dalian Maritime University, Dalian 116026, China
* Correspondence: haijiang@dlmu.edu.cn

Abstract: The dry bulk shipping network is an important carrier of global bulk commodity flow. To better understand the structural characteristics and future development trends of the global dry bulk shipping network (GDBSN), this study proposes a framework for characteristics analysis and link prediction based on complex network theory. The study integrates large-scale heterogeneous data, including automatic identification system data and port geographic information, to construct the GDBSN. The findings reveal that the network exhibits small-world properties, with the Port of Singapore identified as the most influential node. Link prediction results indicate that many potential new shipping routes exist within regions or between neighboring countries, exhibiting clear regional clustering characteristics. The added links mainly influence the local structure, with minimal impact on the overall network topology. This study provides valuable insights for shipping companies in route planning and for port authorities in developing strategic plans.

Keywords: shipping network; dry bulk shipping; automatic identification system; complex network; link prediction

check for updates

Academic Editors: Hui Shan Loh and Qingji Zhou

Received: 13 December 2024
Revised: 14 January 2025
Accepted: 14 January 2025
Published: 16 January 2025

Citation: Li, H.; Zhang, X.; Jia, P.; Ma, Q. Research on the Pattern and Evolution Characteristics of Global Dry Bulk Shipping Network Driven by Big Data. *J. Mar. Sci. Eng.* **2025**, *13*, 147. https://doi.org/10.3390/jmse13010147

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Currently, maritime shipping is the most important mode of transportation in international trade, accounting for nearly 80% of global trade volume. The volume of seaborne trade is mainly based on dry bulk and oil transport, followed by container trade [1]. Dry bulk shipping accounts for over 40% of the global seaborne trade, representing the largest share compared to container, tanker, and other specialized ship types. Although the volume of dry bulk cargo in global shipping accounts for a large proportion, most of the existing studies focus on the field of container transportation, and there are relatively few related studies on dry bulk shipping [2], which needs more attention in academic research.

Dry bulk shipping also plays a vital role in the global economy. Iron ore and coal are the two main types of dry bulk cargo. In 2023, the volume of iron ore transported accounted for 28% of the global dry bulk shipping total, while coal transportation accounted for nearly 24%. These commodities are not only essential raw materials for global industrial production but also integral to energy supply, with their transportation needs closely tied to the stability of the global economy. Instability (due to conflicts, acts of aggression, etc.) can affect the economic activities of both the involved countries as well as those dependent on them [3]. Therefore, in-depth research on the current state of the dry bulk shipping network and its future evolution is essential for understanding changes in global trade patterns and ensuring global economic stability.

AIS data provide data support for the construction of shipping networks. The International Maritime Organization (IMO) stipulates that all passenger ships and cargo ships over 300 gross tons must be equipped with an automatic identification system (AIS). This regulation provides the foundation for the use of AIS data in maritime traffic analysis [4]. AIS data aggregate key information, such as the maritime mobile service identity (MMSI), speed, heading, and position [5]. Based on the characteristics of AIS data, numerous studies have been conducted on themes in the field of maritime traffic, such as ship trajectory prediction [6], collision avoidance [7], and maritime traffic safety assessment [8].

As a complex network, the dry bulk shipping network can be analyzed for its structure using complex network theory. Complex network theory employs mathematical tools such as graph theory and topology to study and analyze complex systems. It can reveal the behavior and essential properties of complex systems and has become the main method for modeling and analyzing complex systems [9]. Many scholars have applied complex network theory to study the maritime network, and the research scope has been gradually expanded from the regional level to the global sea area, analyzing the spatial patterns and network characteristics of shipping networks from both static and dynamic perspectives [10]. However, research on dry bulk shipping networks remains insufficient, and further investigation is needed to enhance the understanding of the structural characteristics of these networks.

In the context of global market fluctuations, the structure and route layout of the dry bulk shipping network may change. While existing studies primarily focus on describing and analyzing the current state of these networks, there is relatively little discussion about their future evolution. Link prediction can use existing topological structure characteristics and attribute information of each node to predict the possible future edge relationships, and it can also predict edge relationships between nodes that have already connected but have not been found yet [11]. Link prediction is often used in social networks to recommend potential friends and achieve personalized social recommendations [12]. Additionally, link prediction can be used to evaluate the accuracy and effectiveness of network evolution models [13]. Based on this, combined with the historical route data between ports in the dry bulk shipping network, link prediction can predict new routes that may appear in the future, thus providing valuable suggestions for shipping companies in route planning and optimization.

Overall, this study aims to investigate the fundamental topological characteristics of the GDBSN and introduces link prediction algorithms to explore potential evolutionary directions of the network. This study fills the research gap in the field, particularly in terms of a deeper exploration of the network's future evolution, and it provides an innovative perspective for understanding the dynamic changes in dry bulk shipping networks. The contributions of this study are outlined as follows:

(1) A comprehensive research framework for the GDBSN is proposed in this study, based on large-scale vessel trajectory data, that integrates "OD network construction, topological structure analysis, and link evolution prediction". A method for constructing the GDBSN based on AIS data is also introduced, and it can provide theoretical support for the analysis of global bulk commodity trade patterns and evolutionary trends.
(2) The topological characteristics of the GDBSN are analyzed in this study based on complex network theory. Through the analysis of complex network metrics and trade link strength, the connectivity features, key nodes, and major shipping routes of the GDBSN are revealed. The dynamic patterns of the global dry bulk shipping trade are also explained in this study, providing theoretical support for the optimization of global trade resource allocation.

(3) Link prediction is introduced in this study to explore the dynamic evolution of the GDBSN. A new route development mechanism is proposed based on historical vessel route data, utilizing link prediction indices. This expands the application of link prediction theory in maritime transport and provides theoretical and decision-making support for the planning and design of new shipping routes.

The rest of this paper is organized as follows: Section 2 reviews the research on shipping networks and link prediction. Section 3 presents the research methodology framework for the GDBSN. Section 4 presents the topological structure of the GDBSN and the results of link prediction. Section 5 presents the study's conclusions.

2. Literature Review

2.1. Shipping Network Studies Based on Complex Network Theory

2.1.1. Construction of Shipping Networks

The primary task of shipping network research, which is based on complex network theory, is to construct shipping networks. It is very important to complete the construction of the shipping network accurately and effectively. At present, there are three methods for constructing shipping networks, each of which has its own characteristics and provides a unique perspective for different research fields.

The first method is to construct a shipping network based on origin–destination (OD) data [14]. This method is based on ship schedules and AIS data. Ports serve as network nodes, and the departure and arrival records of each ship are extracted as OD data to serve as network links, thereby constructing a shipping OD network. In addition, the frequency of shipping routes or ship deadweight can be used as edge weights. The shipping OD network can provide an intuitive representation of the overall distribution pattern [9]. Furthermore, the OD network can reflect the importance of different routes through the weights of the edges [15]. However, the shipping OD network requires a high level of data completeness and accuracy, especially in the case of missing or incomplete data, which may lead to the distortion of network nodes and connecting edge characterization, thus affecting the accuracy of the analysis results. Therefore, ensuring data comprehensiveness and accuracy is the primary prerequisite for constructing shipping OD networks in practical applications.

Secondly, another method is to construct a maritime traffic network based on ship navigation trajectories [16]. Unlike the OD network that only connects port nodes, this method identifies key waypoints as network nodes and defines shipping routes as edges, thereby constructing a maritime traffic network based on actual ship trajectories [17]. This traffic network can more intricately depict the specific sailing routes of vessels and their navigation patterns [18]. However, this method is limited by noise in trajectory data and uneven data density, leading to lower accuracy in network node identification and increasing data processing requirements. In addition, due to the complex structure of this network, complex network theory is difficult to be directly applied. Therefore, the maritime traffic network is more suitable for local path optimization and specific route analysis than for complex network structural analysis.

Finally, the third method is to construct a maritime trade network based on trade data between countries and regions [19]. This network is constructed by modeling the flow of maritime trade goods, with trade volume serving as the weight value of the edges. The maritime trade network reveals the overall structure of global trade from a macro perspective and highlights the positions and interdependencies of countries in maritime trade. This makes it useful for analyzing the structural characteristics of global supply chains and for identifying major trade routes. However, the maritime trade network neglects actual shipping routes, limiting its ability to accurately capture the micro-level

characteristics of maritime networks. Therefore, the maritime trade network is more suitable for large-scale strategic analysis than for detailed shipping logistics research.

In summary, this study finds the method of constructing a shipping network based on OD data to be more suitable. Compared with the traffic network and the trade network, the OD network more effectively represents the direct transportation relationship between ports and accurately reflects the spatial distribution of each port node and route. Meanwhile, the OD network facilitates the analysis of network structure and the assessment of the importance of nodes and edges within the network.

2.1.2. Applications of Complex Network Theory in Shipping Networks

The application of complex network theory has penetrated many different fields. In the field of transportation, complex network theory has been applied in transportation systems such as aviation [20], railways [21], maritime shipping [22], and roads [23]. Specifically in maritime shipping, many researchers have used complex network theory to analyze the topological properties, network structure, and robustness of shipping networks. Existing research on maritime networks can be broadly categorized into studies focusing on container shipping networks, crude oil shipping networks, dry bulk shipping networks, and comprehensive analyses of multi-layer networks.

(1) Container shipping networks

Kaluza et al. [24] initially indicated that complex network theory could be utilized in the study of container shipping networks. Subsequently, scholars began to employ complex network theory to study the topological structure of container shipping networks. For example, Jiang et al. [25] constructed the container shipping network of the Maritime Silk Road, analyzed its topological attributes by using complex network theory, and found that the network exhibits small-world and scale-free properties. Kang et al. [26] analyzed the global liner shipping network using complex network theory and found that both in-degree and out-degree distributions follow power-law distributions. Studies of the topology of container shipping networks have revealed their small-world and scale-free characteristics, and their topology is gradually becoming more well-known. Additionally, more scholars have applied complex network theory to analyze the connectivity [27], robustness [28], and vulnerability [29] of container shipping networks. For instance, Li et al. [30] analyzed the container network before and after public health events and found that hub ports on major shipping routes and cross-regional long-distance routes were vulnerable to emergencies. However, the global container shipping network's configuration remained less affected and more robust than before the event. Moreover, Li et al. [31] reconstructed the global container shipping network using complex network theory. The results indicated that only minor adjustments were needed to enhance the vector centrality of the network. To further enhance the efficiency and sustainability of container shipping networks, researchers have also started to incorporate the multi-port berth allocation problem into their studies [32].

(2) Crude oil shipping network

Currently, the application of complex network theory to the crude oil shipping network primarily focuses on identifying key issues, such as network topological structure characteristics, vulnerability, and resilience assessment. Yan et al. [33] constructed a global oil trade network based on AIS data, revealing the network's typical small-world effect, where a few ports have highly concentrated trade connections. Scholars have also applied complex network theory to assess the robustness and vulnerability of the crude oil network under external shocks. For instance, Wang et al. [34], in their study of China's imported oil shipping network, demonstrated that deliberate attacks inflict much greater damage than random attacks and revealed significant differences in the reliability of different routes.

Mou et al. [35] systematically analyzed the resilience of the crude oil shipping network and found that the network's topological structure significantly affects its resilience: network density and centrality negatively correlate with resilience, while connectivity and scale positively correlate with it. In addition, the influence of oil price volatility on the maritime network structure and traffic dynamics has also been studied. Yu et al. [36] showed that oil price fluctuations alter the network structure of oil import-dependent and export-dependent countries and significantly change traffic flow.

(3) Multi-layer shipping network

Scholars using complex network theory in shipping networks have broadened their focus from single-type to multi-layer network structures and characteristics. Kaluza et al. [24] found that the global cargo shipping network (container, oil, and bulk cargo) exhibits small-world properties with broad degree and weight distributions. Peng et al. [37] analyzed the robustness of container, bulk carrier, and oil tanker networks. The study found that the container network has the worst robustness, while the bulk carrier network has the strongest. Yu et al. [14] developed a spatiotemporal analysis framework of the global maritime network. The research showed that new shipping links between countries such as China, Singapore, Australia, and the UAE have significantly increased traffic flow. Mou et al. [9] analyzed the topological structure of container, tanker, and bulk shipping networks and revealed uneven distribution and regional clustering along the Maritime Silk Road.

Overall, complex network theory has shown broad application prospects and analytical potential in shipping networks. It plays a critical role in revealing network structures, assessing their vulnerability and resilience, and supporting optimization. However, research on dry bulk shipping networks remains limited compared to container and crude oil shipping networks, requiring further studies to enhance understanding of their structural characteristics and dynamic properties.

2.2. Link Prediction Analysis in Shipping Networks

In recent years, link prediction has been applied in shipping network research, particularly in the analysis of network evolution patterns, potential route prediction, and future network structure simulation [38]. Currently, the relevant research mainly focuses on the evolution of container shipping networks and shipping route prediction analysis. Research has shown that new links are typically formed by prioritizing connections between geographically adjacent nodes in shipping networks, thus forming an obvious clustering effect within the region. Meanwhile, adding new links can significantly improve the average degree and clustering coefficient of the network, reduce the average path length and network diameter, and thus improve the overall network connectivity and transportation efficiency, which indicates that the rational layout of new routes has a profound impact on optimizing the network structure [39].

