Sang, L-Z, Wall, A, Mao, Z, Yan, X-P and Wang, J

A novel method for restoring the trajectory of the inland waterway ship by using AIS data

Sang, L-Z, Wall, A, Mao, Z, Yan, X-P and Wang, J (2015) A novel method for restoring the trajectory of the inland waterway ship by using AIS data. OCEAN ENGINEERING, 110. pp. 183-194. ISSN 0029-8018

A novel method for restoring the trajectory of the inland waterway ship by using AIS data

Ling-zhi SANG [a,b], Alan WALL [c], Zhe MAO [a,b], Xin-ping YAN [a,b], Jin WANG [c]

[a] *National Engineering Research Centre for Water Transport Safety (WTSC), 1040, Heping Avenue, Wuhan, Hubei 430063, P.R. China*

[b] *Intelligent Transport Systems Research Centre (ITSC), Wuhan University of Technology, 1040, Heping Avenue, Wuhan, Hubei 430063, P.R. China*

[c] *Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, James Parsons Building, Byrom Street, Liverpool, L3 3AF, UK.*

Corresponding author: Ling-zhi SANG.

E-mail: sanglz@126.com, Mobile phone: 00 44 7419 216822, Address: 1040, Heping Avenue, Wuhan, Hubei 430063, P.R. China.

Abstract: The trajectory of the inland waterway ship is important and useful in analysing the features of the ship behaviour and simulating traffic flows. In the proposed research, a method is designed to restore the trajectory of an inland waterway ship based on the Automatic Identification System (AIS) data. Firstly, three rules are developed to identify and remove the inaccurate data, based on the reception range of the received AIS data and the manoeuvring characteristics of the inland waterway ship. Secondly, the method of restoring the full trajectory incorporating navigational features of the inland waterway ship is proposed to model the ship trajectory. The trajectory is characterised by three types (line, curve and arc) and five steps (line, curve, arc, curve and line) during the turning section. In order to validate the proposed method, the AIS data of two inland waterway ships collected from three AIS-base-stations is selected for the analysis, all inaccurate AIS data is identified and removed by the use of three cleansing rules. The results show that the three developed rules can effectively identify the inaccurate AIS data. The AIS data collected by an AIS-shipboard-unit is then used to: (1) restore the ship trajectory, and (2) validate the proposed method by comparing the reconstituted trajectories with the actual trajectory. This actual trajectory is determined from intermediate higher frequency Global Positioning System (GPS) data and collected from the AIS-shipboard-unit. The residual errors are calculated as the differences between the estimated latitude values of the restored trajectory functions and the real latitude values of the GPS data. Three alternative methods of trajectory restoring are also evaluated. The results show that the proposed method can be used to restore the full trajectory in an effective manner by using AIS data.

Key words: Inland waterway ship; trajectory; ship behaviour; ship safety; Automatic Identification System (AIS)

Abbreviations:

Automatic Identification System: AIS

International Maritime Organization: IMO

The International Convention for the Safety of Life at Sea: SOLAS

Gross Tonnage: GT

Maritime Safety Administration: MSA

Course Over Ground: COG

Speed Over Ground: SOG

Very High Frequency: VHF

Self-Organized Time Division Multiple Access: SOTDMA

Carrier-Sense Time Division Multiple Access: CSTDMA

International Telecommunication Union: ITU

Vessel Traffic System: VTS

Long Range Identification & Tracking: LRIT

Local Area Network: LAN

Ant Colony Algorithm: ACA

Evolutionary Algorithm: EA

The Convention on International Regulations for preventing collisions at sea: COLREGs

Safety Assessment Models for Shipping and Offshore in the North Sea: SAMSON

The Lloyd's Marine Intelligence Unit: Lloyd's MIU

China Classification Society: CCS

Global Positioning System: GPS

1 Introduction

Automatic Identification System (AIS) is a relatively new kind of digital navigational equipment. A series of measures have been undertaken by the International Maritime Organization (IMO) to promote its shipboard applications among which AIS plays an increasingly important role in terms of navigational safety and security. The International Convention for the Safety of Life at Sea (SOLAS) requires AIS to be fitted aboard international voyaging ships with Gross Tonnage (GT) more than 300 and all passenger ships. Moreover, some related authorities, such as Maritime Safety Administration (MSA) in China, also pay special attention to allocate AIS for inland waterway ships. For instance, the Yangtze MSA requires that all inland waterway ships over 100 GT be equipped with AIS by 1st July 2012 at the latest. By all these efforts, AIS has been widely adopted in both international and inland waterway ships.

AIS information usually contains: (1) static information, such as ship name, type, ship length, ship width, call sign and Maritime Mobile Service Identity (MMSI) code, (2) semi-dynamic information such as cargo information and destination, and (3) dynamic information such as position, Course Over Ground (COG), Speed Over Ground (SOG) and turning rate, as well as heading if AIS can receive the heading information from the sensor (International Maritime Organization (IMO), 2002). In addition, AIS allows for additional optional information to be broadcasted, which gives additional ship details including routes and also for shore stations to broadcast some hydro-meteorological and other information. All the above information is broadcasted via the Very High Frequency (VHF) channel with two communication modes:

Self-Organized Time Division Multiple Access (SOTDMA) and Carrier-Sense Time Division Multiple Access (CSTDMA). Currently, two types of ship-borne AIS which are called Class-A AIS and Class-B AIS are widely used by ships. According to the resolution of IMO, Class-A AIS can only adopt the SOTDMA protocol (IMO, 2002 and International Telecommunication Union (ITU), 2014), while Class-B AIS can adopt anyone of these two communication modes (ITU, 2014). Meanwhile, due to the bandwidth limitation, the static and semi-dynamic information of AIS is transmitted with a long period. The transmission period for dynamic information depends on the navigation status of the ship and the type of AIS. According to the requirements of ITU, the transmission periods and further details for Class-A and Class-B are shown in table 1 (ITU, 2014).

The information provided by AIS is very valuable for navigational safety and risk assessment. AIS information can be combined with the data from other maritime surveillance systems, such as the radar based Vessel Traffic System (VTS), to obtain more precise ship positions and trajectories (Harre, 2000). A fuzzy fusion extrapolation method was proposed by Liu and Shi (2009) to integrate data from VTS and AIS systems to improve the accuracy. The method improved the performance and stability of the VTS system by AIS. Cairns (2005) found that AIS might also play a role of ship tracking as a complement to Long Range Identification & Tracking (LRIT). By using AIS and wireless Local Area Network (LAN), Hu, *et al.* (2007) developed a live AIS web system to publish the real-time AIS information to the public.

Furthermore, one of the most important assumptions for the research on collision avoidance (Felski and Jaskolski, 2012), ship trajectory planning and manoeuvring optimization is that the initial speed, course and position of the ship are known beforehand. All these kinds of information can be supplied by AIS data. For instance, a collision avoidance decision-making method was designed by Su, *et al.* (2012). A fuzzy logic theory based system used the AIS information to make the optimal rudder steering decision for the give-way ship. This assisted the give-way ship in avoiding the collision when there was a high collision risk with other ships nearby. Ant Colony Algorithm (ACA) (Tsou and Hsueh, 2010) and Evolutionary Algorithm (EA) (Smierzchalski, 1999; Szlapczynski and Szlapczynska, 2012) were also widely used to find safe and economic trajectories for all involved ships. Under the Convention on the International Regulations for Preventing Collisions at Sea (COLREGs), AIS data was used to determine the decision making point (moment) of the ship in different roles (give-way or stand-on). The evidence theory was then used to assist the ship officer in understanding the role of the own ship when avoiding the collision (Wang, *et al.*, 2013). Consequently, AIS data is necessary for research into the intelligent navigation. Ship domain is an essential issue for the collision avoidance study. AIS data was used to compute the distances and angles between the own ship and other ships, a minimum length and breadth of ship domain were analysed to keep ship safe (Hansen, *et al.*, 2012).

AIS data provides the real-time information that is useful for marine traffic investigation, statistical analysis and theoretical research (Tsou, 2010). The anomalous behaviour of the ship was identified by comparing routes and AIS data (Liu and Chen, 2013). In terms of maritime accident investigation, Ambjorn (2008) found a large number of possible oil spilling ships by AIS data mining in *Seatrack Web,* which could also forecast the diffusion of the spilled oil. Peters and Hammond (2011) designed a method to address probabilistic queries about the location of a vessel in the time interval between two reported positions from AIS data.

