Use of AIS data for performance evaluation of ship traffic with speed control

Likun Wang[1], Yang Li[1], Zheng Wan*[1], Tong Wang[2], Keping Guan[3], Zaili Yang[4]

1. *College of Transport and Communication, Shanghai Maritime University, Shanghai, China*
2. *School of Economics and Management, Tsinghua University, Beijing, China*
3. *College of Navigation, Shanghai Maritime University, Shanghai, China*
4. *Liverpool Logistics, Offshore and Marine Research Institute, Liverpool John Moores University, Liverpool, UK*

Highlights

1. Rational speed control to strike an effective balance between navigation efficiency and system safety.
2. AIS data mining and innovative use for evaluating the performance of shipping traffic.
3. A new framework to explain the process of data acquisition, error elimination and combination of AIS and geocoded data.
4. Analysis of shipping traffic performance from a new perspective with respect to the characteristics of different navigational segments.
5. Real case study to justify the rational ship speed limit in the Shanghai section of the Yangtze River.

Abstract

Speed control in inland water systems needs to strike an effective balance between ship operational efficiency and transport safety. However, speed limit regulations are largely formulated by expert judgment rather than objective evidence-based evaluation, which sometimes leads to their inefficiency

* Corresponding author

due to subjective bias. In this study, a new method is proposed to evaluate the performance of shipping traffic under the current speed limits using automatic identification system (AIS) big data of 4923 ships in the Shanghai section of the Yangtze River, China. Key elements of this method include data acquisition, error elimination, combination of ship AIS and waterway geocoded data to model traffic flow characteristics, and estimation of the correlation between ship speed and congestion level. This study separates the peak and off-peak periods of the investigated waterways, which provides a baseline for future speed control schemes with respect to different timeframes. Shipping traffic performance in different segments is analyzed, and the results reveal that the overall compliance rate of the speed limit is high and only a few over-speed cases are noted in certain segments. Furthermore, we use a normal distribution to model the correlation between ship spot speeds and maximum traffic volumes and between mean speed and aggregated traffic volume of each 0.2-knot bin. The findings support via experimental evidence that the current speed limit in the Shanghai section of the Yangtze River is rational, while providing useful insights for testing the rationality of speed limits in other waterways or shipping channels.

Key words: Yangtze River (Shanghai section) Strait; Speed limit Regulation; Normal distribution fitting; AIS data, Maritime safety

1. Introduction

The Yangtze River in China is the longest and busiest inland waterway in the world. The Shanghai section of the River is 65 nautical miles (nm) long and is located at the downstream of the River towards the East China Sea. Its average annual daily vessel traffic volume was 942.3 ships in 2017, which increased by 13% in comparison with 833.9 ships in 2013 according to Statistical Bulletin of Transportation Industry Development[1]. Given such a ship volume increase, the associated traffic safety attracts growing attention, particularly in the busy segments of the Shanghai section such as Baoshan North channel, Baoshan channel, and Waigaoqiao channel. The Maritime Safety Administration has regulated the speed limit of 12 knots for all the vessels passing through deep-water channels greater than 12.5 m of the Shanghai section since 2011. However, the speed limit was determined by experts based on their experience and it has not been scientifically approved given the difficulty of obtaining vessel speed data in old days. In a broad sense, it raises a critical research issue for ship traffic safety by speed limits in narrow waters from a theoretical perspective. This paper aims to fill this gap by introducing a new paradigm shift regime in setting and justifying the ship speed limit using experimental tests based on automatic identification system (AIS) data.

The Safety of Life at Sea convention stipulates that vessels above 300 gross tonnages on international voyages should be equipped with AIS transceivers [2]. Using AIS, significant time-dependent data such as ship dynamic location and ship speed can be obtained and often used to support research in the marine-related field [3]. In this work, the features of real-time ship speed and ship traffic volumes are analyzed to test the rationality of the speed limit in shipping traffic intensive waters in general and 12-knot limit in the Shanghai section of the Yangtze River in specific.

The remainder of the paper is organized as follows. Section 2 reviews the relevant literature. Section

3 discusses the research methodology, including the selection of key channels, error elimination method in the use of AIS data, and calculation of traffic flow parameters. Section 4 discusses the results and analyzes the correlation between speed limit and ship traffic flow. Finally, conclusions are presented in Section 5.

2. Literature Review

It is crucial to investigate ship speed characteristics to ensure maritime safety. There is no shortage on ship speed studies in the literature. Based on the 135 investigations conducted by the Marine Accident Investigation Branch in the United Kingdom, Tirunagari et al. identified that the traffic density, ship speed, confusion, equipment, bad weather, fatigue, and health were the most frequent causes of maritime accidents [4]. Mazaheri et al. reported that a minor relation existed between traffic distribution and groundings, whereas no relation appeared between groundings and traffic density [5].

Scientific research driven speed limit policies and regulations are more frequently seen from other transport modes (e.g. road) than shipping. Speed limits are successfully imposed on road traffic to enhance safety and reduce fuel consumption [6]. Hosseinlou et al. modeled the relation between the emission of pollutants, fuel consumption, travel time, number of accidents, and vehicle speed [7]. Results showed that the optimal speeds were 73 and 82 km/h from societal and road user perspectives, respectively. Using large-scale data relating to the speed limit increase in western United States, Van Benthem found that a 10 mph speed limit increase on highways led to an increase of 3–4 mph in travel speed, 9%–15% in accidents, 34%–60% in fatal accidents, and 14%–24% in elevated pollutant concentrations (carbon monoxide) around the affected freeways [8]. However, the actual mean speed could be lower or higher than the speed limit. Vadeby and Forsman discovered that the mean speeds on

87 motorways were 119.9 and 115.3 km/h with speed limits of 110 and 120 km/h, respectively, whereas on
88 rural roads they were 84.6 and 84.8 km/h with speed limits of 70 and 80 km/h, respectively [9].

89 Compared to the road transport speed limit research, studies on shipping speed control and its guide
90 to policy making are scanty in the current literature. This study aims to propose an approach for studying
91 the relation between actual mean speed and speed limit in the maritime field. Recently, the use of AIS
92 data has gained considerable attention in the maritime research area. The used approaches are at large
93 classified into two categories in terms of their applications, namely traffic characteristics of ships and
94 navigation safety. In the case of ship traffic, considerable research has been conducted based on AIS
95 data. Wen et al. established a marine traffic complexity model to evaluate traffic situation [10]. Zhang et
96 al. presented an effective analysis method to analyze ship traffic demand and ship traffic spatiotemporal
97 dynamics in port waters [11]. Christian and Kang processed AIS using a Hough Transform algorithm to
98 estimate possible intersections between the shipment route and marine traffic [12]. Furthermore, Kang
99 et al. estimated the speed-density relation based on the Greenshields model, Greenberg model,
100 Underwood model, and Pipes generalized model for analyzing the vessel traffic flow of Singapore Strait
101 [13]. In the case of ship collisions, Kujala et al. used AIS data to calculate the geometric probability of
102 ship collision in the Gulf of Finland [14]. Results revealed that the highest risks were caused by high
103 traffic intensity. Moreover, Mou et al. developed a linear regression model to determine the correlation
104 between the closest point of approach and size, speed, and ship headings; then, they analyzed ship
105 collision using the AIS data of Rotterdam Port [15]. Zhang et al. proposed a method that detected possible
106 near miss ship–ship collisions using AIS data [16]. Furthermore, a good number of studies had
107 undertaken a correlation analysis between ship accidents and traffic data obtained from AIS. Mazaheri
108 et al. conducted a correlation analysis regarding the relation between traffic conditions, complexity of

109 the waterway, and grounding accidents [17]. Goerlandt et al. found that ship collisions occurred when
110 the ship's speed was lower than the average speed of the vessel fleet [3]. Bye and Aalberg used a
111 multivariate logistic regression model to identify the conditions associated with navigation accidents and
112 regarded these conditions as risk indicators based on the Norwegian Maritime Directorate (NMA) data
113 and AIS data [18]. In light of the above, to our best knowledge, no previous studies have investigated
114 the relation between ship actual mean speed and speed limit using AIS big data, despite its important
115 influence on transport safety as indicated in road transport.