In constructing link prediction models, a series of improved models (e.g., KS-Salton, KS-AA) have been proposed by combining traditional similarity indices (Salton, AA, etc.) with network hierarchical structures (e.g., k-shell decomposition) [39]. In addition, a comprehensive prediction index combining endogenous and attractiveness indices has been proposed to improve the accuracy of predicting the potential shipping routes within the network [40]. Overall, link prediction is capable of revealing the evolution of shipping networks and provides a theoretical basis and decision support for the opening of new routes and the adjustment of network structure. However, current research primarily focuses on the analysis of container shipping networks, and link prediction research on dry bulk shipping networks remains in its early stages. Therefore, link prediction is applied in this study to predict routes in the dry bulk shipping network, aiming to provide valuable decision-making references for shipping companies in planning and establishing routes.

This study finds that research on maritime networks using complex network theory predominantly focuses on container shipping, while crude oil and dry bulk shipping networks receive limited attention. Specifically, independent analyses of dry bulk shipping networks are scarce. While some studies compare container, tanker, and bulk carrier sub-networks, most existing research only describes the current state of dry bulk networks, lacking predictive models for network evolution based on future trends and the forecasting of new route formation. This research aims to fill this gap by developing a research framework for the GDBSN. It examines the structural characteristics of the GDBSN across different scales and explores potential link changes. This will enhance the understanding of the complexity of the GDBSN and provide valuable theoretical insights for future network optimization.

3. Methodology

The research framework for analyzing the GDBSN proposed in this study is presented in Figure 1. Based on complex network theory and link prediction, the framework provides a systematic approach to network analysis. Its goal is to comprehensively reveal the topological characteristics, key nodes, and major links of the GDBSN while also predicting potential shipping routes between ports. This methodology offers a novel perspective and a valuable tool for gaining a deeper understanding of the GDBSN.

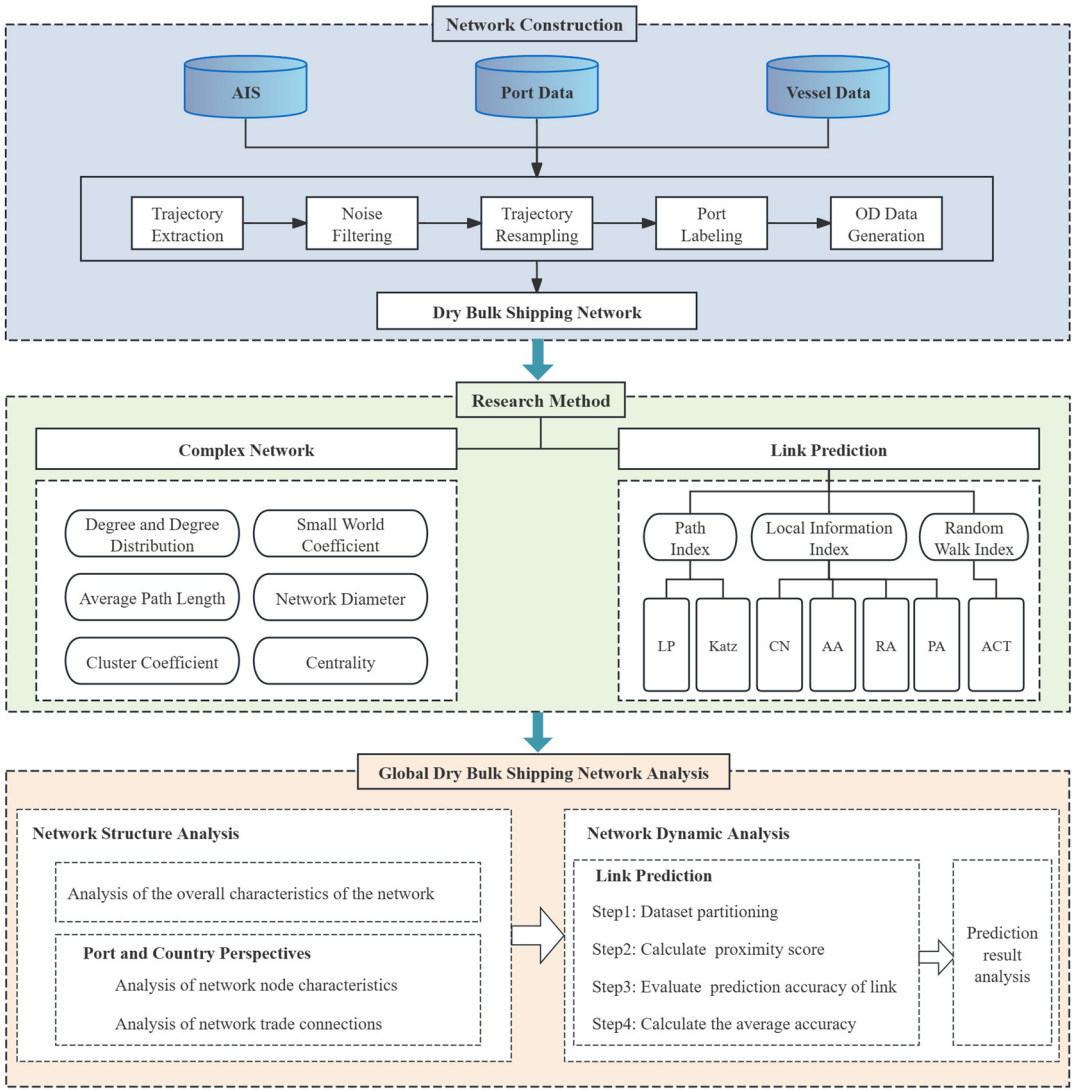


Figure 1. Global dry bulk shipping network analysis framework.

3.1. Data Sources

The core data utilized in this study include AIS data, Lloyd's Register vessel profile data, and global bulk port data. The AIS data utilized in this study comprise voyage records for 10,744 dry bulk vessels that operated worldwide from 1 January 2021 to 31 December 2021. AIS data were provided by Elane Inc. (https://www.elaneglobal.com/) (accessed on 1 March 2024). Figure 2 illustrates the distribution of dry bulk vessel trajectories, which is based on a subset of AIS data. The port data are sourced from the "World Port Index" (27th edition) [41] and from port data provided by Clarkson Research. A total of 1334 bulk ports were selected from port data to identify the origins and destinations of the dry bulk vessels. A visual representation of the global distribution of these ports is provided in Figure 3.

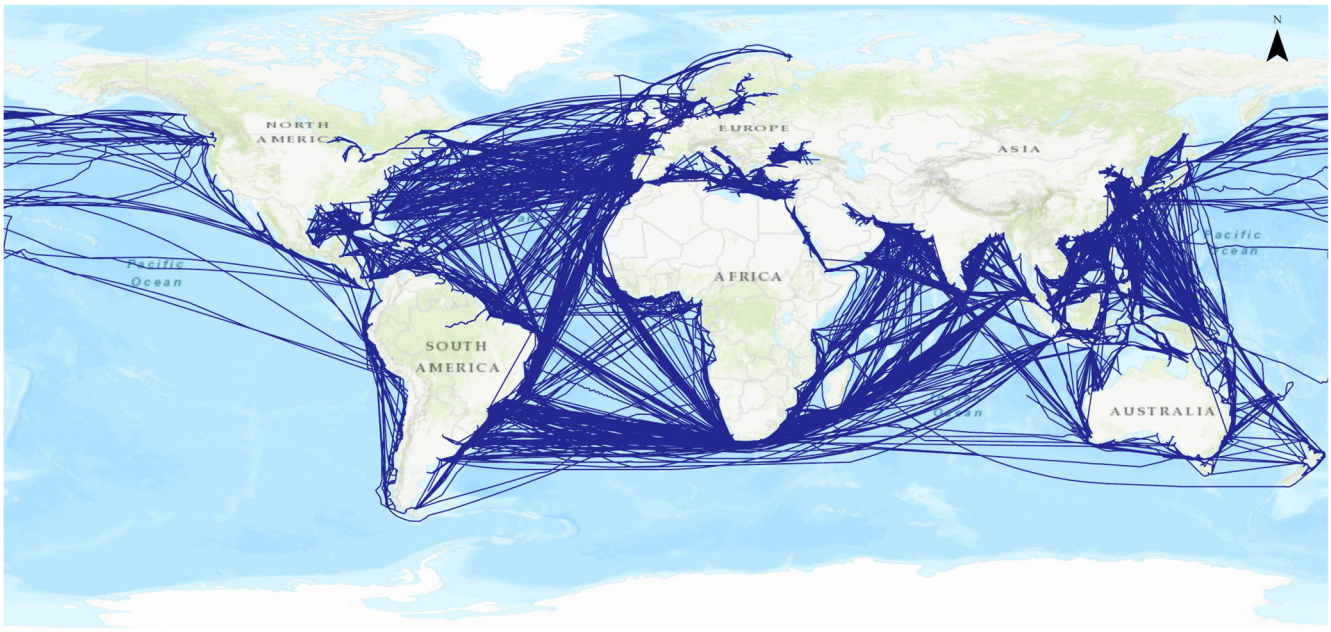


Figure 2. Ship trajectory distribution.

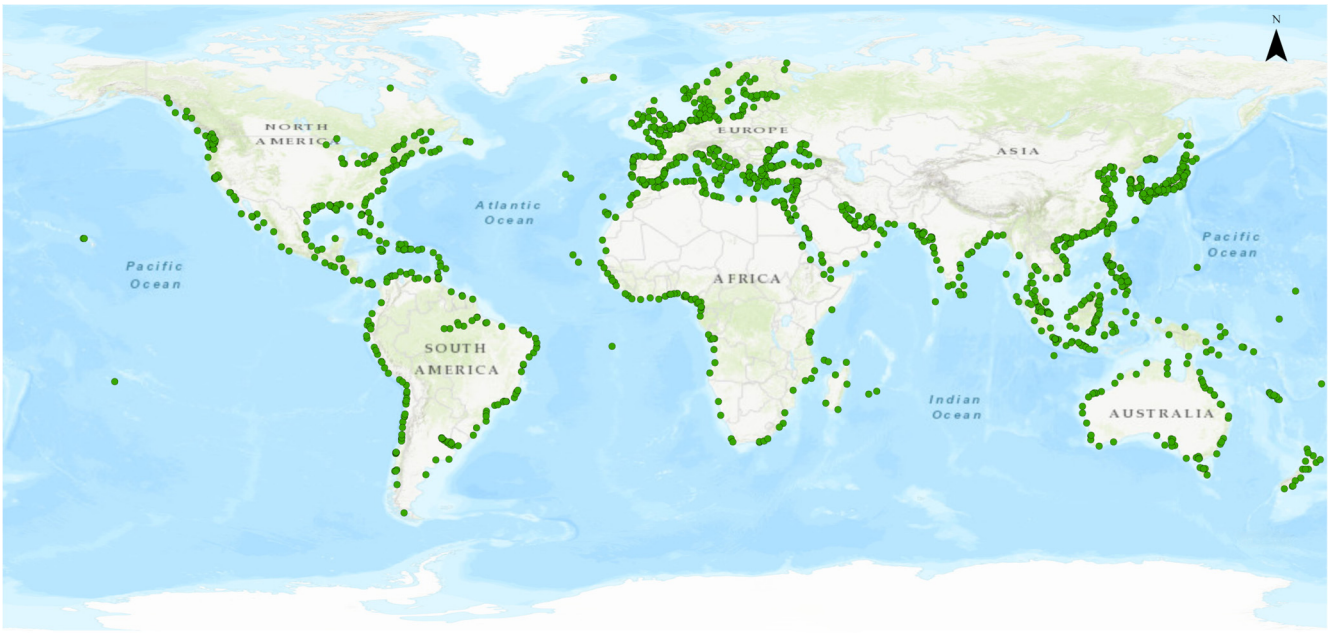


Figure 3. Global dry bulk port distribution.

3.2. Construction of the GDBSN

The primary concept behind the construction of the dry bulk shipping network involves extracting the port entry and exit records of vessels from AIS data and converting each vessel's departure and arrival events into origin–destination (OD) pairs, thereby creating a global OD dataset for dry bulk shipping. Lastly, the Space-L model is applied to construct the GDBSN. The detailed process of constructing the GDBSN is shown in Figure 4.

