

Article

Extracting the Maritime Traffic Route in Korea Based on Probabilistic Approach Using Automatic Identification System Big Data

Jeong-Seok Lee ¹ and Ik-Soon Cho ^{2,*}¹ Graduate School, Korea Maritime and Ocean University, Busan 49112, Korea; shoesy548@kmou.ac.kr² Division of Navigation Convergence Studies, Korea Maritime and Ocean University, Busan 49112, Korea

* Correspondence: ischo@kmou.ac.kr; Tel.: +82-51-410-5072

Abstract: To protect the environment around the world, we are actively developing ecofriendly energy. Offshore wind farm generation installed in the sea is extremely large among various energies, and friction with ships occurs regularly. Other than the traffic designated area and the traffic separate scheme, traffic routes in other sea areas are not protected in Korea. Furthermore, due to increased cargo volume and ship size, there is a risk of collisions with marine facilities and marine pollution. In this study, maritime safety traffic routes that must be preserved are created to ensure the safety of maritime traffic and to prevent accidents with ecofriendly energy projects. To construct maritime traffic routes, the analysis area is divided, and ships are classified using big data. These data are used to estimate density, and 50% maritime traffic is chosen. This result is obtained by categorizing the main route, inner branch route, and outer branch route. The Korean maritime traffic route is constructed, and the width of the route is indicated. Furthermore, this route can be applied as a navigation route for maritime autonomous surface ships.

**Citation:** Lee, J.-S.; Cho, I.-S.Extracting the Maritime Traffic Route in Korea Based on Probabilistic Approach Using Automatic Identification System Big Data. *Appl. Sci.* **2022**, *12*, 635. <https://doi.org/10.3390/app12020635>

Academic Editor: Koji Murai

Received: 23 November 2021

Accepted: 6 January 2022

Published: 10 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.**Copyright:** © 2022 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

The shipping industry is the most efficient mode of long-distance transportation [1]. The sea is used by more than 80% of the world's merchant ships, and various maritime traffic routes are used [2]. As a result, as the shipping industry develops and the demands of marine transportation increase over time, the size of ships has increased [3]. This phenomenon is strengthening the route design standards for the safe operation of ships and the prevention of marine accidents in each country and institution [4]. Moreover, due to the limitation of not being able to designate a route in all sea areas, ships are freely navigating in the sea area where there is no route due to economic considerations. In other words, most ships take similar routes because they are developed to travel the shortest distance possible, which is an economic consideration, and to seek safe and efficient routes with the least amount of risk [5]. Recently, in the area of ship routes, friction is increasing with ecofriendly energy projects, marine ecosystems, and fishing activities. This demonstrates the complexities of marine activities, which can result in marine accidents and environmental pollution, the consequences of which can be catastrophic [6]. In this study, a national main maritime route is developed and presented as a one-of-a-kind route for ship navigation.

Worldwide, ecofriendly energy development is being actively carried out to protect the environment and reduce carbon dioxide emissions [7]. Wind power is one of the fastest-growing ecofriendly energy sources [8]. Wind power generation is being installed on land and offshore, and it is shifting to the sea due to social issues and land development constraints [9]. This movement has an impact on ship safety because it leads to large-scale installation at sea [10]. In addition, the development of Marine Spatial Planning (MSP)

divides spatial and temporal zones using the ocean and preserves sensitive areas [11]. In the case of Korea, there are nine maritime spatial plans. Merchant ships use three types of ports and navigation areas in this maritime space: traffic designated areas, harbor areas, and Traffic Separate Schemes (TSS). Ships' traffic routes outside of legally mandated ports and shipping zones are not preserved [12]. Therefore, it is intended to understand the traffic flow in the coastal waters of Korea and to preserve the main maritime traffic flow. It is expected that this will secure the maritime traffic area in the maritime space and result in safe vessel operation at sea. Furthermore, it is only intended to be used to create a national maritime traffic route map, and it is presented as a methodology.

It utilizes the ship's automatic identification system (AIS) data to create a maritime traffic route. AIS data include Maritime Mobile Service Identity (MMSI), Global Positioning System positions, speed over ground (SOG), course over ground (COG), type of ship, length of all (LOA), breath of ship, and time and place of arrival, and adjacent ships and AIS stations exchange data [13]. The International Maritime Organization (IMO) mandates that cargo ships with a gross tonnage of 300 tons or more engage in international voyages, ships with a gross tonnage of 500 tons or more do not engage in international voyages, and all passenger ships have AIS [14]. AIS data are an important component for analyzing a ship's operational pattern. If the SOG, COG, Rate of Turn, Voyage Data Recorder, etc. of a ship are analyzed by density-based spatial clustering of applications with noise (DBSCAN) clustering, it is possible to identify the actual ship's engine order and use of rudder from the past data and prevent ship accidents. The work in [15] is an example. Recently, a large amount of real-time AIS data have been accumulated, and research using these data is being actively conducted [16]. AIS data are used in a variety of fields, but the parts relevant to this study can be divided into three categories. To begin, Wang et al. and Lee et al. have researched maritime traffic route extraction. In the study of Wang et al., the Douglas–Peucker algorithm was used to simplify the trajectory of the ship, and the hierarchical trajectory was clustered using Hierarchical DBSCAN (HDBSCAN) [17]. The study by Lee et al. extracted density-based routes by analyzing 90% of maritime traffic with kernel density estimation (KDE) based on a geographic information system (GIS). Second, research on maritime traffic networks is being conducted. Arguedas et al. [18] propose a Traffic Route Extraction and Anomaly Detection method that generates a detailed synthetic maritime traffic network based on a ship's historical position data. Frorti et al. [19] proposed graph-based maritime traffic, and an unsupervised approach was developed to extract patterns of life of ships [20,21]. Furthermore, here, by applying Ornstein–Uhlenbeck's mean-reverting stochastic process, we describe the application of graphs to detect waypoints through ship motion modeling and connect them [22,23]. Yan et al. [14] propose a maritime network based on the ship trip semantic object framework. In this case, we employ the method of locating and connecting stop areas and waypoints. According to Wang et al. [24], we propose a network creation method that is distinct from existing ones by defining lane boundaries, lane centerlines, and junctions to extract the maritime road network. Lee et al. classify the wake patterns using the ship's AIS data. In this study, the DBSCAN algorithm was used to classify the patterns according to the arrival and departure of ships and presented guidelines for ship operation [25]. Furthermore, the vessel berthing speed was analyzed using machine learning, and the best receiver operating characteristic curve is presented [26]. As a result, methodologies from other fields are being integrated into the maritime industry, and scientific techniques are actively implemented [27].

In this study, the most optimal data for generating maritime traffic routes are calculated using one year of AIS data. The goal is not to construct detailed routes for all ships but rather to structure main routes that are commonly used and formed in Korea. For route calculation, it is divided into 16 areas. Furthermore, the data were classified by ship type and ship size. The density analysis method is used to select the vessel passage based on the data generated in this manner. Vessel passageways are created by dividing them into three categories: main route, inner branch route, and outer branch route. In addition, the GIS-based Korean maritime traffic route is constructed, and the width of the route is

calculated and displayed. In addition, the main maritime traffic route developed with the ship's unique characteristics and operational factors can be chosen as the navigation route for the [maritime autonomous surface ship \(MASS\)](#) and can be used as a specific content area for the MSP's navigation and port. This result suggests the possibility of developing into a maritime network, and the optimal route and will contribute to the safe navigation of ships and the prevention of accidents [28].

2. Materials and Methods

2.1. Analysis Area

The coastal waters of Korea, located in northeastern Asia, are the study's analysis area. As a result of Korea's geographical characteristics, the maritime transportation industry plays an important role, and the development of the Korea Autonomous Surface Ship is being carried out actively by the 4th industrial revolution [29]. This type of industrial development has been combined with artificial intelligence (AI), big data, and the Internet of Things to develop safe navigation and maritime transport networks for ships [30]. Maritime traffic routes should be adequately required for MASS's autonomous operation. This is why a route based on actual AIS data from the nearby Korean Sea is required. Furthermore, recent efforts have been made to develop new and renewable energy sources in Korea's coastal waters. In particular, the composition of the offshore wind farm (OWF) is rapidly increasing, which is a dangerous factor in the customary route of ships [8]. It is even expanding to distant seas with good wind conditions with floating offshore wind power. At a time when securing passageways is becoming increasingly important due to the increasing size of ships, the subject of different marine uses collides. As shown in Figure 1, the neighboring waters of Korea were chosen as the analysis area to solve this problem and find an acceptable compromise point.