116 The above literature review reveals that 1) ship speed is among the most important attributes
117 influencing marine navigational safety; 2) actual vehicle speeds are not always perfectly in line with the
118 speed limit accordingly to the lesson learnt from other transport modes (e.g. road); 3) research on ship
119 speed limits is scarce in the current literature; 4) AIS data are often used to obtain the traffic
120 characteristics of ships and analyze their navigation safety, however, rarely on the solution to ship speed
121 limit. This study therefore proposes a methodology to quantify the characteristics of actual ship speed
122 and traffic volume based on AIS data and demonstrate it using the speed limit of the Shanghai section of
123 the Yangtze River. The results will aid to develop rational speed limit regulations to optimize ship traffic
124 flow while minimizing the associated maritime risks. .

125

126 **3. Observations from AIS data**

127 **3.1 AIS data acquisition**

128 In total, 83140629 AIS records (from June 1, 2014, 8:00:00 to June 30, 2014, 8:00:00) involving
129 4923 ships are obtained from the Shanghai section of the Yangtze River. The geographic scope used for
130 the AIS data collection is shown by the blue dotted line in Fig. 1. The corresponding longitudes and

latitudes of the four vertices are (121°18′E, 30°N), (121°18′E, 31°N), (122°30′E, 30°N), and (122°30′E, 31°N), respectively. The border of the Shanghai section of the Yangtze River is a line between point (31°30' 52.4″N, 121°18′56.9″E), point (31°37′34.4″N, 121°22′33.0″E), and line of longitude 122°29′38.6″E, as depicted by red dotted line in Fig. 1.

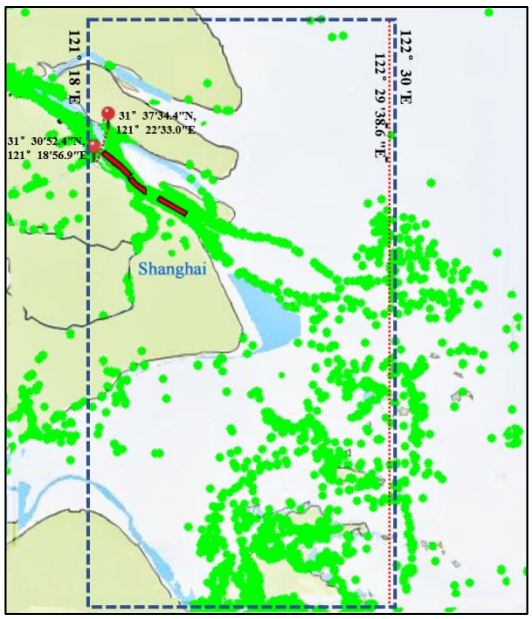


Fig. 1 Geographic scope of the AIS data.
Note: The green dots represent the ships.

3.2 AIS data error elimination

The cleaning of big data is recommended by detecting and removing errors and inconsistencies from the data to improve the data quality [19]. In the case of AIS big data, data cleaning methods include deleting duplicate data, handling abnormal MMSI number, location, speed data, and disposing out-of-range data [13, 20]. For missing data, interpolation processing will be employed [11]. There are four steps to clean the AIS data:

Step 1: Delete duplicate AIS records.

If records are repetitive, only one record is kept. Duplicate data refer to two or more AIS data that

have the same field values, such as MMSI number, receiving time stamp of field data, ship dynamic location, and ship speed.

Step 2: Delete abnormal AIS data records.

Abnormal AIS data refer to the records where the associated MMSI numbers are outside the permissible range [100000000, 999999999] [13], or the speed over ground (SOG) is outside the range of the AIS receiver (e.g. $SOG < 0$ and $SOG > 102.2$), or the record where absolute longitude is $|X| > 180^{o}$, or the absolute latitude is $|Y| > 90^{o}$ [20], or the geographic coordinates exceed the expected AIS receiving range [13, 20].

Herein, all the abnormal AIS records are deleted.

Step 3: Delete unreasonable ship speed and unreasonable ship acceleration/deceleration.

Based on the traffic restrictions in the Singapore straits, Kang et al. set the minimum ship speeds as 4 knots [13]. Meanwhile, Qu et al. found that 25% of the ship speed was greater than the maximum speed limit (15 knots) for traffic restrictions in the Singapore straits [21]. Therefore, they expanded the reasonable maximum speed to 20 knots, and the reasonable ship speeds are defined in the range from 4 to 20 knots. Moreover, Qu et al. and Zhang et al. calculated the average speed by the ratio of distance to the time interval based on Newton's laws of motion to determine whether the ship speed is reasonable with the ship's position and acceleration/deceleration [11, 21].

The ship speed regulation in the Yangtze River sets up the speed limit at 12 knots excluding high speed vessels. From the AIS data records of the Yangtze River, Fig. 2 reveals that only 0.06% of speeds are greater than 30 knots. It is difficult to decide whether the corresponding data is from high speed ships or speed errors. Hence, the reasonable speed range is defined from 0 to 30 knots and the out-of-scope data are deleted.

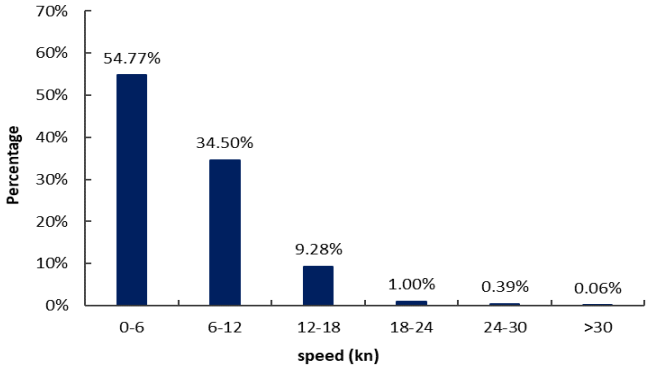


Fig. 2 Ship speed distribution.

Step 4: Correction of missing data.

Salama filled missing data by statistical analysis and dynamic similarity matching [22]. However, Zhang et al. set an interval of 30s for AIS data linear interpolation [11]. Moreover, Shelmerdine interpolated ship speed, length, draft, and tonnage using a natural neighbor function [23].

In this work, the AIS data are aggregated into 20-min bins to obtain the average ship speed and ship traffic volume. The AIS records are then sorted over time from small to large for each vessel, when the time between two AIS records of a specific vessel is larger than 20 min. Meanwhile, the interpolated method is used to cope with the supplementary data. Given that the proportion of time interval greater than 20 min is only 0.021%, the interpolation is not taken into account in the investigated case of the Shanghai section.

Table 1 lists the AIS data cleaning steps of the Yangtze River from June 1, 2014 to June 30, 2014.

Table 1 Summary of data cleaning results.

