

Article

Oil Flow Analysis in the Maritime Silk Road Region Using AIS Data

Yijia Xiao ^{1,2}, Yanming Chen ^{1,2,*} , Xiaoqiang Liu ^{1,2}, Zhaojin Yan ^{1,2}, Liang Cheng ^{1,2} and Manchun Li ^{1,2}

¹ Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Key Laboratory for Land Satellite Remote Sensing Applications of Ministry of Natural Resources, School of Geography and Ocean Science, Nanjing University, Nanjing 210023, China; MG1827078@smail.nju.edu.cn (Y.X.); MG1727062@smail.nju.edu.cn (X.L.); DG1727032@smail.nju.edu.cn (Z.Y.); lcheng@nju.edu.cn (L.C.); limanchun@nju.edu.cn (M.L.)

² Collaborative Innovation Center of South China Sea Studies, Nanjing 210023, China

* Correspondence: chenyanming@nju.edu.cn

Received: 7 March 2020; Accepted: 17 April 2020; Published: 20 April 2020



Abstract: Monitoring maritime oil flow is important for the security and stability of energy transportation, especially since the “21st Century Maritime Silk Road” (MSR) concept was proposed. The U.S. Energy Information Administration (EIA) provides public annual oil flow data of maritime oil chokepoints, which do not reflect subtle changes. Therefore, we used the automatic identification system (AIS) data from 2014 to 2016 and applied the proposed technical framework to four chokepoints (the straits of Malacca, Hormuz, Bab el-Mandeb, and the Cape of Good Hope) within the MSR region. The deviations and the statistical values of the annual oil flow from the results estimated by the AIS data and the EIA data, as well as the general direction of the oil flow, demonstrate the reliability of the proposed framework. Further, the monthly and seasonal cycles of the oil flows through the four chokepoints differ significantly in terms of the value and trend but generally show an upward trend. Besides, the first trough of the oil flow through the straits of Hormuz and Malacca corresponds with the military activities of the U.S. in 2014, while the second is owing to the outbreak of the Middle East Respiratory Syndrome in 2015.

Keywords: automatic identification system data; 21st Century Maritime Silk Road region; oil flow analysis; maritime oil chokepoint; Middle East Respiratory Syndrome

1. Introduction

“The Belt and Road Initiative”, comprising the “Silk Road Economic Belt” and “21st Century Maritime Silk Road” (MSR), was proposed by China to promote economic trade and deepen the connection between China and its associated countries [1–4]. The MSR region is not only the lifeline of China’s energy transportation [5] but also a key region in terms of maritime oil transportation [6]. The maritime oil chokepoint (MOC) is defined as the narrow channels along the widely used global sea routes by the U.S. Energy Information Administration (EIA) [7,8]. Within the MSR and its surrounding regions, the straits of Malacca, Hormuz, Bab el-Mandeb, and the Cape of Good Hope are four MOCs. These four MOCs control the oil import of the largest oil importer, China [6,9], and the oil export of the Persian Gulf, as well as other major oil importers and exporters. Therefore, the oil flow through these four chokepoints can show the real energy transportation within the MSR and its surrounding regions directly. Furthermore, transportation systems are often exposed to risks ranging from natural disasters to hazardous events caused by man [4,10,11]. Therefore, the monitoring of oil flow through those MOCs can show anomalies and help countries to respond promptly. Presently, the data on the

available average daily oil flows through MOCs from 2011 to 2016, released by the EIA, are calculated from the annual oil flow [7,8]. However, the information on the EIA data is too limited to meet the requirements for analysis on a smaller timescale. Therefore, obtaining and analyzing the oil flow in these four MOCs on multiple timescales is significant for ensuring China's maritime oil transportation security and strategic energy reserves.

To determine and analyze the oil flow, it is necessary to first assess the number of oil tankers passing through the MOCs at a certain time and generate statistics accordingly. In existing research, statistical methods for the ship traffic volume are divided into two categories: statistics based on imaging systems and statistics based on laser sensors. Statistics based on imaging systems employ images covering MOCs from radar, closed-circuit television (CCTV), infrared, and other imaging systems. Although the radar imaging system [12,13] presents the advantages of a large observation range and is not affected by the weather, not all areas can be observed by the system. In addition, the radar image can only provide spatial information about the ship. As supplements to the radar imaging system, the CCTV imaging system [14,15] and infrared imaging system [16,17] can obtain intuitive images, but they are extremely limited when it comes to monitoring wide ranges. For the statistics based on laser sensors [18,19], a laser beam is emitted to the target by a laser-ranging sensor, and it is received by the photoelectric element after being reflected by the target. Therefore, the position of the target is calculated based on the time elapsed from the emission to the reception of the laser beam and the speed of laser propagation, which is key to detecting ships through the MOCs. This method presents the advantages of round-the-clock operation, low cost, and high accuracy. However, it only obtains the spatial information about the ship and not the load information. Although the aforementioned methods can provide statistics regarding ship traffic volume, the imaging and laser sensors are only deployed near the port and can only obtain statistical information near the port. Furthermore, they present a common problem regarding the lack of load information in the calculation of the oil flow.

In 2000, the International Maritime Organization (IMO) adopted a regulation stating that internationally voyaging cargo ships of 300 gross tonnage or more, non-internationally voyaging cargo ships of 500 gross tonnage or more, and all passenger ships regardless of size are required to be equipped with an automatic identification system (AIS) [20–22]. Being a new type of spatio-temporal data, AIS data provide the potential for the oil flow analysis of the MOCs across four timescales (daily, monthly, seasonal, and annual). The AIS data can be received via shore-based facilities or satellite-AIS [20,23], which enables the worldwide monitoring of ship activities. These data, which include information such as the deadweight tonnage, IMO number (ship number issued by IMO), and the Maritime Mobile Service Identity (MMSI) number (ship number issued by MMSI), can be used to acquire unique ship identification and oil flow statistics. Previous studies mainly focused on ship spatial feature mining [6,24,25], ship anomaly detection [26–28], ship collisions [29–31], ship main route extraction [5,32–34], geospatial pattern analysis [35,36], catching assessment [37], and environmental pollution assessment [38–41]. However, studies on the oil flow statistics of MOCs are rare.

Considering the aforementioned problem, a maritime oil flow analysis technical framework is proposed herein. This framework promotes the statistics of oil flow to the scale of a single ship, and the temporal resolution and spatial resolution are improved when compared with the statistical data. Furthermore, the framework proposed in this study can calculate the transport volume of oil, which addresses the disadvantage that the previous methods have in only being able to calculate the deadweight of the oil tankers. We apply the framework to the straits of Malacca, Hormuz, Bab el-Mandeb, and the Cape of Good Hope by using the AIS data from 2014 to 2016. Therefore, we calculated and analyzed the oil flow through the MOCs, thus making available statistical data for ensuring the security and stability of oil transportation within the MSR and its surrounding region. Furthermore, with the support of real-time data, the oil flow through these choke points can be monitored for countries within the MSR and its surrounding regions to master and respond to special situations.

2. Study Area and Data

2.1. Study Area

The study area covers the MSR and its surrounding region, including the straits of Malacca, Hormuz, Bab el-Mandeb, and the Cape of Good Hope (see Figure 1). These four MOCs are our study targets due to their important strategic positions.