Meanwhile, the historical AIS information is a very valuable source for navigational risk assessment. The following can all be studied by using AIS data: statistical analysis of the historical data, traffic flow characteristics such as traffic volume, and distributions of the size and type of ships (Gunnar Aarsæther and Moan, 2009). Mou, *et al.* (2010) used AIS data to identify the correlations between the Closest Point of Approach (CPA) and the size/speed/course of the ship. Linear regression models were then developed. In addition, a dynamic method based on Safety Assessment Models for Shipping and Offshore in the North Sea (SAMSON) was formulated to assess the risk. In the Singapore Strait, under a data-cleansing procedure, AIS data was used to compute three navigational risk indices: speed dispersion, degree of acceleration and deceleration, and the number of ship domain overlaps. After that, the dangerous areas and ship types were identified (Qu, *et al.*, 2011).

Last but not least, many traffic simulation models involving AIS have been built to assess maritime accidents (e.g. collisions, groundings and oil spills). Several traffic simulation platforms were developed in the San Francisco Bay (Merrick, *et al.*, 2003), the Gulf of Finland (Goerlandt and Kujala, 2011), the port of Koper (Perkovic, *et al.*, 2012) and the coast of Portugal (Silverira, *et al.*, 2013). AIS can provide reliable data for maritime risk assessment. Weng, *et al.* (2012) estimated the vessel collision frequency by using the software *@risk* in the Singapore Strait with AIS data. This data was obtained from the Lloyd's Marine Intelligence Unit (Lloyd's MIU) database. The collision frequency was obtained as the product of the number of vessel conflicts and the causation probability. Goerlandt and Kujala (2014) simulated the voyage and obtained ship trajectories to estimate the risk of the ship-ship collision using AIS data quantitatively.

Currently, Class-B AIS is widely used by inland waterway ships. It is worth noting that the rate of turning is not contained in Class-B AIS messages. To reduce the chance of data overload, the transmitting power of Class-B is lower than Class-A, so is the data rate. The Class-B AIS signal which uses the CSTDMA protocol to broadcast messages sometimes generates garbled messages (Norris, 2006). The signal can be interfered by many factors, such as the occlusion by river banks and buildings, bad weather conditions, especially restricted inland waterways. As a result, the information can easily be incorrectly coded or lost. At the same time, the ship trajectory is commonly restored by connecting the location points sequentially in terms of time series (Tang, *et al.*, 2012). However, if positions extracted from AIS data are incorrect, the estimated ship trajectory is also wrong. Missing or lost data can reduce the accuracy of the trajectory restored. Under either circumstance, the restored trajectory would affect seafarers' judgment on traffic situations and furthermore increase navigation risk significantly. Consequently, screening out incorrect AIS data and restoring a realistic trajectory are meaningful for enhancing the applicability of AIS data and improving the safety of inland waterway ships. This research focuses on how to identify inaccurate AIS data. The restoration of the inland waterway ship trajectory is then investigated.

2 Data collection

Two parts of AIS data are used by case studies in this research. The data was collected by three AIS-base-stations and an AIS-shipboard-unit.

For Part 1 AIS data, three AIS-base-stations are located beside the *Tianxingzhou Yangtze River Bridge*, *Wuhan Yangtze River Bridge* and *Baishazhou Yangtze River Bridge* respectively. These AIS-base-stations can receive AIS data of the whole Wuhan section of the Yangtze River. All of the longitude, latitude and time information of AIS is stored into an AIS database after being parsed. As shown in figure 1, the three red pentagrams on the map are the locations of the AIS-base-stations.

For Part 2 AIS data, a Class-B AIS-shipboard-unit is upgraded and approved by the China Classification Society (CCS) to save its own messages when broadcasting. In addition, this AIS-shipboard-unit can save its own information in the middle of each broadcasting interval to obtain more navigation information. Thus the time interval of AIS data collected in this research is half of the standard interval (approximate 15 seconds when SOG is larger than 2 knots). Meanwhile, the Global Positioning System (GPS) data from this equipment's embedded GPS module is stored once per second. This equipment was installed in a merchant ship with the name *xinpingjiang-1025* (MMSI: 413762554) navigating on the Yangtze River to collect AIS and GPS data. In this research, when *xinpingjiang-1025* was about to pass *Shishou Bend* that is a noticeably curved waterway of the Yangtze River, 15 sets of AIS data and 213 sets of GPS data in a same time period were selected to analyse the effect of restored trajectories. As the position and time information of AIS is given by the embedded GPS modular, AIS data has the same position and time information with corresponding ship GPS data when broadcasting, as shown in table 2.

The AIS positions when the ship was about to pass *Shishou Bend* are shown in figure 2.

3 Restoration method of the inland waterway ship trajectory

Class-B AIS ship-borne mobile equipment is commonly noted in inland waterways. As a result, the AIS data from inland waterway ships is not immune to errors. Inaccurate data can significantly influence the restored trajectory. Therefore a cleansing procedure for the data is necessary.

3.1 Method of cleansing the AIS data of inland waterway ships

3.1.1 Categories of AIS inaccurate data

AIS data of inland waterway ships nearby can be acquired by the AIS-base-stations. The trajectory can then be obtained by connecting ship positions after data parsing. According to the reception range of AIS-base-stations and the manoeuvrability of the vessel, inaccurate AIS data can be grouped into the following three categories:

(1) Range anomaly: the longitude or latitude exceeds the expected received range. In this research, AIS stations used for data acquisition are laid in Wuhan beside the Yangtze River. Therefore there exists a normal range of the ship position. The normal range of latitude is from 29°N to 32°N while the normal longitude range is from 113°E to 116°E. It would be considered as inaccurate data if the latitude or longitude is outside the range.

(2) Velocity anomaly: the latitude and longitude are within the normal range, but the values of latitude and longitude significantly deviate from a normal trajectory with an abnormal change of SOG. The trajectory near the anomaly point is observed to be thorn-like (see figure 5).

(3) Course anomaly: while the latitude and longitude are within the normal range and there is no

velocity anomaly, the change of COG is abnormal and the trajectory near the anomaly point is not smooth.

Based on these three categories of inaccurate AIS data, three corresponding rules are designed to pre-process AIS data before restoring the trajectory.

3.1.2 Cleansing method

According to the state of the own ship, the AIS equipment broadcasts the message at a certain approximate time interval as shown in table 2. Therefore the received ship positions are in a discontinuous form of time-series.

If there are m sets of AIS data and the ship positions are a time series of discontinuous points within a certain time interval following the requirements of ITU, the longitude values are $x_1, x_2, x_3, \cdots, x_{m-1}, x_m$ and the latitude values are $y_1, y_2, y_3, \cdots, y_{m-1}, y_m$. The ship status matrix is set as:

$$\begin{bmatrix} longitude \\ latitude \\ SOG \\ COG \end{bmatrix} = \begin{bmatrix} x_1, x_2, x_3, \cdots, x_{m-1}, x_m \\ y_1, y_2, y_3, \cdots, y_{m-1}, y_m \\ v_1, v_2, v_3, \cdots, v_{m-1}, v_m \\ c_1, c_2, c_3, \cdots, c_{m-1}, c_m \end{bmatrix}$$

Inaccurate data of the first category can be identified and removed directly from the data sequence by detecting whether the longitude and latitude are within the valid range. The range is determined by the coverage of data acquisition devices. Suppose the range of longitude is $[x_{min}, x_{max}]$, and the range of latitude is $[y_{min}, y_{max}]$.

Rule 1 is:

$$\begin{matrix} x_i \le x_{min} \ or \ x_i \ge x_{max}, \ or \\ y_i \le y_{min} \ or \ y_i \ge y_{\max} \end{matrix} \qquad i = 1,2,3, \cdots, m-1, m \tag{1}$$

If the value of x_i or y_i satisfies anyone of these four inequalities, it means that this position is out of the normal range and will be identified as inaccurate data.

The rules of other two categories of inaccurate data are considered to be based on the navigational characteristics of the inland waterway ship. According to the communication transfer protocol of AIS, Class-B AIS messages including the position and time information would be transmitted from the ship with a certainty period. During the interval, SOG and COG of the inland waterway ship should be relatively steady or with a smooth change. The change values of SOG and COG can be calculated based on two adjacent positions. If anyone of SOG and COG changes sharply at a point, the point would then be considered as inaccurate data. It is important to note that when measuring the changes of SOG and COG, the obvious inaccurate data needs to be cleansed by Rule 1 first.

The basic idea of the method to identify the second and third categories of inaccurate data is as follows. The inaccurate bound is defined as W, the sampled value at the time t is set as x_t. The predicted value x'_t at the time t can then be calculated:

$$x'_t = x_{t-1} + (x_{t-1} - x_{t-2}) \tag{2}$$

If $|x_t - x'_t| > W$, x_t will be considered as inaccurate data and be removed from the data sequence. The inaccurate data will be replaced with x'_t to identify next value x_{t+1}.

With the longitude, latitude and interval time information, the value of SOG v_i can be calculated to identify the second category of inaccurate data. According to equation (2), the predicted SOG v'_i can be obtained.