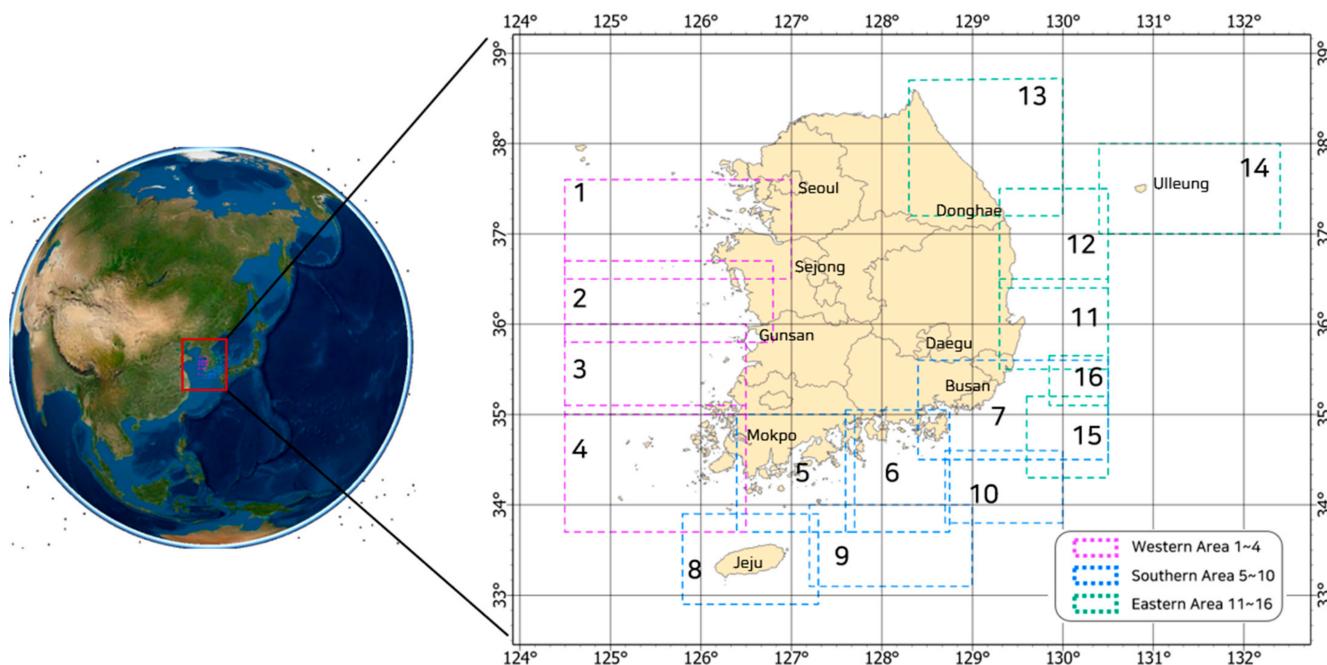


Figure 1. Analysis of Korean maritime traffic route location and area division details.

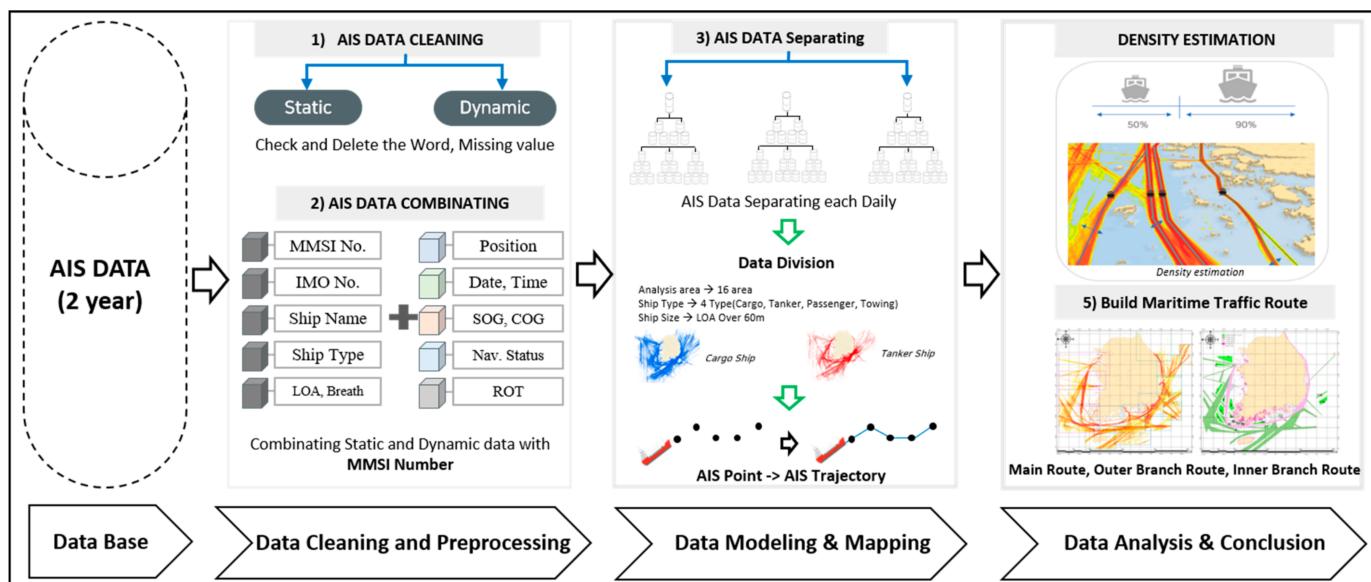
In Korea, the area under analysis was divided into a total of 16 sections to effectively analyze the vast amount of AIS data and understand the overall maritime traffic route. The West Sea region was divided into 4 areas, the South Sea region into 6, and the East Sea region into 6. The coordinate system used to divide regions is depicted in Table 1.

Table 1. Coordinate system of the analysis area.

No.	Latitude	Longitude	No.	Latitude	Longitude
Area 1	36.50° N–37.60° N	124.50° E–127.00° E	Area 9	33.10° N–34.00° N	127.20° E–129.00° E
Area 2	35.80° N–36.70° N	124.50° E–126.80° E	Area 10	33.80° N–34.60° N	128.70° E–130.00° E
Area 3	35.00° N–36.00° N	124.50° E–126.80° E	Area 11	35.50° N–36.50° N	129.30° E–130.50° E
Area 4	33.70° N–35.10° N	124.50° E–126.50° E	Area 12	36.40° N–37.50° N	129.30° E–130.50° E
Area 5	33.70° N–35.00° N	126.40° E–127.70° E	Area 13	37.20° N–38.70° N	128.30° E–130.00° E
Area 6	33.70° N–35.05° N	127.60° E–128.75° E	Area 14	37.00° N–38.00° N	130.40° E–132.40° E
Area 7	34.50° N–35.60° N	128.40° E–130.50° E	Area 15	34.30° N–35.20° N	129.60° E–130.50° E
Area 8	32.90° N–33.90° N	125.80° E–127.30° E	Area 16	35.10° N–35.65° N	129.85° E–130.50° E

2.2. Overview of Extracting the Korean Maritime Traffic Route

In this study, we performed the process of preprocessing the data and divided the data to be suitable for analysis consisting of analyzing and comprehending the traffic flow in Korea's coastal waters. This protects the main maritime traffic flow and considers improving the analysis processing speed by segmenting the AIS data and applying it to the MASS route. The maritime traffic route created in this way was created based on GIS to be used by the public and to increase the understanding of those involved [4]. Figure 2 shows the flow chart of the study.

**Figure 2.** Research of the flow chart of Korean maritime traffic route.

The analysis begins with a preprocessing operation that allows a large amount of AIS data to be analyzed appropriately. The ship's AIS data are displayed in the form of points at regular time intervals, and when the points are connected using the MMSI, a single ship trajectory is created [31]. The trajectory created in this way has a unique value for each vessel. The trajectory can check in which space the ships are concentrated through the density analysis, and through this, the sailing patterns of the ships can be analyzed [32]. Density analysis can be done in a variety of ways, but the density analysis method is used to choose the vessel route. The polygonal navigation area is determined by connecting the edges of the density analysis result and is divided into three categories. It is suggested that the domain of the unique maritime traffic route be preserved.

2.3. Preprocessing and Statistical Analysis of AIS Data Used for Analysis

The term AIS data refers to a device that allows for the transmission and reception of various information based on the vessel's movement. This information is navigational

equipment that aids in the prevention of ship collisions as well as equipment that aids in receiving ship information on land and advising on arrival and departure. The AIS data generated here are stored in real time in a vast amount and help decision making by deriving meaningful results [33]. This database stores AIS data for about two years, from 1 January 2018 to 31 December 2018 and from 1 September 2019 to 31 August 2020. The data capacity is saved in the form of a 1.1 TB comma-separated variable file. Static and dynamic data are the two types of data. Static data represent the ship's unique information and includes MMSI, name, type, IMO number, callsign, length, draft, and gross tons (GTs). In addition, dynamic data means data that changes when a ship operates, and it includes MMSI, date, latitude, longitude, SOG, COG, and heading. The MMSI value is used to combine these two types of data. For preprocessing, Python was used. The total capacity of all AIS data passing through Korea per day is approximately 1.5 GB, with approximately 20 million data points. To increase the efficiency of the data capacity and analysis speed, the area was divided into 16, and to secure the main routes of commercial ships, it was extracted into four categories: cargo ship, tanker ship, passenger ship, and towing ship. Furthermore, in the case of small ships and fishing vessels, the ship's operation pattern is free and irregular, so the ship's size was classified as 60 m or greater [4]. As the distance from the land increases, the data loss rate occurs owing to the limitations of AIS. Table 2 displays the database capacity as well as the results of dividing the vessel type and vessel length for use in the analysis.

Table 2. AIS data characteristics.

Categorization	AIS
Data Period	1 January 2018–31 December 2018 1 September 2019–31 August 2020
Data Volume	Approx. 1.1 TB
Analysis Area	Republic of Korea
Ship Type	Cargo Ship, Tanker Ship, Passenger Ship, Towing ship
Ship Length	60 m or Over of Ship's Type (≥ 60 m)

The characteristics of ship operation vary according to the season. Typhoons primarily occur and pass through the seas of Northeast Asia during the summer and autumn seasons, and some factors, such as tropical cyclones, have an impact on ship operations. As a result, one week of data is extracted for each of the four seasons of 2018: spring, summer, autumn, and winter. Here, the days when typhoons and tropical cyclones do not occur are selected. As a result, traffic volumes were extracted from 1 to 7 March, from 1 to 7 June, from 1 to 7 September, and from 1 to 7 December. The Republic of Korea's sea area was divided into 16 areas based on the data extracted in this manner to efficiently use it for analysis. Table 3 displays the statistical characteristics of the data used in this study.

