Heliyon 10 (2024) e28953

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

# Spatial information extraction of fishing grounds for light purse seine vessels in the Northwest Pacific Ocean based on AIS data

Lijun Wan [a,b], Tianfei Cheng [b,c,1], Wei Fan [b,c], Yongchuang Shi [b,c], Heng Zhang [b,c], Shengmao Zhang [b,c,d], Linlin Yu [b,e], Yang Dai [b,c,d], Shenglong Yang [b,c,*]

[a] School of Navigation and Naval Architecture, Dalian Ocean University, Dalian, 116023, China
[b] Key Laboratory of East China Sea & Oceanic Fishery Resources Exploitation and Utilization, Ministry of Agriculture, East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Shanghai, 200090, China
[c] Key and Open Laboratory of Remote Sensing Information Technology in Fishing Resource, Chinese Academy of Fishery Sciences, Shanghai, 200090, China
[d] Laoshan Laboratory, Qingdao, 266237, China
[e] College of Information, Shanghai Ocean University, Shanghai, 201306, China

ARTICLE INFO

*Keywords:*
Northwest pacific ocean
Light purse seine vessels
AIS
GMM
KDE
HSA

ABSTRACT

Ecological fishery management requires high-precision fishery information to support resource management and marine spatial planning. In this paper, the Automatic Identification System (AIS) was adopted to extract the spatial information on the fishing grounds of light purse seine vessels in the Northwest Pacific Ocean. The spatial distributions of fishing grounds mapped by the data mining, kernel density analysis and hotspot analysis methods were compared. The spatial similarity index was applied to determine the spatial consistency between the computed spatial information and fisheries resource information. Finally, the spatial information derived by the best method was used to investigate the characteristics of fishing activity. The results showed that: the speed of light purse seine vessels related to operations was lower than 1.6 knots. The spatial information extracted by the three methods was consistent with the catch data distribution, and the spatial similarity between the fishing effort and catch data was the highest. The spatial variation in fishing activity was similar to that in the chub mackerel migration route. AIS data could be used to provide high-resolution fishery information. Light purse seine fishing vessels typically operate and travel along the exclusive economic zone boundary, and increased attention must be given to fishing vessel operation supervision. A comprehensive supervision system can be employed to monitor the operations of fishing vessels more effectively. The results of this study can provide technical support for the management of fishing activities and conservation of marine resources in this region using AIS data.

## 1. Introduction

Investigating and understanding the distribution of fishing grounds are crucial for promoting sustainable fishery management and

* Corresponding author: Key Laboratory of East China Sea & Oceanic Fishery Resources Exploitation and Utilization, Ministry of Agriculture, East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Shanghai, 200090, China.
*E-mail addresses:* fanwee@126.com (W. Fan), ysl6782195@126.com (S. Yang).
[1] The author contributed to the work equally and should be regarded as co-first authors.

https://doi.org/10.1016/j.heliyon.2024.e28953
Received 29 September 2023; Received in revised form 19 March 2024; Accepted 27 March 2024
Available online 1 April 2024

supporting the long-term health of both marine ecosystems and human communities. Commercial fishing data are limited in terms of temporal and spatial scales, are not shared publicly and are uncertain, especially on the high seas [1], which limits the understanding and management of fishery resources. The public trajectory information provided by AIS data can be used to mine high-precision fishing information, which has been applied to create high-precision fishing intensity maps of fishing grounds, evaluate fishing pressure [2], and determine the relationship between the spatial distribution of fishing activity and related marine factors [3,4,5,6,7]. Compared with catch data, AIS has the advantages of high spatiotemporal accuracy, public access and real-time performance.