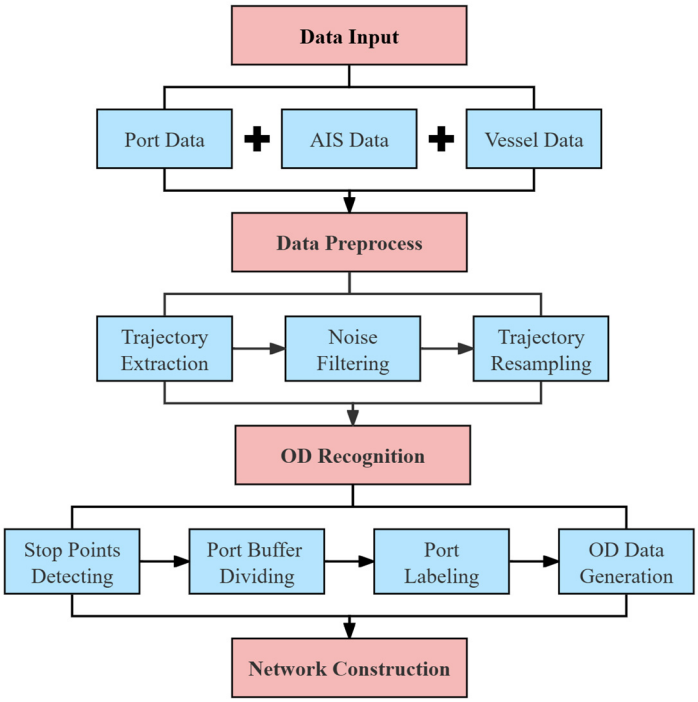


Figure 4. Dry bulk shipping network construction process.

3.2.1. Data Preprocessing

Due to the large volume and inherent noise in AIS data, preprocessing is necessary. In order to address the noise in the AIS data, a median filtering method was applied to denoise the latitude and longitude data, thereby ensuring the accuracy of subsequent analyses. Figure 5a shows the original trajectories, and Figure 5b shows the denoised trajectories. This method effectively removed noise from trajectory data, making subsequent analysis more reliable.

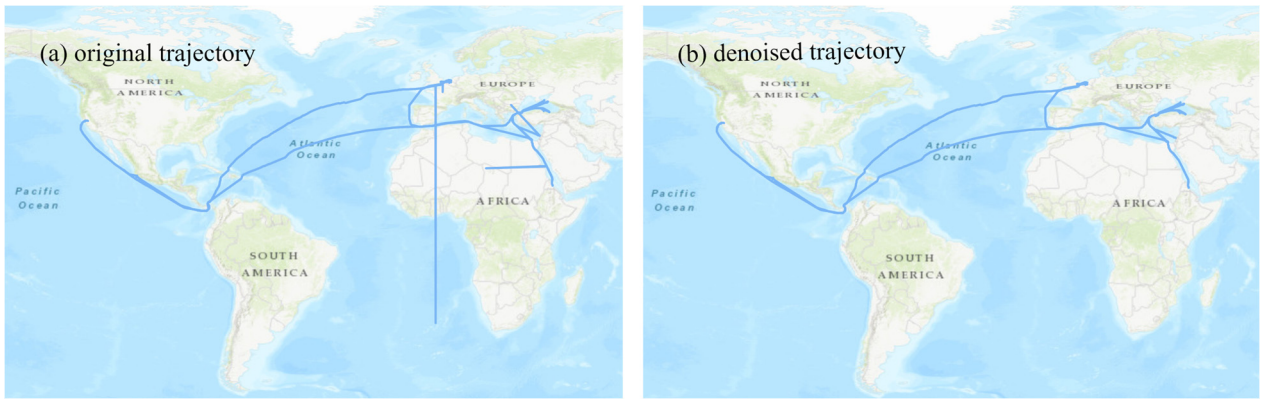


Figure 5. Comparison of ship trajectory denoising effects.

In addition, frequent updates of AIS data often result in consecutive trajectory points containing repetitive or identical information. To improve processing efficiency, a uniform sampling algorithm [42] was employed to compress the denoised data, ensuring that the trajectory shape was preserved while reducing the data volume. The method effectively minimized the data size, thereby enhancing the efficiency of subsequent analyses.

3.2.2. Origin–Destination Recognition

After preprocessing the data, this study further extracted OD data from the AIS data. OD recognition requires port labeling to identify the ports where vessels dock. The port labeling process in this study was completed through the following steps:

(1) Detection of stop points: Stop points occur when a vessel stays at a specific location for a long time. Detecting stop points helps narrow down the vessel's docking locations. In this study, the ST-DBSCAN algorithm [43] was applied to detect stop points.
(2) Division of port buffers: To determine whether a stop point was within a port area, a buffer zone was generated for each port. The location of the buffer zone was based on the port's latitude and longitude, and the radius of the buffer zone was set according to the size of the port to ensure the entire port area was covered.
(3) Port labeling: This step was performed to determine whether a stop point fell within a port area. A spatial intersection between the stop points and the port buffer zones was performed to identify whether a stop point belonged to a particular port. Ultimately, stop points within the port buffer zones were labeled. In this study, the labeling results for a vessel are shown in Figure 6, with blue markers representing the ports visited by the vessel.

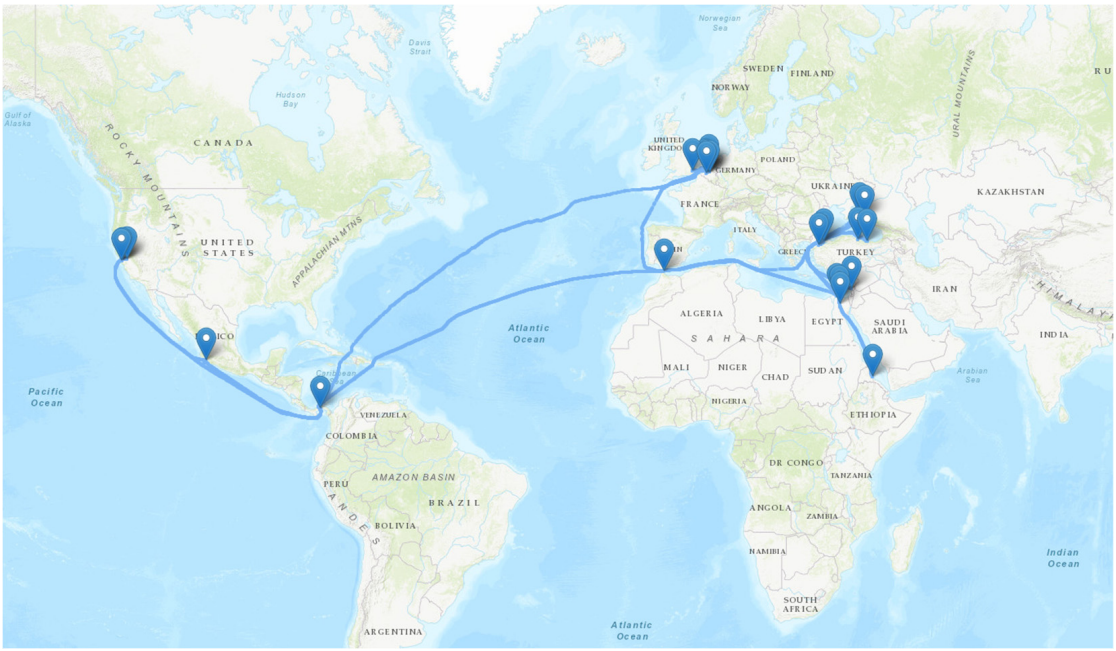


Figure 6. Diagram of port labeling results.

After the completion of port labeling, the OD dataset was generated using the port labeling results. Table 1 presents a sample of the generated OD data, with variable definitions provided in Table 2. Each retrieved entry represents a voyage, including information such as the vessel's departure and arrival port names, port coordinates, and other relevant details.

Table 1. Example of origin–destination (OD) data.

MMSI	S-Port	S-Lng	S-Lat	E-Port	E-Lng	E-Lat	Time
205087000	Liverpool	−3.022	53.447	Odesa	30.764	46.501	12 May 2021 18:28
205096000	Damietta	31.756	31.480	Santo Domingo	−69.864	18.374	17 June 2021 5:10
206039000	Cilacap	109.080	−7.766	Au Prince	−72.408	18.572	13 July 2021 4:38

Table 2. Definitions of the variables.

Parameter	Description
MMSI	Maritime mobile service identity
S-Port	Vessel origin port name
S-Lng	Longitude of vessel origin port
S-lat	Latitude of vessel origin port
E-Port	Vessel destination port name
E-lng	Longitude of vessel destination port
E-lat	Latitude of vessel destination port
Time	Vessel arrival time

3.2.3. Network Construction

Two primary approaches are commonly used in existing studies on shipping network modeling: the Space-L model and the Space-P model [44]. Both models define network nodes similarly, representing ports as nodes in the network. However, they differ significantly in the method used to establish edges. The Space-L model treats two ports consecutively visited on the same shipping route as adjacent nodes and creates a connecting edge between them to represent the direct link between ports. In contrast, the Space-P model adopts a broader perspective, considering two ports connected as long as they appear on the same route, regardless of their relative positions. Consequently, the Space-P model reveals all possible connections between ports along a shipping route. Figure 7 illustrates the network topologies for both models. Theoretically, the Space-L model more accurately represents the actual connections between ports, while the Space-P model emphasizes the identification of potential route associations. Since this study focuses on the actual connections between ports, the Space-L model was selected for network construction.

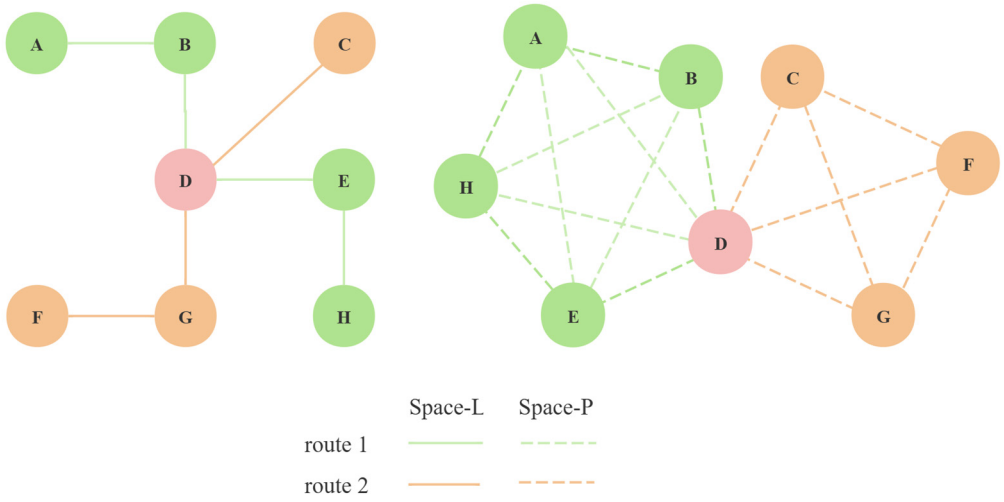


Figure 7. Network topologies for the Space-L and Space-P models.

3.3. Analytical Methods

3.3.1. Complex Network Metrics

Complex network theory is a key method for studying shipping networks. This approach treats ports as nodes and shipping routes between ports as edges, thus revealing the topological structure of the shipping network. In this study, complex network metrics are employed to analyze the topological characteristics of an undirected, unweighted dry bulk shipping network. The relevant metrics are provided in Table 3.

Table 3. Complex network metrics and formula.

Metric	Formula	Description	Meaning
Degree [45]	$k_i = \sum_{j=1, i \neq j}^{N} a_{ij}$	a_{ij} is an element of the adjacency matrix, where $a_{ij} = 1$ if there is an edge between node i and node j, and $a_{ij} = 0$ otherwise. N is the total number of nodes in the network.	Defined as the number of nodes directly connected to the node.
Average path length [45]	$L = \frac{2}{N(N-1)} \sum_{i \neq j} d_{ij}$	d_{ij} represents the shortest distance between node i and node j.	Defined as the average distance between any two nodes in the network. It is used to measure the difficulty of connecting network nodes.
Clustering coefficient [45]	$C_i = \frac{2M_i}{N_i(N_i-1)}$	N_i represents the number of nodes connected to node i; M_i represents the actual number of edges between these N_i nodes.	Defined as the possibility that the adjacent nodes of any node are connected. It can measure the aggregation degree of each port in the network.
Network diameter [45]	$D = \max(d_{ij})$	d_{ij} epresents the distance between any two nodes i and j.	Defined as the maximum distance between any two nodes in the network.
Small-World coefficient [46]	$\sigma = \frac{C/L}{C_R/L_R}$	C and L represent the average clustering coefficient and the average shortest path length, respectively. C_R and L_R represent the clustering coefficient and average topological path length of an E-R random network with the same size as the existing network, respectively.	If $\sigma > 1$, the network is a small-world network.
Degree centrality [47]	$C_D(i) = \frac{k_i}{N-1}$	k_i is the degree of node i.	The greater the degree centrality of a node, the more important the node is in the network.
Betweenness centrality [47]	$C_B(i) = \sum_{j \neq i} \frac{n_{jk}(i)}{n_{jk}}$	n_{jk} is the number of shortest paths between nodes j and k; $n_{jk}(i)$ is the number of shortest paths between nodes j and k that pass through the node i.	It represents the number of times that the shortest path between all node pairs passes through a given node.
Closeness centrality [47]	$C_C(i) = \frac{N-1}{\sum_{j=1, j \neq i}^{N} d_{ij}}$	d_{ij} represents the shortest distance between node i and node j.	It measures the ease with which a node can reach other nodes in the network.
Eigenvector centrality [34]	$C_E(i) = \lambda^{-1} \sum_{j=1}^{N} a_{ij} E_j$	λ is the eigenvalue corresponding to the matrix $E = (E_1, E_2, \cdots, E_n)$.	It assumes that the importance of a node is determined by the importance of its neighboring nodes.