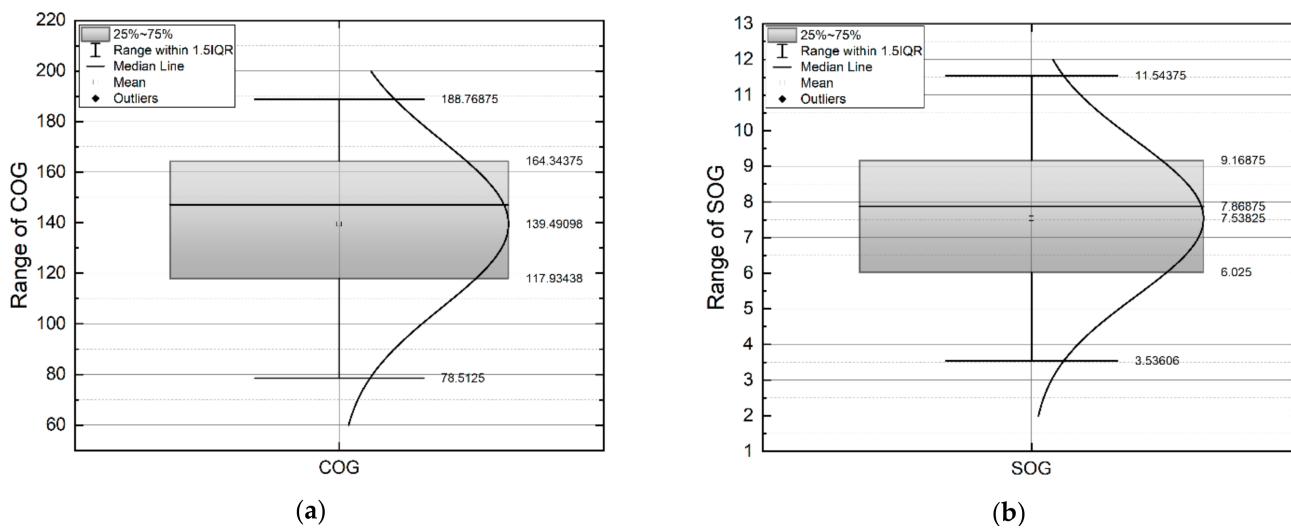
Table 3. AIS data statistics by area.

Area	Data Counting	COG (°)		SOG (kts)		GT (ton)		Length (m)	
		Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Area 1	7,125,269	171.1	5.9	13,403.4	131.3				
Area 2	2,424,427	140.4	9.2	17,477.7	144.8				
Area 3	2,360,251	158.0	8.8	14,927.5	129.5				
Area 4	7,249,111	188.8	9.6	26,748.6	159.0				
Area 5	3,712,294	154.1	8.4	10,769.9	116.8				
Area 6	8,023,942	170.4	6.1	12,811.6	111.2				
Area 7	14,654,929	169.2	3.8	14,419.0	122.2				
Area 8	2,954,921	134.5	8.5	17,348.6	128.6				
Area 9	2,744,786	119.2	11.5	41,535.5	172.3				
Area 10	1,464,791	116.7	11.0	33,849.4	160.8				

Table 3. Cont.

Area	Data Counting	COG (°)	SOG (kts)	GT (ton)	Length (m)
		Mean	Mean	Mean	Mean
Area 11	4,664,982	152.0	3.5	73,756.7	120.8
Area 12	1,322,681	159.5	7.9	16,218.0	118.3
Area 13	959,212	142.2	6.2	6844.9	96.7
Area 14	931,028	80.9	6.5	14,566.8	99.3
Area 15	857,881	96.4	6.4	19,763.3	109.8
Area 16	661,153	78.5	6.0	21,415.2	110.9

The total number of data from analysis Areas 1 to 16 is about 62 million. Although some analysis areas have overlapping regions, it is used to ensure the smooth connection of maritime traffic routes. In Korea, Analysis Area 7 has the most data. The sea area with representative ports including Busan and Ulsan has the largest amount of data. In addition, this sea area contains Korea's representative container port, oil tanker port, and car carrier port. COG is primarily used to examine ship navigation patterns. Although there is a limit to suggesting the movement of all ships as an element that informs the ship's direction, the pattern can be outlined in a large frame. AIS data have a different transmission/reception interval depending on the ship's SOG and status [34]. The SOG of a ship expresses the relative speed of the ship relative to the Earth's surface. Vessels have relatively low SOG for entering and leaving major ports, and sea areas with long passageways have large SOG values. There are no ports in neighboring waters in Areas 9 and 10, and the pattern of ships passing with high SOG can be confirmed. Anchored or drifting ships are not properly considered in this task, but it is possible to determine which sea area has the fastest speed. A ship's gross tonnage is one of the most widely used types of volumetric tonnage and represents the ship's cargo hold in volume units. As it shows the largest value in Area 11, it is possible to check which sea area mainly large ships passed through. Area 11 is Korea's largest steel company, and bulk ships transport iron ore and coal. Length overall (LOA) is the most commonly used term for length, which is an element that represents a ship's length. The longest vessel can be seen passing through Area 9. Figure 3 depicts the results of the data as a box plot.

**Figure 3.** Cont.

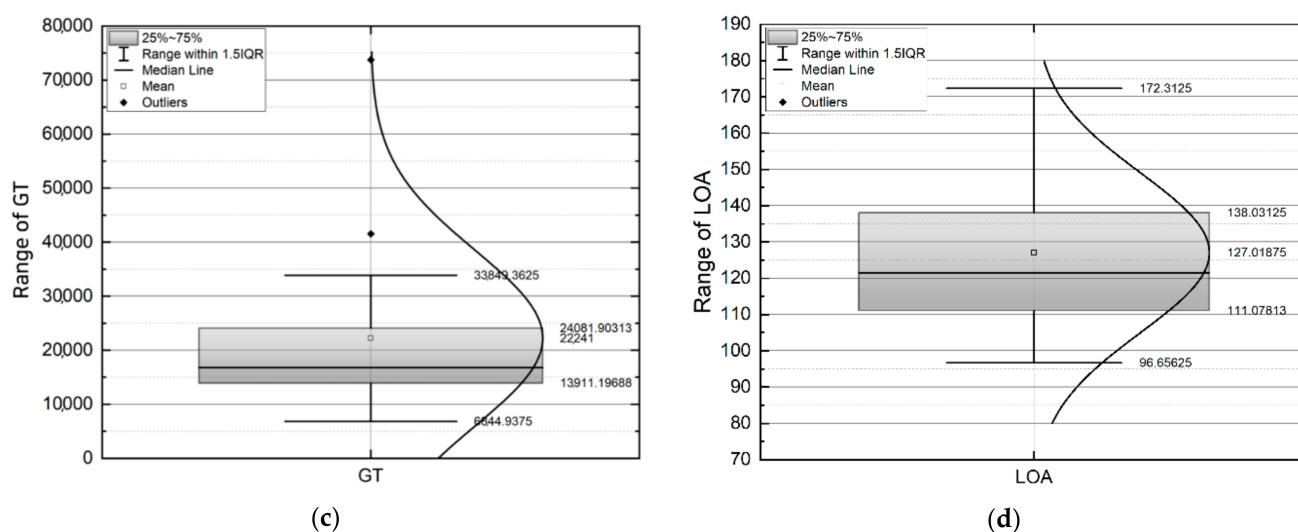


Figure 3. This figure shows a box plot and a normal distribution form to identify the statistical characteristics of the data used for analysis. (a) As a result of showing the COG, the mean value is about 139.5° , 25% and 75% of the data distribution are analyzed from about 118° to 164° . (b) As a result of representing SOG, the mean value was about 7.5 kts, 25% and 75% of the data distribution were analyzed from about 6 kts to about 9 kts. (c) GT of the ship by showing the results; it can be seen that the mean value is analyzed as about 22,241 tons, and the outlier appears as a value of 73,756 tons. (d) As a result of showing the LOA, the distribution of percentage was analyzed from about 111 to 138 m.

2.4. Density Analysis for Extracting Main Maritime Traffic Routes

The AIS data collected consist of data in the form of points that constitute a coordinate system converted to World Geodetic System 84 (WGS) coordinates. One point is made up of spatial information and ship attribute data, and it can be used for a variety of purposes, including weight application. The ship's point data are represented as a line to represent the ship's trajectory. The ship trajectory is created by listing the points of the ship in chronological order based on the same MMSI. The exceptional trajectory of a single vessel can clarify the vessel's navigation pattern, and the density can be used to determine which sea area the vessel has passed through the most. The density analysis based on the wake equator includes the estimation of line density and kernel density (KDE). As a result of the characteristics of AIS data, data are continuously accumulated in berthing ships and drafting ships at almost no speed, resulting in a difference in data amount in some areas. However, if the density analysis is performed based on the ship's track record, the limitation of concentration in some sea areas can be solved. In addition, KDE analysis was performed to explain the difference between point and line analysis results [35]. As such, KDE is a nonparametric method for estimating data distributions. We could add all of the formed kernel functions and then total them before dividing this figure by the number of data [36] after generating the kernel function with a focus on the value of the data for each observed dataset. Equation (1) is used to calculate KDE.

$$f_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (1)$$

x = random variable; x_i = obsevation; K = kernel function; h = bin width

Here, h is replaced with the bin size of the histogram and used as a parameter for adjusting the kernel; the kurtosis of the kernel is determined according to the size of h . With the kernel function applied to the total number of observations, K is divided by the sum of the observations. In other words, Equation (1) describes a method for calculating the average distribution after it has been mapped to the kernel. It is important to select an

appropriate bandwidth for KDE because over-smoothing may occur in areas with high data density and under-smoothing may occur in areas with low data density. In short, the resulting bandwidth is a “free parameter” that has a significant impact on the estimates.