Steps	Type of data	Records before cleaning	Records of cleaned	% of deleted records
1	Duplicate data	83,140,629	1,622,296	1.95
2	Abnormal MMSI	81,518,333	274,388	0.33
	Abnormal SOG	81,243,945	152,838	0.18
	Abnormal latitude and longitude	81,091,107	0	—
3	Speed out-of-scope data	81,091,107	48,164	0.06
4	Missing data	81,042,943	0	—

3.3 Geocoding AIS data to deep-water channels

(1) Selection of key deep-water channels (lanes)

There are five deep-water channels in the Shanghai section, namely Waigaoqiao channel, Baoshan channel, Baoshan North channel, South channel, and Yangtze estuary deep-water channel. Based on their geographical distribution, this study considers, based on the trajectory intensity, three main channels, namely Waigaoqiao channel, Baoshan channel, and Baoshan North channel. The three key dual-lane channels are numbered as lane 1, lane 2, …, lane 6, and coded as L_1, …, L_6. Herein, L_1, L_3, and L_5 are the lanes of traffic towards the upstream of the River (e.g. river direction), whereas L_2, L_4, and L_6 are the lanes of traffic towards the downstream of the River (i.e. sea direction), as shown in Fig. 3.

(2) Selection of small segments in the middle of waterway lanes

Six small segments were selected from the middle of the six lanes $\{L_1, \ldots, L_6\}$ and recorded as S_k ($k \in \{1,2,...,6\}$) correspondingly (Fig. 3). S_k^+ and S_k^- represent the upstream and downstream areas of S_k, respectively, as shown in Fig. 4. The length of the observed middle area was set as 0.25 nm. Based on the hypothesis that a ship can travel 0.33 nm within 20 min if its minimum speed is 1 knot, this study selects 0.25 nm as the length of observation sections to ensure that the ship can travel through it within

20 min.

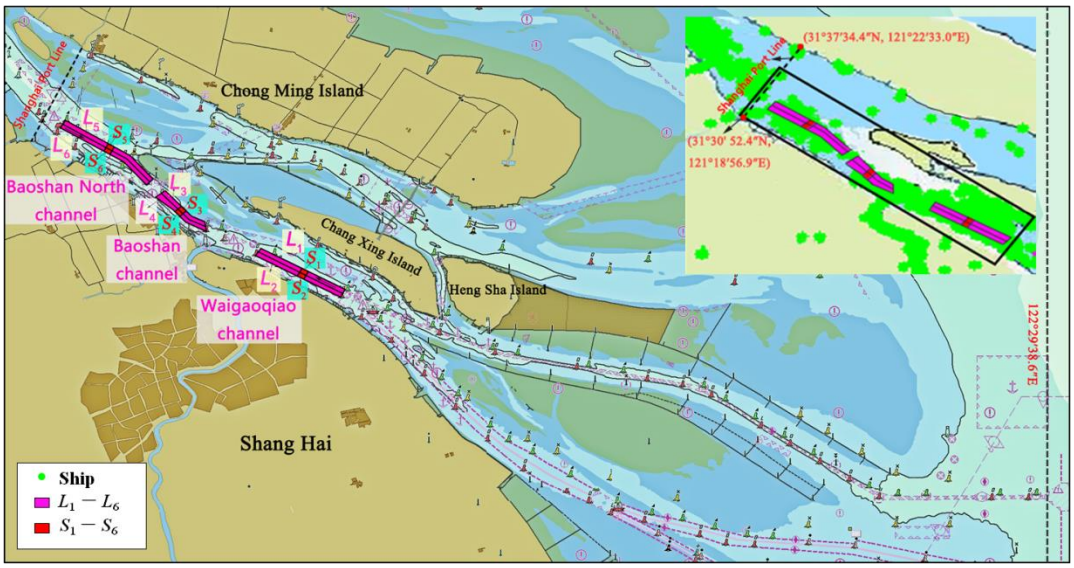


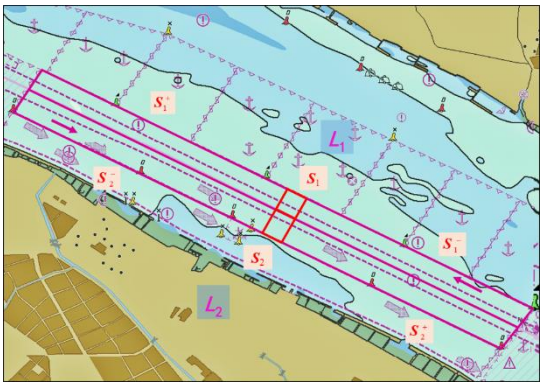
Fig. 3 Key lanes in the Yangtze River (Shanghai section)

Notes: In the rectangular area in the small window, the green point represents the ship, and the higher density of the green points represents a higher density of ship trajectory.

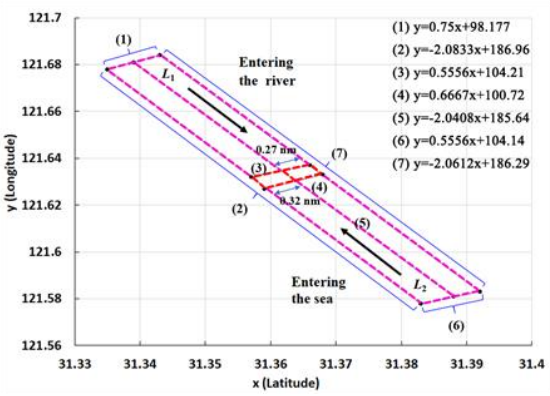
(3) Geocoding of selected lanes and segments

Based on the landmarks along the boundary of each lane, such as geographical marking points, buoys, and lighthouses, the range of the lanes $\{L_1, \ldots, L_6\}$ (Fig. 3) is shown in the nautical map (Figs. 4 a–c). A linear function is then adopted to express each edge of the polygon boundary of the key lanes, as shown in Figs. 4(d–f).

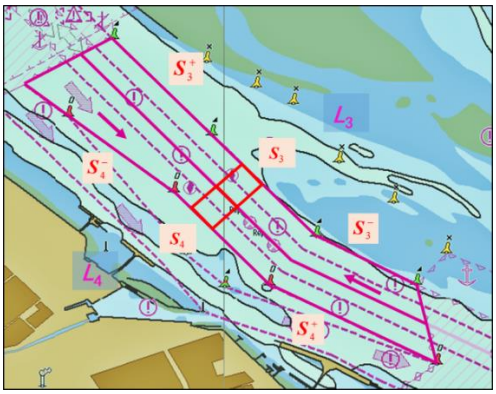
To obtain the red lined segments $\{S_1, \ldots, S_6\}$, we randomly selected some points near the middle part of each lane as the boundary points. The width of each section is defined as 0.25 nm, and length of each segment is shown in Figs. 4(d–f). Each edge of the middle segment is also geocoded using linear functions.