3.3.2. Link Prediction Indices

Link prediction is applied to a given network $G(V, E)$, where V represents the set of nodes and E represents the set of edges, to compute the probability of potential links based on the observed network topology. These potential links are subsequently ranked in descending order, facilitating the prediction of the most likely existing links. In the GDBSN, link prediction aims to estimate the likelihood of shipping routes between port pairs that have not yet established connections, based on the known network topology. Link prediction currently includes those based on node attribute similarity, network structure similarity, and likelihood analysis. This study primarily employs methods based on network structure similarity, utilizing the following key similarity indices:

(1) Local similarity indices

i. Common Neighbors (CN)

The basic assumption of the CN index is that two unconnected nodes are more inclined to be connected if they have more common neighbors. The CN index is defined as follows: For node v_x in the network, its neighbor set is denoted as $\Gamma(x)$, and the similarity of two nodes v_x and v_y is defined as their common number of neighbors, as shown in the following formula:

$$s_{xy}^{CN} = |\Gamma(x) \cap \Gamma(y)| \quad (1)$$

where $|\Gamma(x) \cap \Gamma(y)|$ represents the cardinality of the set.

Based on the concept of common neighbors, the influence of node degrees at both ends is incorporated, resulting in the generation of six additional similarity indices from various perspectives, as presented in Table 4.

Table 4. Link prediction indices based on local information.

No.	Index	Formula [4]
1	Salton [48]	$s_{xy}^{Salton} = \frac{\lvert\Gamma(x)\cap\Gamma(y)\rvert}{\sqrt{k_x k_y}}$
2	Jaccard [48]	$s_{xy}^{Jaccard} = \frac{\lvert\Gamma(x)\cap\Gamma(y)\rvert}{\lvert\Gamma(x)\cup\Gamma(y)\rvert}$
3	Sorenson [48]	$s_{xy}^{Sorenson} = \frac{2\times\lvert\Gamma(x)\cap\Gamma(y)\rvert}{k_x+k_y}$
4	Hub promoted index (HPI) [48]	$s_{xy}^{HPI} = \frac{\lvert\Gamma(x)\cap\Gamma(y)\rvert}{min\{k_x,k_y\}}$
5	Hub promoted index (HDI) [48]	$s_{xy}^{HDI} = \frac{\lvert\Gamma(x)\cap\Gamma(y)\rvert}{max\{k_x,k_y\}}$
6	LHN-I [48]	$s_{xy}^{LHN-I} = \frac{\lvert\Gamma(x)\cap\Gamma(y)\rvert}{k_x k_y}$

[4] $\Gamma(x)$ and $\Gamma(y)$ represent the neighbor sets of nodes x and y, respectively, and k_x and k_y represent the degrees of nodes x and y, respectively.

ii. Adamic–Adar (AA)

The AA index considers the degree information of the common neighbors between two nodes. The underlying principle is that common neighbors with smaller degrees contribute more than those with larger degrees. The AA index assigns a weight value to each common neighbor based on its degree, where the weight is the inverse of the logarithm of the degree of the node. The formula is given as follows:

$$s_{xy}^{AA} = \sum_{z\in|\Gamma(x)\cap\Gamma(y)|} \frac{1}{\log k_z} \quad (2)$$

iii. Resource Allocation (RA)

The RA index assumes that each medium has a unit of resources and distributes them equally to its neighbors, so the amount of resource received by node v_y is defined as the similarity between v_x and v_y. The formula is given as follows:

$$s_{xy}^{RA} = \sum_{z \in |\Gamma(x) \cap \Gamma(y)|} \frac{1}{k_z} \tag{3}$$

iv. Preferential Attachment (PA)

According to the PA index, the higher the degrees of two nodes, the greater the likelihood of their connection. The index is defined as follows:

$$s_{xy}^{PA} = k_x k_y \tag{4}$$

(2) Path similarity indices

i. Local Path (LP)

The LP index takes into account the effects of second-order and third-order paths between nodes. The index is defined as follows:

$$s_{xy}^{LP} = (A^2)_{xy} + \alpha \cdot (A^3)_{xy} \tag{5}$$

where α is a variable parameter used to control the local range, and as α decreases, the contribution of third-order paths also decreases. $(A^n)_{xy}$ represents the number of paths of length n between nodes v_x and v_y.

ii. Katz

The Katz index considers all paths in the network. The index is defined as follows:

$$s_{xy}^{Katz} = \sum_{l=1}^{\infty} a^l \cdot \left| paths_{x,y}^{<l>} \right| \tag{6}$$

where $\alpha > 0$ is a tunable parameter that controls the weight of the paths, and $\left| paths_{x,y}^{<l>} \right|$ represents the number of paths of length l that connect nodes v_x and v_y.

(3) Random walk similarity indices

i. Average Commute Time (ACT)

The ACT index defines the average first-passage time $m(x, y)$ as the average number of steps a random walk particle needs to take to travel from node v_x to node v_y. The average commuting time between nodes v_x and v_y is defined as $n(x, y) = m(x, y) + m(y, x) = M(l_{xx}^{+} + l_{yy}^{+} - 2l_{zy}^{+})$. According to the ACT index, the shorter the commuting time between two nodes, the closer the nodes are to each other. The formula is given as follows:

$$s_{xy}^{ACT} = \frac{1}{l_{xx}^{+} + l_{yy}^{+} - 2l_{xy}^{+}} \tag{7}$$

where l_{xy}^{+} represents the element corresponding to the position in the x-th row and y-th column of the network's Laplacian matrix.

The AUC index is used in this study to measure the accuracy of the link prediction algorithm. AUC is one of the most widely used evaluation indices. The underlying principle is to randomly select an edge from the test set and, for each selected edge, randomly choose a non-existent edge. If the score of an edge in the test set exceeds that of a non-existent

edge, 1 point is added; if the two scores are equal, 0.5 points are added. The formula is given as follows:

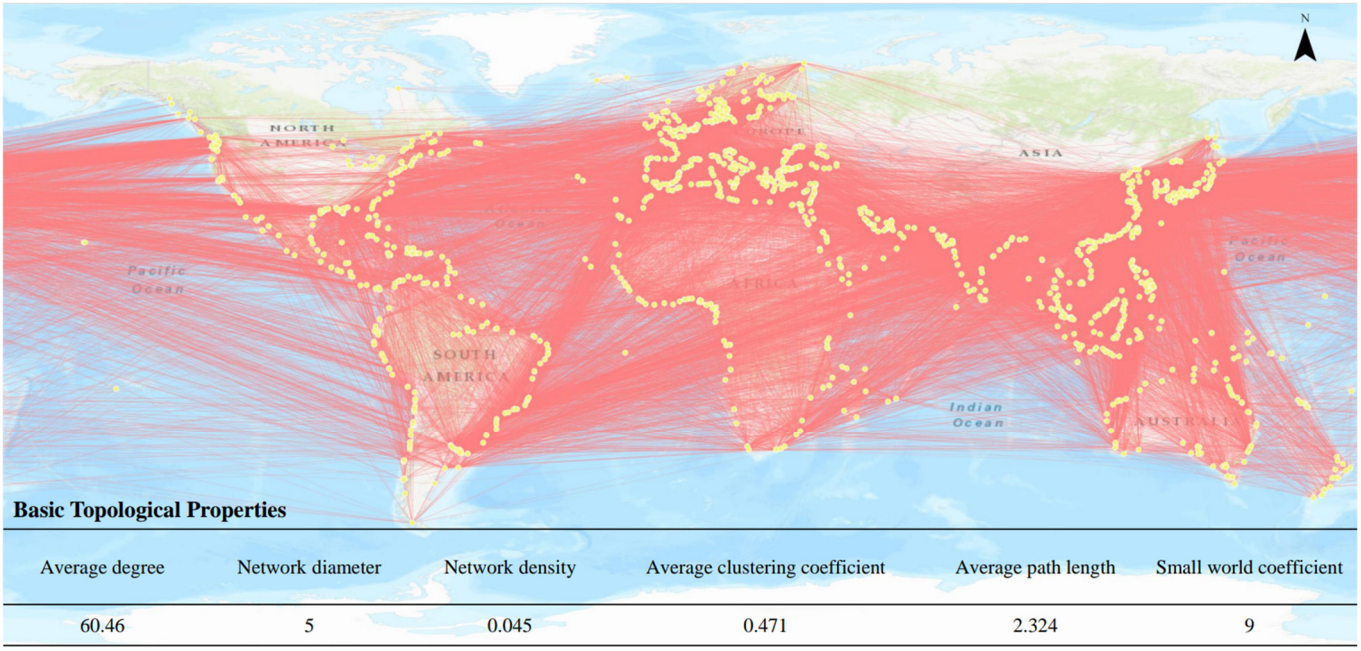
$$AUC = \frac{n' + 0.5 \times n''}{n} \quad (8)$$

where n represents the number of independent comparisons (i.e., the sample size); n' denotes the number of times the score of an edge in the test set exceeds that of a non-existent edge; and n'' represents the number of times the scores are equal. The closer the AUC is to 1, the higher the accuracy of the link prediction.

4. Results and Discussion

4.1. Topological Characteristics Analysis of the GDBSN

Figure 8 illustrates the GDBSN constructed in this study. The topological characteristics of the undirected, unweighted network were analyzed using complex network metrics. The results indicate that the average degree of the GDBSN is 60.46, suggesting that each port is connected to approximately 60 other ports on average. Additionally, the network's average path length is 2.324, the average clustering coefficient is 0.471, and the network density is 0.045. The diameter of the network is 5, suggesting that starting from any port, at most four intermediate ports are needed to reach any other port, reflecting the high connectivity and efficiency of the GDBSN. Further analysis reveals that the small-world coefficient of the network exceeds 1, indicating that the network exhibits typical small-world characteristics. Despite its large scale and complex structure, any two ports can be connected through only a small number of intermediate ports on average.



Average degree	Network diameter	Network density	Average clustering coefficient	Average path length	Small world coefficient
60.46	5	0.045	0.471	2.324	9

Figure 8. Global dry bulk shipping network and key topological properties.

A statistical analysis of the node degrees in the GDBSN was performed, and the degree distribution is presented in Figure 9. After fitting the degree distribution characteristics to a power-law distribution, the residual value obtained was 0.4921. This value suggests that although the network somewhat exhibits the characteristics of a scale-free distribution, its scale-free nature is relatively weak. A prominent characteristic of scale-free networks is the existence of a small number of highly connected nodes. In the GDBSN, only 4.5% of the ports have a degree greater than 200, including Singapore Port and Ningbo-Zhoushan Port.

However, the majority of nodes have relatively few connections, with 32.9% of the ports having a degree smaller than 20.

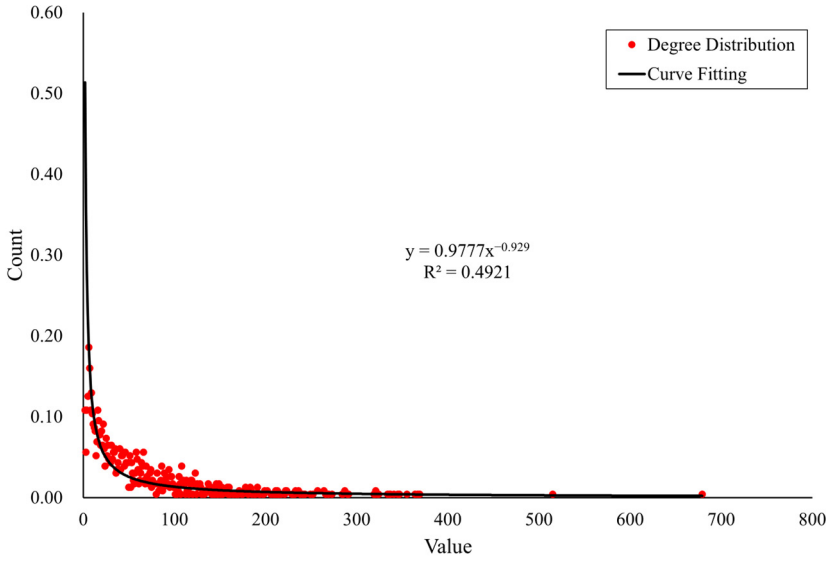


Figure 9. Node degree distribution.

4.2. Node Characteristics Analysis of the GDBSN

In the study of the GDBSN, a more detailed analysis of port node characteristics was conducted to identify ports with significant influence within the network. Building upon this, the present study analyzes node characteristics at both the port and country levels to uncover the core positions and roles of various ports and countries within the global network.