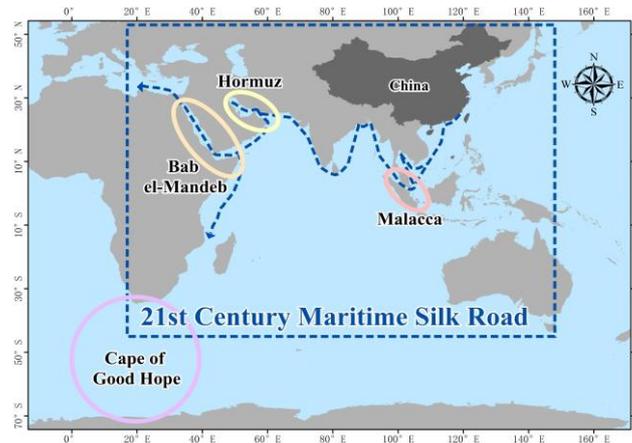


Figure 1. The study area.

The Strait of Malacca, which links the Indian Ocean and the South China Sea, is an important channel for West Asia to transport oil to East Asia. China, Japan, and South Korea, and some other countries regard it as the “lifeline” of energy transportation. Flow through the Strait of Malacca rose to 16 million barrels per day (b/d) in 2016, whereby the Strait of Malacca retained its position as the second busiest MOC [8]. The Strait of Hormuz, which links the Persian Gulf and the Arabian Sea, controls the oil export from the Persian Gulf. It is the busiest MOC worldwide, through which millions of barrels of oil travel globally every day [8]. The Bab el-Mandeb Strait, which links the Red Sea and the Arabian Sea, is an important waterway for maritime traffic and trade between Europe, Asia, and Africa. More than 20,000 ships pass through it annually, making it one of the most important and busiest straits. Suez tankers traveling from the Persian Gulf to Europe generally choose to travel through that strait. The Cape of Good Hope, which links the Atlantic Ocean and the Indian Ocean, was the best choice for ships to travel between Asia and Europe before the Suez Canal was navigated. It still works for supertankers that are unable to travel through the Suez Canal today.

2.2. Study Data

Limited by our existing research data, this study used the AIS data from 1 January 2014 to 31 December 2016. This study stored 1096 database files for an interval of one day. The number of the original AIS data is more than 7.1 billion, which accounts for 1594G. The data contain information for 30 attributes (presented in Table 1), which can be divided into three categories: static information, dynamic information, and voyage-related information [23]. Static information relates to the fixed physical characteristics of the ship itself. This information is recorded manually; thus, it is prone to missing data and errors. Dynamic information is the information that changes over time on a voyage. The navigation status information in the dynamic information is manually inputted, which has the values of “underway using the engine”, “at anchor”, “moored”, “underway sailing”, etc. The rest is generated automatically via the sensor connected to the AIS; therefore, it is of high reliability. The longitude and latitude information are in full precision (1/10,000 degree) [42]. Voyage-related information refers to the information that must be manually inputted before each voyage. This includes the details of the “estimated arrival time”, “destination”, and “draft.” Such information is generally reported to countries along the way via ship-shore data exchange.

Table 1. Field of automatic identification system (AIS) data.

Categories	Attribute		
Static information	1. MMSI	2. IMO	3. Vessel_name
	4. Callsign	5. Vessel_type	6. Vessel_type_code
	7. Vessel_type_cargo	8. Vessel_class	9. Length
	10. Width	11. Flag	12. Flag_code
	13. Vessel_type_main	14. Vessel_type_sub	
Dynamic information	1. Longitude	2. Latitude	3. Speed overground
	4. Course overground	5. Rate of turning	6. Heading
	7. Nav_status	8. Nav_status_code	9. Source
	10. Ts_pos_utc	11. Ts_static_utc	12. DT_pos_utc
	13. DT_static_utc		
Voyage-related information	1. Destination	2. Estimated arrival time	3. Draft

The load information was obtained from <https://www.myshiptracking.com/>, <https://www.vesselfinder.com/>, <https://www.marinetraffic.com/>, <http://ship.chinaports.com/>, and <http://marinelike.com/en/vessels/>. The obtained data were stored in a separate table, including three attributes: MMSI, Vessel_Type, and DeadWeight.

3. Study Method

In this study, a maritime oil flow analysis technical framework, which consists of data preprocessing, extraction of the ship point pairs, and discrimination of the oil tanker load condition, is proposed (see Figure 2). To develop this framework, this study used the AIS spatio-temporal massive data. The framework is applied to the straits of Malacca, Hormuz, Bab el-Mandeb, and the Cape of Good Hope.

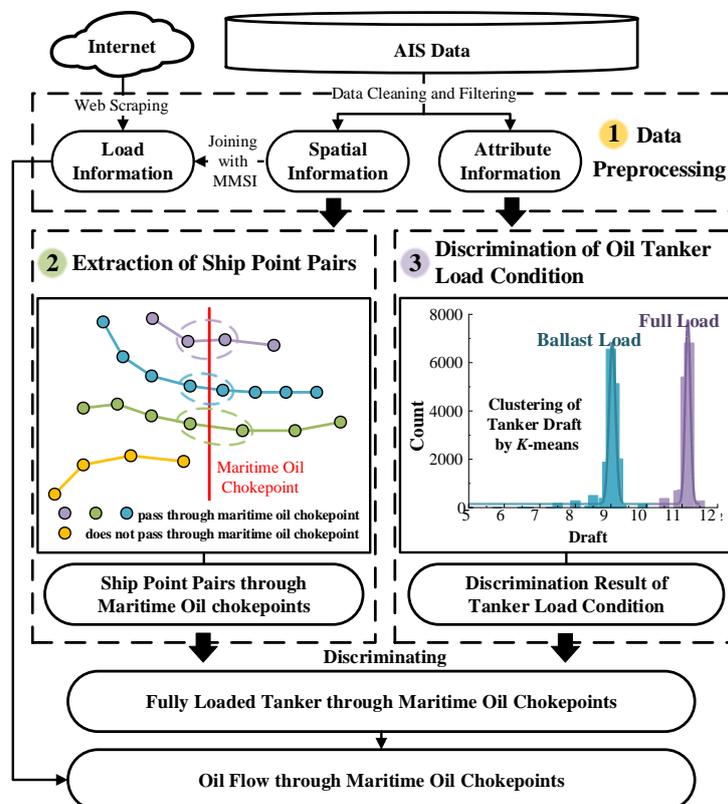


Figure 2. Main technical framework.

3.1. Data Preprocessing

The original AIS data inevitably contain errors and missing information due to the manual input of the data. Therefore, data cleaning is necessary to ensure the quality of the AIS data. The data that need to be cleaned are collectively called “dirty data.” Dirty data consist of three types of data: incomplete data, erroneous data, and duplicate data [43–45]. The standards for data cleaning are presented in Table 2.

Table 2. Data cleaning standards.

Data Type	Cleaning Standard
Incomplete data	Data whose IMO, MMSI, Draft, or TS_pos_utc are empty
Erroneous data	Data whose MMSI is not 9 digits long
Duplicate data	Data with the same property values for all fields

The AIS data contain non-tanker data and data outside the MOCs, which are not required for this study. Therefore, after data cleaning, data filtering is required to eliminate any redundant information. In this study, the criterion of data filtering is that the value of “Vessel_type_sub” is “crude oil tanker” or “oil products tanker” and that the AIS data is within four MOCs. The result of attribute filtering is used in Section 3.3. Because the oil flow statistics require load information, the result of attribute and spatial filtering is connected to the load information through the public key of the MMSI number, which is then used in Section 3.2. The process of data preprocessing is shown in Figure 3.