In addition, line density analysis is a method of calculating the density of linear characteristics near each output grid cell. It is possible to apply a weight to the ship property data to indicate the density of each track map. Equation (2) is the calculation formula.

$$\text{Line Density} = \frac{((L_n \times V_n) + (L_{n+1} \times V_{n+1}))}{\text{Area of Circle}} \quad (2)$$

By applying a search radius to any grid cell, a circle is formed around the center of each raster cell. We multiplied the field value by the trajectory of ships in this circle. Here, L_n means the ship’s track record of n , and L_{n+1} represents the ship’s track chart of $n + 1$. In addition, V_n denotes an attribute value of a vessel of n , and V_{n+1} denotes an attribute value of a vessel of $n + 1$. Ship density analysis is used and applied in a variety of contexts. High and low levels of traffic can be identified in the relevant sea area as a measure of the vessel’s traffic volume. The density analysis result of 60 m or more, excluding small ships and fishing vessels, is a suitable method for selecting the main route rather than analyzing the routes in all sea areas.

2.5. Related Study and Application to Extract Vessel Passage

In various studies, AIS data were used to find the route of a ship and to apply it to the network. Furthermore, various studies are being conducted to calculate the safety distance with OWFs by understanding the vessel’s traffic volume and applying it to the MSP. Since the methods of securing maritime traffic differ depending on the characteristics of each country, this study will also take Korea’s characteristics into account. First, the maritime facility installed in the artificial island and exclusive economic zone of the sea, etc., must pass through with a safety distance of 500 m [37]. As a result, a search radius of 500 m was chosen and analyzed in this study based on the individual ship’s track map. Furthermore, as a result of the density analysis that appears, each cell was chosen to be 100 m. The debate over how much data to collect to calculate traffic volume from the perspective of maritime traffic is primarily focused on MSP. In the UK and Belgium, 90% of the total sea traffic was applied and selected as a sea transport route, and it was preserved as a unique shipping area with the highest priority [38,39]. In this study, it was conducted in the form of removing the lower value by applying the density analysis result, not 90% of the UK’s total maritime traffic. To select Korea’s maritime traffic route in this study, a quantile method was used, which divides all density into equal values ranging from 0% to 100% based on the analysis results. As a result, as shown in Figure 4, the total sea traffic was divided into 90%, 75%, and 50%.

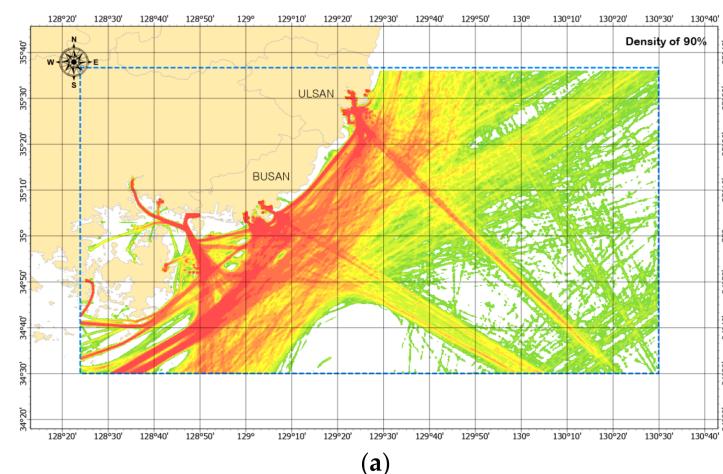


Figure 4. Cont.

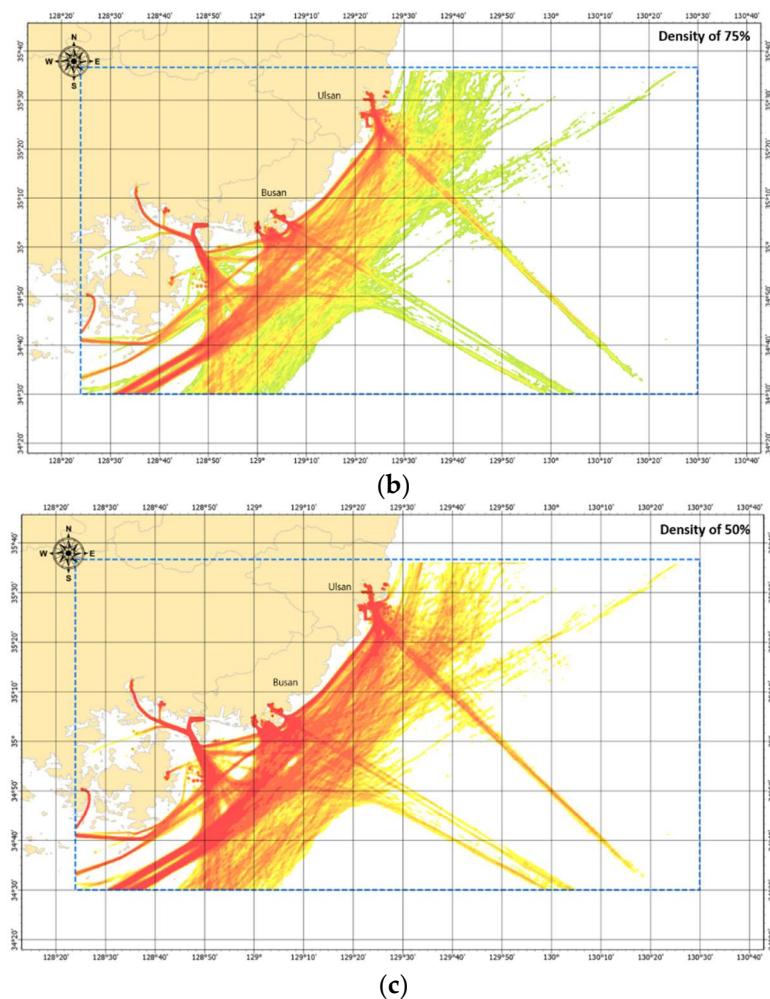


Figure 4. As the result of analysis of maritime traffic density in sea Area 7, the sea area with the highest traffic volume in Korea (a) shows 90% of the total sea traffic volume and shows that all sea areas are covered (b) of the total sea traffic volume. The result shows 75%, and the result is relatively less than 90%. (c) The result shows 50% of the total maritime traffic and shows the main maritime routes.

Figure 4 depicts the analysis Area 7 results in three different formats. It is necessary to select a percentage of an appropriate value to create Korea's main maritime traffic route while determining the type of vessel, size of the vessel, and seasonal characteristics. The route for navigating the origins and destinations of ships considers economical and safe routes as operational factors first [40]. As a result, the majority of merchant ships follow similar paths and can be classified as a single traffic route. Figure 4a depicts the safety distance between OWFs and ships in the United Kingdom, where 90% of traffic volume is [38]. The majority of the sea area has a traffic volume and a traffic volume limit that includes a small traffic volume. If it is impossible to use it for other purposes for the safe operation of the ship, a sustainable interest will not be established, so the result is excluded. Figure 4b depicts the results, which show that 75% of the total sea traffic volume is represented. According to Chebyshev's inequality, this figure results from using a range of 2σ . Chebyshev's inequality is more general, stating that for a wide range of different probability distributions, a minimum of 75% of values must lie within two standard deviations of the mean [41]. In other words, Chebyshev's inequality indicates that at least 75% of the data distribution exists within the 2σ of the probability distribution under any circumstances [42]. Therefore, it is the result of visually confirming whether the application of this study to the maritime traffic route is possible. To extract the main

maritime traffic route, it shows a result that is relatively less than 90% of the result. Figure 4c is the result of applying 50% of the total sea traffic volume. It is the result of selecting the nearest percentage of the most similar value after calculating the standard deviation value of the result shown in the analysis Area 7. Among the three percentages, 50% was chosen as the result that could represent the main route while preserving the ships' traditional passage route. Table 4 shows the threshold value of each percentage of sea traffic. Since 90% has the largest number of included data, the threshold value is the smallest, and it gradually increases as it approaches 50%. The density result value distributed in this way calculates the standard deviation (SD) value, and it is possible to derive the mean value. The data are presented in this direction for each percentage, but the value most similar to the SD value is 50%, which is an appropriate value for selecting the main route. Some analysis areas produce results that are similar to the SD value of 75%, but the same percentage is used.

Table 4. Probability-based total sea traffic percentage result.

No.	Threshold of 90%	Threshold of 75%	Threshold of 50%	Standard Deviation	Mean
Area 1	166	332	664	863	224
Area 2	52	104	234	275	124
Area 3	72	145	217	228	132
Area 4	86	173	281	373	239
Area 5	132	308	551	549	236
Area 6	207	345	552	810	443
Area 7	177	444	799	796	403
Area 8	90	180	316	279	233
Area 9	131	196	268	208	258
Area 10	57	173	385	353	293
Area 11	73	184	332	305	233
Area 12	47	78	118	108	93
Area 13	48	96	144	105	36
Area 14	26	52	79	50	40
Area 15	55	105	161	177	163
Area 16	104	160	225	116	188

As previously stated, the methods used to select maritime traffic routes differ by country. Apply various standards derived from data analysis results as acceptable standards for each country. It has been explained how much data will be used to select major routes from a large number of maritime traffic routes. Furthermore, it is suggested that 50% of the total marine traffic volume be selected as the main route that should be kept as a minimum for harmonization with the entities that use other marine spaces such as shipping routes, ecofriendly energy development, due to the Republic of Korea's characteristics, and fishing activities. There may be differences between countries, and there is a point where it has a different percentage depending on topographical characteristics.

3. Results of the Korean Maritime Traffic Route Analysis

3.1. Results of Main Maritime Traffic Routes Based on Density Analysis

The maritime traffic routes must be chosen to ensure the safety of ships and maritime activities. Except for the traffic designated area, harbor area, and TSS, no stipulated shipping and port areas exist in Korea. This is increasing conflict between OWFs and other factors such as the development of marine energy. Marine accidents on ships are very risky in terms of scale and environmental pollution. Therefore, we propose a density analysis method to select the main route that should be kept to a minimum and use it as a methodology. The density analysis method used to select the essential sea area for ship navigation is a result value based on 50% of total maritime traffic volume.