4.2.1. Port-Level Analysis

The degree of a port directly reflects the extent of its trade connections. Specifically, a higher degree indicates that the port is connected to a greater number of other ports. Figure 10 illustrates the degree value of each port, with the top 10 ports highlighted. It can be observed that ports with higher degree values are frequently located in key maritime transportation hubs, connecting critical international shipping routes and trade corridors, such as Said Port and Gibraltar Port. Additionally, ports with higher degree values, such as Singapore Port and Shanghai Port, are typically economic centers of their respective countries or regions, suggesting that the prosperity of a port is closely tied to regional economic development. This study uses the entropy-weighted TOPSIS method [49] to find out how important each port is in the GDBSN. It does this by using four centrality metrics. The result is shown in Table 5.

The research findings indicate that Singapore Port performs exceptionally well in all metrics, with its overall score ranking first, making it the most influential port in the GDBSN. Additionally, Gibraltar Port has a betweenness centrality value of 48,316.50, ranking second. This highlights its significant role as an intermediary connecting other ports, particularly as a crucial hub on transatlantic and Mediterranean shipping routes. Similarly, the betweenness centrality of Istanbul Port, Said Port, and Las Palmas Port is relatively high, indicating that they also play a certain intermediary role within the GDBSN. At the same time, Gibraltar Port also demonstrates high closeness centrality, which means that it has strong accessibility and capacity to connect to other ports quickly within the network. Ningbo-Zhoushan Port and Rizhao Port have high scores in eigenvector centrality, indicating that they are connected to other significant ports and they have high influence in

the GDBSN. In general, most of the top 10 ports are located in important global trade routes and strategic geographical locations. This phenomenon reveals that geographical location has a decisive influence on the importance of ports. Finally, it is noteworthy that Chinese ports, including Rizhao Port, Ningbo-Zhoushan Port, and Tianjin Port, are all among the top 10. These ports perform well in terms of centrality metrics, further reflecting China's prominent position in the GDBSN.

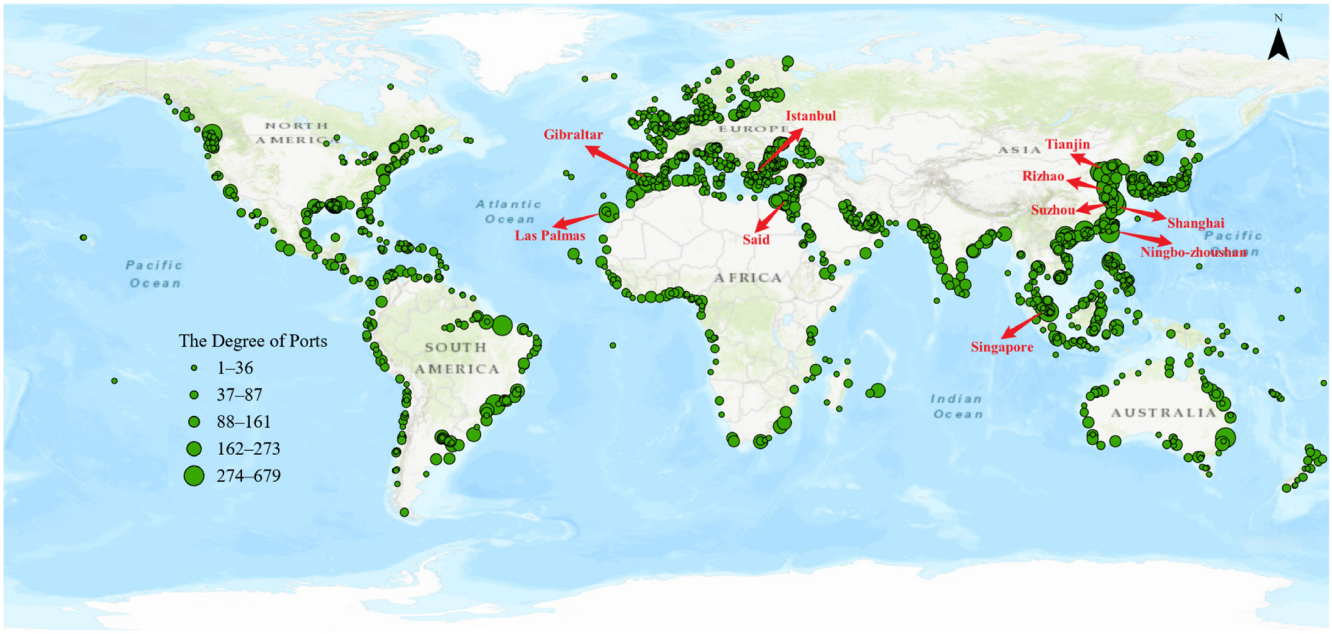


Figure 10. The degree distribution of ports in the GDBSN, with the top 10 high-degree ports highlighted.

Table 5. Port importance rankings in the GDBSN using entropy-weighted TOPSIS.

Rank	Port	Degree Centrality	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality	Score Index
1	Singapore	0.509	90,929.19	0.666	1.000	1.000
2	Gibraltar	0.386	48,316.50	0.612	0.677	0.599
3	Istanbul	0.259	28,946.93	0.561	0.474	0.390
4	Said	0.277	19,266.68	0.572	0.628	0.372
5	Rizhao	0.275	15,294.30	0.566	0.677	0.362
6	Ningbo-Zhoushan	0.273	11,501.32	0.567	0.684	0.347
7	Las Palmas	0.266	20,097.37	0.565	0.496	0.346
8	Tianjin	0.260	10,016.09	0.562	0.668	0.333
9	Kaohsiung	0.252	14,043.55	0.554	0.604	0.332
10	Santos	0.251	14,706.86	0.564	0.588	0.332

4.2.2. Country-Level Analysis

Figure 11 illustrates the distribution of the GDBSN degree value with the location of the national capital as the node. This distribution reflects the positions and roles of different countries in the GDBSN. Countries with higher degrees are distributed across multiple continents, particularly in North America, Europe, and Asia. Higher degree values indicate that the ports in these countries maintain extensive connections with those in other countries globally and highlight their significant influence in the GDBSN. Specifically, European countries have generally high node degree values, which indicates that most countries within the European region are closely connected to other ports in the GDBSN.

Similarly, countries in Asia, such as China, India, and Japan, also display relatively high degree values, reflecting their crucial roles in global trade. Additionally, countries located along key trade routes, such as Singapore, the UAE, and Egypt, despite having relatively small land areas, also exhibit high degree values, which shows the importance of their strategic geographical locations. In contrast, countries with low degree values are more dispersed, mainly in Africa and some small island states, indicating that these countries are relatively marginal in the GDBSN and have weak connectivity to the major shipping networks. On the whole, the countries with a high degree value are usually the countries with relatively developed economies or significant economic influence in the region, such as the United States, China, and Australia, indicating that the country degree value of the GDBSN has a certain correlation with the country's economic strength and foreign trade capacity.

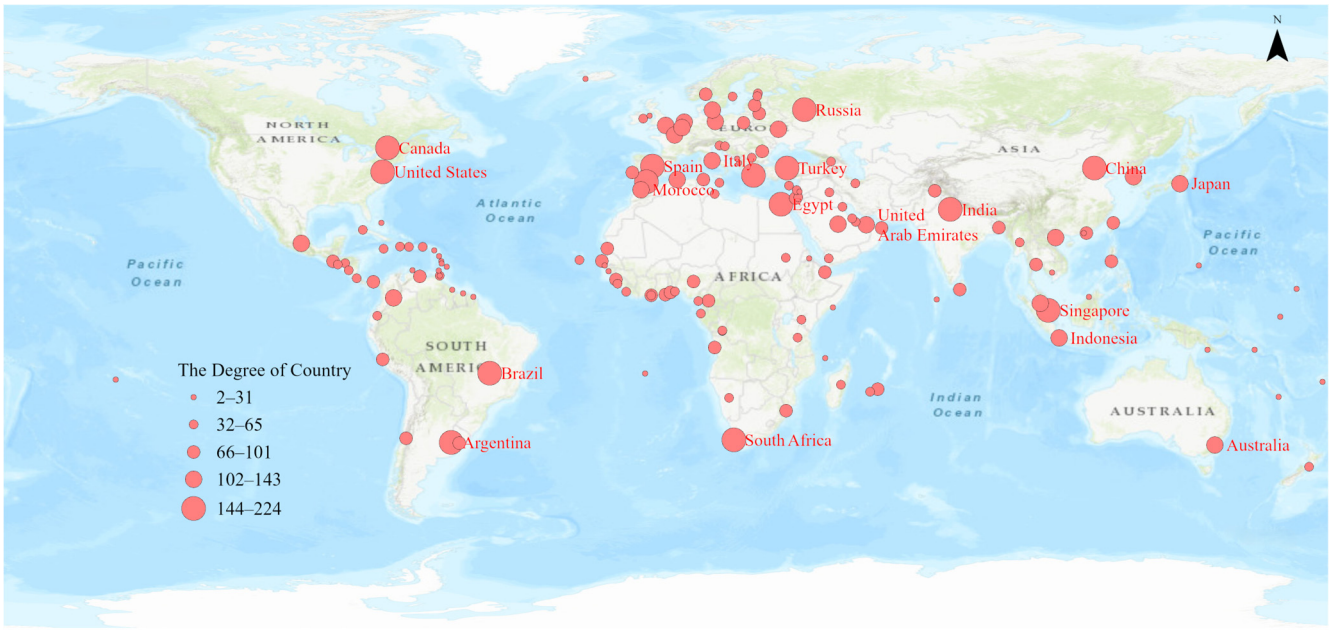


Figure 11. The degree hierarchy of countries.

4.3. Trade Link Analysis of the GDBSN

4.3.1. Port-Level Analysis

This study quantifies the connection frequency between pairs of ports in the GDBSN, with the results shown in Table 6, which lists the top 20 port pairs with the highest connection frequencies. The analysis shows that domestic ports in China display particularly strong connectivity, with the trade frequency between Ningbo-Zhoushan Port and Shanghai Port reaching as high as 1609, making them the most frequently connected pair. Furthermore, frequent connections are observed between Shanghai Port, Tangshan Port, Suzhou Port, and Ningbo-Zhoushan Port, indicating robust dry bulk transport demand within these regions, which reflects the active trade exchanges between coastal economic zones in China. Further analysis shows a clear regional clustering effect in the GDBSN, with many high-frequency shipping routes connecting geographically proximate ports, such as the route from Kaohsiung Port to Taichung Port in Taiwan and from Dalian Port to Shanghai Port. These short-distance and high-frequency shipping routes not only reflect the frequent flow of goods within the region but also demonstrate the strong geographic closeness of these ports. In addition, the GDBSN also shows a clear international characteristic. For example, the emergence of transnational routes such as Davao to Mariveles and Ningbo-Zhoushan Port to Singapore Port shows that the network has the ability to connect across

borders. These routes not only span geographical distances but also connect trade between different economies. In summary, the connection frequency distribution of the GDBSN shows a complex structure based on regional agglomeration and global extension. This structure not only ensures the transport of goods in the region but also effectively meets the needs of international trade.

Table 6. Top 20 port pairs by link weight.

Order	Origin	Destination	Weight	Order	Origin	Destination	Weight
1	Ningbo-Zhoushan	Shanghai	1609	11	Davao	Mariveles	706
2	Shanghai	Tangshan	1550	12	Qinhuangdao	Tangshan	651
3	Shanghai	Suzhou	1381	13	Nantong	Suzhou	600
4	Ningbo-Zhoushan	Suzhou	1305	14	Nantong	Shanghai	573
5	Cangzhou	Shanghai	1056	15	Kaohsiung	Taichung	573
6	Ningbo-Zhoushan	Tangshan	945	16	Dalian	Shanghai	565
7	Suzhou	Tangshan	874	17	Ningbo-Zhoushan	Qinhuangdao	552
8	Taizhou Jiangsu	Zhenjiang	827	18	Ningbo-Zhoushan	Singapore	552
9	Esmeraldas	Misurata	822	19	Shanghai	Zhenjiang	541
10	Qinhuangdao	Shanghai	781	20	Jiangyin	Shanghai	532

4.3.2. Country-Level Analysis

This study further analyzes trade connections between countries and identifies the top 20 country pairs based on trade frequency, as shown in Table 7. The highest trade frequency is observed between China and Australia, reflecting the close cooperation between the two nations in bulk commodities. This is likely due to Australia's role as a resource exporter, supplying raw materials such as coal and ores to the Chinese market. In addition, China maintains high-frequency trade with Singapore, Japan, and South Korea, with Singapore serving as a key transit hub in Southeast Asia and playing a central role in the flow of goods across the region. Moreover, the frequent trade interactions between China and resource-exporting countries like Indonesia further highlight the driving influence of resource demand on the shipping network's structure. Overall, high-frequency trade routes tend to be concentrated between resource demand and supply countries, reflecting the clear economic complementarity within the GDBSN.

Table 7. Top 20 country pairs by link weight.