High-resolution fishing activity information can indirectly reflect the spatial distribution of fishery resources and fishing grounds [8]. The aggregation ranges of the operation positions and fishing intensity of vessels are usually defined as fishing grounds [9,10,11]. However, defining fishing grounds based on AIS data does not exhibit a unified standard. Fishing areas with the highest catch or highest potential commercial value have also been defined as fishing grounds [12]. The distance attenuation effect of the kernel density estimation (KDE) method can suitably reflect the distance attenuation effect of the fishing activity point density and the difference in the spatial position, which has been used to calculate fishing grounds [13,7]. KDE method does not define the exceedance value of the kernel density value for identifying true hotspots, which can easily cause problems due to the subjective determination of a threshold for the fishing activity aggregation area. As a local spatial statistical analysis method, hotspot spatial analysis (HSA) provides a clear threshold for the identification of fishing activity hotspots [14]. However, the density change at the grid connection of the analysis results is significant, which can easily lead to spatial discontinuity in geographic phenomena [15]. Chen combined the KDE and HSA methods (KDHSA) by introducing geographic units with kernel density values into the has approach to identify hotspots and cold spots of fishing activities with statistical significance in space, extracting hotspots of fishing activities, and generating high-precision fishing ground maps [14].

The utilization of AIS data mining for fishing vessel operation information, as an alternative approach to fishery resource management, necessitates a close spatial correspondence with the spatial distribution of catch data. However, the above studies focused only on the boundary and gathering blocks of maps and the operational efficiency of the method [13,14], neglecting that the spatial information calculated based on AIS points should be consistent in terms of the spatial distribution of fishery data. Which spatial information mining method achieves better spatial agreement with fishery data should be examined and could provide a systematic basis for the use of AIS technology in fishery management [16,17].

The Northwest Pacific Ocean is one of the most important fishing grounds worldwide. The light purse seine fishery sector, which emerged in 2013, mainly targets Japanese sardine (*Sardinops melanostictus*) and chub mackerel (*Scomber japonicus*), and incidentally targets Pacific sardine (*Sardinops sagax*), neon flying squid (*Ommastrephes bartrami*), and saury (*Cololabis saira*) [18,19]. The Northwest Pacific Ocean fishing ground is crucial for the economy and food security of coastal countries and significantly impacts the sustainable development of global fish resources and marine ecosystems [2]. Several studies have focused on single-species fishing grounds for Japanese sardine (*Sardinops melanostictus*) [20] and chub mackerel (*Scomber japonicus*) [21]. However, there has been no research on fishing vessel behaviour and fishing pressure in this region. Understanding the spatial distribution of light purse seine vessel activity, can not only helps researchers better understand fishing pressure and the distribution of fishery resources, but also support more effective fishery management, ensuring sustainable utilization of resources.

In this paper, we adopted trajectory information for light purse seine vessels in the Northwest Pacific Ocean considering the Chinese mainland vessels from 2020 to 2022. Various methods were applied to mine spatial distribution information on fishing activity. The spatial similarity index was calculated between the fishing intensity, kernel density, and hotspot spatial information and catch data from the same period. The results of three methods were comprehensively compared using multiple indicators. The spatial characteristics were analysed based on the optimal method. These results could provide a systematic basis for the application of AIS technology in fishery management and analysis.

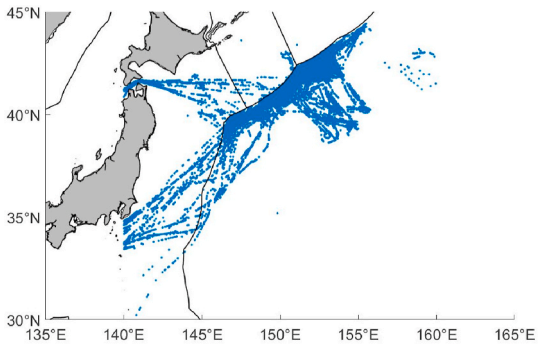


**Fig. 1.** 2020 light purse seine vessel track map.

## 2. Materials and methods

### 2.1. Data and processing

#### 2.1.1. Data sources

AIS data were obtained from the SPIRE satellite. Information on light purse seine vessels in the Northwest Pacific Ocean was collected from the Light Purse Seine Vessels Science and Technology Group of the Chinese mainland. According to the MMSI number, all trajectory data of fishing vessels from 2020 to 2022 were extracted, and each track point included the date, longitude, latitude, heading, and speed. Finally, a total of 45,38 and 37 light purse seine vessels with track data in the Northwest Pacific Ocean were selected for each year.