As explained in Section 2, the analysis area is divided into 16 areas to reduce the data volume and improve analysis efficiency. Following that, the data from each area are

combined into a single layer by dividing the data by season, ship type, and ship size. The 16 analysis areas overlap to form a single connected traffic stream. If this process is applied to 16 analyzed sea areas, one major maritime traffic route in Korea is generated, and the result is shown in Figure 5.

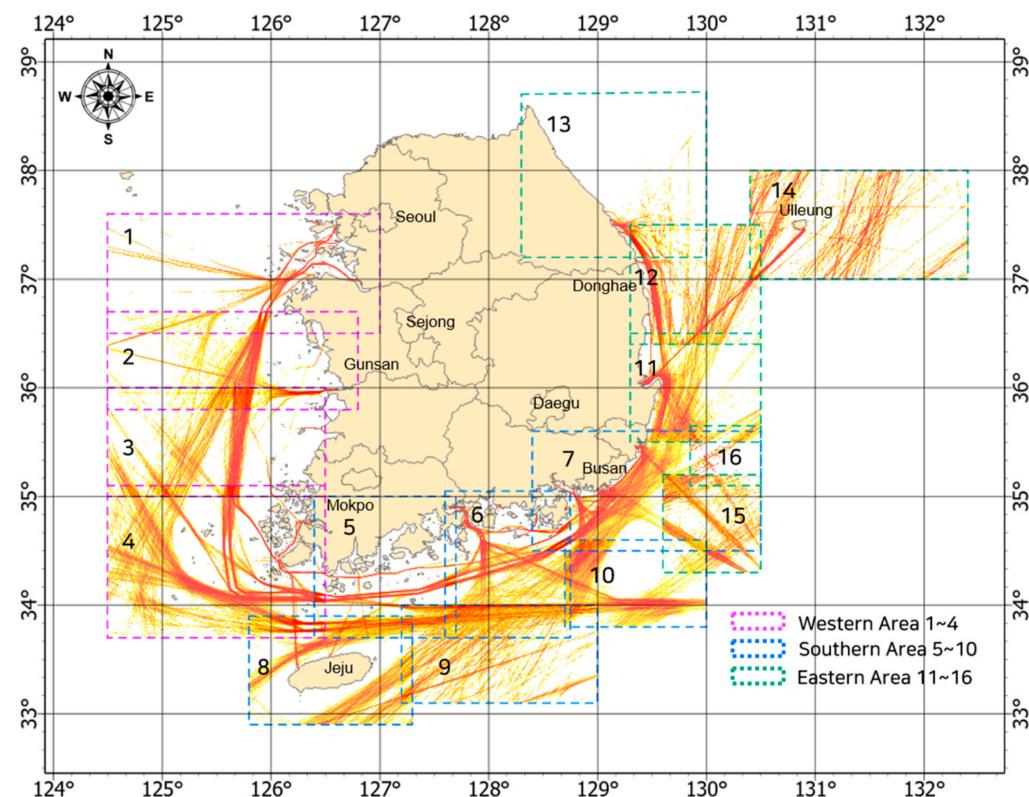


Figure 5. Final result of the Korean main maritime traffic route.

3.2. The Korean Maritime Traffic Route Classification and Definition

The maritime route is not separated from the land road. In terms of roads on land, the highway that connects cities is representative, and it is referred to by various terms such as the main road or artery road, and branch road, which refers to a small road. Currently, the maritime route is not defined in detail and is given different names depending on the purpose of use. In this study, Korea's maritime traffic routes are classified into three types.

3.2.1. Definition of the Main Route

As a result of this, the location with the highest traffic volume and flow was chosen as the main route. It is specifically composed in the form of circulation around the Republic of Korea. The development of marine technology, the acceleration of marine resource development, and the development of new and renewable energy are rapidly increasing. As a result of the rapid development of OWFs, it is especially important to secure a safe main route for ships. Figure 6 depicts the main route selection result. The main route is a one-of-a-kind area that must be preserved first for the vessel's safety. A total of 52 vertex points were extracted and composed of one area.

3.2.2. Definition of the Outer Branch Route

An outer branch route is a route that branches off or exists outside of the main route in Korea's coastal waters. As the outer branch route is far from the land, the density is low, as opposed to the sea area adjacent to the land due to the data reception limit.

However, the density is widely distributed, because it has the characteristic of approaching ports from various seas to enter domestic ports. It also has the characteristic of

becoming narrower as it approaches the main route. Although density analysis results are shown in some sea areas, it is difficult to select all sea areas as routes, so only clear routes were selected as outer branch routes. Figure 7 shows the analysis results.

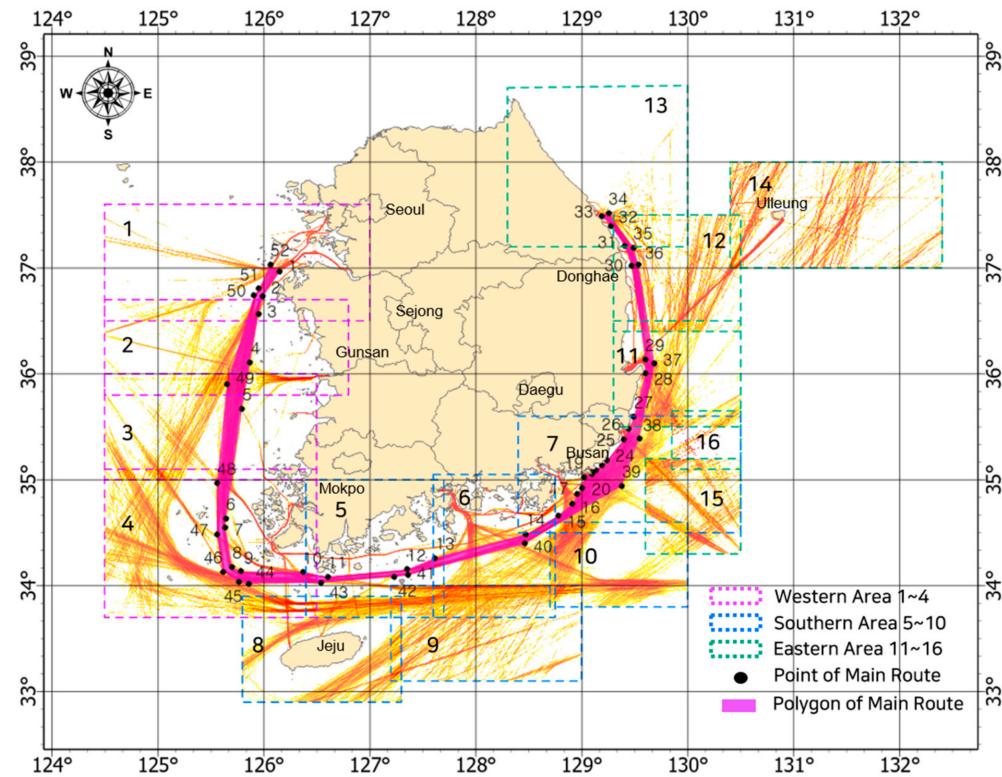


Figure 6. Result of the main route selection.

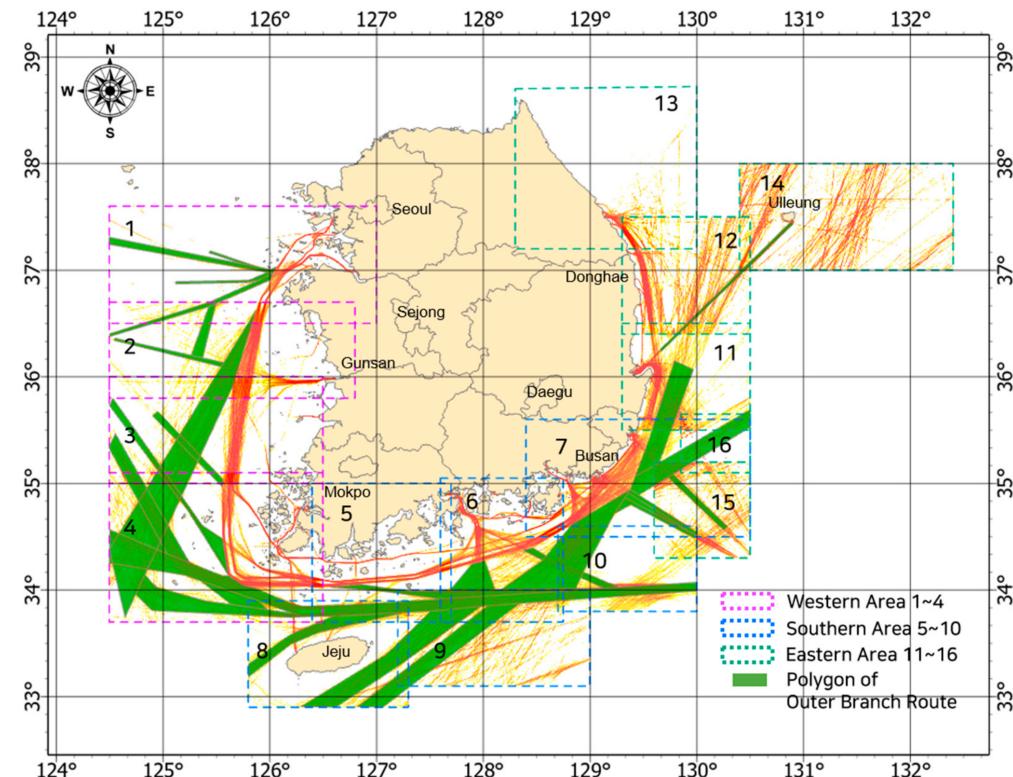


Figure 7. Result of the outer branch route selection.

The outer branch route is notable for being the primary route for large ships. Small ships primarily use the main route or an inner branch route, but some routes are frequently used by ships for the safety of the ship's navigation. However, the outer branch route is not a sea area that preserves the flow of marine traffic. The passage width of the vessel is wide, and it shows a relatively free pattern. It is necessary to reduce friction with other users by implementing appropriate systems and regulations. A total of 27 outer branch routes were extracted, and each of them has unique characteristics.