The inaccurate bound of the second category W_1 is set as: $W_1 = k_1 \times v_{i-1}$ where k_1 is the parameter of the SOG change. Then Rule 2 for identifying SOG anomaly is:

$$|v_i - v_i'| > W_1 \qquad i = 2,3,\cdots,m-1,m \tag{3}$$

In the case of this research, for the real AIS data of the inland waterway ship, $k_1 \in [0.3, 0.7]$.

If v_i at the i^{th} AIS point satisfies equation (3), this AIS point can be identified as inaccurate data.

In order to identify the third category of inaccurate data, an essential criterion on COG is constructed based on the ship manoeuvrability. When the ship changes the course, the maximum value of the course change between two adjacent AIS positions can be calculated by simplifying the motion as a steady circle as shown in figure 3.

On the basis of the above assumptions, the maximum threshold of the angle change of the inland waterway ship is developed as follows. Figure 3 shows three consecutive points: A, B and C. The corresponding centre point of arc $\widehat{AB}$ is point O, the radius is R_1 and the average SOG is $\bar{v}_1$, $\bar{v}_1 = (v_A + v_B)/2$.The corresponding centre point of arc $\widehat{BC}$ is point O' which is close to point O, the radius is R_2 and the average SOG is $\bar{v}_2$, $\bar{v}_2 = (v_B + v_C)/2$. In arcs $\widehat{AB}$ and $\widehat{BC}$, there exist:

$$\begin{cases} T_1 = \dfrac{2\pi R_1}{\bar{v}_1} \\ T_2 = \dfrac{2\pi R_2}{\bar{v}_2} \end{cases} \tag{4}$$

where:

T_1 is the time period when the ship completed an entire circle motion with speed $\bar{v}_1$.

T_2 is the time period when the ship completed an entire circle motion with speed $\bar{v}_2$.

According to the arcs $\widehat{AB}$ and $\widehat{BC}$, there are:

$$\begin{cases} \dfrac{\Delta t_1}{T_1} = \dfrac{\angle 1}{2\pi} \\ \dfrac{\Delta t_2}{T_2} = \dfrac{\angle 2}{2\pi} \end{cases} \tag{5}$$

where:

Δt_1 is the time interval of the ship moving from A to B.

Δt_2 is the time interval of the ship moving from B to C.

$\angle 1$ is the centre angle of arc $\widehat{AB}$.

$\angle 2$ is the centre angle of arc $\widehat{BC}$.

Therefore the angles $\angle\alpha$ and $\angle\beta$ shown in figure 3 can be obtained as follows:

$$\begin{cases} \angle\alpha = \dfrac{\pi}{2} - \dfrac{\Delta t_1}{T_1}\pi \\ \angle\beta = \dfrac{\pi}{2} - \dfrac{\Delta t_2}{T_2}\pi \end{cases} \tag{6}$$

Then the COG change $\angle x$ of point B can be computed by the following equation:

$$\angle x = \pi - \angle\alpha - \angle\beta = \frac{\Delta t_1 \bar{v}_1}{2R_1} + \frac{\Delta t_2 \bar{v}_2}{2R_2} \tag{7}$$

Given the length of the inland waterway ship as l, the relationship between the length l and the radius of the steady circle R can be approximately expressed as:

$$R = k_2 \times l \tag{8}$$

where:

$R \leq \min(R_{1,}R_2)$.

k_2 is the parameter between the ship length l and radius of the steady circle R.

Then ∠x can be computed by the following equation:

$$\angle x \leq \frac{\Delta t_1 \bar{v}_1}{2k_2 l} + \frac{\Delta t_2 \bar{v}_2}{2k_2 l} \tag{9}$$

For an inland waterway ship, the tactical diameter ($2R$) is generally two to four times as long as the length of the ship (Wu, 1999), therefore $k_2 \in [2,4]$. Suppose the distance between points A and B is s_1, $s_1 = \Delta t_1 \bar{v}_1$, and the distance between points B and C is s_2, $s_2 = \Delta t_2 \bar{v}_2$. Then:

$$\angle x \leq \frac{1}{4l}(s_1 + s_2) \tag{10}$$

According to the above equations, the inaccurate data bound of steering angle can be set as W_2:

$$W_2 = \max \angle x = \frac{s_1 + s_2}{4l} \tag{11}$$

The longitude and latitude of two adjacent points, the COG c_i and the corresponding predicted value c_i' can then be calculated to identify the third category of inaccurate data.

Rule 3 is:

$$|c_i - c_i'| > W_2 \qquad i = 2,3,\cdots,m-1,m \tag{12}$$

If the COG c_i at the i^{th} AIS point satisfies equation (12), it means that the change of COG exceeds the value when the ship is turning the course with the largest helm angle. This AIS point can then be identified as inaccurate data.

3.2 Method of restoring the inland waterway ship trajectory

There are two situations of the inland waterway ship motion during the voyage: steady course and turning course. The whole trajectory is a combination of both:

(1) The first one considered as a straight line trajectory when the ship is navigating on a steady course.

(2) The second one considered as a curve trajectory when the ship is navigating on a turning course.

At the point of intersection, these two trajectories are tangent to each other.

The trajectory can be simply restored as a straight line when the course is steady and the ship is in linear navigation. However, when the inland waterway ship is turning the course, the motion is changing constantly until reaching the next linear steady course. As a result, the trajectory is a complex curve in this stage.

When steering to change the course, the inland waterway ship will get into a steady circle period after the manoeuvring is completed. Based on the navigational features, the ship has a starting curve trajectory before settling into the steady circle and then has a circular trajectory of constant radius. This circular trajectory is referred to as an arc in this work. After the course turning is completed, the helm will be eased and the course will be adjusted to stay on a new certain course, the ship will then move ahead along as a straight line.

By using the navigational features of the ship, the second type of trajectory when changing the course can be divided into three parts:

(1) The centre part is an arc of constant radius which is tangent to the adjacent trajectories at the endpoints.

(2) The other two parts on each side cannot be restored by circles or lines and can only be restored by curves. These two parts are referred to as curves in this research.

Therefore, the trajectory can be restored separately by line, curve and arc. Meanwhile, if the course turning time is too short to let the inland waterway ship enter the steady turning stage, the arc trajectory will not be generated. Then the trajectory while the ship is navigating on a turning course can be restored by curve directly.

3.2.1 Line trajectory on the steady course

If m sets of the position data are received when navigating on a steady course, the trajectory can be restored by the method of piecewise linear interpolation. Every two adjacent points in time series within the time interval required by the ITU can be linked by a straight line, the trajectory $f(x)$ satisfies:

$$\begin{cases} f(x) = \sum_{i=1}^{m-1} f_i \\ f_i = y_i + (x - x_i)\dfrac{y_{i+1} - y_i}{x_{i+1} - x_i} \end{cases} \qquad i = 1,2,3,\cdots,m-1 \qquad (13)$$

The slope of the straight line k_i is:

$$k_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \qquad i = 1,2,3,\cdots,m-1 \qquad (14)$$

3.2.2 Arc trajectory on the turning course

In the course turning stage, after steering and before easing the helm, the ship is in a steady circle period and the trajectory is an arc of a circle. The points of the steady circle need to be processed first. In general, the turning radius R and the length of ship l have the following relation:

$$R = k_2 \times l \qquad k_2 \in [2,4] \qquad (8)$$

According to ship positions, one circle can be obtained by every three sequential positions. Therefore m positions can form $m-2$ circles.

If there are three sequential points (x_i, y_i), (x_{i+1}, y_{i+1}), (x_{i+2}, y_{i+2}), the position of the circle centre (p_i, q_i) can be calculated by the following equations:

$$\begin{cases} (x_i - p_i)^2 + (y_i - q_i)^2 = (x_{i+1} - p_i)^2 + (y_{i+1} - q_i)^2 \\ (x_i - p_i)^2 + (y_i - q_i)^2 = (x_{i+2} - p_i)^2 + (y_{i+2} - q_i)^2 \end{cases} \qquad (15)$$

The radius R_i satisfies:

$$R_i = \sqrt{(x_i - p_i)^2 + (y_i - q_i)^2} \qquad (16)$$

The ship can be considered as being in the steady circle stage when passing these three sequential points if the radius R_i of the corresponding steady circle satisfies:

$$R_i \le \max(k_2) \times l \qquad (17)$$

Then the trajectory can be restored as an arc passing those three points.