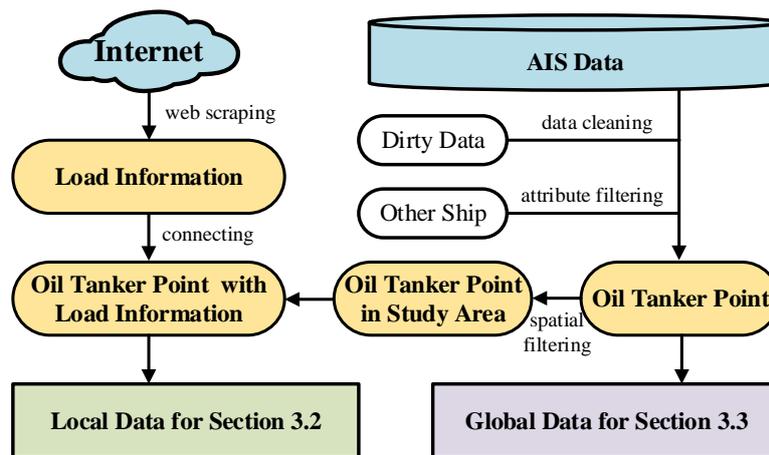


Figure 3. Data preprocessing.

3.2. Extraction of Ship Point Pairs Based on Ship Trajectory

It is necessary to generate the ship trajectory to determine whether the ship passes through a MOC or not. The ship trajectory is formed by connecting the set of ship points with the same MMSI number in chronological order (see Figure 4a). Its mathematical expression is $S_{MMSI} = \{V_1, V_2, \dots, V_k, \dots, V_{m-1}, V_m\}$, where MMSI in S_{MMSI} can identify the ship trajectory uniquely and V_k is the shipping point reported at time t_k .

The generated ship trajectory does not distinguish voyages, and it still has redundant information. Therefore, it is necessary to extract the ship point pair around the MOC from the trajectory. If adjacent trajectory points V_k and V_{k+1} are on both sides of the chokepoint, then they are extracted as the ship point pair $\{V_k, V_{k+1}\}$; otherwise, they would be dropped (refer to Figure 4b). As each ship point pair corresponds to the ship passing the MOC in a voyage, extracting the ship point pair can identify the voyage. This can simultaneously eliminate redundant information.

The adjacent points in some ship point pairs have significant time differences, which cannot be considered to be the same voyage. Therefore, we set the threshold as 24 h to remove these anomalies [5].

If the time difference between the adjacent points is more than 24 h, the ship point pair will be deleted; otherwise, it will be retained (see Figure 4c). The retained ship point pairs will be used for the subsequent oil flow statistics (see Figure 4d).

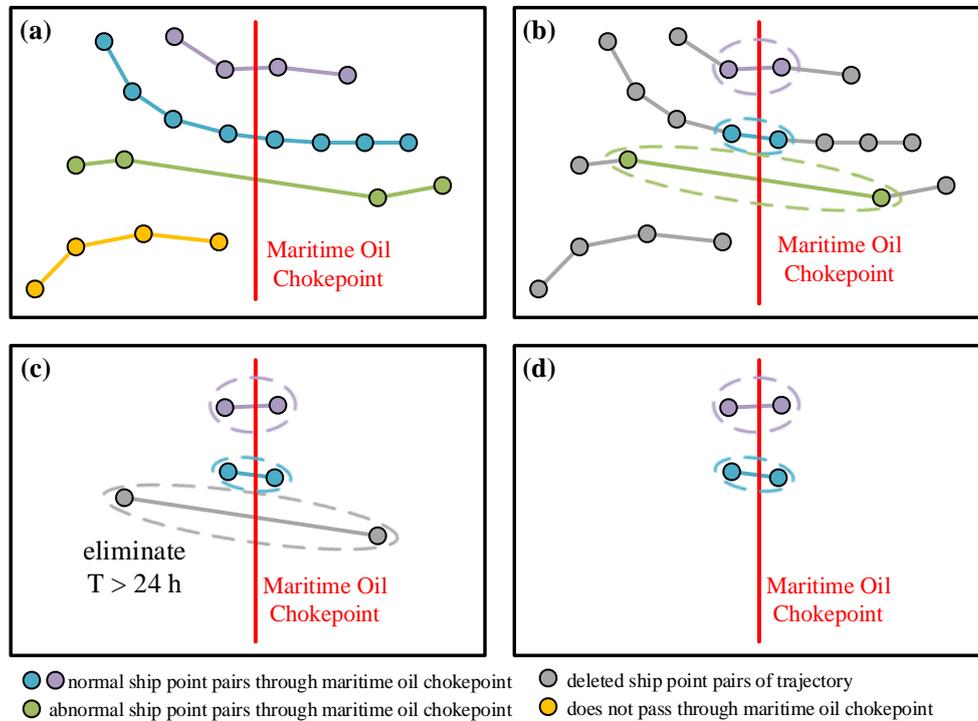


Figure 4. The extraction of ship point pairs. (a) Original ship trajectories; (b) The process of extracting the ship point pairs; (c) The process of eliminating abnormal ship point pairs; (d) The extracting result of ship point pairs.

3.3. Discrimination of the Oil Tanker Load Condition by the K-Means Clustering Method

The load condition of the oil tanker while passing through the MOC needs to be determined to develop the statistics for the oil flow. Park et al. mentioned that for the tanker and bulk carriers, the two most common operating conditions are the full load and ballast conditions [46]. Therefore, we only considered the full load and the ballast load of the tanker in the calculations, regardless of the other conditions.

For every oil tanker, we counted the record number under the different drafts and obtained the statistical results. The frequency histogram would be drawn by taking the draft as the abscissa and the record number as the ordinate (see Figure 5b). The frequency distribution map of the oil tanker draft presents a bimodal distribution structure, and the two peaks correspond to the ballast load and full load, respectively.

This fact enables us to discriminate the load condition of every oil tanker using the k-means clustering method [47,48]. The k-means clustering method is conducted by taking the draft as the clustering distance and the record number as the weight. The category with a small value of clustering results is considered to be the load condition of the “ballast load”, while the other category is considered to be the “full load” (see Figure 5c). There is a large difference between the draught under the full load and that under the ballast load. By using this method, the draught with a small deviation will not lead to a misjudgment of the load condition, which ensures that the result will not be affected by the unreliable draught.

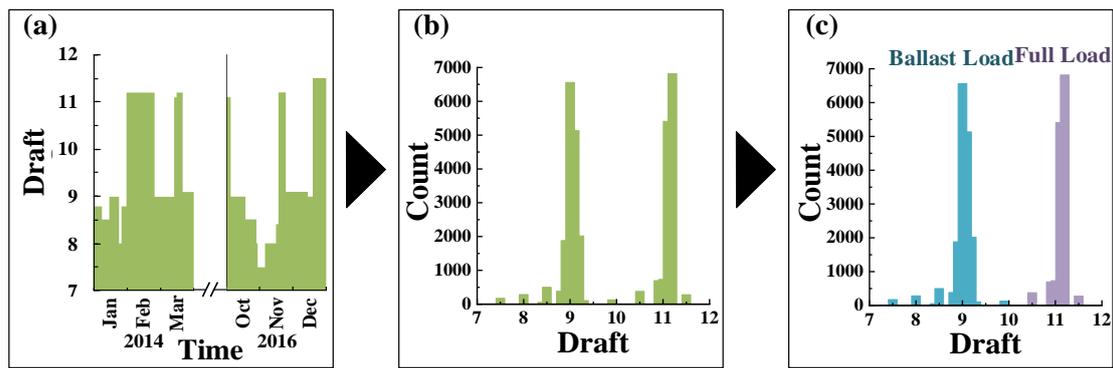


Figure 5. Discrimination of the oil tanker load condition. (a) The variation of the oil tanker draft over time; (b) The frequency distribution map of the oil tanker draft; (c) The clustering result of the oil tanker draft.

4. Results and Discussion

Combined with the discriminant result of the oil tanker load condition, the oil tanker point pair passing through the MOC can be identified as a ballast load or full load. For the condition of the full load, we used a deadweight to indicate the capacity of the tankers in a voyage. By contrast, for the condition of the empty load, zero was used instead. Therefore, the oil flow through the chokepoint can be obtained and analyzed across four timescales: the day, month, season, and year.