3.2.3. Definition of the Inner Branch Route

The route divided from the main route of the coastal waters of Korea to the inner sea area is defined as the inner branch route. The inner branch route does not take into account legal routes and is primarily used by ships at domestic ports. It demonstrates results that are distorted toward coastal waters in particular. Furthermore, it is examined how it is formed to enter and exit the port. The sea area where the sea route for entering and leaving the port is formed is excluded from this result. In some sea areas, there is traffic for ships, but sea areas where the density is cut off or disappears were excluded from the analysis. Figure 8 shows the result of the inner branch route.

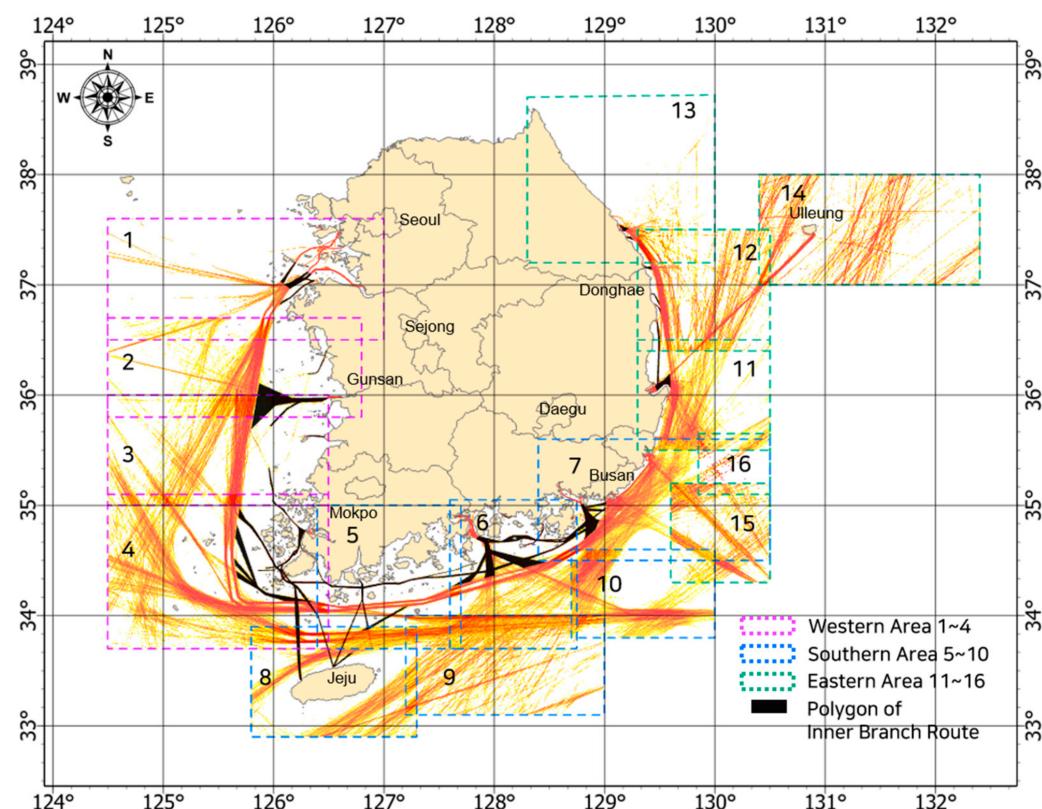


Figure 8. Result of the inner branch route selection.

This decision was made in response to the passenger ship that connects Jeju Island, and it was chosen as the inner branch route. Since the inner branch route, similar to the main route, is primarily used by ships, it is a sea area that must be preserved for the safe passage of ships. Moreover, the inner branch route is smaller in size than the ship using the main route or is mainly used as a passenger ship and towing ship connecting between islands and is used as a very customary route. A total of 58 inner branch routes were analyzed.

3.3. The Result of Korean Maritime Traffic Main Route as Polygon

Three types of routes were developed for ships passing through Korea's coastal waters. The main route has the most vessel traffic and circulation in the country; the outer branch

route has access to domestic ports from overseas or leaves Korea for foreign countries, and the inner branch route has access to ports and routes that are biased toward the coast. These are distinguished by the route. Figure 9 shows the final results of analyzing the data for one week by season in 2018 for cargo ships, tanker ships, passenger ships, and towing ships of 60 m or more. Some limitations were excluded from the analysis because small ships and fishing vessels exhibit a variety of and irregular navigation patterns. Furthermore, it does not show the passage routes of all ships, but it is useful in calculating Korea's main routes.

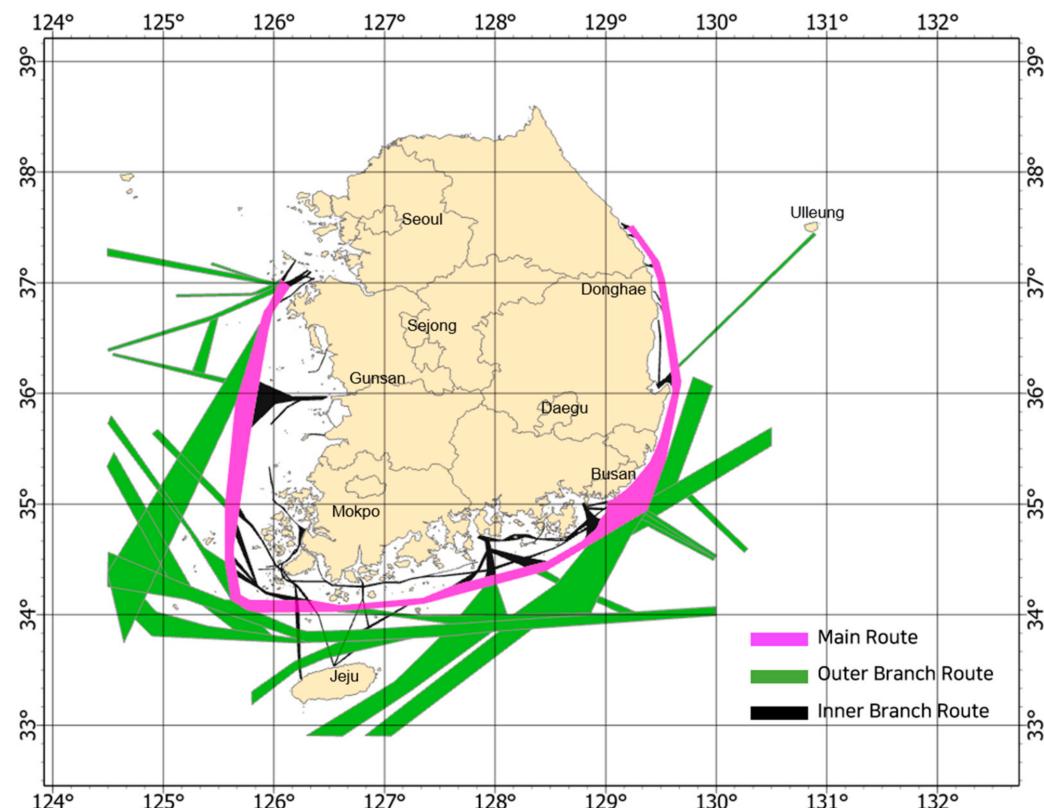


Figure 9. Result of the Korean maritime traffic main route as polygon creation.

4. Discussion

A sufficient water route is required for a vessel's safe navigation. While a ship is sailing, there is a regulation that requires an appropriate route width for emergencies or collision avoidance. The harbor and fishery design criteria are used to create the route in Korea. Here, the standard for the width of the route is 0.5–2.0 length overall (L), and restrictions are depending on the length of the passing vessel [43]. Furthermore, the World Association for Waterborne Transport Infrastructure (PIANC) is a global route design guideline centered on Europe [44]. The interaction of OWFs and maritime navigation of PIANC working group 161 is focused on resolving MSP disputes. In particular, it deals with the conflict between the OWF and ship navigation. In PIANC working group 161, the space required for the Starboard side wind farm and shipping route is 0.3 Nautical Mile + 6 Ship lengths + 500 m, while the space required for the Port side wind farm and shipping route is 6 Ship lengths + 500 m. According to the data, the largest sailing vessel is 399 m, which is calculated as 3.5 km and 2.9 km by the Starboard side and Portside, respectively, and it is calculated to have 6.4 km of free water on both sides.

As a result of the ship's steering characteristics, there is a difference in the safety distance between the Starboard and Port sides, as shown in Figure 10. The value chosen for both sides is 6.4 km, which is compared to the width of the main route calculated in this study. It can be seen that the place where the width is smaller than the value of PIANC does not have an appropriate distance with OWFs, and the location that is greater than the

PIANC value appears to have an appropriate distance. Figure 11 is a graph that extracts the route width of the main route for each vertex point.

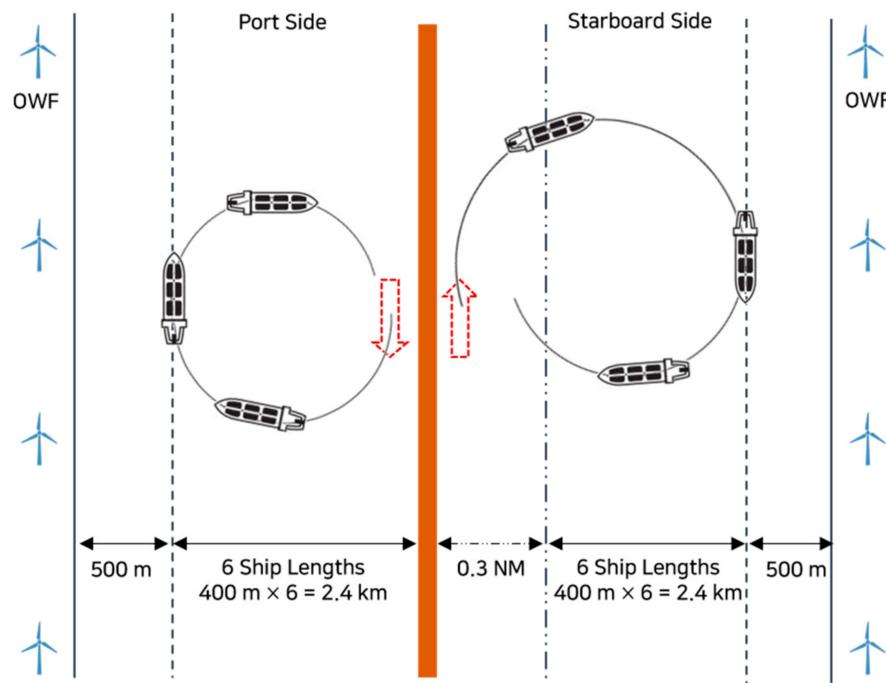


Figure 10. Safety distance between PIANC's ship navigation and OWFs.