#### 2.1.2. Data processing

The trajectory data of the selected fishing vessels were preprocessed based on the MMSI number. First, the data with fewer than 100 track points were removed. Then, any speed outliers were eliminated, and only data with speeds between 0 and 15 knots were retained. The data were chronologically sorted, and any duplicate track data were removed. After applying these filters, a total of 14,891, 31,094, and 367,774 track points could be obtained in 2020, 2021, and 2022, respectively. Subsequently, the brightness of each track point was calculated using the 'calcSol' function of the solaR package. A track point map for 2020 is shown in Fig. 1. According to Fig. 1, the 145°E–160°E–35°N-45°N region was selected as the study area in this paper.

### 2.2. Methods

#### 2.2.1. Technical workflow

The study in this paper was based on AIS data for the Chinese mainland light purse seine vessels in the Northwestern Pacific and focused on extracting information on fishing grounds. The process is shown in Fig. 2. ① The raw AIS data were preprocessed to remove any outliers. ② Fishing vessel fishing behaviour was identified, and the status of the vessels during fishing was determined, along with extracting of fishing track points. ③ Three methods (the data mining, KDE, and KDHSA methods) were used to extract fishing grounds, and the results were compared and analysed. ④ The extraction results of the three methods were compared with a fishing distribution map derived from fisheries data for spatial similarity analysis. ⑤ The best method was determined, and spatiotemporal feature analysis of the extracted fishing ground information was conducted using the optimal method.

#### 2.2.2. Operation behaviour of light purse seine vessels

Light purse seine vessels primarily target small to medium-sized fish in the upper and middle water layers. They employ a specific

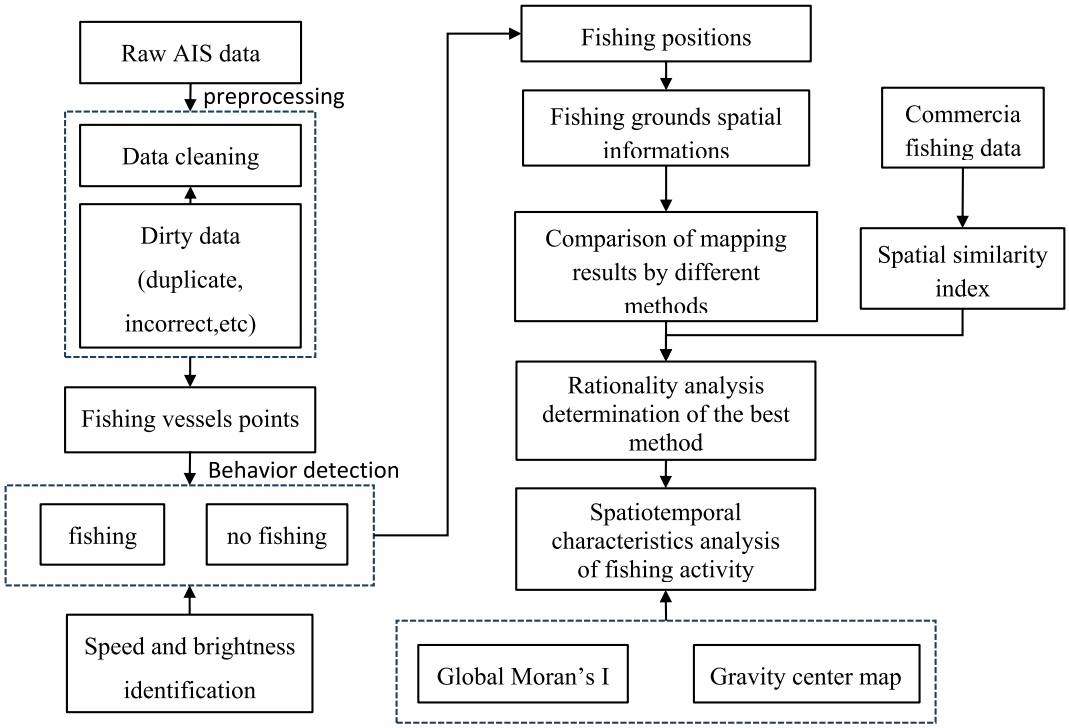


**Fig. 2.** Technical workflow.

fishing technique, namely, drifting light lure, which capitalizes on the nocturnal phototactic behavior of fish. This method involves the using of lights at night to attract and surround fish with fishing vessels, allowing quick capture. During fishing, the vessels maintain a low speed. During the day, the fishing vessels either remain stationary or transit to a different fishing area at high speed. As a result, light purse seine vessels typically search for fish stocks or rest during the day, while their fishing operations are conducted only at night.