4.1. Annual Variation in Oil Flow

We compared the annual oil flow through the four MOCs estimated from the AIS data with the annual oil flow released by the EIA (see Figure 6 and Table 3). The comparison shows that the two sets of data are highly consistent. This indicates that the framework proposed in this study is reliable.

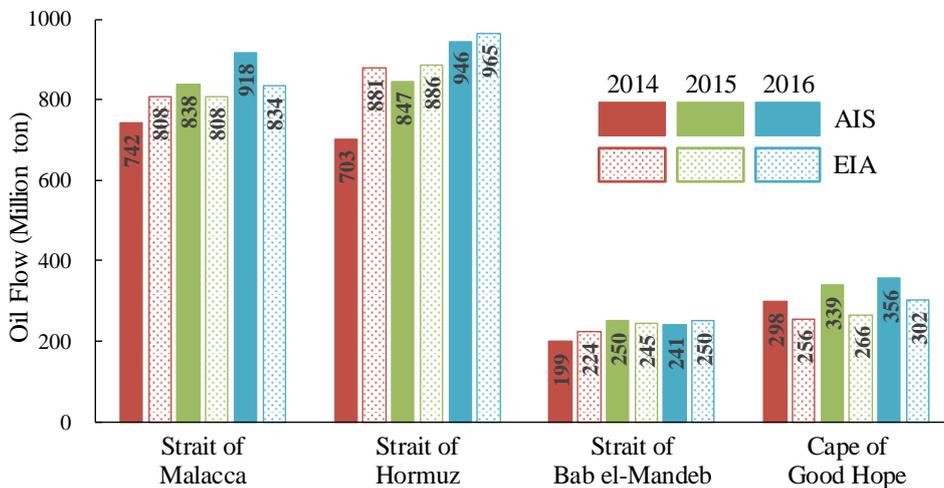


Figure 6. Annual oil flow through maritime oil chokepoints (MOCs) compared with the Energy Information Administration (EIA) data.

Table 3. Ratio (%) of the annual oil flow estimated in this study compared to the EIA data of the MOCs.

Year	Strait of Malacca	Strait of Hormuz	Strait of Bab el-Mandeb	Cape of Good Hope
2014	91.80	79.82	88.82	116.66
2015	103.68	95.53	102.09	127.29
2016	109.98	98.02	96.17	117.67

The estimated annual oil flow values through the straits of Malacca, Hormuz, and Bab el-Mandeb do not differ much from the EIA data. Except for the value in the Strait of Hormuz in 2014, which is approximately 80% of the EIA data, the rest are approximately between 90% and 110% of the EIA data (See Table 3). In the Cape of Good Hope, because it is a different statistical region, the estimated annual oil flow value is at least 116% of the EIA data, which is significantly higher than the EIA data. We statistically analyzed all of the oil tankers passing between the Cape of Good Hope and Antarctica in this study, while the EIA data only considered the oil tankers within a certain range around the Cape of Good Hope. Therefore, the annual oil flow in the Cape of Good Hope estimated in this study has a higher value.

The annual oil flow data estimated in this study and the EIA data show an upward trend from 2014 to 2016. In the Strait of Malacca, Strait of Hormuz, and the Cape of Good Hope, the annual oil flow data estimated and the EIA data both show an upward trend year by year. In the Strait of Bab el-Mandeb, the annual oil flow data estimated shows an overall increase from 2014 to 2016 with a similar growth trend to the EIA data. They both increased significantly from 2014 to 2015, and they stabilized from 2015 to 2016, with only a slight increase or decrease.

The scatter plot is shown in Figure 7 by taking the EIA data as the abscissa and the oil flow estimated from the AIS data as the ordinate. As shown in Figure 7, the slope of the fit line is close to 1. The mean of the estimated oil flow is 560 million tons, while that of the EIA data is 536 million tons, showing a small difference of 4%. R^2 has a large value of 0.9517, and the root mean square error (RMSE) is 68.6889 million tons, which is very small relative to the annual oil flow. Furthermore, we calculated the correlation coefficient between the estimated annual oil flow and the EIA data, which is 0.9756. All of the statistical results show that the oil flow estimated from the AIS data and EIA data are similar and that they have a strong correlation.

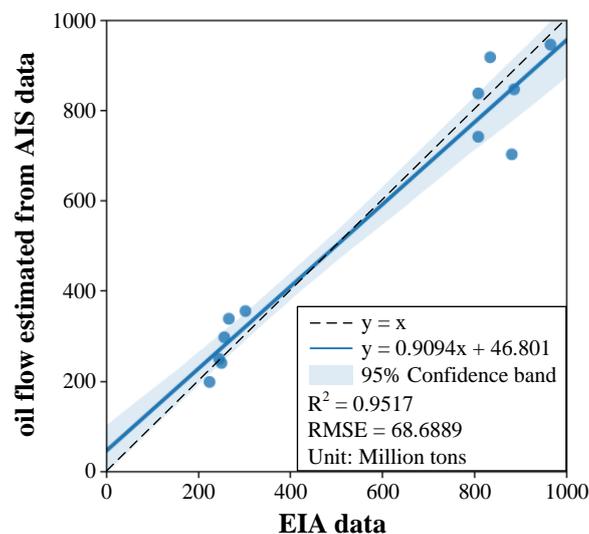


Figure 7. A scatter plot and linear fitting of the estimated oil flow and the EIA data.

4.2. Two-Way Annual Average Oil Flow

With the aim of checking the direction of crossing through the MOC and the correlation with the load, the two-way annual average oil flow of the four MOCs from 2014 to 2016 is calculated (see Figure 8). This study revealed that the oil flow is different for different directions and the general direction of the oil flow calculated in this study is consistent with the actual direction of oil flow. This shows that the framework proposed in this study is reliable.

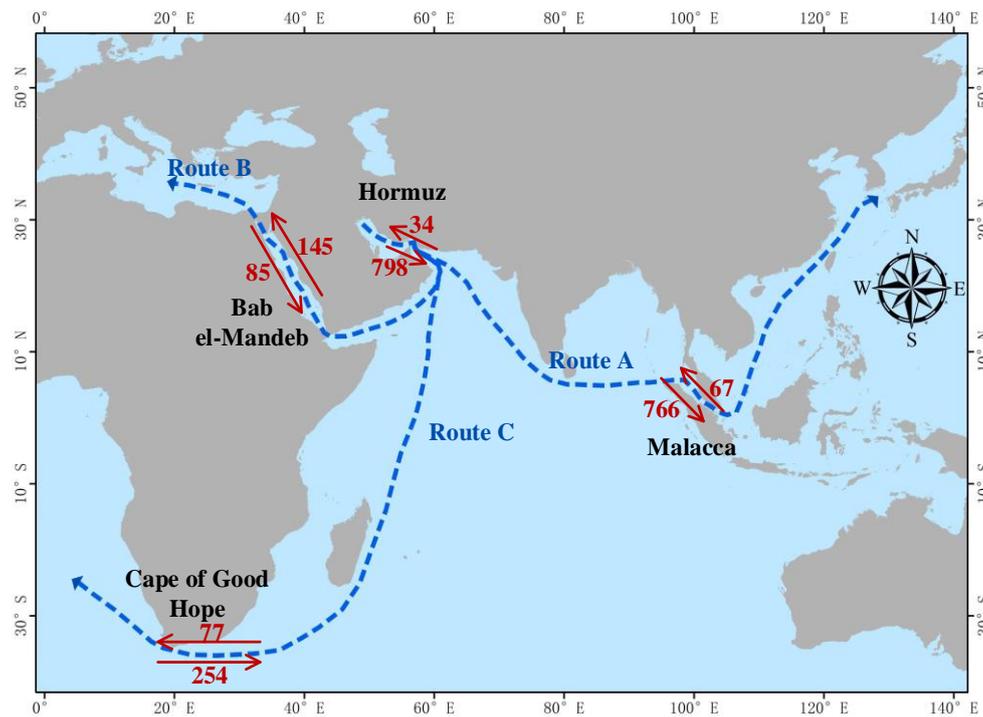


Figure 8. Two-way annual average oil flows of the four MOCs (units: 1 million tons).