Order	Country A	Country B	Weight	Order	Country A	Country B	Weight
1	Australia	China	8962	11	India	Singapore	1443
2	China	Singapore	4136	12	Australia	South Korea	1400
3	Australia	Japan	2969	13	China	Taiwan (China)	1363
4	China	Indonesia	2961	14	Mexico	United States	1328
5	China	South Korea	2385	15	Japan	South Korea	1240
6	China	Philippines	2101	16	Russia	Turkey	1219
7	China	Russia	1749	17	Brazil	Singapore	1191
8	Turkey	Ukraine	1644	18	China	United States	1191
9	China	Japan	1601	19	Brazil	China	1177
10	China	Vietnam	1511	20	Japan	Russia	1144

The global dry bulk trade network is illustrated in Figure 12, highlighting a trade pattern dominated by a few major countries and regions. The analysis reveals that a significant number of trade connections center around China, particularly with Asian countries as well as with Europe, North America, and South America, indicating dense trade exchanges. In terms of intercontinental trade, the routes between Europe and North America are also

quite dense, reflecting strong dry bulk trade between these two regions. Additionally, South America—especially Brazil and Argentina—exhibits very close trade connections with Asia, North America, and Europe. As a major exporter of commodities such as iron ore and grains, Brazil supplies large quantities of raw materials to Asian and European markets via transoceanic routes, fulfilling the resource demands of large importing countries for iron ore and food. The frequent connections between Australia and East Asia further underscore Australia's significant position in the export of minerals and coal. Overall, the global dry bulk trade is primarily concentrated among a few major countries and regions, reflecting the close ties between exporting and importing nations of bulk commodities such as iron ore, coal, and grains. From a structural perspective, the global bulk trade network is complex and diverse, with various regions closely interconnected through shipping routes.

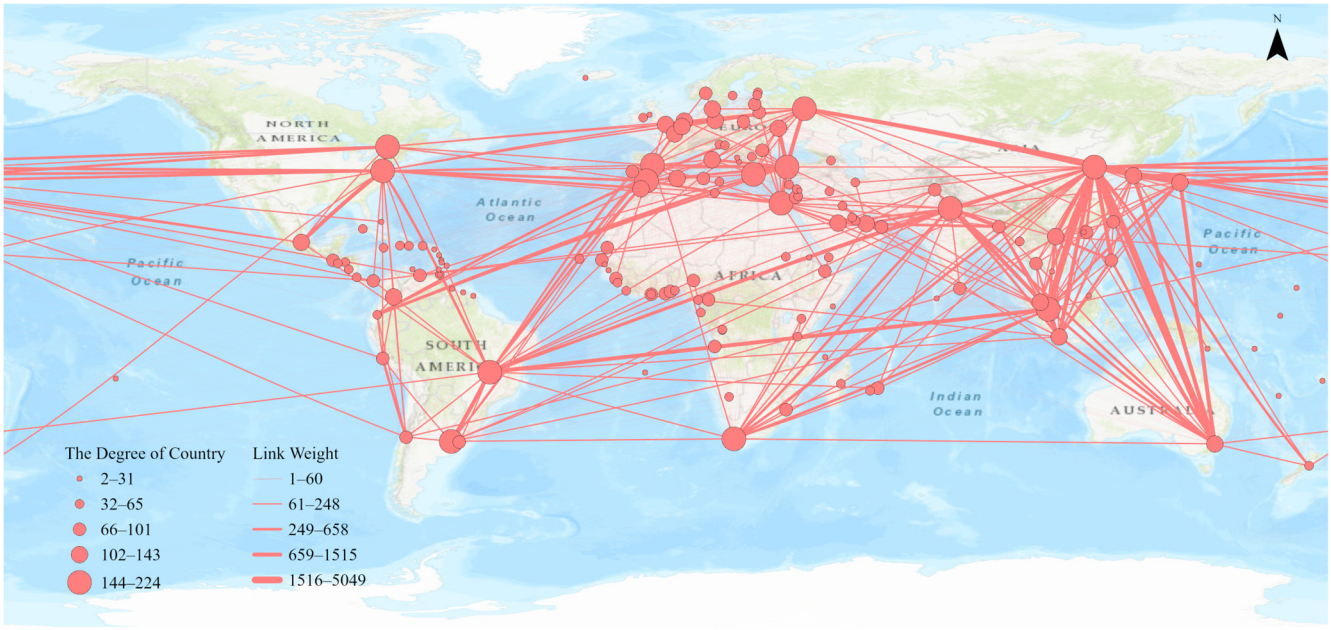


Figure 12. The spatial distribution of the national bulk trade network.

4.4. Link Prediction Analysis of the GDBSN

Based on the analysis of the topological structure, node importance, and trade connections of the GDBSN, this study further investigates the potential connections within the network. Link prediction aims to utilize existing topological structure information to predict edges that do not currently exist in the network—specifically, predicting direct connections between ports that have not yet been established in the GDBSN. In this study, link prediction indices outlined in Section 3.3.2 were employed to predict the potential routes in the GDBSN. The average AUC values for each similarity index were calculated after 100 independent experiments. The results are presented in Figure 13. Clearly, the RA index achieved the highest prediction accuracy, with an AUC of 0.946. Therefore, the analysis in this study focuses primarily on the prediction results based on the RA index.

Through the sorting of the potential link similarity scores obtained via the RA index, Table 8 presents the top 20 predicted new shipping routes. It was found that the majority of the newly predicted routes are concentrated within regions or between neighboring countries, such as those between Santos Port and Vitoria Port in Brazil, Hong Kong Port and Zhenjiang Port in China, and Mizushima Port and Hirohata Port in Japan. These routes are primarily located within the same continent or between neighboring nations, which may be attributed to geographical proximity and strong trade partnerships, reflecting the evident principle of preferential connection [50]. Additionally, the connectivity between Asia and

the Far East regions has been enhanced. Several routes involving China (Hong Kong, Guangzhou) and Japan (Chiba, Mizushima) are shown in the table. These regions, as major global manufacturing hubs and consumer markets, have experienced sustained growth in shipping demand, which contributes to the development of these routes. Cross-border routes, such as those between Istanbul Port (Turkey) and St. Petersburg Port (Russia), as well as between La Plata Port (Argentina) and Said Port (Egypt), could indicate the growing economic cooperation and trade exchanges between these countries. The link prediction results provide insights into the potential new shipping routes that could emerge within the GDBSN. However, the actual decision to establish new routes or strengthen existing ones is contingent upon various practical factors, such as trade cooperation between countries, market demand, and other logistical considerations.

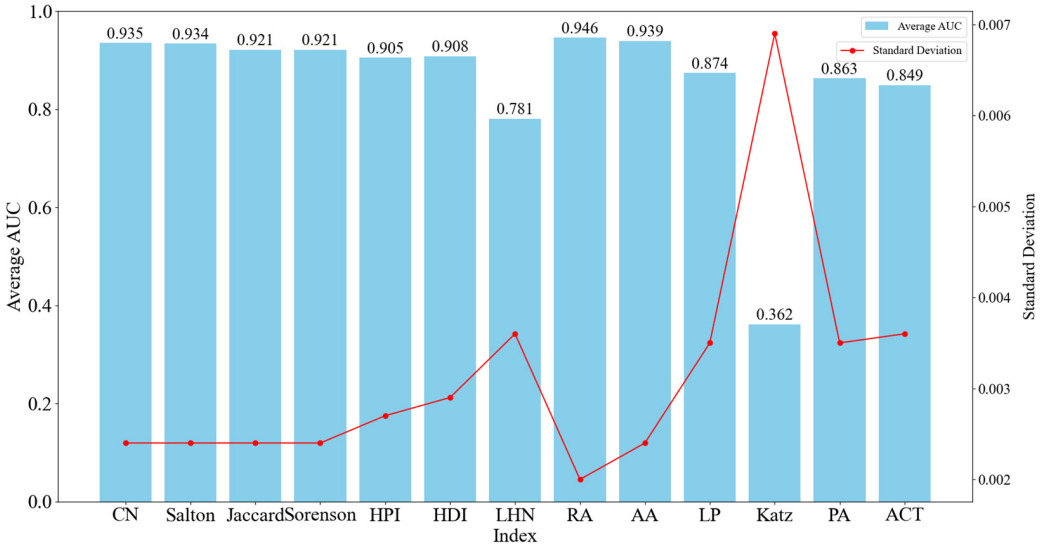


Figure 13. AUC values and standard deviation of AUC values for different indices.

Table 8. Top 20 potential links and their scores based on RA index.

Rank	Potential Link	Score
1	Santos (Brazil)-Vila Do Conde (Brazil)	1.343
2	La Plata (Argentina)-Said (Egypt)	1.204
3	Hong Kong (China)-Zhenjiang (China)	1.169
4	Houston (United States)-Ponta da Madeira (Brazil)	1.114
5	Mizushima (Japan)-Tobata (Japan)	1.077
6	Gibraltar (Gibraltar)-Mykolayiv (Ukraine)	1.074
7	Istanbul (Turkey)-La Plata (Argentina)	1.072
8	Hay Point (Australia)-Vanino (Russia)	1.070
9	Gresik (Indonesia)-Manila (Philippines)	1.065
10	Chiba (Japan)-Manila (Philippines)	1.040
11	Casablanca (Morocco)-Damietta (Egypt)	1.020
12	Arroyo Seco (Argentina)-Ponta da Madeira (Brazil)	1.018
13	Fujairah (United Arab Emirates)-Manila (Philippines)	1.011
14	Ceuta (Spain)-Skagen (Denmark)	1.011
15	Said (Egypt)-Tubarao (Brazil)	1.010
16	Gresik (Indonesia)-Kaohsiung (Taiwan)	1.002
17	Istanbul (Turkey)-St Petersburg (Russia)	0.998
18	Fujairah (United Arab Emirates)-Guangzhou (China)	0.997
19	Chiba (Japan)-Yantai (China)	0.994
20	Casablanca (Morocco)-Nemrut Bay (Turkey)	0.986

The top 416 potential new shipping routes (scores over 0.6), 977 routes (scores over 0.5), and 2498 routes (scores over 0.4) were further incorporated into the GDBSN. The resulting spatial distribution pattern is presented in Figure 14. The red lines represent existing routes in the GDBSN, and the blue lines represent potential new routes. As shown in the figure, the potential new routes exhibit significant regional clustering, primarily concentrated in core areas with dense port networks, particularly in global trade hubs such as Europe, East Asia (particularly China and Japan), North America, and Southeast Asia. These regions are already major global trade centers with dense routes. Further analysis indicates that the potential new routes not only involve connections within regions but also across geographic boundaries, notably establishing new links between Asia and South America, as well as between Africa and Asia. This suggests that future global dry bulk trade is likely to become more globalized, with an increase in intercontinental trade links. By analyzing the distribution of existing and potential new routes, it can be observed that most of the predicted new routes extend existing major trade corridors while also introducing new connections in some peripheral or underdeveloped regions. Overall, link prediction results indicate that the dry bulk trade network is expected to become more globalized, potentially strengthening links within existing major trade centers and expanding to some underdeveloped regions.

To further understand the evolution of the network structure, this study incorporated the predicted edges into the original GDBSN and systematically analyzed the impact of these changes on the network's topological characteristics. The results are shown in Table 9, which outlines the changes in the network's topological characteristic values after adding the predicted edges. The following trends can be observed:

Table 9. Parameter comparison between the original network and the network after adding edges.

	Average Degree	Diameter	Density	Average Clustering Coefficient	Average Path Length
Original network	60.460	5	0.045	0.471	2.324
416 added connections	61.084	5	0.046	0.483	2.322
977 added connections	61.925	5	0.046	0.498	2.318
2498 added connections	64.205	5	0.048	0.530	2.309

Firstly, the average degree of the original network was 60.460 and increased gradually with the addition of the predicted edges. This change suggests that the newly predicted connections improved the average number of links between ports, leading to a denser network of port nodes. As a result, the overall connectivity of the network was strengthened, allowing each port to establish connections with more ports.

Meanwhile, the average clustering coefficient was 0.471 in the original network, and it reached 0.530 after the addition of 2498 edges. The significant improvement in the clustering coefficient indicates that the newly added connections enhanced the local clustering effect between ports, facilitating the formation of "closed-loop" structures between neighboring ports, where the connectivity among adjacent nodes is strengthened. The addition of these edges contributes to enhancing the tightness of local networks and leading to more stable connections within smaller port groups. This is beneficial for fostering close cooperation in regional trade and transportation.

Additionally, the density of the original network increased gradually from 0.045 to 0.048. With the addition of potential new shipping routes, the ratio of actual connections to the potential connections in the network increased. The change in density was relatively small, but this trend indicates that the newly added connections have contributed to making

the network structure more compact. It also results in more frequent connections between ports. This contributes to the enhancement of the network's robustness.