### 2.2.3. *Identification of the fishing vessel operation status*

Based on the fishing behaviour, light purse seine vessels generally operate at night at low speeds. The speed and brightness are two crucial characteristics used to identify the fishing status of these vessels. In this paper, a data mining model was developed to determine the operational status of fishing vessels.

The speed threshold was calculated using a Gaussian mixture model (GMM) [22]. The calculation equation is as formula 1:

$$G(x) = \sum_{i=1}^{I} \pi_i N(x|\mu_i, \omega_i) \tag{1}$$

where $N(x|\mu_i, \omega_i)$ is the *i-th* component of the mixture model; $\mu_i$ and $\omega_i$ denote the mean and covariance, respectively, of the *i-th* component; and $\pi_i$ is the mixing coefficient.

The time status can be defined as nighttime when the brightness is zero, and as daytime when the brightness is higher than zero [23]. The fishing status is denoted as 1, while the nonfishing status is denoted as 0, which can be calculated as formula 2:

$$P = \begin{cases} 1, if\ Bo0 = 0\ and\ v_1 < v < v_2 \\ 0, otherwise \end{cases} \tag{2}$$

where Bo0 is the brightness value; v is the fishing speed; and $v_1$ and $v_2$ denote the lower and upper limits, respectively, of the speed threshold.

### 2.2.4. *Fishing effort and fishing intensity calculation*

Based on fishing detection, the fishing effort and fishing intensity were calculated.

In this paper, fishing effort was defined as the amount of time that light purse seine fishing vessels engage in fishing operations at night, which reflects the level of effort invested by these vessels in catch acquisition.

Each AIS point was assigned a timestamp. Except for the first and last points of a trajectory series, the time of each AIS point was defined as half of the sum of the time differences between this point and the front and after trajectory points. The first and last track points determine the time difference between adjacent points, referred to as the duration [2,24], The process is shown in Fig. 3, which