As shown in figure 4, if $t \in [t_i, t_{i+2}]$, in this same period, the angle θ_t can be calculated by the following equations:

$$\begin{cases} \sin\theta_t = (x - p_i)/R_i \\ \cos\theta_t = (y - q_i)/R_i \end{cases} \quad t \in [t_i, t_{i+2}] \tag{18}$$

If $t = t_i$, the angle θ_i satisfies:

$$\begin{cases} \sin\theta_i = (x_i - p_i)/R_i \\ \cos\theta_i = (y_i - q_i)/R_i \end{cases} \tag{19}$$

If $t = t_{i+2}$, the angle θ_{i+2} satisfies:

$$\begin{cases} \sin\theta_{i+2} = (x_{i+2} - p_i)/R_i \\ \cos\theta_{i+2} = (y_{i+2} - q_i)/R_i \end{cases} \tag{20}$$

The position (x, y) of the point on the arc can be obtained by the following equations:

$$\begin{cases} x = R_i \times \sin\left[\frac{t - t_i}{t_{i+2} - t_i}(\theta_{i+2} - \theta_i)\right] + p_i \\ y = R_i \times \cos\left[\frac{t - t_i}{t_{i+2} - t_i}(\theta_{i+2} - \theta_i)\right] + q_i \end{cases} \quad t \in [t_i, t_{i+2}] \tag{21}$$

If the time spent in the steady circle is relatively long, the radius R_{i+1} of new arc passing next three points $(x_{i+1}, y_{i+1}), (x_{i+2}, y_{i+2}), (x_{i+3}, y_{i+3})$ meets equation (17) as well. Then a new position of circle centre (p_{i+1}, q_{i+1}) can be obtained, and the new radius R_{i+1} is:

$$R_{i+1} = \sqrt{(x_{i+1} - p_{i+1})^2 + (y_{i+1} - q_{i+1})^2} \tag{22}$$

For the four sequential points $(x_i, y_i), (x_{i+1}, y_{i+1}), (x_{i+2}, y_{i+2}), (x_{i+3}, y_{i+3})$, if $t \in [t_i, t_{i+3}]$, the ship trajectory can be simply considered as a sequential arc with the radius changing from R_i to R_{i+1} linearly and the course angle changing from θ_i to θ_{i+1} linearly.

If $t \in [t_i, t_{i+3}]$, the centre (p_t, q_t) satisfies:

$$\begin{cases} p_t = p_i + \frac{t - t_i}{t_{i+3} - t_i}(p_{i+1} - p_i) \\ q_t = q_i + \frac{t - t_i}{t_{i+3} - t_i}(q_{i+1} - q_i) \end{cases} \quad t \in [t_i, t_{i+3}] \tag{23}$$

The radius R_t satisfies:

$$R_t = R_i + \frac{t - t_i}{t_{i+3} - t_i}(R_{i+1} - R_i) \quad t \in [t_i, t_{i+3}] \tag{24}$$

The course θ_t satisfies:

$$\theta_t = \theta_i + \frac{t - t_i}{t_{i+3} - t_i}(\theta_{i+1} - \theta_i) \quad t \in [t_i, t_{i+3}] \tag{25}$$

The position (x, y) of the point on the arc can be obtained by the following equations:

$$\begin{cases} x = R_t \times \sin\theta_t + p_t \\ y = R_t \times \cos\theta_t + q_t \end{cases} \quad t \in [t_i, t_{i+3}] \tag{26}$$

If the next radius R_{i+n} of new arc passing $(x_{i+n-2}, y_{i+n-2}), (x_{i+n-1}, y_{i+n-1}), (x_{i+n}, y_{i+n})$ also meets equation (17), the final functions of the position (x, y) on the arc are:

$$\begin{cases} x = R_t \times \sin\theta_t + p_t \\ y = R_t \times \cos\theta_t + q_t \end{cases} \quad t \in [t_{i+n-2}, t_{i+n}] \tag{27}$$

where:

$$\begin{cases} R_t = R_{i+n-3} + \frac{t - t_{i+n-3}}{t_{i+n} - t_{i+n-3}}(R_{i+n} - R_{i+n-1}) \\ \theta_t = \theta_{i+n-3} + \frac{t - t_{i+n-3}}{t_{i+n} - t_{i+n-3}}(\theta_{i+n} - \theta_{i+n-1}) \\ p_t = p_{i+n-3} + \frac{t - t_{i+n-3}}{t_{i+n} - t_{i+n-3}}(p_{i+n} - p_{i+n-1}) \\ q_t = q_{i+n-3} + \frac{t - t_{i+n-3}}{t_{i+n} - t_{i+n-3}}(q_{i+n} - q_{i+n-1}) \end{cases} \quad t \in [t_{i+n-2}, t_{i+n}]$$

3.2.3 Curve trajectory on the turning course

In the course turning stage, before and after the steady circle period, the surge speed and rate of turning is changing nonlinearly. The generated trajectory at this time is a nonlinear curve. The generation of ship trajectory partly depends on manoeuvrability parameters of the ship, such as the steering indices K and T (Nomoto, 1957; Sutulo, *et al.*, 2002). To simplify the modelling process, each of the two trajectory curves is obtained by connecting adjacent points with the piecewise cubic spline interpolation method (Hasberg, *et al.*, 2008) in this paper. The two curves not only have continuous tangents, but also have continuous curvatures at the intersection points. The curve trajectory can be constructed by the following method.

In m sets of AIS data, if there exists the longitude range $[x_{min}, x_{max}]$, the function $s(x)$ of the cubic spline trajectory is:

$$s(x) = \sum_{i=1}^{m-1} s_i \tag{28}$$

where:

$i = 1,2,\cdots,m-1$.

s_i is a third degree polynomial in interval $[x_i, x_{i+1}]$.

Setting the piecewise trajectory function $s_i = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + y_i$, there are three parameters in each piecewise trajectory among the $m - 1$ piecewise trajectories. In total, $3(m - 1)$ parameters need to be solved to obtain $s(x)$.

For the cubic spline trajectory, it is continuous everywhere, thus the AIS position (x_{i+1}, y_{i+1}) satisfies the previous function s_i:

$$y_{i+1} = a_i(x_{i+1} - x_i)^3 + b_i(x_{i+1} - x_i)^2 + c_i(x_{i+1} - x_i) + y_i \quad i = 1,2,\cdots,m-1 \tag{29}$$

Both the first derivative $s'(x)$ and second derivative $s''(x)$ should be continuous everywhere and at the AIS positions. To achieve the condition, there are:

$$\begin{cases} s_i'(x_{i+1}) = s_{i+1}'(x_{i+1}) \\ s_i''(x_{i+1}) = s_{i+1}''(x_{i+1}) \end{cases} \quad i = 1,2,\cdots,m-2 \tag{30}$$

Both the first derivatives and second derivatives on the endpoint with this curve trajectory and the adjacent line trajectory are equal, so is the endpoint with this curve trajectory and the adjacent arc trajectory.

The boundary condition on the endpoint with the line satisfies:

$$s_1'' = 0 \tag{31}$$

The boundary condition on the endpoint with the arc satisfies:

$$s_2'' = \frac{R_i^2}{(q_i - y_i)^2} \quad (32)$$

$3(m-1)$ equations can be obtained from equations (29), (30), (31) and (32) in order to solve all the parameters and to obtain each piecewise trajectory function s_i (Liu, *et al.*, 2011). The function $s(x)$ of the cubic spline trajectory can then be obtained.

4 Case studies

4.1 AIS data cleansing

The real AIS data collected by three AIS-base-stations is used in the case study to identify the inaccurate data. Based on the collected data of part 1, two inland waterway ships named *pinganda-99* (MMSI: 413800000) and *yuxinhuo-11379* (MMSI: 413984805) are selected when passing the Wuhan waterway of the Yangtze River. All inaccurate AIS data is identified and cleansed by the proposed three cleansing rules.

The raw positions of these two ships are shown in figure 5. Figure 5 illustrates some irregularities in the AIS data. Incorrect AIS information greatly affects the display of the trajectories. The positions after cleansing are shown in figure 6. However, it is shown in figure 6 that nearly all inaccurate data (red points) is identified using the described algorithms, and the inaccurate data can be removed to obtain a clearer vessel trajectory (blue trajectory).

Only after being processed by cleansing all inaccurate data, the AIS data can be used to restore the trajectory. Therefore, data cleansing is the first and important step to obtain the trajectory of an inland waterway ship.

4.2 Restoration of the inland waterway ship trajectory using AIS data

This case study uses the following type of data: (1) the AIS data collected by the Class-B AIS-shipboard-unit; and (2) higher frequency GPS data collected from the embedded GPS module. The trajectory can be restored by the position information of AIS obtained from the embedded GPS module. This restored trajectory can be assessed by comparison with the more frequent GPS data collected on the ship.

When the inland waterway ship is navigating on a steady course, the trajectory can be considered as a linear function and restored by a straight line using equation (13) directly. However, when the ship is turning the course, this ship can have both nonlinear trajectory and linear trajectory. This trajectory restoration is analysed in detail in the following case study.