As shown in Figure 8, in the Strait of Hormuz, the average annual amount of oil exported from the Persian Gulf is 798 million tons, while the average annual amount of oil transported into the Persian Gulf is only 34 million tons. This is reasonable, because the strait of Hormuz controls oil export from the Persian Gulf, which is the world's largest oil exporter. There are three main routes for transporting oil from the Persian Gulf (see Route A, Route B, and Route C in Figure 8). In the Strait of Bab el-Mandeb and Malacca, the main direction of the oil flow calculated in this study is consistent with the routes. Meanwhile, in the Cape of Good Hope, this is not consistent with the routes.

The reasons for inconsistency in the Cape of Good Hope are as follows: ① Route C is the least busy of the three main routes; hence, only 77 million tons of oil are transported from the Indian Ocean to the Pacific Ocean in the Cape of Good Hope. ② There are many big oil exporters in western Africa, such as Angola and Nigeria, which rank ninth and tenth in the world. These countries will export oil to other countries such as China through the Cape of Good Hope. Therefore, 254 million tons of oil are transported from the Pacific Ocean to the Indian Ocean in the Cape of good Hope. As a result, the general direction of oil flow from the Pacific Ocean to the Indian Ocean in the Cape of Good Hope is reasonable.

4.3. Daily, Monthly, and Seasonal Variation in Oil Flow

The daily oil flows through the four MOCs estimated from the AIS data are decomposed into the cycle, trend, and residual flows based on the monthly and seasonal cycles by using STL (Seasonal and Trend decomposition using Loess) (see Figure 9). STL is a decomposition method based on locally weighted regression. By setting different parameters, STL can be used to decompose the data according to the different cycles [49]. As illustrated in Figure 9, the red time series diagram represents the daily oil flow, the green time series decomposition diagram represents the monthly oil flow, and the blue time series decomposition diagram represents the seasonal oil flow. They have a smaller timescale than the EIA data; therefore, they can provide more information for analysis.

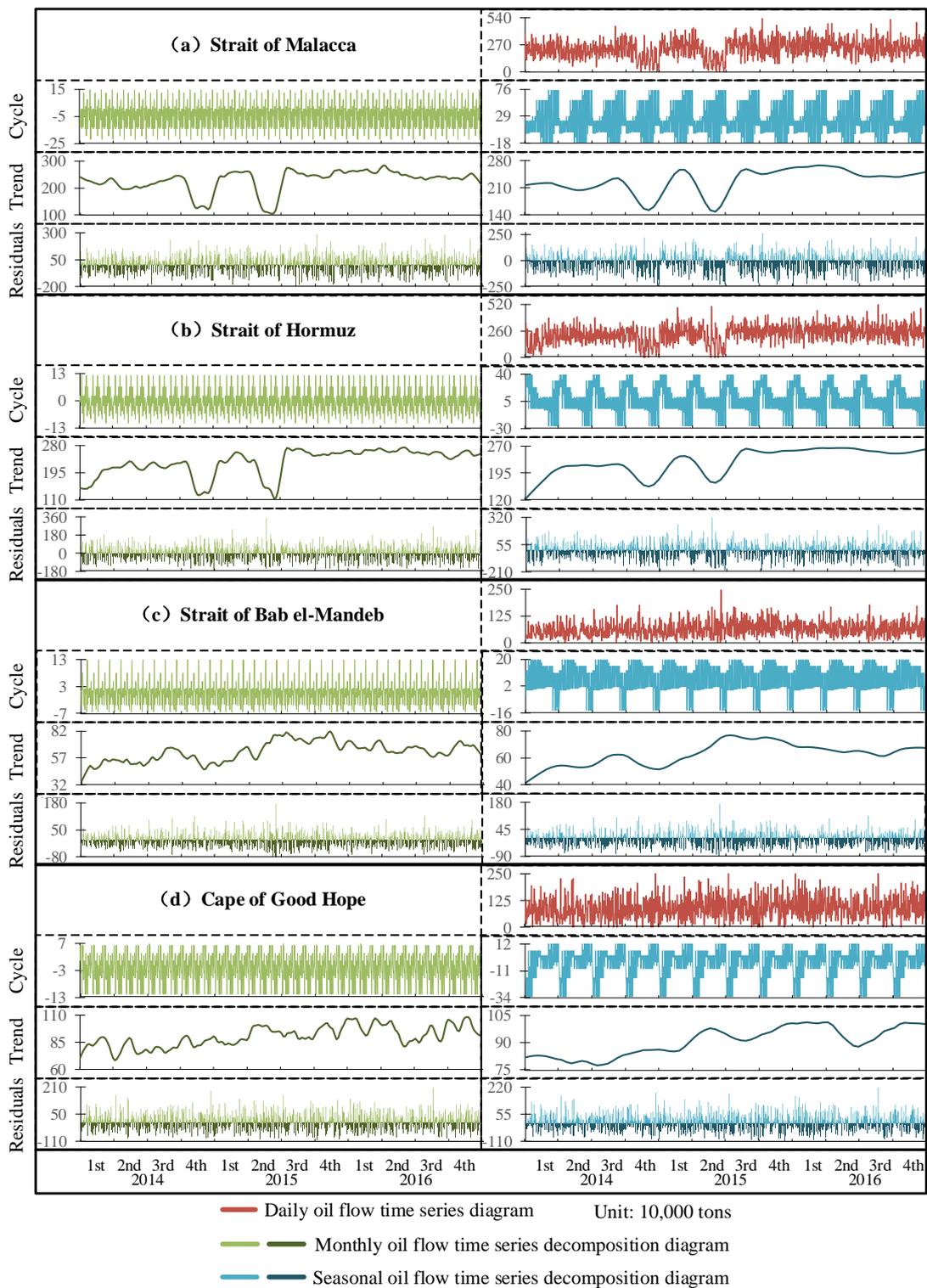


Figure 9. Daily oil flow time series diagram and the monthly and seasonal oil flow time series decomposition diagram in the four MOCs.

The cycles of these four chokepoints are significantly different in terms of the value and variation trend (see the cycle in Figure 9). The maximum values for the cycles of the four chokepoints in the seasonal cycle range from 12 to 76, the minimum values range from -34 to -18 , and the variation range is from 36 to 94. All these statistical values show the large numerical differences between the cycles of the four chokepoints. The seasonal cycle of the oil flow through the Strait of Malacca shows an overall

upward trend, with violent fluctuations in one cycle. The Strait of Hormuz decreases at first and then stabilizes and continues to decrease later in one cycle. The Strait of Bab el-Mandeb fluctuates at first and then decreases suddenly in one cycle. Finally, the Cape of Good Hope increases at first and then stabilizes and continues to increase later in one cycle. The variation trends of the seasonal cycles of the four chokepoints have different forms. The monthly cycles show a pattern similar to that of the seasonal cycles; that is, significant differences exist between the cycles of the four chokepoints.