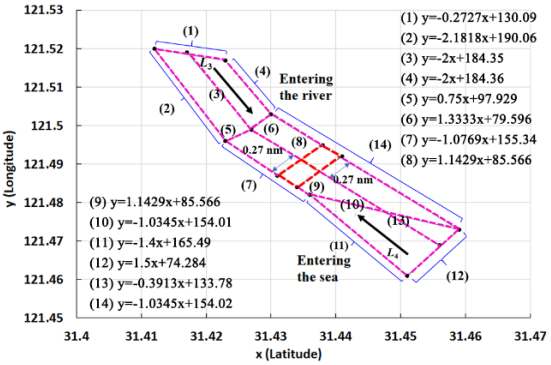
(a) Boundary of Waigaoqiao channel



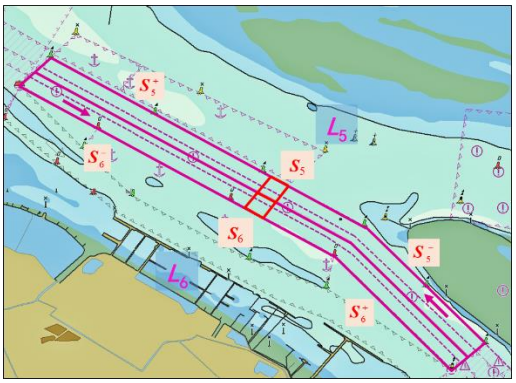
(d) Polygon of Waigaoqiao channel and function expressions



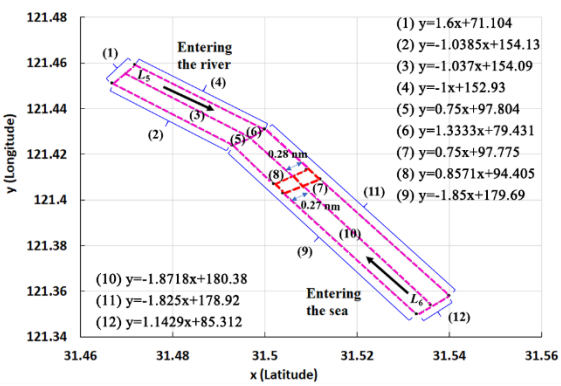
(b) Boundary of Baoshan channel



(e) Polygon of Baoshan channel and function expressions



(c) Boundary of Baoshan North channel



(f) polygon of Baoshan North channel and function expressions

Fig. 4 Location and function expressions of each channel.

(4) Geocoding AIS data of the selected lanes and segments

Using the functions in Fig. 4, we extract the AIS data involving the selected lanes and segments. A

228 total of 4923 ships and 5217849 AIS records are obtained. The detailed data are listed in Table 2.

229 Table 2 AIS data of the key channels.

Channel/Middle section of channel	Symbol	Inland river direction	Symbol	Sea direction
Waigaoqiao channel	L_1	1,200,405	L_2	1,416,215
Baoshan channel	L_3	728,077	L_4	360,956
Baoshan North channel	L_5	1,015,448	L_6	496,748
Middle segment of Waigaoqiao channel	S_1	47438	S_2	55779
Middle segment of Baoshan channel	S_3	51102	S_4	16389
Middle segment of Baoshan North channel	S_5	30387	S_6	9326

230

231 **4. Calculation steps of ship traffic flow parameters**

232 The AIS data are aggregated into 20-min bins for each channel to obtain the traffic volume and ship
233 speed. The total time is divided into intervals, such as $[t_0,t_1)$, …, $[t_i,t_{i+1})$, …, $[t_{n-1},t_n)$, where t_0 =
234 08:00:00, June 1, 2014, t_n = 08:00:00, June 30, 2014, and $t_{i+1} = t_i + 20$ min.

235 The traffic volume is calculated as the number of the ships passing through S_k. The ship speed is a
236 spot speed, calculated as the average value of the real time speeds of all ships passing through S_k. The
237 traffic volume and ship speed of each channel $L_1, L_2, \ldots, L_6$ are calculated using the following three steps:

238 **Step 1**: Calculation of the traffic volume $Q^{S_k,[t_i,t_{i+1})}$.

239 For each lane L_k (k = 1,2,…,6), the ship passes the middle of the lane in the sequence of S_k^-, S_k,
240 and S_k^+. The set $Q^{S_k,[t_i,t_{i+1})}$ (ship/20 min) represents the traffic volume of S_k during $[t_i,t_{i+1})$. Herein,
241 we define a new parameter $Z_{d_j}^{S_k,[t_i,t_{i+1}]}$ as an evaluation indicator to examine whether the ship d_j appears
242 in the segment S_k during time $[t_i,t_{i+1})$.

243 a. If the trajectory of ship d_j during time $[t_i,t_{i+1})$ appear only in segment S_k^- or S_k^+, ship d_j
244 does not pass through segment S_k. Therefore, $Z_{d_j}^{S_k,[t_i,t_{i+1}]} = 0$.

245 b. If the trajectory of ship d_j during time $[t_i,t_{i+1})$ appear in any of the following sets of S_k, (S_k^- &
246 S_k), (S_k^- & S_k & S_k^+), (S_k & S_k^+), or (S_k^- & S_k^+), ship d_j passes through S_k during $[t_i,t_{i+1})$.

Therefore, $Z_{d_j}^{S_k,[t_i,t_{i+1}]}=1$.

c. $Q^{S_k,[t_i,t_{i+1})}$ is accordingly calculated as follows:

$$Q^{S_k,[t_i,t_{i+1})}=\sum_{d_j=1}^{n}(Z_{d_j}^{S_k,[t_i,t_{i+1})}=1), \tag{1}$$

where n is the total number of ships passing through S_k during $[t_i,t_{i+1})$.

Step 2: Calculation of the ship speed $v^{S_k,[t_i,t_{i+1})}$.

Since the change of ship speed is very small, the spot speed is considered as the average speed at every 20-min interval, which is denoted as $v^{S_k,[t_i,t_{i+1})}$ and calculated as the average value of all ship speeds over S_k during a time slot of 20 mins, as follows:

$$v^{S_k,[t_i,t_{i+1})}=\frac{v_1^{S_k,[t_i,t_{i+1})}+\mathrm{L}\ +v_{d_j}^{S_k,[t_i,t_{i+1})}+\mathrm{L}\ v_n^{S_k,[t_i,t_{i+1})}}{n}, \tag{2}$$

where $v_{d_j}^{S_k,[t_i,t_{i+1})}$ refers to the average speed of ship d_j passing through S_k during $[t_i,t_{i+1})$.

4. Observations from AIS data

4.1 Observations from ship numbers in the lanes

Figure 5 shows the numbers of vessels in the six channels per hour based on AIS data. The vessel numbers of each channel (L_1, L_2, L_3, L_4, L_5, and L_6) range in the intervals of [12.28, 24.41], [17.17, 28.59], [9.45, 21.38], [8.07, 16.62], [7.76, 25.69], and [6.72, 17.72], respectively. Moreover, two traffic periods are shown in Fig. 5 with a peak period from 2:00 to 6:00 p.m. and an off-peak period from 3:00 to 7:00 a.m. Besides, the ship number between 8:00 a.m. and 6:00 p.m. is slightly larger than that of the period between 6:00 p.m. and 8:00 a.m., indicating that the waterway traffic is slightly more congested during the morning time. Identifying the peak and low peak periods can provide a baseline for future speed control schemes under different time periods.