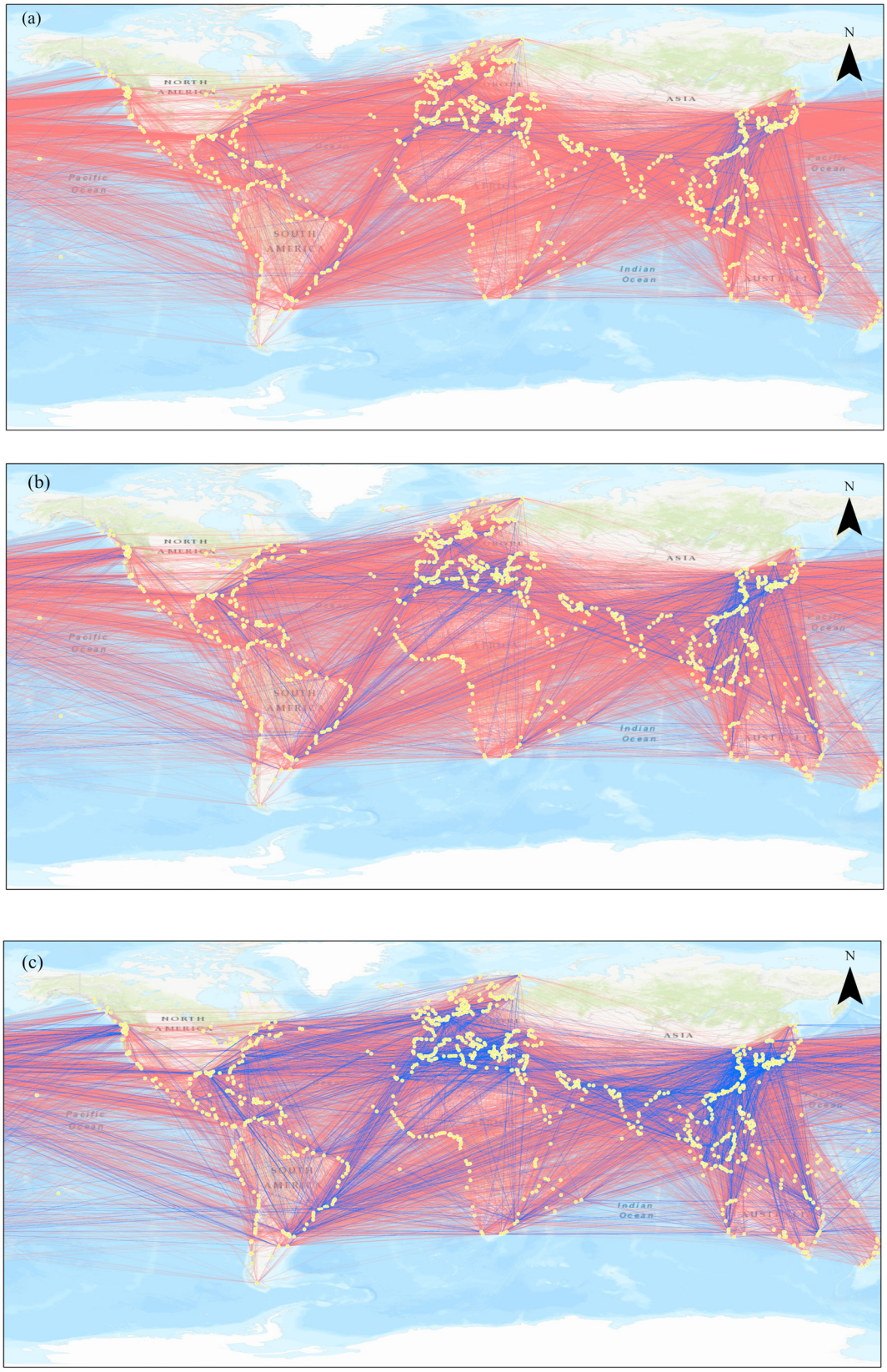


Figure 14. The GDBSN after adding edges: (**a**) 416 routes; (**b**) 977 routes; (**c**) 2498 routes.

Furthermore, the original network had an average path length of 2.324, and it decreased to 2.309 with the addition of new connections. This change indicates that the newly added connections shortened the average path between ports, although the effect was modest. The slight reduction in the average path length suggests that, while some port access paths have become shorter, the overall global connectivity of the network has not been significantly affected, and it remains in a stable state.

Finally, despite the increase in the number of connections, the network's diameter still remained at 5. This indicates that the maximum distance between any two ports in the network was not altered. The fact that the diameter remained unchanged suggests that the newly added connections were mainly focused on the local structure and did not influence the global shortest path length.

The changes in the aforementioned network characteristics suggest that adding potential connections to the original network increases the average degree and clustering coefficient, thereby improving the network's local clustering and strengthening port-to-port connections. However, the network's diameter and average path length remain largely unchanged, indicating that the newly added edges primarily affect the local network structure without substantially altering its global topology. Therefore, these results imply that optimizing the GDBSN should prioritize improving the connectivity of local network structures. Although the overall connectivity efficiency remains stable after the addition of numerous connections, enhancing local clustering and inter-port connectivity can improve the network's robustness.

The above study employs a link prediction method based on structural similarity to reveal the future evolutionary trends of the GDBSN and potential new shipping routes. However, the actual operation of the predicted routes is still significantly influenced by policy factors. Globally, the structure of international trade and the patterns of dry bulk shipping are often deeply affected by policy changes. Trade policies such as trade agreements, tariff policies, or economic sanctions can alter the flow of goods between major economies, thereby impacting the supply–demand dynamics of the dry bulk market. For example, free trade agreements can promote the flow of goods, while trade frictions may lead to supply chain disruptions, affecting port connectivity. Additionally, environmental regulations, such as the IMO 2020 sulfur cap, are prompting shipping companies to adjust fuel choices, optimize shipping routes, and update vessel technologies, which may result in changes to route and port configurations. Therefore, while this study provides valuable insights into the future evolution of the shipping network and potential new routes, its practical application requires careful evaluation and adjustment according to the specific policy environment to ensure the feasibility and effectiveness of the findings under different policy contexts.

4.5. Implications

The dry bulk market is not static and may undergo changes due to factors such as geopolitical events, regulatory management, and global economic trends. These factors can affect the demand and supply for transportation, which in turn impacts aspects such as market prices and shipping routes. For example, the impact of the Russia–Ukraine conflict on the global energy market has not only disrupted energy supplies but also altered energy transportation routes. Similar geopolitical factors have led to increased market uncertainty, thereby affecting the supply–demand dynamics of dry bulk commodities. Therefore, understanding the structure of the current global dry bulk shipping network and its potential development directions is crucial for forecasting market trends and formulating effective response strategies. Based on the research findings, this study proposes several

policy recommendations aimed at providing theoretical support and practical guidance for the development of global trade and the operation of related shipping enterprises.

Firstly, the analysis of port importance reveals that several highly influential hub ports exist within the GDBSN. These ports exhibit high centrality within the network and play a crucial role. Factors such as their operational efficiency and cargo handling capacity directly impact the network's overall efficiency. These ports are indeed important hubs for various countries or regions in actual operations. It is recommended that governments and port operators increase their investments in infrastructure development at these key hub ports, particularly in areas such as smart port technologies, automation equipment, and shipping information platforms. This will enhance port operational efficiency and further promote the smooth functioning of the global shipping network. Concurrently, some ports within the network exhibit relatively poor connectivity. These ports exhibit weaker self-development and expansion capabilities and are situated at the periphery of the network structure. In establishing trade links with other ports, they are constrained by geographical distance and often rely on nearby central hub ports to sustain trade activities. For these ports, it is recommended that efforts be made to prioritize the improvement of port capacity and infrastructure, as well as to enhance connectivity with surrounding regional ports. This will enhance their competitiveness within the GDBSN and promote their better integration into the global trade system.

Secondly, based on the results of the trade connection analysis, it is evident that the GDBSN exhibits a significant regional clustering effect, with high-frequency routes primarily concentrated between geographically proximate ports. To promote balanced development in the GDBSN, it is recommended that governments and port operators strengthen regional cooperation, particularly by advancing the construction and optimization of intercontinental routes. Meanwhile, developing countries and emerging economies should advocate for the signing of multilateral shipping agreements. Multilateral cooperation can not only improve inter-regional connectivity in the shipping network but also enhance global shipping efficiency and reduce the negative impacts of geographical and economic disparities. In addition, the study findings indicate that trade relations between countries display significant resource complementarity, which aligns with the current realities of shipping operations, particularly in the transportation of bulk commodities. The shipping cooperation between resource-exporting and resource-importing countries is of critical importance. Therefore, the study recommends focusing on enhancing shipping cooperation between these countries, especially through policy coordination and collaboration, to ensure the stable transportation of bulk commodities.

Finally, the link prediction results reveal some potential new shipping routes, some of which are situated in marginal markets and underdeveloped regions. These routes may serve as key growth points for the GDBSN in the future. Therefore, it is recommended that shipping companies proactively plan and evaluate the potential of these new routes while simultaneously exploring opportunities to expand existing routes. Meanwhile, governments can encourage investment in these areas by providing subsidies for shipping infrastructure development, offering tax incentives for routes, and implementing other policies aimed at enhancing port and route infrastructure. The research also indicates that the newly added connections have significantly enhanced the local clustering and density of the network, with a particularly pronounced impact on local structures. Therefore, governments and port operators should place greater emphasis on strengthening collaborative efforts among regional ports. This includes fostering additional regional partnerships, particularly in areas with high potential. By increasing investment in infrastructure and optimizing shipping services, regional trade flow and transportation efficiency can be further enhanced, thus ensuring the smooth operation of the entire shipping network.

In conclusion, governments and shipping companies should work together to support the sustainable growth of global shipping, especially in the areas of port infrastructure development, regional cooperation, and shipping network optimization. Adopting flexible and forward-looking policies will be important to facing future challenges as the global shipping network continues to develop. These policies will help improve shipping efficiency, enhance network stability, and foster balanced global economic growth.

5. Conclusions

A comprehensive analysis of the GDBSN was conducted using the proposed framework for shipping network analysis. Various aspects of the network, including its topological characteristics, port significance, and trade connections, were explored. Additionally, link prediction was applied to forecast potential shipping routes. The following conclusions were drawn:

(1) The GDBSN demonstrates a high average degree and a short average path length, signifying robust connection and transport efficiency, which reflects the significant activity and extensive interconnectedness of the global dry bulk market. The network exhibits conventional small-world traits: despite its extensive size, only a limited number of intermediary ports are necessary to link any two ports. The study shows that the GDBSN's node degree distribution is scale-free, meaning that only a few key ports have a high degree while most ports have fewer connections. This phenomenon reflects the importance of hub ports in the global shipping market. These ports are not only key collection points for trade flows but may also play a strategic role in the global shipping network.

(2) At the port level, Singapore Port ranks first in terms of its comprehensive centrality, highlighting its pivotal position in the GDBSN. Ports that rank highly are predominantly located at key transportation hubs, underscoring the significant role of geographical factors in global trade. At the country level, economic powerhouses such as the United States, China, and several European countries show greater connectivity within the GDBSN, reflecting their pivotal roles in global trade. In contrast, the ports of most African countries and certain small island nations are located more peripherally within the network. This reflects the regional imbalances present in the GDBSN.

(3) The shipping network exhibits a clear regional clustering effect in trade connections, with numerous high-frequency routes linking geographically proximate ports, highlighting their strong geographical proximity. This phenomenon also reflects the close economic ties between regions. At the country level, trade connections within the GDBSN are predominantly concentrated among a few major countries and regions, reflecting the close ties between commodity-exporting and -importing nations.

(4) Based on the link prediction results, it was found that new shipping routes are primarily concentrated within regional areas and between neighboring countries, reflecting the continuous expansion of the GDBSN in localized regions. The majority of potential new routes extend existing major trade corridors and introduce new connections in peripheral or underdeveloped regions. Moreover, the link prediction results indicate that the newly added connections significantly increased the average degree and average clustering coefficient of the GDBSN. However, the network diameter and average path length remained relatively stable, suggesting that the newly added edges primarily affected the local structure rather than significantly impacting the global topology.

With the steady growth of global trade [51], the findings of this study have significant implications for enterprises and countries involved in dry bulk shipping. By analyzing the

network topology characteristics, this study provides valuable insights for policymakers, enabling a deeper understanding of the current dry bulk transport landscape and supporting evidence-based policy development. For instance, governments could increase investments in infrastructure at key ports to enhance their global connectivity. At the same time, shipping companies can leverage link prediction algorithms to assess the potential of new shipping routes and explore opportunities for expanding existing ones. These measures will not only improve the operational efficiency of shipping enterprises but also contribute to a more efficient and stable development of global trade.

Although this study partially fills the gap in research on dry bulk shipping networks, there are still several areas that require further improvement in future studies. Specifically, due to data limitations, this study only analyzes the overall trade patterns of dry bulk commodities and does not explore the trade dynamics of specific commodity types, such as coal or ores. Future research could conduct a more detailed analysis of different commodity categories to uncover their specific trade linkages, providing more precise insights for strategic decision-making by relevant enterprises and ports.