Although the cycles of the four MOCs are different, their trends are consistent with each other's. Two obvious troughs exist in the Strait of Malacca and Strait of Hormuz, which will be explained in Section 4.3. Ignoring the troughs, the oil flows through the four chokepoints and there is an increase in the fluctuations at first. Then, they reach their peaks in mid-late 2015 without a further increase or with a slight decrease. In 2014, the oil flow through the Strait of Malacca decreases at first and then it increases. In 2016, the flow through the straits of Malacca and Bab el-Mandeb shows a slight downward trend. In 2016, the oil flows through the Strait of Hormuz and the Cape of Good Hope, but these flows do not show an upward or downward trend. In contrast, the flow through the Strait of Hormuz is more stable, while the flow through the Cape of Good Hope fluctuates more widely.

4.4. Events Corresponding to the Troughs in the Oil Flow

As demonstrated in Figure 9, two obvious troughs exist in the Strait of Malacca and the Strait of Hormuz, and we were able to find some events that correspond to these troughs. We believe that the first trough in the Strait of Hormuz in the fourth quarter of 2014 has a connection with the military activities of the U.S. During this period, the U.S. successively sent two aircraft carriers on 18 October 2014, and laser artillery warships on 10 December 2014, to the Persian Gulf. They were sent to combat the extremist armed forces during the civil wars in Iraq and Syria (see Figure 10a,b). These series of military activities are highly related to the first trough in the Strait of Hormuz in space and time. The second trough in the Strait of Hormuz in the second quarter of 2015 is associated with the Middle East Respiratory Syndrome (MERS) outbreak in 2015 [50] (see Figure 10c). The MERS originated in the Middle East, broke out on 18 May 2015, and ended approximately on 14 July 2015. Since it was the origin of the outbreak, the Middle East was naturally affected by it; thus, this caused the formation of the trough during the outbreak. The troughs in the Strait of Malacca relate to the troughs in the Strait of Hormuz, which can be explained as follows.

The oil flows through the straits of Malacca and Hormuz have two obvious troughs, while those in the Bab el-Mandeb Strait and the Cape of Good Hope have none. Furthermore, the troughs in the straits of Malacca and Hormuz are highly consistent in terms of the duration and the extent. This indicates that the events corresponding to the troughs had a considerable impact on Route A but that they had little impact on Routes B and C (see Figure 10c). We have developed the following hypotheses in terms of why the two troughs mainly affected Route A. (1) The oil passing through the Strait of Hormuz is transported more to Asian markets via Route A than to Britain and the U.S. via Routes B and C. As estimated by the EIA, 76% of the crude oil and condensate that moved through the Strait of Hormuz went to the Asian markets in 2018 [7]. (2) The aircraft carriers and warships sent to the Persian Gulf by the U.S. influenced the oil tankers of the other countries to a certain extent. Therefore, Route A was affected greatly, while the remaining two routes, i.e., Routes B and C, to Britain and the U.S., were less affected. (3) During the MERS epidemic in 2015, apart from the Middle East, only countries in Southeast Asia—including the Republic of Korea and China—found MERS cases. More information is listed in Table 4. All of these countries would more or less reduce their oil import. Therefore, the volume of oil transportation to Southeast Asian countries via Route A would be reduced, while Routes B and C would not be affected by the MERS epidemic. Thus, Route A is highly sensitive to the changes in the oil flow through the Strait of Hormuz. As a result, the two troughs of oil flow in the Strait of Malacca are highly consistent with those in the Strait of Hormuz in terms of time and duration.

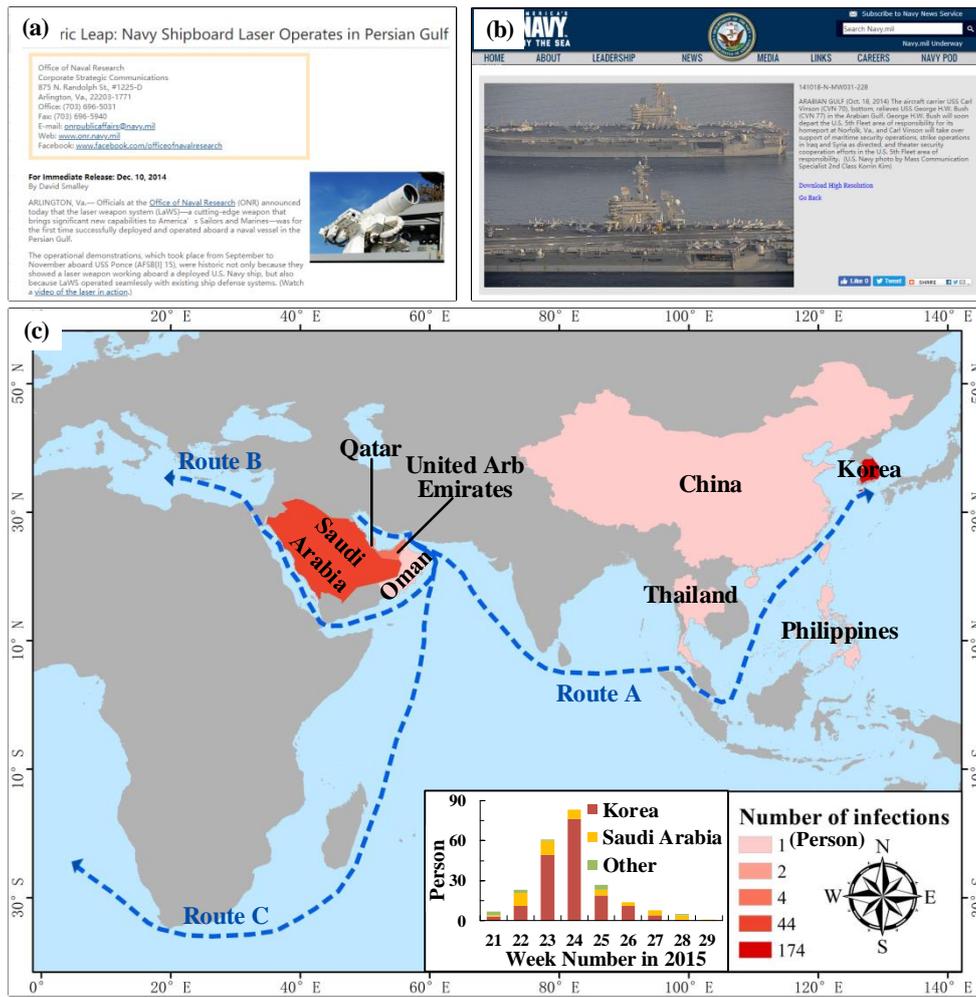


Figure 10. Events corresponding to the abnormal troughs. (a) Military activities in the Persian Gulf on 18 October 2014 (<https://www.onr.navy.mil/en/Media-Center/Press-Releases/2014/LaWS-shipboard-laser-uss-ponce>); (b) Military activities in the Persian Gulf on 10 December 2014 (https://www.navy.mil/view_image.asp?id=186243); (c) The situation of countries involved in the MERS outbreak in 2015 (https://www.who.int/csr/don/archive/disease/coronavirus_infections/en/).

Table 4. Number of cases in various countries during the MERS epidemic in 2015 (https://www.who.int/csr/don/archive/disease/coronavirus_infections/en/).