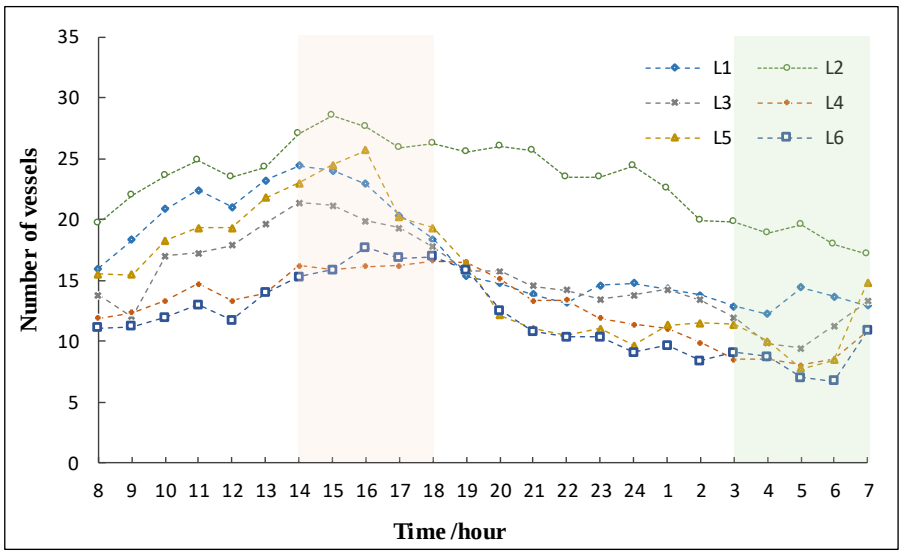
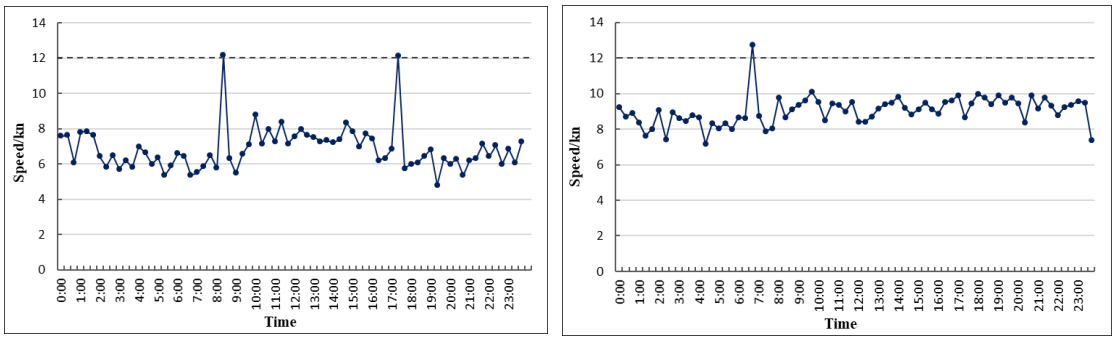


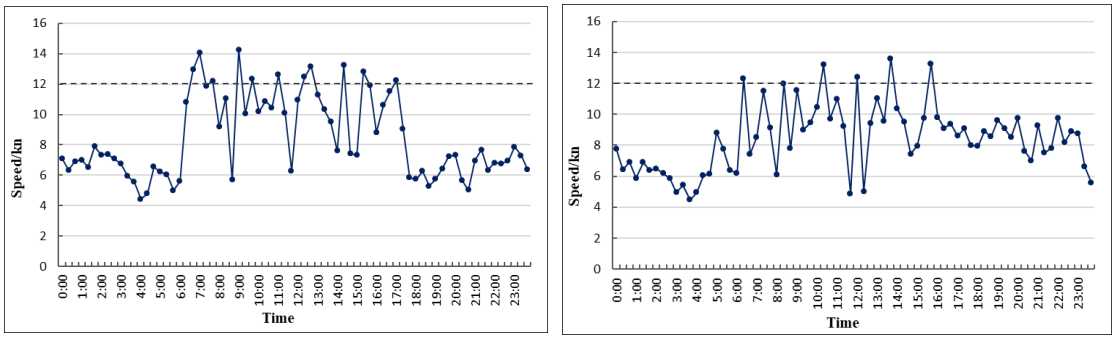
Fig. 5 Average numbers of vessels in the six channels per hour.

4.2 Characteristics of ship traveling speed in the investigated segments

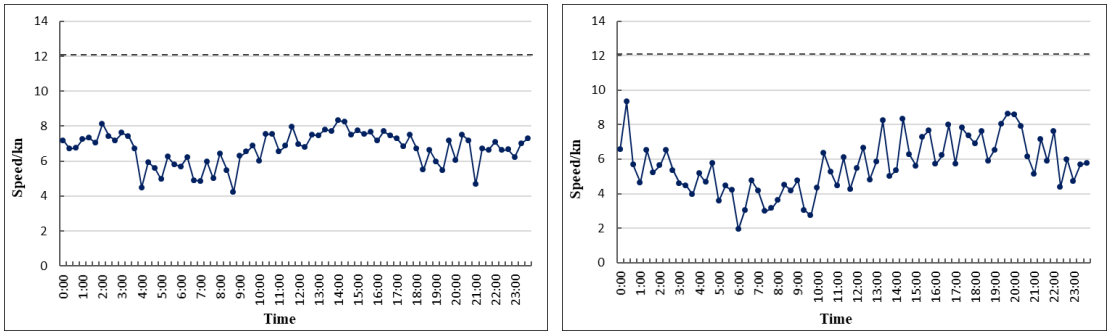
Characteristics of real ship traveling speed in the six chosen segments are shown in Fig. 6. For the Waigaoqiao channel segment (S_1, S_2), the average ship speed passing through the segment is slightly higher than 12 knots in some special periods. However, the majority of all average speeds are lower than the speed limit ranging between 6 and 8 knots. For the Baoshan North channel key segment (S_5, S_6), all the speeds are much lower than 12 knots ranging from 4 to 8 knots, whereas approximate 25% of the speeds in the Baoshan channel key segments (S_3, S_4) are slightly larger than the speed limit. Most of the average speeds range from 8 to 12 knots.



(a) Inland river direction of Waigaoqiao channel (S_1) (b) Sea direction of Waigaoqiao channel (S_2)



(c) Inland river direction of Baoshan channel (S_3) (d) Sea direction of Baoshan channel (S_4)

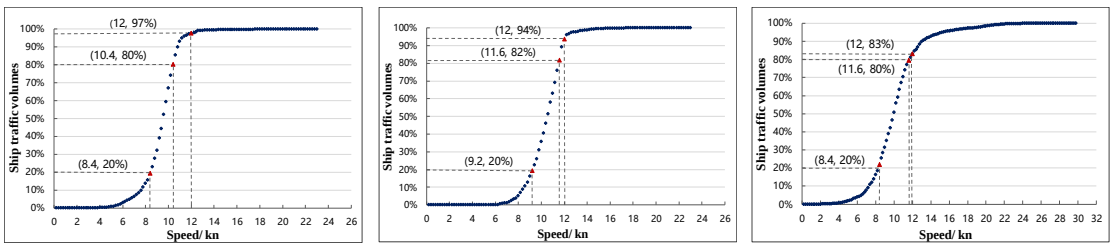


(e) Inland river direction of Baoshan North channel (S_5) (f) Sea direction of Baoshan North channel (S_6)

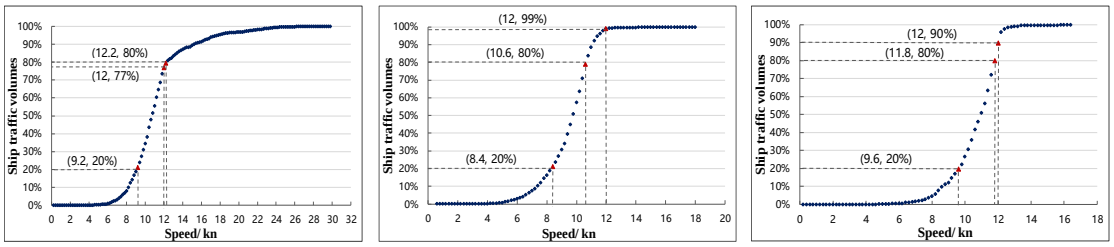
Fig. 6 Average speeds at every 20-min interval.