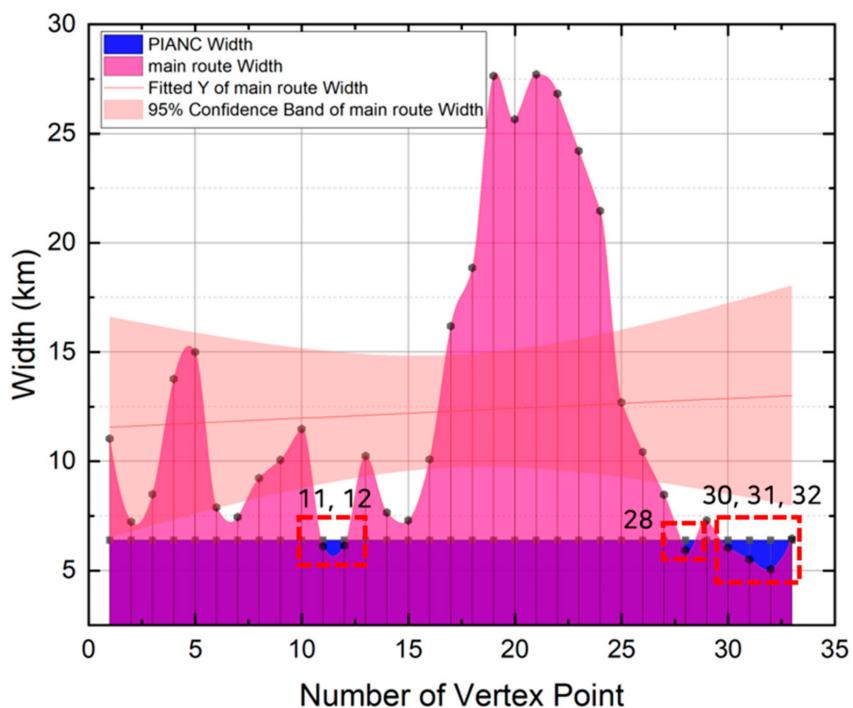


Figure 11. Comparison graph of width value of PIANC and width of the main route and 95% confidence interval of the main route.

It was stated that the outer branch route can be moved and changed through appropriate regulations, whereas the inner branch route has limitations in applying PIANC regulations because many places pass through very narrow sea areas. Therefore, in any case, only the main route that cannot be changed or moved was compared with the route

width. As a result, a location with a width less than 6.4 km appears to have a total of 6 armies. Table 5 displays the summarized result values.

Table 5. Summary of the main route width and PIANC calculation value.

Vertex No.	Width of Main Route (km)	Difference (km)	Vertex No.	Width of Main Route (km)	Difference (km)
1	11.04	+4.64	18	18.86	+12.46
2	7.23	+0.83	19	27.64	+21.24
3	8.49	+2.09	20	25.66	+19.26
4	13.77	+7.39	21	27.71	+21.31
5	15.00	+8.60	22	26.83	+20.43
6	7.89	+1.49	23	24.21	+17.81
7	7.45	+1.05	24	21.46	+15.06
8	9.23	+2.83	25	12.69	+6.29
9	10.06	+3.66	26	10.43	+4.03
10	11.48	+5.08	27	8.47	+2.07
11	6.11	-0.29	28	5.93	-0.47
12	6.14	-0.26	29	7.29	+0.89
13	10.24	+3.84	30	6.04	-0.36
14	7.65	+1.25	31	5.50	-0.90
15	7.29	+0.89	32	5.05	-1.35
16	10.09	+3.69	33	6.46	+0.06
17	16.19	+9.79			

The result of the safety distance with PIANC shown here is that the value means that it has a safety distance with OWFs, and the closer it is to the value, the less the safety distance decreases, which becomes a risk factor for the safe navigation of the vessel. This water area appears to be a phenomenon in which traffic is dense due to many nearby islands in the selection of the main route, with vertex points 11, 12, 28, 30, 31, and 32 (6 locations) showing a negative value. However, it is judged that it is necessary to understand the characteristics of the relevant sea area in detail or to apply safety regulations for the safe operation of the vessel. Based on a 95% confidence bend, it appears that all values do not overlap with the PIANC when analyzing the entire dataset. As a result, some sea areas' traffic characteristics must be reviewed, but in terms of reliability, they all appear to satisfy 6.4 km or more. In this study, the main route was verified with the value compared with the main route and PIANC working group 161. The width of the route, measured vertically at one vertex point, is the largest in the Busan coastal waters, and it becomes narrower toward the eastern coastal waters. It can be explained as an indirect representation of the complexity of maritime traffic. For the sustainable growth of offshore wind farms and maritime routes, area comparison is effectively performed.

5. Conclusions and Future Work

The purpose of this research is to extract maritime traffic routes carried out at the national level and to create major routes through which many ships pass. So far, there has been no outcome from the creation of maritime route maps. There are various studies on creating waypoints and legs to create a network of maritime traffic, but the research data extracted to the area where ships navigate are limited to the application of MSP. Furthermore, passage plans for vessel navigation are developed and implemented based on the navigation officers' and masters' experience. As ships execute their passage plans, a massive amount of AIS data is accumulated, and if expressed as a ship trajectory, it is possible to determine which route the ship chose and followed. In other words, a lot of AIS data are data from the past that safely completed the operation, so the result value itself can be reliable data. There are two reasons for establishing Korea's main maritime traffic route.

- First, it is to protect the area of the new and renewable energy zone that is being actively developed recently and the area of the ship passage within the MSP. So far, the

domestic shipping area has only been protected in a very small area, which is restricted to the vicinity of ports. To keep ship routes to a minimum, a 50% density analysis was used for cargo ships, tanker ships, passenger ships, and towing ships larger than 60 m, but not all ships. This result does not use all the data, and it is essential to keep the area to which these ships navigate. Based on the density analysis result, it is divided into three routes: the main routes, the outer branch route, and the inner branch route. Once again, it is divided into routes that cannot be moved or changed and routes that allow movement and change through appropriate regulations.

- Second, domestic maritime traffic routes are designed to respond to the emergence of MASS in the future. The massive amount of AIS data used in this analysis is a historical record of ships operating safely and efficiently. If a ship's navigation route is chosen with such conviction based on data, it is determined that it can be applied to the route of future ships. If a country's maritime route is selected and major waypoints are presented, it is expected that MASS will be able to use a route that connects the world beyond one country.

One limitation is that data from four different types of ships larger than 60 m were used to determine the most irreversible main route for this study. As a result of the characteristics of its cargo, each vessel's operation pattern and port differ. Furthermore, the density analysis result was set at 50%, and the maritime traffic route was chosen. Although there is a point in using different characteristics of Korea in the results shown using different values from previous studies, it is necessary to analyze them by applying various percentages. The limitations are planned to be carried out as part of a future study, and we will attempt to extract maritime traffic routes based on various data that do not limit other ship types and lengths.

Author Contributions: Conceptualization, I.-S.C. and J.-S.L.; methodology, I.-S.C. and J.-S.L.; software, J.-S.L.; validation, I.-S.C. and J.-S.L.; formal analysis, J.-S.L.; investigation, J.-S.L.; resources, J.-S.L.; data curation, J.-S.L.; writing—original draft preparation, J.-S.L.; writing—review and editing, I.-S.C. and J.-S.L.; visualization, J.-S.L.; supervision, I.-S.C.; project administration, I.-S.C.; funding acquisition, I.-S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This study was conducted with data provided by the Ministry of Oceans and Fisheries (MOF, Korea) for Ship traffic route research in 2021.