Author Contributions: Conceptualization, H.L. and X.Z.; methodology, H.L. and X.Z.; formal analysis, X.Z. and Q.M.; investigation, H.L.; data curation, X.Z. and Q.M.; writing—original draft preparation, X.Z.; writing—review and editing, H.L.; visualization, X.Z. and P.J.; supervision, H.L. and P.J.; funding acquisition, H.L. and P.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China, grant number 2023YFB4302200, and the Key Project of the Liaoning Province Social Science Planning Fund, grant number L21ACL002.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data will be made available on request. The processed data are available at GitHub: https://github.com/xinzhang922/network-data (accessed on 14 January 2025).

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

References

1. UNCTAD. Review of Maritime Transport 2023. United Nations Conference on Trade and Development. 2023. Available online: https://unctad.org/publication/review-maritime-transport-2023 (accessed on 22 October 2024).
2. Ducruet, C. The geography of maritime networks—A critical review. *J. Transp. Geogr.* **2020**, *88*, 102824. [CrossRef]
3. Michail, N.A.; Melas, K.D. Geopolitical Risk and the LNG-LPG Trade. *Peace Econ. Peace Sci. Public Policy* **2022**, *28*, 243–265. [CrossRef]
4. Yang, J.; Bian, X.; Qi, Y.; Wang, X.; Yang, Z.; Liu, J. A spatial-temporal data mining method for the extraction of vessel traffic patterns using AIS data. *Ocean Eng.* **2024**, *293*, 116454. [CrossRef]
5. Bai, X.; Cheng, L.; Iris, Ç. Data-driven financial and operational risk management: Empirical evidence from the global tramp shipping industry. *Transp. Res. Part E Logist. Transp. Rev.* **2022**, *158*, 102617. [CrossRef]
6. Chen, X.; Wei, C.; Zhou, G.; Wu, H.; Wang, Z.; Biancardo, S.A. Automatic Identification System (AIS) Data Supported Ship Trajectory Prediction and Analysis via a Deep Learning Model. *J. Mar. Sci. Eng.* **2022**, *10*, 1314. [CrossRef]
7. Wang, W.; Liu, K.; Huang, L.; Xin, X.; Wu, X.; Yuan, Z. Generation and complexity analysis of ship encounter scenarios using AIS data for collision avoidance algorithm testing. *Ocean Eng.* **2024**, *312*, 119034. [CrossRef]
8. Ma, Q.; Tang, H.; Liu, C.; Zhang, M.; Zhang, D.; Liu, Z.; Zhang, L. A big data analytics method for the evaluation of maritime traffic safety using automatic identification system data. *Ocean Coast. Manag.* **2024**, *251*, 107077. [CrossRef]
9. Yang, K.; Li, J.; Liu, M.; Lei, T.; Xu, X.; Wu, H.; Cao, J.; Qi, G. Complex systems and network science: A survey. *J. Syst. Eng. Electron.* **2023**, *34*, 543–573. [CrossRef]
10. Mou, N.; Liu, C.; Zhang, L.; Fu, X.; Xie, Y.; Li, Y.; Peng, P. Spatial Pattern and Regional Relevance Analysis of the Maritime Silk Road Shipping Network. *Sustainability* **2018**, *10*, 977. [CrossRef]

11. Zhang, Z.; Wu, X.; Zhu, G.; Qin, W.; Liang, N. A Graph Attention Network-Based Link Prediction Method Using Link Value Estimation. *IEEE Access* **2024**, *12*, 34–45. [CrossRef]
12. Daud, N.N.; Ha, S.; Anuar, N.B. Applications of link prediction in social networks: A review. *J. Netw. Comput. Appl.* **2020**, *166*, 102716. [CrossRef]
13. Kumari, A.; Behera, R.K.; Sahoo, B.; Sahoo, S.P. Prediction of link evolution using community detection in social network. *Computing* **2022**, *104*, 1077–1098. [CrossRef]
14. Yu, H.; Fang, Z.; Lu, F.; Murray, A.T.; Zhao, Z.; Xu, Y.; Yang, X. Massive Automatic Identification System Sensor Trajectory Data-Based Multi-Layer Linkage Network Dynamics of Maritime Transport along 21st-Century Maritime Silk Road. *Sensors* **2019**, *19*, 4197. [CrossRef]
15. Mou, N.; Ren, H.; Zheng, Y.; Chen, J.; Niu, J.; Yang, T.; Zhang, L.; Liu, F. Traffic Inequality and Relations in Maritime Silk Road: A Network Flow Analysis. *ISPRS Int. J. Geo Inf.* **2021**, *10*, 40. [CrossRef]
16. Liu, Z.; Gao, H.; Zhang, M.; Yan, R.; Liu, J. A data mining method to extract traffic network for maritime transport management. *Ocean Coast. Manag.* **2023**, *239*, 106622. [CrossRef]
17. Onyango, S.O.; Owiredu, S.A.; Kim, K.-I.; Yoo, S.-L. A Quasi-Intelligent Maritime Route Extraction from AIS Data. *Sensors* **2022**, *22*, 8639. [CrossRef]
18. Liu, L.; Shibasaki, R.; Zhang, Y.; Kosuge, N.; Zhang, M.; Hu, Y. Data-driven framework for extracting global maritime shipping networks by machine learning. *Ocean Eng.* **2023**, *269*, 113494. [CrossRef]
19. Mesa-Arango, R.; Narayanan, B.; Ukkusuri, S.V. The Impact of International Crises on Maritime Transportation Based Global Value Chains. *Netw. Spat. Econ.* **2019**, *19*, 381–408. [CrossRef]
20. Zhao, S.; Sun, S. A study on centrality measures in weighted networks: A case of the aviation network. *AIMS Math.* **2024**, *9*, 3630–3645. [CrossRef]
21. Yin, L.; Wang, Y. Spatiotemporal Evolution Analysis of the Chinese Railway Network Structure Based on Self-Organizing Maps. *ISPRS Int. J. Geo Inf.* **2023**, *12*, 161. [CrossRef]
22. Cong, L.; Zhang, H.; Wang, P.; Chu, C.; Wang, J. Impact of the Russia–Ukraine Conflict on Global Marine Network Based on Massive Vessel Trajectories. *Remote Sens.* **2024**, *16*, 1329. [CrossRef]
23. Wen, Z.; Weng, X.; Zhang, P. Evaluating the Connectivity and Imbalance Contribution of New Sections towards Highway Network: A Complex Network Perspective. *J. Adv. Transp.* **2023**, *2023*, 6616512. [CrossRef]
24. Kaluza, P.; Kölzsch, A.; Gastner, M.T.; Blasius, B. The complex network of global cargo ship movements. *J. R. Soc. Interface* **2010**, *7*, 1093–1103. [CrossRef]
25. Jiang, L.; Jia, Y.; Zhang, C.; Wang, W.; Feng, X. Analysis of topology and routing strategy of container shipping network on "Maritime Silk Road". *Sustain. Comput. Inform. Syst.* **2019**, *21*, 72–79. [CrossRef]
26. Kang, L.; Wu, W.; Yu, H.; Su, F. Global Container Port Network Linkages and Topology in 2021. *Sensors* **2022**, *22*, 5889. [CrossRef] [PubMed]
27. Guerrero, D.; Letrouit, L.; Pais-Montes, C. The container transport system during COVID-19: An analysis through the prism of complex networks. *Transp. Policy* **2022**, *115*, 113–125. [CrossRef]
28. Xu, M.; Deng, W.; Zhu, Y.; Lü, L. Assessing and improving the structural robustness of global liner shipping system: A motif-based network science approach. *Reliab. Eng. Syst. Saf.* **2023**, *240*, 109576. [CrossRef]
29. Jin, L.; Chen, J.; Chen, Z.; Sun, X.; Yu, B. Impact of COVID-19 on China's international liner shipping network based on AIS data. *Transp. Policy* **2022**, *121*, 90–99. [CrossRef]
30. Li, Z.; Li, H.; Zhang, Q.; Qi, X. Data-driven research on the impact of COVID-19 on the global container shipping network. *Ocean Coast. Manag.* **2024**, *248*, 106969. [CrossRef]
31. Li, J.; Zhao, H.; Xu, B. Optimization of Container Shipping Network Reconfiguration under RCEP. *J. Mar. Sci. Eng.* **2022**, *10*, 873. [CrossRef]
32. Venturini, G.; Iris, Ç.; Kontovas, C.A.; Larsen, A. The multi-port berth allocation problem with speed optimization and emission considerations. *Transp. Res. Part D Transp. Environ.* **2017**, *54*, 142–159. [CrossRef]
33. Yan, Z.; He, R.; Yang, H. The small world of global marine crude oil trade based on crude oil tanker flows. *Reg. Stud. Mar. Sci.* **2022**, *51*, 102215. [CrossRef]
34. Wang, J.; Wang, X.; Feng, Y.; Cao, Y.; Guo, Z.; Liu, Z. Assessing the Connectivity Reliability of a Maritime Transport Network: A Case of Imported Crude Oil in China. *J. Mar. Sci. Eng.* **2023**, *11*, 1597. [CrossRef]
35. Mou, N.; Sun, S.; Yang, T.; Wang, Z.; Zheng, Y.; Chen, J.; Zhang, L. Assessment of the Resilience of a Complex Network for Crude Oil Transportation on the Maritime Silk Road. *IEEE Access* **2020**, *8*, 181311–181325. [CrossRef]
36. Yu, H.; Fang, Z.; Lu, F.; Murray, A.T.; Zhang, H.; Peng, P.; Mei, Q.; Chen, J. Impact of oil price fluctuations on tanker maritime network structure and traffic flow changes. *Appl. Energy* **2019**, *237*, 390–403. [CrossRef]
37. Peng, P.; Cheng, S.; Chen, J.; Liao, M.; Wu, L.; Liu, X.; Lu, F. A fine-grained perspective on the robustness of global cargo ship transportation networks. *J. Geogr. Sci.* **2018**, *28*, 881–889. [CrossRef]

38. Shibuya, K.; Shibasaki, R. Modeling Structural Changes in Intra-Asian Maritime Container Shipping Networks Considering Their Characteristics. *Sustainability* **2023**, *15*, 10055. [CrossRef]
39. Ge, J.; Wang, X.; Shi, W. Link prediction of the world container shipping network: A network structure perspective. *Chaos Interdiscip. J. Nonlinear Sci.* **2021**, *31*, 113123. [CrossRef]
40. Zhu, J.; Yang, Y.; Jin, Y.; Liu, W. Improved Link Prediction Method for Maritime Silk Road Shipping Network Using Composite Index. *Transp. Res. Rec. J. Transp. Res. Board* **2024**, *2678*, 674–688. [CrossRef]
41. U.S. National Geospatial-Intelligence Agency. *National Geospatial-Intelligence Agency, World Port Index*, 27th ed.; U.S. National Geospatial-Intelligence Agency: Springfield, VI, USA, 2021.
42. Liu, Z.; Yuan, W.; Liang, M.; Zhang, M.; Liu, C.; Liu, R.W.; Liu, J. An online method for ship trajectory compression using AIS data. *J. Navig.* **2024**, 1–22. [CrossRef]
43. Li, H.; Jia, P.; Wang, X.; Yang, Z.; Wang, J.; Kuang, H. Ship carbon dioxide emission estimation in coastal domestic emission control areas using high spatial-temporal resolution data: A China case. *Ocean Coast. Manag.* **2023**, *232*, 106419. [CrossRef]
44. Hu, Y.; Zhu, D. Empirical analysis of the worldwide maritime transportation network. *Phys. A Stat. Mech. Its Appl.* **2009**, *388*, 2061–2071. [CrossRef]
45. Tsiotas, D. Measuring the effect of distance on the network topology of the Global Container Shipping Network. *Sci. Rep.* **2021**, *11*, 21250. [CrossRef] [PubMed]
46. Xu, M.; Pan, Q.; Muscoloni, A.; Xia, H.; Cannistraci, C.V. Modular gateway-ness connectivity and structural core organization in maritime network science. *Nat. Commun.* **2020**, *11*, 2849. [CrossRef]
47. Wan, Z.; Su, Y.; Li, Z.; Zhang, X.; Zhang, Q.; Chen, J. Analysis of the impact of Suez Canal blockage on the global shipping network. *Ocean Coast. Manag.* **2023**, *245*, 106868. [CrossRef]
48. Lü, L.; Zhou, T. *Link Prediction*, 1st ed.; Higher Education Press: Beijing, China, 2013; pp. 59–60.
49. Wan, C.; Zhao, Y.; Zhang, D.; Yip, T.L. Identifying important ports in maritime container shipping networks along the Maritime Silk Road. *Ocean Coast. Manag.* **2021**, *211*, 105738. [CrossRef]
50. Mou, N.; Fang, Y.; Yang, T.; Zhang, L. Assortative Analysis of Bulk Trade Complex Network on Maritime Silk Road. *IEEE Access* **2020**, *8*, 131928–131938. [CrossRef]
51. Bilican, M.S.; Iris, Ç.; Karatas, M. A collaborative decision support framework for sustainable cargo composition in container shipping services. *Ann. Oper. Res.* **2024**, *342*, 79–111. [CrossRef]

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