4.3 Characteristics of ship traffic volumes in the segments

From the aggregated traffic volume at different speeds (Fig. 7), approximately 20% of vessels pass through the chosen six segments at 8.4, 9.2, 8.4, 9.2, 8.4, and 9.6 knots, respectively, whereas 80% of vessels pass through the chosen six segments at 10.4, 11.6, 11.6, 12.2, 10.6, and 11.8 knots, respectively. Furthermore, the results indicate that most vessels travel below the preset speed limit of 12 knots in S_1, S_2, S_5, and S_6, while a few over-speed cases occur in S_3, and S_4.



S_1 S_2 S_3



S_4 S_5 S_6

Fig. 7 Aggregated traffic volume vs speed.

4.4 Ship speed *vs* traffic volume curve fitting

(1) Ship speed *vs* max traffic volume

The kurtosis and skewness of traffic volume distributions for the chosen six segments are similar. Moreover, the distributions of traffic volumes and speeds are well fitted with a normal distribution given as follows:

$$Q = f(v) = a\exp\left(-\frac{(v-b)^2}{c}\right), \quad (3)$$

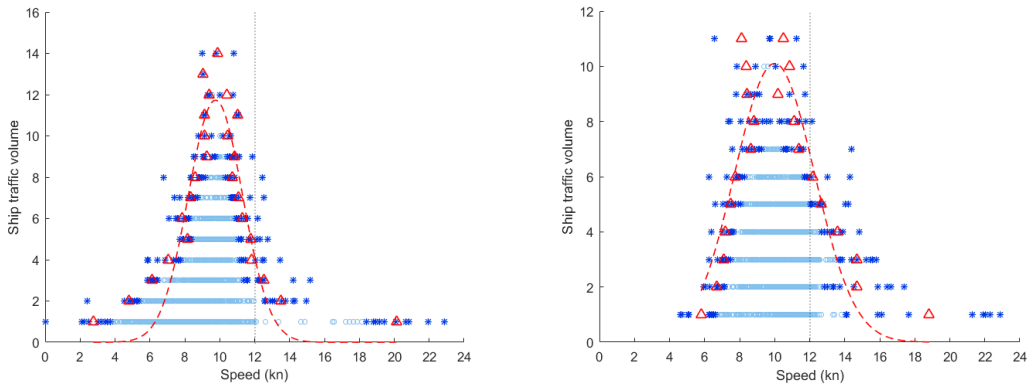
where v is the value of $v^{S_k,[t_i,t_{i+1})}$, each $v^{S_k,[t_i,t_{i+1})}$ corresponds to several $Q^{S_k,[t_i,t_{i+1})}$, Q approximates to the maximum value of $Q^{S_k,[t_i,t_{i+1})}$ (see notes in Fig. 8 for details), and a, b, and c are parameters. Table 3 lists the values of the parameters and goodness of fit (R^2). The result shows that the fitting degree (R^2) of S_k is 0.8624, 0.8145, 0.8672, 0.822, 0.7546, and 0.9138, respectively, indicating that the fitting result is appropriate.

Figure 8 and Table 3 provide some insightful observations. Using the traffic volume of Waigaoqiao channel (for the direction of entering the Yangtze River) as an example, when the speed is below

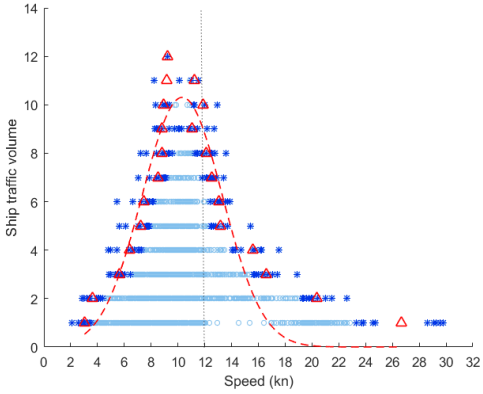
approximately 10 knots, the traffic volume will gradually increase with an increase in speed. When the speed approaches to approximately 10 knots, the traffic volume reaches its maximum value. However, when the speed continues to increase steadily after exceeding 10 knots, the traffic volume decreases. Meanwhile, the other five sections have a similar tendency. Hence, the ship traffic volumes of the chosen six sections increase with an increase in the ship mean speed until they reach a peak stage and then they decrease. Additionally, the traffic volume of S_k (k = 1, 2, …, 6) will reach its peak value when the speed reaches 9.757, 9.995, 10.32, 11.42, 9.832, and 10.14 knots respectively, suggesting that the speed corresponding to the peak value of the traffic volume is lower than the speed limit (12 knots).

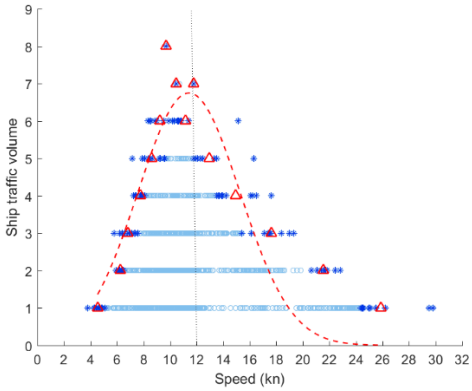
Table 3 Normal distribution fitting results under medium density.

Segment	*a*	*b*	*c*	The optimal speed	R^2
S_1	11.730	9.575	2.199	9.757	0.8624
S_2	10.100	9.995	3.223	9.995	0.8145
S_3	10.310	10.320	4.233	10.320	0.8672
S_4	6.749	11.420	5.441	11.420	0.8220
S_5	9.894	9.832	2.311	9.832	0.7546
S_6	4.935	10.140	2.954	10.14	0.9138

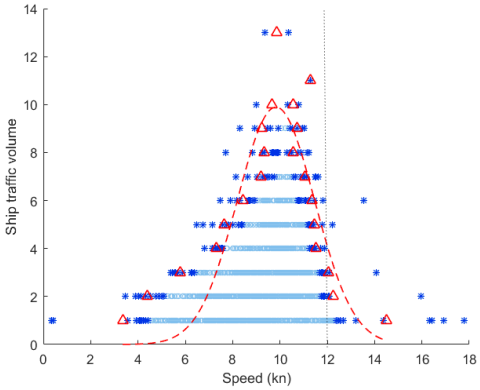


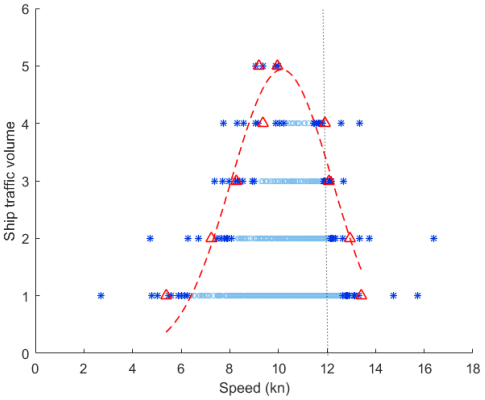
(a) Inland river direction of Waigaoqiao channel (S_1) (b) Sea direction of Waigaoqiao channel (S_2)





(c) Inland river direction of Baoshan channel (S_3) (d) Sea direction of Baoshan channel (S_4)





(e) Inland river direction of Baoshan North channel (S_5) (f) Sea direction of Baoshan North channel (S_6)

Fig. 8 Ship speed vs traffic volume fitting

Note:

The blue point (○) represents baseline data. The dark blue star (*) represents the value of the first 10 minimum (or maximum) values data of a certain ship traffic volume. The red triangle (△) represents a calculated point whose value is equal to the average value of 10 minimum speed value for a certain ship traffic volume, or the average of the corresponding 10 maximum speed value, and the red imaginary line (---) represents the normal fitted curve under medium density.