Date	Country	Number	Date	Country	Number
18 May 2015	United Arb Emirates	1	20 May 2015	Korea	2
21 May 2015	Qatar	1	21 May 2015	Korea	1
22 May 2015	Qatar	2	24 May 2015	Saudi Arabia	2
26–29 May 2015	Korea	8	26–30 May 2015	Saudi Arabia	9
29 May 2015	China	1	29 May 2015	Oman	1
30 May 2015	Korea	1	31 May 2015	Korea	2
1–3 June 2015	Korea	15	1–4 June 2015	Saudi Arabia	5
3 June 2015	United Arb Emirates	1	4 June 2015	Korea	6
5 June 2015	Korea	5	5–8 June 2015	Saudi Arabia	8
6 June 2015	Korea	9	7 June 2015	Korea	14
8–12 June 2015	Korea	62	9–12 June 2015	Saudi Arabia	3
13–16 June 2015	Korea	28	13–17 June 2015	Saudi Arabia	5
15 June 2015	United Arb Emirates	1	17–19 June 2015	Korea	1
18 June 2015	Thailand	1	19–30 June 2015	Saudi Arabia	6
20–23 June 2015	Korea	9	21 June 2015	United Arb Emirates	2
24–26 June 2015	Korea	4	27–30 June 2015	Korea	1
1–3 July 2015	Korea	2	1–14 July 2015	Saudi Arabia	6
4–7 June 2015	Korea	2	6 July 2015	Philippines	1

5. Conclusions

Oil flow assessment through the MOCs can help increase the security and stability of energy transportation and also enable the discovery of anomalies. To generate statistics on oil flow through MOCs, a maritime oil flow analysis technical framework is proposed. Using the AIS data from 1 January 2014 to 31 December 2016, the framework was applied to four MOCs within the MSR and its surrounding regions. The oil flows through the four chokepoints across four timescales (daily, monthly, seasonal, and annual) were determined and analyzed. The following conclusions were drawn from this study:

1. The annual oil flow estimated from the AIS data is similar to the EIA data in terms of the value and variation trends, except for the Cape of Good Hope. The statistics, such as R^2 and RMSE, show that the two sets of data are similar and have a strong correlation. The general directions of the oil flow in the four MOCs are consistent with the actual directions. Therefore, the framework proposed in this study is reliable.
2. Compared with the EIA data, the daily, monthly, and seasonal oil flow estimated in this study has smaller timescales, which can provide more information. The cycles of the four MOCs differed greatly, but their trends are consistent with each other's. Furthermore, there were two obvious troughs in the oil flow through the straits of Malacca and Hormuz.
3. The two troughs in the straits of Malacca and Hormuz were consistent in terms of the duration and extent. We believe that the first trough in the Strait of Hormuz is related to the military activities of the U.S. within the Persian Gulf. The second trough is associated with the worldwide MERS outbreak that occurred in 2015. The troughs in the Strait of Malacca are related to the troughs in the Strait of Hormuz.

By analyzing the above data, we discovered certain patterns as well as anomalies. These discoveries can provide support for national energy transportation strategy formulation, and the prevention and control of abnormalities. However, immense potential exists that can still be mined to provide significant value. Furthermore, with the support of real-time data, the real-time oil flow through the MOCs can be determined. This would be valuable regarding the detection of anomalies in time to aid national emergency responses, among other uses. The framework is suitable for ships carrying only one type of cargo in all MOCs in the world. However, it still has shortcomings, such as not using new technology. In our future research, we will introduce a new time series method to mine deeper information. We will also consider additional variables such as the nationality of the ship and the ship size to expand the breadth of the analysis.

Author Contributions: Conceptualization, Yijia Xiao, Yanming Chen, and Manchun Li; Data curation, Yijia Xiao and Zhaojin Yan; Formal analysis, Xiaoqiang Liu, Zhaojin Yan and; Funding acquisition, Yanming Chen, Liang Cheng, and Manchun Li; Investigation, Yanming Chen, Liang Cheng, and Manchun Li; Methodology, Yijia Xiao, Xiaoqiang Liu, and Zhaojin Yan; Project administration, Yanming Chen and Manchun Li; Resources, Liang Cheng; Supervision, Yanming Chen and Manchun Li; Visualization, Yijia Xiao and Zhaojin Yan; Writing—original draft, Yijia Xiao, Xiaoqiang Liu, and Zhaojin Yan; Writing—review & editing, Yanming Chen, Liang Cheng and Manchun Li. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key R&D Plan of China (2017YFB0504205), and the Science and Technology Innovation Project of Nanjing for Overseas Scholars.

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

References

1. Jia, H. Scientific collaborations shine on Belt and Road. *Natl. Sci. Rev.* **2017**, *4*, 652–657. [[CrossRef](#)]
2. Huang, Y. Understanding China's Belt & Road initiative: Motivation, framework and assessment. *China Econ. Rev.* **2016**, *40*, 314–321.
3. Lai, L.; Guo, K. The performance of one belt and one road exchange rate: Based on improved singular spectrum analysis. *Phys. A Stat. Mech. Appl.* **2017**, *483*, 299–308. [[CrossRef](#)]

4. Wan, C.; Yang, Z.; Zhang, D.; Yan, X.; Fan, S. Resilience in transportation systems: A systematic review and future directions. *Transp. Rev.* **2018**, *38*, 479–498. [[CrossRef](#)]
5. Cheng, L.; Yan, Z.J.; Xiao, Y.J.; Chen, Y.M.; Zhang, F.L.; Li, M.C. Using big data to track marine oil transportation along the 21st-century Maritime Silk Road. *Sci. China Technol. Sci.* **2019**, *62*, 677–686. [[CrossRef](#)]
6. 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](#)]
7. EIA. The Strait of Hormuz Is the World's Most Important Oil Transit Chokepoint. Available online: <https://www.eia.gov/todayinenergy/detail.php?id=39932> (accessed on 24 November 2019).
8. EIA. World Oil Transit Chokepoints. Available online: <https://www.eia.gov/beta/international/regions-topics.php?RegionTopicID=WOTC> (accessed on 24 November 2019).
9. Zhang, Z.X. China's energy security, the Malacca dilemma and responses. *Energy Policy* **2011**, *39*, 7612–7615. [[CrossRef](#)]
10. Faturechi, R.; Miller-Hooks, E. Measuring the performance of transportation infrastructure systems in disasters: A comprehensive review. *J. Infrastruct. Syst.* **2015**, *21*, 04014025. [[CrossRef](#)]
11. Omer, M.; Mostashari, A.; Nilchiani, R.; Mansouri, M. A framework for assessing resiliency of maritime transportation systems. *Marit. Policy Manag.* **2012**, *39*, 685–703. [[CrossRef](#)]
12. Leng, X.; Ji, K.; Yang, K.; Zou, H. A Bilateral CFAR Algorithm for Ship Detection in SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1536–1540. [[CrossRef](#)]
13. Wang, Y.; Liu, H. A hierarchical ship detection scheme for high-resolution SAR images. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 4173–4184. [[CrossRef](#)]
14. Xiao, L.; Xu, M.; Hu, Z. Real-Time Inland CCTV Ship Tracking. *Math. Probl. Eng.* **2018**, *2018*, 1205210. [[CrossRef](#)]
15. Lee, J.M.; Lee, K.H.; Nam, B.; Wu, Y. Study on Image-Based Ship Detection for AR Navigation. In Proceedings of the 2016 6th International Conference on IT Convergence and Security (ICITCS), Prague, Czech Republic, 26 September 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–4.
16. Wang, H.; Zou, Z.; Shi, Z.; Li, B. Detecting ship targets in spaceborne infrared image based on modeling radiation anomalies. *Infrared Phys. Technol.* **2017**, *85*, 141–146. [[CrossRef](#)]
17. Jiang, B.; Ma, X.; Lu, Y.; Li, Y.; Feng, L.; Shi, Z. Ship detection in spaceborne infrared images based on Convolutional Neural Networks and synthetic targets. *Infrared Phys. Technol.* **2019**, *97*, 229–234. [[CrossRef](#)]
18. Misović, D.S.; Milić, S.D.; Durović, Ž.M. Vessel Detection Algorithm Used in a Laser Monitoring System of the Lock Gate Zone. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 430–440. [[CrossRef](#)]
19. Kim, Y.-K.; Kim, Y.; Jung, Y.S.; Jang, I.G.; Kim, K.-S.; Kim, S.; Kwak, B.M. Developing accurate long-distance 6-DOF motion detection with one-dimensional laser sensors: Three-beam detection system. *IEEE Trans. Ind. Electron.* **2012**, *60*, 3386–3395.
20. Kroodsmas, D.A.; Mayorga, J.; Hochberg, T.; Miller, N.A.; Boerder, K.; Ferretti, F.; Wilson, A.; Bergman, B.; White, T.D.; Block, B.A.; et al. Tracking the global footprint of fisheries. *Science* **2018**, *359*, 904–908. [[CrossRef](#)]
21. Goerlandt, F.; Montewka, J.; Zhang, W.; Kujala, P. An analysis of ship escort and convoy operations in ice conditions. *Saf. Sci.* **2017**, *95*, 198–209. [[CrossRef](#)]
22. Boerder, K.; Miller, N.A.; Worm, B. Global hot spots of transshipment of fish catch at sea. *Sci. Adv.* **2018**, *4*, eaat7159. [[CrossRef](#)]
23. Cervera, M.A.; Ginesi, A.; Eckstein, K. Satellite-based vessel Automatic Identification System: A feasibility and performance analysis. *Int. J. Satell. Commun. Netw.* **2011**, *29*, 117–142. [[CrossRef](#)]
24. Etienne, L.; Devogele, T.; Buchin, M.; McArdle, G. Trajectory Box Plot: A new pattern to summarize movements. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 835–853. [[CrossRef](#)]
25. Wen, Y.; Zhang, Y.; Huang, L.; Zhou, C.; Xiao, C.; Zhang, F.; Peng, X.; Zhan, W.; Sui, Z. Semantic modelling of ship behavior in harbor based on ontology and dynamic bayesian network. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 107. [[CrossRef](#)]
26. Pallotta, G.; Vespe, M.; Bryan, K. Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction. *Entropy* **2013**, *15*, 2218–2245. [[CrossRef](#)]
27. Mascaro, S.; Nicholson, A.; Korb, K. Anomaly detection in vessel tracks using Bayesian networks. *Int. J. Approx. Reason.* **2014**, *55*, 84–98. [[CrossRef](#)]