In order to compare the proposed method with other restoration methods, firstly, the trajectory of the ship is restored by the piecewise linear interpolation method that connects two adjacent AIS points in the form of time series. Secondly, the piecewise cubic interpolation method and the piecewise cubic spline interpolation method (Hasberg, *et al.*, 2008) are also used to restore the trajectories. Finally, the trajectory restored by using the proposed method based on the ship navigational features and the trajectories restored by using the three existing interpolation methods are compared.

4.2.1 Restoration of the trajectory with complete AIS data

When *xinpingjiang-1025* was about to pass *Shishou Bend*, which is a typical and noticeable winding channel of the Yangtze River, 15 sets of AIS data broadcasted by its own are used to restore the trajectory, while 213 sets of GPS data in the same period are recorded to analyse the effect of the restoration.

Based on the ship positions of these 15 sets of AIS data, three traditional interpolation methods and the new proposed method are used to restore the trajectory separately. The estimated values of latitude can be obtained from functions of the four restored trajectories. Residual errors are computed by the differences between the estimated latitude values of functions of the four restored trajectories and real latitude values of GPS.

The first trajectory restored by the piecewise linear interpolation method is shown in figure 7. The residual errors are shown in figure 8. As shown in figure 8, the residual errors are discrete. The mean value is 2.7276×10^{-6}, and the standard deviation is 6.3394×10^{-6}.

The second trajectory restored by the piecewise cubic interpolation method and its residual errors are shown in figures 9 and 10. As shown in figure 10, the residual errors follow a normal distribution. The mean value is 0.2613×10^{-6} while the standard deviation is 3.2231×10^{-6}.

The third trajectory restored by the piecewise cubic spline interpolation method and its residual errors are shown in figures 11 and 12. As shown in figure 12, the residual errors follow a normal distribution. The mean value is -0.1422×10^{-6} while the standard deviation is 2.6654×10^{-6}.

The fourth trajectory restored by using ship navigational features is shown in figure 13. From point 1 to point 3 and from point 13 to point 15, the ship is approximately navigating straight ahead, and from point 6 to point 9, the ship is in the steady circle period with an arc trajectory. Its residual errors are shown in figure 14 and they follow a normal distribution. The mean value is 0.1736×10^{-6} while the standard deviation is 2.2760×10^{-6}.

It is shown that the first trajectory restored by the piecewise linear interpolation method has the largest residual errors, the distribution of residual errors is discrete in the range of $[-1.0\times10^{-5}, 2.0\times10^{-5}]$. The second trajectory restored by the piecewise cubic interpolation method has a better effect than the first trajectory, the residual errors are in the range of $[-0.8\times10^{-5}, 1.1\times10^{-5}]$. The third trajectory restored by the piecewise cubic spline interpolation method has a better effect than the second one, the residual errors are in the range of $[-10.7\times10^{-6}, 6.4\times10^{-6}]$. Compared with the first two trajectories, the magnitude of residual errors decreases from 10^{-5} to 10^{-6}, thus the accuracy of the trajectory has been improved.

However, the last trajectory restored by using ship navigational features has the best effect of all. The residual errors are in the range of $[-5.1\times10^{-6}, 6.8\times10^{-6}]$ following a normal distribution. Compared to the third one, this trajectory has a similar level of residual errors and a smaller standard deviation.

The specific comparisons of these four trajectories' residual errors are shown in table 3.

Every trajectory has 213 residual errors in terms of the differences between estimated latitude values of the restored trajectory functions and real latitude values of GPS (see table 2), as the four sets of residual errors are shown in figure 15. The residual errors of the first trajectory (linear interpolation) are shown by blue dashed lines, the residual errors of the second (cubic interpolation) by blue lines, the third residual errors (cubic spline interpolation) by green lines, and the residual errors of the fourth trajectory (proposed by this work) by red lines.

The errors of two trajectories restored by the piecewise linear interpolation method (blue dashed lines) and by using ship navigational features (red lines) from GPS No.1 to No.32 and from No.182 to No.213 are overlapped. That is because when using ship navigational features to restore the trajectory, AIS points from 1 to 3 (from GPS No.1 to No.32) and AIS points from 13 to 15 (from GPS No.182 to No.213) are deemed to be straight lines.

From GPS No.17 to No.32, the residual errors of the fourth trajectory are larger than those of the second and third ones. It may be because the trajectory is not so straight under wind drift or current drift.

Within the last three AIS points (from GPS No. 182 to 213), the fourth trajectory is much better than the third one. It can be shown that the cubic spline interpolation method is not suitable to restore a straight line trajectory. The cubic spline function with a power three polynomial suffers slightly from the *Runge's Phenomenon* (Süli and Mayers, 2003) which reduces the accuracy in this case.

Meanwhile, from AIS point 6 to AIS point 9 (from GPS No.77 to No.122), the ship is deemed to be in a steady circle period and has an arc trajectory, leading to most residual errors of the fourth trajectory being significantly smaller compared to the first two, and being slightly smaller compared to the third trajectory.

4.2.2 Restoration of the trajectory with missing AIS data

Some inaccurate data in the real AIS data exists, thus the inaccurate data needs to be identified and removed by using the AIS data cleansing algorithm before restoring the trajectory. In order to simulate the data after the cleansing process, for 15 sets of AIS data referred in table 2, the analysis is repeated with one AIS data point removed. With the exception of the end points (AIS points 1 and 15), a total of thirteen situations are simulated in which only an identical AIS point is lost in each situation. For example, point 2 is lost in situation 1, point 3 is lost in situation 2, *et al*. Therefore, in each case, the remaining and useful data is not regular or continuous, but has one AIS point missing.

Using this imperfect AIS data which simulates an inaccurate data point removed or a lost data point, the trajectory is restored again by the four restoring methods, the four restored trajectories can then be analysed by the residual errors.

From the second set to the fourteenth set, whatever data is lost, three trajectories can be restored directly with the three interpolation methods, and residual errors of GPS's latitude can still be calculated.

When restored by the proposed method, the manoeuvring procedure is not clear as before, especially when the missing data is at an intersection point of two different kinds of navigational features. The several different situations are as follows:

(1) If the lost data is point 3, it is still a line between point 1 and point 2. The line between point 2 and point 3 cannot be identified. Therefore it is a curve from point 2 to point 6.

(2) If the lost data is point 6, the arc between points 6, 7 and 8 cannot be identified any more. Therefore it is a curve between point 3 and point 7, and there is a unique circle from point 7 to point 9.

(3) If the lost data is point 7, the arc is unique between points 6, 8 and 9.

(4) If the lost data is point 8, the arc is unique between points 6, 7 and 9.

(5) If the lost data is point 9, the arc between points 7, 8 and 9 cannot be identified. It is a curve trajectory between point 8 and point 13, and there is a unique circle from point 7 to point 8.

(6) If the losing data is point 13, it is still a line between point 14 and point 15, the line between point 13 and point 14 cannot be identified. It is a curve from point 9 to point 14.

When the seventh set of AIS data is missing, the residual errors are shown in figure 16 as an example. The changes of residual errors of the four trajectories are shown by the blue cross symbol, the blue circle symbol, a green asterisk symbol and a red spot symbol respectively.

It is shown in figure 16 that the new restoring method using ship navigational features has a tiny change of residual errors, indicating that it is more robust than the other three methods.

Means of the absolute values of all 213 residual errors can be calculated in order to compare the effect when the lost AIS data set is one of the points from point 2 to point 14. The results are shown in figure 17.

It is shown that when one set of AIS data is missing, the restored trajectory by using ship navigational features still has the best effect among these four methods when considering the mean of the residual errors. The accuracy of the trajectory restored by the piecewise linear interpolation method is the worst.

Although the effects of trajectory restored by the piecewise cubic interpolation method and the piecewise cubic spline interpolation method are good, both of their accuracies are lower than that by using ship navigational features. The proposed restoring method by using the ship navigational features is robust to obtain the trajectory of the inland waterway ship.

It is worth mentioning that the analysis carried for comparisons above is limited to situations where only one lost AIS data point is considered. As for situations with more than one lost data point, further work needs to be conducted. Meanwhile, the ship navigating along the river is always drifting due to wind and current. These two factors are important for safe navigation, so are the manoeuvrability parameters as mentioned previously (such as the steering indices K and T). Further research taking into account these factors/parameters may need to be conducted in order to restore more accurate trajectories. Furthermore, more validating research should be done with a larger sample of AIS data to: (1) develop a real-time traffic monitoring system which can present navigation behaviours of the ship and other details, and (2) apply the proposed method in some collision-avoidance scenarios to further test its application.

5 Conclusions

(1) Because Class-B AIS equipped by inland waterway ships sometimes generates garbled messages, some inaccurate data may be received. The trajectory restored by traditional methods becomes unreliable for seafarers and managers. As a result, the data needs to be cleansed before restoring the trajectory of an inland waterway ship.