(2) Ship speed *vs* aggregated traffic volume per 0.2-knot bin

The ship speeds $v^{S_k,[t_i,t_{i+1})}$ are sorted from small to large and aggregated into 0.2-knot bin, as [0,

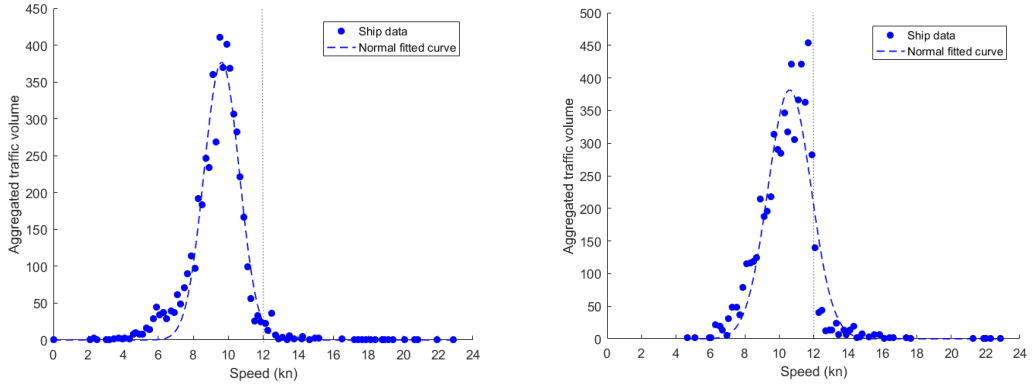
0.2), [0.2, 0.4), ….]. For each bin, we computed the mean value of the corresponding speed $v^{S_k,[t_i,t_{i+1})}$ and sum of the corresponding traffic volume $Q^{S_k,[t_i,t_{i+1})}$. Using Equation (3) to fit the relation between mean speed and aggregated traffic volume of each 0.2-knot bin, the results are indicated by $\hat{v}^{S_k,[t_i,t_{i+1})}$ and $\hat{Q}^{S_k,[t_i,t_{i+1})}$. Herein, v represents $\hat{v}^{S_k,[t_i,t_{i+1})}$ and Q represents $\hat{Q}^{S_k,[t_i,t_{i+1})}$ in function (3).

Moreover, for the aggregated traffic volume at different mean speeds, the relation between ship speed and aggregated traffic volume is well fitted using a normal distribution. The R^2 values are 0.9618, 0.8905, 0.95, 0.9142, 0.9104, and 0.7712 respectively, as listed in Table 4. The fitting result indicates a better performance than the fitting model of total traffic volume with mean speed, and better performance than the model proposed by Kang et al. [13], wherein average R^2 value was 0.691.

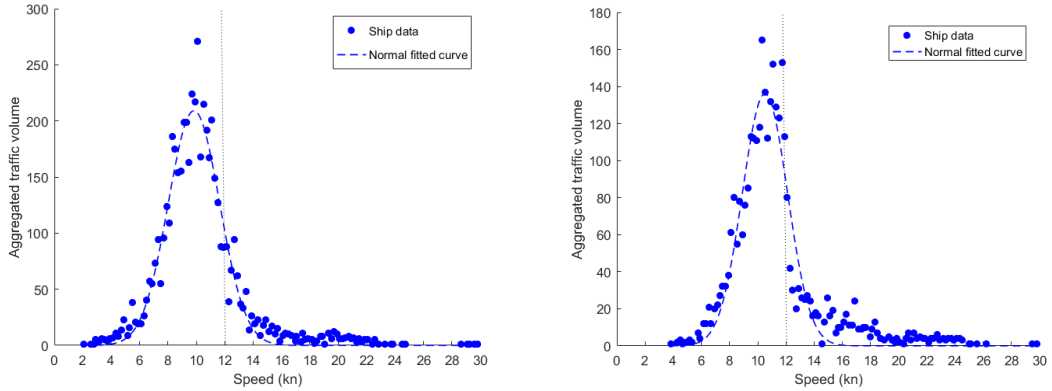
Figure 9 shows the same conclusion as the one derived from Figure 8. The results show the aggregated traffic volume before the speed approaches the limit level (12 knots). The optimal speed corresponding to the maximum flow of ships in different sections is slightly different, but all around 10 knots, slightly below the specified speed limit of 12 knots. Furthermore, the optimal speed in the Yangtz River direction (9.63, 9.83, and 9.23 knots) is slightly lower than the optimal speed in the sea direction of the Yangtz River (e.g. 10.61, 10.51, and 11.14 knots).

Table 4 Normal distribution fitting results under medium density.

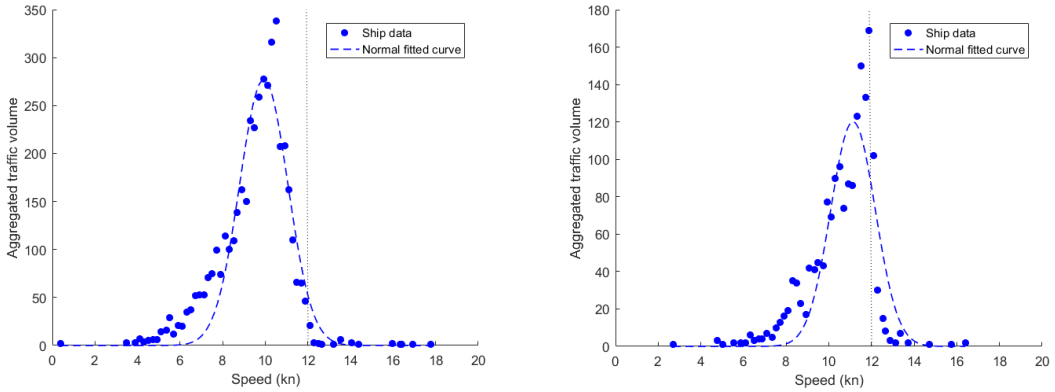
Segment	*a*	*b*	*c*	The optimal speed	R^2
S_1	376.5	9.628	1.427	9.63	0.9618
S_2	381.4	10.610	1.774	10.61	0.8905
S_3	209.0	9.834	2.559	9.83	0.9500
S_4	136.1	10.510	2.264	10.51	0.9126
S_5	276.9	9.929	1.604	9.93	0.9104
S_6	120.1	11.14	1.483	11.14	0.7712



(a) Inland river direction of Waigaoqiao channel (S_1) (b) Sea direction of Waigaoqiao channel (S_2)



(c) Inland river direction of Baoshan channel (S_3) (d) Sea direction of Baoshan channel (S_4)



(e) Inland river direction of Baoshan North channel (S_5) (f) Sea direction of Baoshan North channel (S_6)

Note: The blue line (—) presents the normal fitted curve

Fig. 9 Ship mean speed vs aggregated traffic volume fitting (0.2-knot bin).