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

References

1. Tu, E.; Zhang, G.; Rachmawati, L.; Rajabally, E.; Huang, G.-B. Exploiting AIS Data for Intelligent Maritime Navigation: A Comprehensive Survey from Data to Methodology. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 1559–1582. [[CrossRef](#)]
2. United Nations. *U.N. Conference on Trade and Development, Review of Maritime Transport*; United Nation Publication: New York, NY, USA; Geneva, Switzerland, 2019; Volume 2019, pp. 14–20.
3. Weng, J.; Yang, D. Investigation of Shipping Accident Injury Severity and Mortality. *Accid. Anal. Prev.* **2015**, *76*, 92–101. [[CrossRef](#)] [[PubMed](#)]
4. Lee, J.S.; Son, W.J.; Lee, H.T.; Cho, I.S. Verification of Novel Maritime Route Extraction Using Kernel Density Estimation Analysis with Automatic Identification System Data. *J. Mar. Sci. Eng.* **2020**, *8*, 375. [[CrossRef](#)]
5. Yan, Z.; Xiao, Y.; Cheng, L.; Chen, S.; Zhou, X.; Ruan, X.; Li, M.; He, R.; Ran, B. Analysis of Global Marine Oil Trade Based on Automatic Identification System (AIS) Data. *J. Transp. Geogr.* **2020**, *83*, 102637. [[CrossRef](#)]
6. Redfern, J.V.; Mckenna, M.F.; Moore, T.J.; Calambokidis, J.; Deangelis, M.L.; Becker, E.A.; Barlow, J.; Forney, K.A.; Fiedler, P.C.; Chivers, S.J. Assessing the Risk of Ships Striking Large Whales in Marine Spatial Planning. *Conserv. Biol.* **2013**, *27*, 292–302. [[CrossRef](#)] [[PubMed](#)]
7. Kardakaris, K.; Boufidi, I.; Soukissian, T. Offshore Wind and Wave Energy Complementarity in the Greek Seas Based on ERA5 Data. *Atmosphere* **2021**, *12*, 1360. [[CrossRef](#)]

8. Lee, J.S.; Son, W.J.; Lee, H.T.; Cho, I.S. A Study on the Factors Affecting Optimal Site of Offshore Wind Farm from the Perspective of Maritime Traffic Using Spatial Analysis. *Korea Soc. Coast. Disaster Prev.* **2020**, *7*, 85–96. [[CrossRef](#)]
9. Soukissian, T. Use of Multi-Parameter Distributions for Offshore Wind Speed Modeling: The Johnson SB Distribution. *Appl. Energy* **2013**, *111*, 982–1000. [[CrossRef](#)]
10. Lee, J.S.; Son, W.J.; Lee, B.K.; Cho, I.S. Optimal Site Selection of Floating Offshore Wind Farm Using Genetic Algorithm. *J. Korean Soc. Mar. Environ. Saf.* **2019**, *25*, 658–665. [[CrossRef](#)]
11. Douvere, F. The Importance of Marine Spatial Planning in Advancing Ecosystem-Based Sea Use Management. *Mar. Policy* **2008**, *32*, 762–771. [[CrossRef](#)]
12. Lee, J.S. A Study on Navigable Spatial Planning Using Big Data on Maritime Traffic in Geographic Information System Environment. Master’s Thesis, Korea Maritime and Ocean University, Busan, Korea, 2020.
13. Son, W.J.; Lee, J.S.; Lee, H.T.; Cho, I.S. An Investigation of the Ship Safety Distance for Bridges across Waterways Based on Traffic Distribution. *J. Mar. Sci. Eng.* **2020**, *8*, 331. [[CrossRef](#)]
14. Yan, Z.; Xiao, Y.; Cheng, L.; He, R.; Ruan, X.; Zhou, X.; Li, M.; Bin, R. Exploring AIS Data for Intelligent Maritime Routes Extraction. *Appl. Ocean Res.* **2020**, *101*, 102271. [[CrossRef](#)]
15. Hu, Y.; Zhang, A.; Tian, W.; Zhang, J.; Hou, Z. Multi-Ship Collision Avoidance Decision-Making Based on Collision Risk Index. *J. Mar. Sci. Eng.* **2020**, *8*, 640. [[CrossRef](#)]
16. Murray, B.; Perera, L.P. An AIS-Based Deep Learning Framework for Regional Ship Behavior Prediction. *Reliab. Eng. Syst. Saf.* **2021**, *215*, 107819. [[CrossRef](#)]
17. Wang, L.; Chen, P.; Chen, L.; Mou, J.; Ship, A.I.S. Ship AIS Trajectory Clustering: An HDBSCAN-Based Approach. *J. Mar. Sci. Eng.* **2021**, *9*, 566. [[CrossRef](#)]
18. Arguedas, V.F.; Pallotta, G.; Vespe, M. Maritime Traffic Networks: From Historical Positioning Data to Unsupervised Maritime Traffic Monitoring. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 722–732. [[CrossRef](#)]
19. Froti, N.; Millefiori, L.M.; Braca, P. *Unsupervised Extraction of Maritime Patterns of Life from Automatic Identification System Data*; OCEANS: Marseille, France, 2019.
20. Millefiori, L.M.; Braca, P.; Bryan, K.; Willett, P. Modeling Vessel Kinematics Using a Stochastic Mean-Reverting Process for Long-Term Prediction. *IEEE Trans. Aerosp. Electron. Syst.* **2016**, *52*, 2013–2330.
21. Millefiori, L.M.; Braca, P.; Arcieri, G. Scalable Distributed Change Detection and Its Application to Maritime Traffic. In Proceedings of the IEEE International Conference on Big Data, Boston, MA, USA, 11–14 December 2017.
22. Millefiori, L.M.; Braca, P.; Willett, P.; Millefiori, L.M.; Braca, P.; Willett, P. Consistent Estimation of Randomly Sampled Ornstein–Uhlenbeck Process Long-Run Mean for Long-Term Target State Prediction. *IEEE Signal Process. Lett.* **2016**, *23*, 1562–1566. [[CrossRef](#)]
23. Ester, M.; Kriegel, H.P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *International Conference on Data Mining and Knowledge Discovery*; University of Munich Oettingenstr: Munich, Germany, 1996; Volume 4, pp. 226–231.
24. Wang, G.; Meng, J.; Han, Y. Extraction of Maritime Road Networks from Large-Scale AIS data. *IEEE Access* **2019**, *7*, 123035–123048. [[CrossRef](#)]
25. Lee, H.T.; Lee, J.S.; Yang, H.; Cho, I.S.; An, A.I.S. An AIS Data-Driven Approach to Analyze the Pattern of Ship Trajectories in Ports Using the DBSCAN Algorithm. *Appl. Sci.* **2021**, *11*, 799. [[CrossRef](#)]
26. Lee, H.T.; Lee, J.S.; Son, W.J.; Cho, I.S. Development of Machine Learning Strategy for Predicting the Risk Range of Ship’s Berthing Velocity. *Mar. Sci. Eng.* **2020**, *8*, 376. [[CrossRef](#)]
27. Serry, A. The Automatic Identification System (AIS): A Data Source for Studying. In *Maritime Transportmarit*; IDEES: Barcelone, Spain, 2016.
28. Zhang, X.; Wang, C.; Liu, Y.; Chen, X. Decision-Making for the Autonomous Navigation of Maritime Autonomous Surface Ships Based on Scene Division and Deep Reinforcement Learning. *Sensors* **2019**, *19*, 4055. [[CrossRef](#)]
29. Lee, C.H.; Kim, Y.S.; Shin, Y.R. Data Usage and the Legal Stability of Transactions for the Commercial Operation of Autonomous Vessels Based on Digital Ownership in Korean Civil Law. *Sustainability* **2021**, *13*, 8134. [[CrossRef](#)]
30. Jović, M.; Tijan, E.; Žgaljic, D.; Aksentijević, S. Improving Maritime Transport Sustainability Using Blockchain-Based Information Exchange. *Sustainability* **2020**, *12*, 8866. [[CrossRef](#)]
31. Sheng, P.; Yin, J.b. Extracting Shipping Route Patterns by Trajectory Clustering Model Based on Automatic Identification System Data. *Sustainability* **2018**, *10*, 2327. [[CrossRef](#)]
32. Lu, N.; Liang, M.; Zheng, R.; Liu, R.W. Historical AIS Data-Driven Unsupervised Automatic Extraction of Directional Maritime Traffic Networks. In Proceedings of the IEEE 5th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA), Chengdu, China, 10–13 April 2020.
33. Le, T.M.; Le, G.D.; Gourmelon, F.; Queffelec, B. How Can Automatic Identification System Data Be Used for Maritime Spatial Planning? *Ocean Coast. Manag.* **2018**, *166*, 18–30.
34. Hamada, K.; Wada, Y.; Nanao, J.; Watanabe, D.; Majima, T. Long-Term Ship Position Prediction Using Automatic Identification System (AIS) Data and End-to-End Deep Learning. *Sensors* **2021**, *21*, 7169.
35. Lampe, O.D.; Hauser, H. Interactive Visualization of Streaming Data with Kernel Density Estimation. In *IEEE Pacific Visualization Symposium*; IEEE: Hong Kong, China, 2011.

36. Botev, Z.I.; Grotowski, J.F.; Kroese, D.P. Kernel Density Estimation via Diffusion. *Ann. Statist.* **2010**, *38*, 2916–2917. [[CrossRef](#)]
37. United Nations. *U.N. Convention on the Law of the Sea*; Part V, Article 60; United Nations Publicaiton: New York, NY, USA; Geneva, Switzerland, 1982; p. 45.
38. UK Maritime & Coastguard Agency. MGN 543(M+F) Safety of Navigation: Offshore Renewable Energy Installations(ORELs). In *Safety of Navigation: Offshore Renewable Energy Installations (ORELs)—Guidance on UK Navigational Practice, Safety and Emergency Response*; MCA: Southampton, UK, 2016; pp. 7–23.
39. FPS (The Federal Public Service). *Marine Spatial Plan for the Belgian Part of the North Sea*; FPS: Brussels, Belgium, 2016; pp. 42–43.
40. Zhang, L.; Meng, Q.; Fang Fwa, T.F. Big AIS Data Based Spatial-Temporal Analyses of Ship Traffic in Singapore Port Waters. *Transp. Res. E* **2019**, *129*, 287–304. [[CrossRef](#)]
41. Kvanli, A.H.; Pavur, R.J.; Keeling, K.B. *Concise Managerial Statistics*; Cengage Learning: Boston, MA, USA, 2006; pp. 81–82, ISBN 9780324223880.
42. Chernick, M.R. *The Essentials of Biostatistics for Physicians, Nurses, and Clinicians*; John & Wiley Sons: New York, NY, USA, 2011; pp. 49–50, ISBN 9780470641859.
43. MOF. *Harbour and Fishery Design Criteria*; Ministry of Oceans and Fisheries of Korea: Sejong, Korea, 2017.
44. PIANC (The world Association for Waterborne Transport Infrastructure). *MarCom Working Group 161: Interaction between Offshore Wind Farms and Maritime Navigation-Design Guidelines*; PIANC: Brussels, Belgium, 2018.