(2) Based on both the range of AIS signals received and the manoeuvring characteristics of the inland waterway ship, three rules are built to identify the inaccurate data and cleanse AIS data as the data pre-processing of the trajectory restoration. In the case studies, two inland waterway ships' real AIS data is effectively cleansed by these rules.

(3) Based on navigational features of the inland waterway ship, a novel restoring method is designed to obtain the ship trajectory. In order to demonstrate the proposed method, case studies are conducted where the trajectories are restored using real AIS data and GPS data from an inland waterway ship. Furthermore, a comparison analysis with other three traditional methods is carried out through calculating

the residual errors. These residual errors are the differences between estimated latitude values of the restored trajectory functions and real latitude values of GPS. The results show that the trajectory restored by the proposed method is the best.

(4) Effects of trajectories restored by the new provided method and other three traditional methods are analysed when one AIS data point is lost or removed. The results show that the proposed restoring method using the ship navigational features is robust in obtaining the correct trajectory.

The trajectory restored by the ship navigational features will make the analysis of the ship navigation more effective. This type of ship trajectory will also assist in analysing the pattern and other essential details of the ship traffic flow.

Acknowledgements

This work is supported by the Natural Science Foundation of China (Grant No. 51309187) and FP7 Marie Curie IRSES of European Union (Grant No.314836). Grateful acknowledgement is made to Prof. Xiu-min Chu and Miss Wen-juan Zhang who provided us with considerable help when collecting the data.

References

Ambjorn C. 2008. Seatrack web forecasts and backtracking of oil spills-an efficient tool to find illegal spills using AIS. In: IEEE/OES, US/EU-Baltic International Symposium. Tallinn, Estonia, 1-9.

Cairns WR. 2005. AIS and long range identification & tracking. *Journal of Navigation*, 58(2), 181-189.

Felski A, Jaskolski K. 2012. Information unfitness as factor constraining Automatic Identification System (AIS) application to anti-collision manoeuvring. *Polish Maritime Research*, 19(3), 60-64.

Goerlandt F, Kujala P. 2011. Traffic simulation based ship collision probability modeling. *Reliability Engineering & System Safety*, 96(1), 91-107.

Goerlandt F, Kujala P. 2014. On the reliability and validity of ship-ship collision risk analysis in light of different perspectives on risk. *Safety Science*, 62, 348-365.

Aarsæther GK, Moan T. 2009. Estimating navigation patterns from AIS. *Journal of Navigation*, 62(4), 587-607.

Hansen MG, Jensen TK, Lehn-Schiøler T, Melchild K, Rasmussen FM. 2012. Empirical ship domain based on AIS data. *Journal of Navigation*, 66(6), 931-940.

Harre I. 2000. AIS adding new quality to VTS systems. *Journal of Navigation*, 53(3), 527-539.

Hasberg C, Hensel S, Westenkirchner M, Bach K. 2008. Integrating spline curves in road constraint object tracking. In: 11th IEEE International Conference on Intelligent Transportation Systems. Beijing, China, 1009-1014.

Hu Q, Chen J, Shi C. 2007. Bringing live AIS information on the web sea charts by using Ajax. In: 7th International Conference on Intelligent Transport Systems Telecommunications. Sophia Antipolis, France, 1-5.

International Maritime Organization. 2002. Resolution A.917(22): Guidelines for the on-board operational use of ship-borne Automatic Identification System (AIS).

International Telecommunication Union. 2014. Technical characteristics for an Automatic Identification System using time division multiple access in the VHF maritime mobile frequency band.

Liu C, Shi X. 2009. Study of data fusion of AIS and radar. In: International Conference of Soft Computing and Pattern Recognition. Malacca, Malaysia, 674-677.

Liu Z, Chen R, Ding D, Le L. 2011. GTD model based cubic spline interpolation method for wide-band frequency-and angular-sweep. *Applied Computational Electromagnetics Society Journal*, 26(8), 696-704.

Liu C, Chen X. 2013. Inference of single vessel behaviour with incomplete satellite-based AIS data. *Journal of Navigation*, 66(6), 813-823.

Merrick JRW, Dorp JR Van, Blackford JP, Shawb GL, Harraldb J, Mazzuchib TA. 2003. A traffic density analysis of proposed ferry service expansion in San Francisco Bay using a maritime simulation model. *Reliability Engineering & System Safety*, 81(2), 119-132.

Mou JM, Tak C, Ligteringen H. 2010. Study on collision avoidance in busy waterways by using AIS data. *Ocean Engineering*, 37(5-6), 483-490.

Nomoto R. 1957. On the steering qualities of ships. *International Shipbuilding Progress*, 35(4), 75-82.

Norris A. 2006. Automatic Identification Systems - the Effects of Class B on the Use of Class A Systems. *Journal of Navigation*, 59: 335-347.

Perkovic M, Gucma L, Przywarty M, Gucma M, Petelin S, Vidmar P. 2012. Nautical risk assessment for LNG operations at the port of Koper. *Journal of Mechanical Engineering*, 58(10), 607-613.

Peters DJ, Hammond TR. 2011. Interpolation between AIS reports: Probabilistic inferences over vessel path space. *Journal of Navigation*, 64(4), 595-607.

Qu X, Meng Q, Suyi L. 2011. Ship collision risk assessment for the Singapore Strait. *Accident Analysis & Prevention*, 43(6), 2030-2036.

Dahlquist G, Björk Å. 1974. *Numerical methods.* Prentice Hall, New Jersey.

Smierzchalski R. 1999. Evolutionary trajectory planning of ships in navigation traffic areas. *Journal of Marine Science and Technology*, 4(1), 1-6.

Su CM, Chang KY, Cheng CY. 2012. Fuzzy decision on optimal collision avoidance measures for ships in vessel traffic service. *Journal of Marine Science and Technology*, 20(1), 38-48.

Süli E, Mayers D. 2003. An introduction to numerical analysis. Cambridge University Press, Cambridge.

Sutulo S, Moreira L, Guedes Soares C. 2002. Mathematical models for ship path prediction in manoeuvring simulation systems. *Ocean Engineering*, 29(1), 1-19.

Szlapczynski R, Szlapczynska J. 2012. On evolutionary computing in multi-ship trajectory planning. *Applied Intelligence*, 37(2), 155-174.

Tang C, Shao Z, Tang Q, Pan J, Ji X. 2012. Vessel track distribution algorithm based on AIS. *Journal of Jimei University (Natural Science)*, 17(2), 109-112.

Tsou MC. 2010. Discovering knowledge from AIS database for application in VTS. *Journal of Navigation*, 63(3), 449-469.

Tsou MC, Hsueh CK. 2010. The study of ship collision avoidance route planning by ant colony algorithm. *Journal of Marine Science and Technology*, 18(5), 746-756.

Wang Y, Zhang J, Chen X, Chu X, Yan X. 2013. A spatial-temporal forensic analysis for inland-water ship

collisions using AIS data. *Safety Science*, 57: 187-202.
Weng J, Meng Q, Qu X. 2012. Vessel collision frequency estimation in the Singapore Strait. *Journal of Navigation*, 65(2), 207-221.
Wu XH. 1999. *Operability and Seakeeping of Vessel (2th Edition)*. China Communication Press, Beijing.