5. Conclusion

We develop an AIS data based method for evaluating the performance of shipping traffic under speed limits. The Shanghai section of the Yangtze River is considered as an example. The detailed steps of the proposed method include 1) cleaning the data to filter out problematic and duplicated data, 2) geocoding the AIS data to waterway segments, 3) calculating the ship traffic characteristics, and 4) estimating ship speed and traffic volume to analyze the rationality of the preset speed limit.

Our study seperates the peak and off-peak periods to provide a baseline for future speed control schemes within different time periods. Approximately 80% of vessels pass through the Yangtze River (Shanghai section) at speeds lower than 12 knots; however, a few over-speed cases are noted in Segments 3 and 4. Finally, we use a normal distribution to test the fitness of the relation between spot speeds and corresponding maximum traffic volumes and between and mean speed and aggregated traffic volume of each 0.2-knot bin. The results show that the current speed limit of 12 knots is rational.

This research aid maritime safety policymakers to determine rational ship speed limits based on real performance data. The findings can also help both maritime safety authorities and ship owners to adjust the ship speeds to avoid heavy traffic congestion in narrow waterways. The proposed methods can be tailored to analyze and adjust the speed limits in other waterways in the world.

Acknowledgements: We would like to thank the anonymous reviewers for their insightful comments to improve this article.

Funding: This research is sponsored by the National Science Foundation of China [Grant nos. 71904117, 71704103, 71573172]. Shanghai Pujiang Program [Grant No. 5PJC060]. The authors also acknowledge the funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823904 (ENHANCE).

Declaration of Interest: There are no conflicts of interest to declare.

References

[1] Ministry of Transport of the People's Republic of China, 2017. Statistical Bulletin of Transportation Industry Development. http://xxgk.mot.gov.cn/jigou/zhghs/201806/t20180622_3036269.html

[2] IMO, L.E., 2001. SOLAS. International Convention for the Safety of Life at Sea, 1974, and 1998 Protocol relating thereto.

[3] Goerlandt, F., Goite, H., Valdez Banda, O.A., Höglund, A., Ahonen-Rainio, P., Lensu, M., 2017. An analysis of wintertime navigational accidents in the Northern Baltic Sea. Saf. Sci. 92, 66-84. https://doi.org/10.1016/j.ssci.2016.09.011

[4] Tirunagari, S., Hänninen, M., Ståhlberg, K., Kujala, P., 2012. Mining causal relations and concepts in maritime accidents investigation reports. In: Proceedings of the International Conference cum Exhibition on Technology of the Sea.

[5] Mazaheri, A., Montewka, J., Kujala, P., 2013. Correlation between the ship grounding accident and ship traffic- a case study based on the statistics of the Gulf of Finland. Trans. Nav. Int. J. Mar. Navig. Saf. Sea Transp. 7, 119-124. https://doi.org/10.12716/1001.07.01.16

[6] Yang, H., Wang, X., Yin, Y., 2012. The impact of speed limits on traffic equilibrium and system performance in networks. Transp. Res. Part B Methodol. 46, 1295-1307. https://doi.org/10.1016/j.trb.2012.08.002

[7] Hosseinlou, M.H., Kheyrabadi, S.A., Zolfaghari, A., 2015. Determining optimal speed limits in traffic networks. IATSS Res. 39, 36-41. https://doi.org/10.1016/j.iatssr.2014.08.003

419 [8] Van Benthem, A., 2015. What is the optimal speed limit on freeways? J. Public Econ. 124, 44-62.
420 https://doi.org/10.1016/j.jpubeco.2015.02.001
421 [9] Vadeby, A., Forsman, Å., 2018. Traffic safety effects of new speed limits in Sweden. Accid. Anal.
422 Prev. 114, 34-39. https://doi.org/10.1016/j.aap.2017.02.003
423 [10] Wen, Y., Huang, Y., Zhou, C., Yang, J., Xiao, C., Wu, X., 2015. Modelling of marine traffic flow
424 complexity. Ocean Eng. 104, 500-510. https://doi.org/10.1016/j.oceaneng.2015.04.051
425 [11] Zhang, L., Meng, Q., Fang Fwa, T., 2017. Big AIS data based spatial-temporal analyses of ship
426 traffic in Singapore port waters. Transp. Res. Part E Logist. Transp. Rev.
427 https://doi.org/10.1016/j.tre.2017.07.011
428 [12] Christian, R., Kang, H.G., 2017. Probabilistic risk assessment on maritime spent nuclear fuel
429 transportation (Part II: Ship collision probability). Reliab. Eng. Syst. Saf. 164, 136-149.
430 https://doi.org/10.1016/j.ress.2016.11.017
431 [13] Kang, L., Meng, Q., Liu, Q., 2018. Fundamental diagram of ship traffic in the Singapore Strait.
432 Ocean Eng. 147, 340-354. https://doi.org/10.1016/j.oceaneng.2017.10.051
433 [14] Kujala, P., Hänninen, M., Arola, T., Ylitalo, J., 2009. Analysis of the marine traffic safety in the Gulf
434 of Finland. Reliab. Eng. Syst. Saf. 94, 1349-1357.
435 https://doi.org/10.1016/j.ress.2009.02.028
436 [15] Mou, J.M., van der Tak, C. Ligteringen, H., 2010. Study on collision avoidance in busy waterways
437 by using AIS data. Ocean Eng. 37, 483-490.
438 https://doi.org/10.1016/j.oceaneng.2010.01.012
439 [16] Zhang, W., Goerlandt, F., Kujala, P., Wang, Y., 2016. An advanced method for detecting possible
440 near miss ship collisions from AIS data. Ocean Eng. 124, 141-156.
441 https://doi.org/10.1016/j.oceaneng.2016.07.059
442 [17] Mazaheri, A., Montewka, J., Kotilainen, P., Sormunen, O-VE., Kujala, P., 2015. Assessing
443 grounding frequency using ship traffic and waterway complexity. J. Navig. 68, 89-106.
444 https://doi.org/10.1017/S0373463314000502
445 [18] Bye, R.J., Aalberg, A.L., 2018. Maritime navigation accidents and risk indicators: An exploratory
446 statistical analysis using AIS data and accident reports. Reliab. Eng. Syst. Saf. 176, 174-186.
447 https://doi.org/10.1016/j.ress.2018.03.033
448 [19] Kumar, R., Chadrasekaran, R., 2011. Attribute correction-data cleaning using association rule and

clustering methods. Int. J. Data Min. Knowl. Manag. Process 1, 22-32.
https://doi.org/10.5121/ijdkp.2011.1202
[20] Sang, L., Wall, A., Mao, Z., Yan, X., Wang, J., 2015. A novel method for restoring the trajectory of the inland waterway ship by using AIS data. Ocean Eng. 110, 183-194.
https://doi.org/10.1016/j.oceaneng.2015.10.021
[21] Qu, X., Meng, Q., Li, S., 2011. Ship collision risk assessment for the Singapore Strait. Accid. Anal. Prev. 43, 2030-2036.
https://doi.org/10.1016/j.aap.2011.05.022
[22] Salama, A.S., 2010. Topological solution of missing attribute values problem in incomplete information tables. Inf. Sci. 180, 631-639.
https://doi.org/10.1016/j.ins.2009.11.010
[23] Shelmerdine, R.L., 2015. Teasing out the detail: How our understanding of marine AIS data can better inform industries, developments, and planning. Mar. Policy 54, 17-25.
https://doi.org/10.1016/j.marpol.2014.12.010