28. Rong, H.; Teixeira, A.P.; Guedes Soares, C. Data mining approach to shipping route characterization and anomaly detection based on AIS data. *Ocean Eng.* **2020**, *198*, 106936. [[CrossRef](#)]
29. Zhang, W.; Goerlandt, F.; Montewka, J.; Kujala, P. A method for detecting possible near miss ship collisions from AIS data. *Ocean Eng.* **2015**, *107*, 60–69. [[CrossRef](#)]
30. Silveira, P.A.M.; Teixeira, A.P.; Soares, C.G. Use of AIS data to characterise marine traffic patterns and ship collision risk off the coast of Portugal. *J. Navig.* **2013**, *66*, 879–898. [[CrossRef](#)]
31. Zhang, W.; Goerlandt, F.; Kujala, P.; Wang, Y. An advanced method for detecting possible near miss ship collisions from AIS data. *Ocean Eng.* **2016**, *124*, 141–156. [[CrossRef](#)]
32. Vettor, R.; Soares, C.G. Detection and analysis of the main routes of voluntary observing ships in the North Atlantic. *J. Navig.* **2015**, *68*, 397–410. [[CrossRef](#)]
33. Zhen, R.; Jin, Y.; Hu, Q.; Shao, Z.; Nikitakos, N. Maritime anomaly detection within coastal waters based on vessel trajectory clustering and Naïve Bayes Classifier. *J. Navig.* **2017**, *70*, 648–670. [[CrossRef](#)]
34. Chen, R.; Chen, M.; Li, W.; Wang, J.; Yao, X. Mobility modes awareness from trajectories based on clustering and a convolutional neural network. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 208. [[CrossRef](#)]
35. Zhang, X.; Chen, Y.; Li, M. Research on geospatial association of the urban agglomeration around the South China Sea based on marine traffic flow. *Sustainability* **2018**, *10*, 3346. [[CrossRef](#)]
36. Gourmelon, F.; Le Guyader, D.; Fontenelle, G. A dynamic GIS as an efficient tool for integrated coastal zone management. *ISPRS Int. J. Geo-Inf.* **2014**, *3*, 391–407. [[CrossRef](#)]
37. Natale, F.; Gibin, M.; Alessandrini, A.; Vespe, M.; Paulrud, A. Mapping fishing effort through AIS data. *PLoS ONE* **2015**, *10*, e0130746. [[CrossRef](#)] [[PubMed](#)]
38. Coello, J.; Williams, I.; Hudson, D.A.; Kemp, S. An AIS-based approach to calculate atmospheric emissions from the UK fishing fleet. *Atmos. Environ.* **2015**, *114*, 1–7. [[CrossRef](#)]
39. Merico, E.; Dinoi, A.; Contini, D. Development of an integrated modelling-measurement system for near-real-time estimates of harbour activity impact to atmospheric pollution in coastal cities. *Transp. Res. Part D Transp. Environ.* **2019**, *73*, 108–119. [[CrossRef](#)]
40. Jia, H.; Adland, R.; Prakash, V.; Smith, T. Energy efficiency with the application of Virtual Arrival policy. *Transp. Res. Part D Transp. Environ.* **2017**, *54*, 50–60. [[CrossRef](#)]
41. Wang, C.; Hao, L.; Ma, D.; Ding, Y.; Lv, L.; Zhang, M.; Wang, H.; Tan, J.; Wang, X.; Ge, Y. Analysis of ship emission characteristics under real-world conditions in China. *Ocean Eng.* **2019**, *194*, 106615. [[CrossRef](#)]
42. Tu, E.; Zhang, G.; Rachmawati, L.; Rajabally, E.; Huang, G. Bin Exploiting AIS Data for Intelligent Maritime Navigation: A Comprehensive Survey from Data to Methodology. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 1559–1582. [[CrossRef](#)]
43. Wu, X.; Zhu, X.; Wu, G.-Q.; Ding, W. Data mining with big data. *IEEE Trans. Knowl. Data Eng.* **2013**, *26*, 97–107.
44. Yin, Y.; Gong, G.; Han, L. Theory and techniques of data mining in CGF behavior modeling. *Sci. China Inf. Sci.* **2011**, *54*, 717–731. [[CrossRef](#)]
45. Zhu, J.; Chen, J.; Hu, W.; Zhang, B. Big learning with Bayesian methods. *Natl. Sci. Rev.* **2017**, *4*, 627–651. [[CrossRef](#)]
46. Park, D.M.; Kim, Y.; Seo, M.G.; Lee, J. Study on added resistance of a tanker in head waves at different drafts. *Ocean Eng.* **2016**, *111*, 569–581. [[CrossRef](#)]
47. MacQueen, J. Some Methods for Classification and Analysis of Multivariate Observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 21 June–18 July 1965 and 27 December 1965–7 January 1966; University of California Press: Berkeley, CA, USA, 1967; Volume 1, pp. 281–297.
48. Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A k-means clustering algorithm. *J. R. Stat. Soc. Ser. C Appl. Stat.* **1979**, *28*, 100–108. [[CrossRef](#)]
49. Cleveland, R.B.; Cleveland, W.S.; McRae, J.E.; Terpenning, I. STL: A seasonal-trend decomposition. *J. Off. Stat.* **1990**, *6*, 3–73.
50. Donnelly, C.A.; Malik, M.R.; Elkholy, A.; Cauchemez, S.; Van Kerkhove, M.D. Worldwide Reduction in MERS Cases and Deaths since 2016. *Emerg. Infect. Dis.* **2019**, *25*, 1758. [[CrossRef](#)]