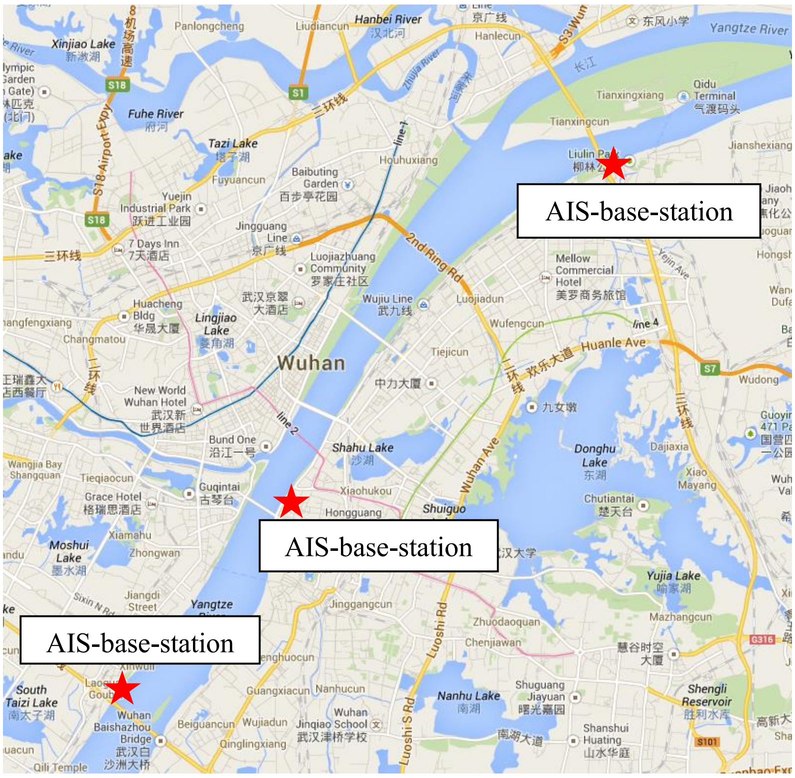

Figure 1 Locations of AIS-base-stations

Figure 2 Ship positions of AIS data

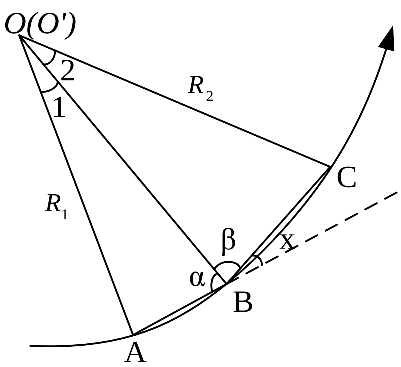

Figure 3 Diagram when ship is altering course

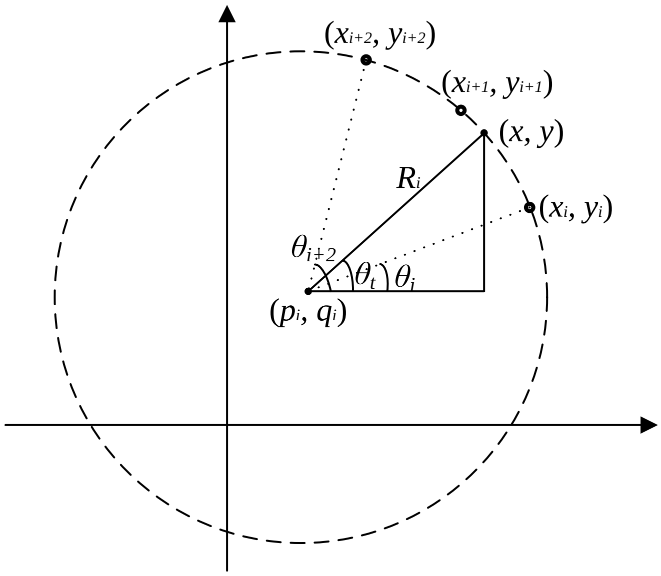

Figure 4 Arc trajectory when the ship is in a steady circle stage

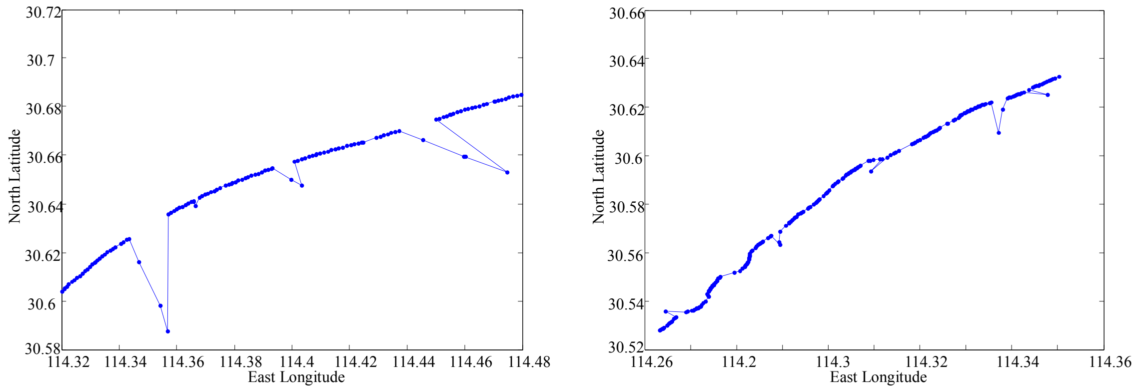

Figure 5 Raw AIS position data (left plot: *pinganda-99*, right plot: *yuxinhuo-11379*) (Unit: degree)

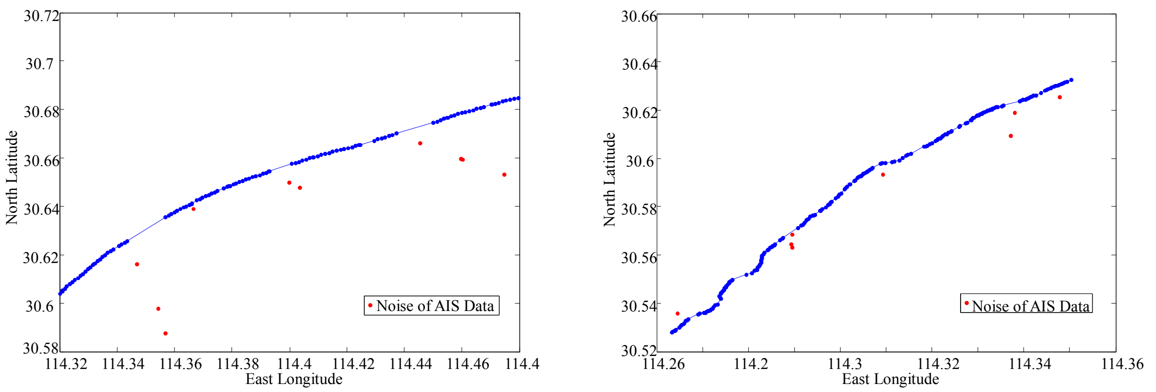

Figure 6 Cleansed AIS position data (left plot: *pinganda-99*, right plot: *yuxinhuo-11379*) (Unit: degree)

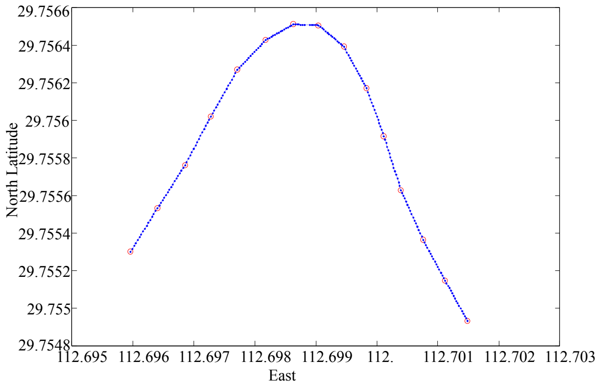

Figure 7 Trajectory restored by the piecewise linear interpolation method (Unit: degree)

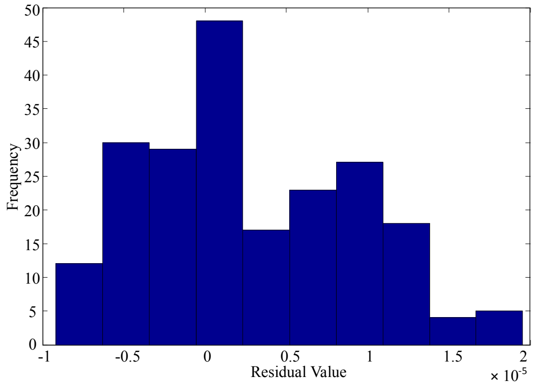

Figure 8 Distribution of the residual errors (The horizontal axis unit: degree)

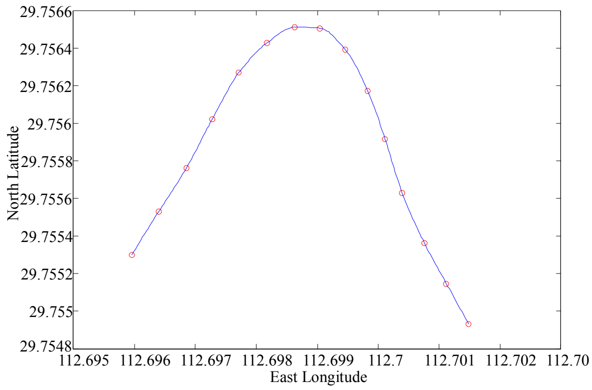

Figure 9 Trajectory restored by the piecewise cubic interpolation method (Unit: degree)

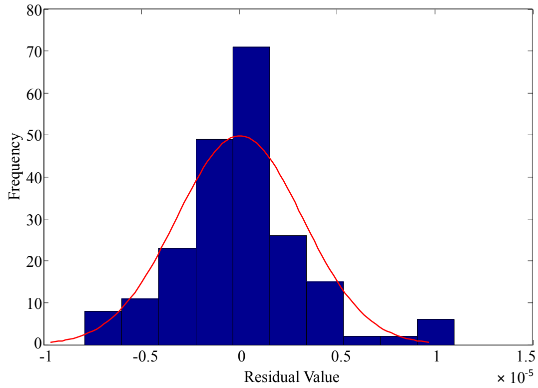

Figure 10 Distribution of the residual errors (The horizontal axis unit: degree)

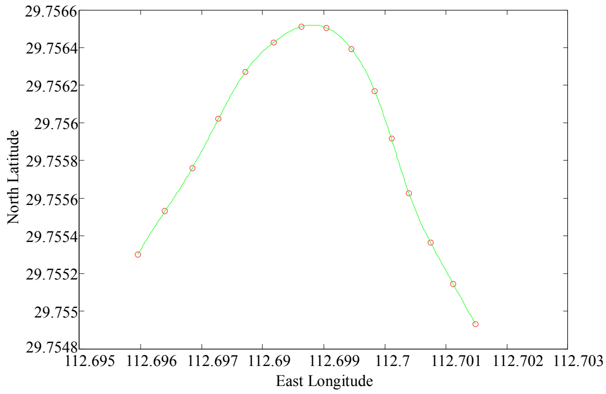

Figure 11 Trajectory restored by the piecewise cubic spline interpolation method (Unit: degree)

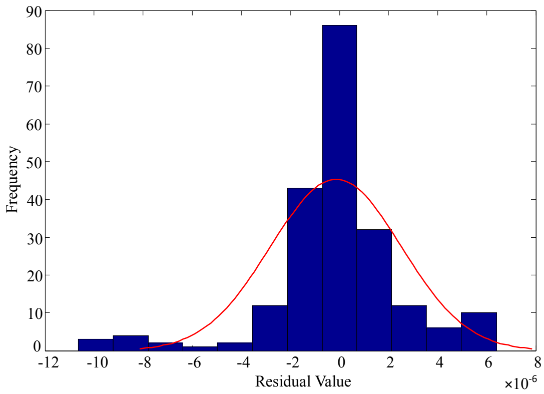

Figure 12 Distribution of the residual errors (The horizontal axis unit: degree)

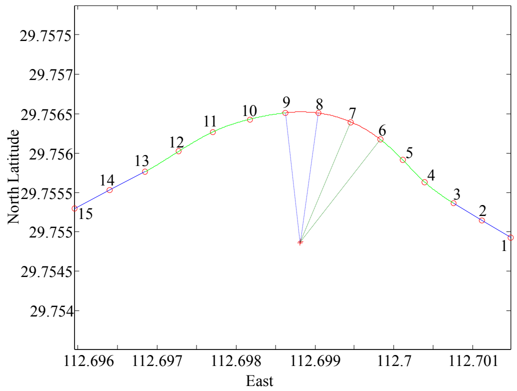

Figure 13 Trajectory restored by using ship navigational features (Unit: degree)

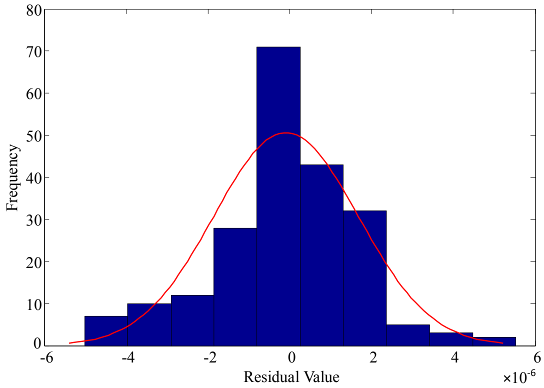

Figure 14 Distribution of the residual errors (The horizontal axis unit: degree)

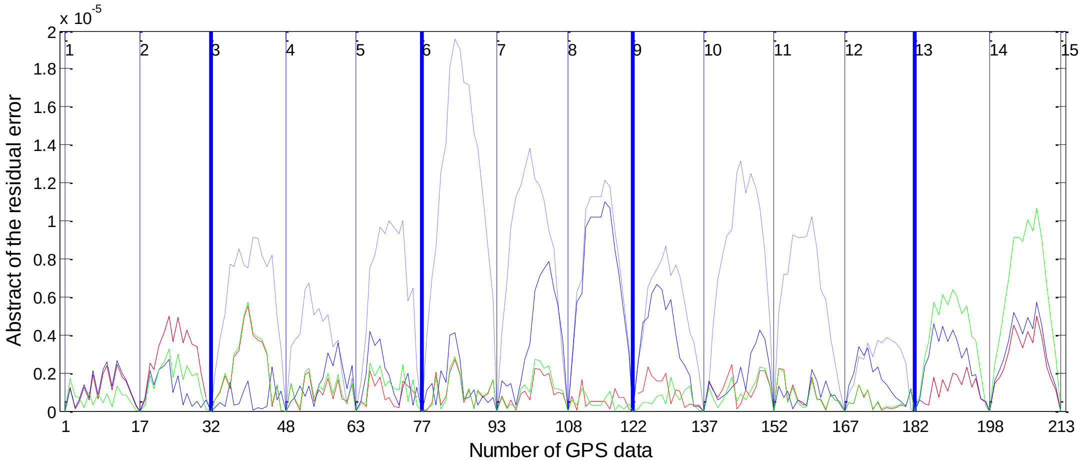

Figure15 Distributions of four sets of the residual errors (The vertical axis unit: degree)

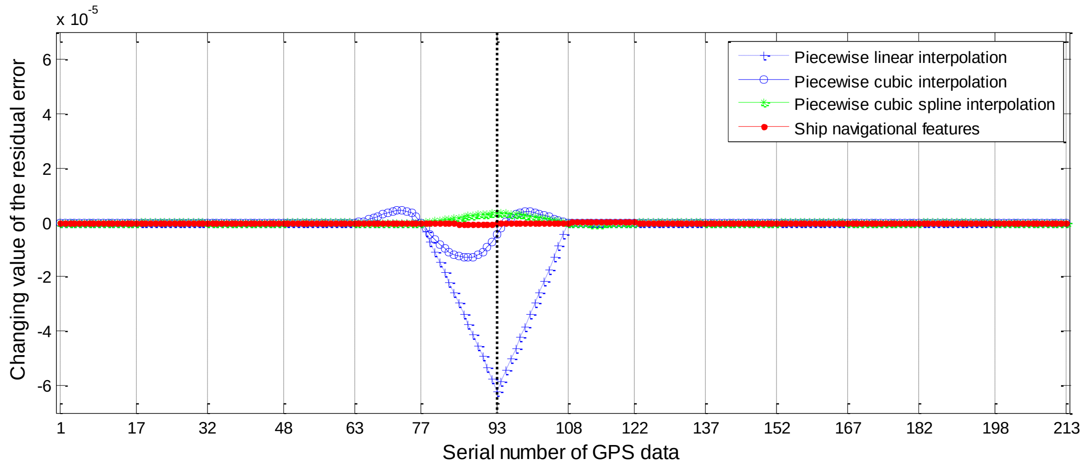

Figure16 Changing values when the seventh set of AIS is removed (The vertical axis unit: degree)

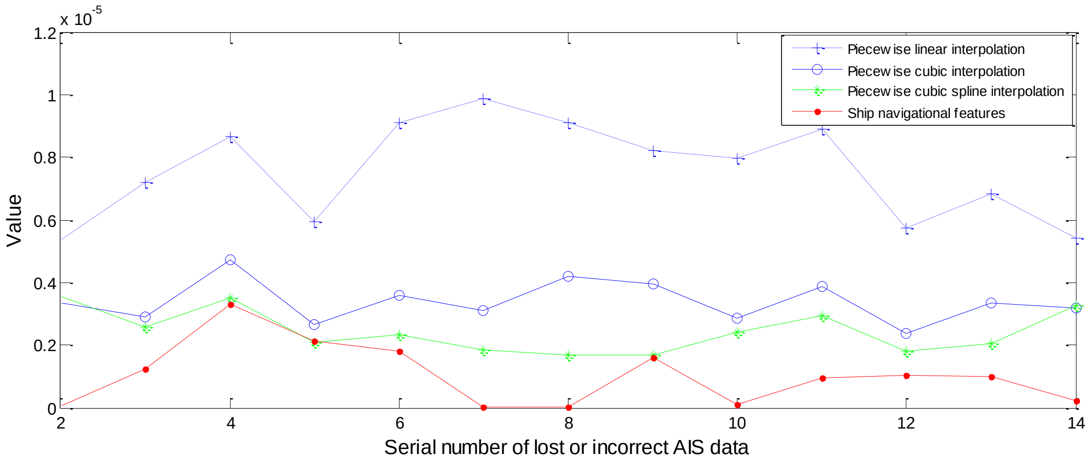

Figure17 Means of the absolute of residual errors for the four different methods of restoration (The vertical axis unit: degree)