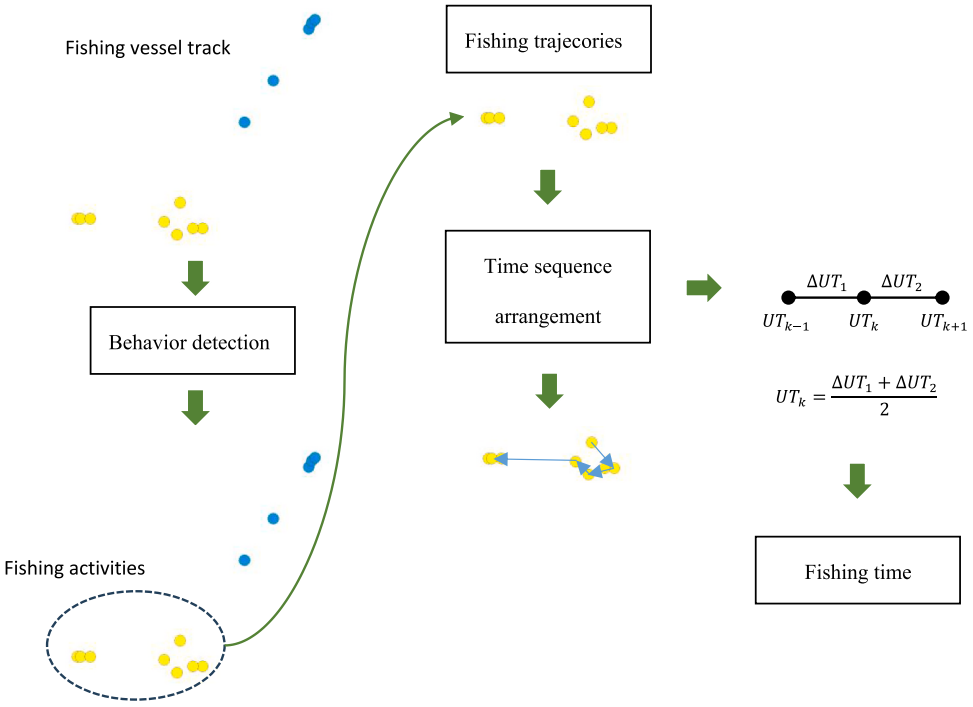


**Fig. 3.** Calculation process for the fishing time of a fishing vessel.

can be expressed as formula 3:

$$T_k = \begin{cases} UT_{k+1} - UT_k & k = 1 \\ \dfrac{(UT_{k+1} - UT_k) + (UT_k - UT_{k-1})}{2} & 1 < k < n \\ UT_k - UT_{k-1} & k = n \end{cases} \tag{3}$$

where $T_k$ is the duration of the k-th point in the extracted fishing vessel track point sequence, and $UT_k$ is the update date of the k-th track point in the extracted fishing vessel track point sequence. Due to the artificial shutdown and signal loss of AIS, segments with time intervals between adjacent track points greater than 10 h were divided into two segments for separate calculations.

The fishing intensity was defined as the fishing effort per unit area and unit time. A monthly temporal resolution was used in this paper, and the spatial resolution was $0.1^\circ \times 0.1^\circ$. The fishing intensity can be calculated as formula 4:

$$E_{lon,lat} = \sum T_{lon,lat} \tag{4}$$

where lon and lat are the longitude and latitude grid coordinates, respectively, after rasterization, and $T_{lon,lat}$ is the cumulative fishing effort of the track points of the fishing vessel in the grid.

To better understand the spatial variation in the fishing effort of light purse seine vessels in each month, the centre of gravity of fishing effort was calculated as formula 5 [25]:

$$FX = \frac{\sum_{i=1}^{n}(F_i \times X_i)}{\sum_{i=1}^{n} F_i} \quad FY = \frac{\sum_{i=1}^{n}(F_i \times Y_i)}{\sum_{i=1}^{n} F_i} \tag{5}$$

where FX and FY are the longitude and latitude, respectively of the centre of gravity of the fishing effort in each month; $X_i$ is the longitude of the i-th grid; $Y_i$ is the latitude of the i-th grid; n is the total number of grids; and $F_i$ is the fishing effort in the i-th grid.

### 2.2.5. Kernel density estimation

Kernel density estimation (KDE) is a nonparametric method for estimating the probability density function of a random variable based on kernels as weights. This method can be used to calculate the density of features in their surrounding neighbourhoods [26,6]. The KDE method is closely related to the first law of geography, and its analysis results indicate that the closer the objects are, the higher their correlation. The KDE calculation can be expressed as formula 6:

$$f(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{1}{h}(x - x_i)\right) \tag{6}$$

where $K$ is a kernel function; h is the bandwidth; n is the number of known points within the bandwidth range; and d is the dimension of the data. In this article, the kernel density tool in ArcGIS software was used to calculate the fishing activity intensity and decay with distance. The spatial resolution is $0.01^\circ \times 0.01^\circ$ and the count value for the population field, which is the number of times each point appears, was selected. The Search radius parameter (bandwidth) can be calculated as formula 7:

$$SearchRadius = 0.9 * \min\left(SD, \sqrt{\frac{1}{\ln(2)}} * D_m\right) * n^{-0.2} \tag{7}$$

where $D_m$ is the (weighted) median distance of the average centre, and n is the sum of the values in the population field. Moreover, SD is the standard distance.

### 2.2.6. Hotspot extraction by the KDHSA method

In this paper, a method was employed that combines the kernel density estimation (KDE) and hotspot analysis (HSA) techniques to extract activity aggregation areas [27,28,14]. Based on the KDE calculation results, the geographic units with kernel density values were introduced into the HSA method to identify statistically significant spatial hot spots in fishing activities and to extract fishing vessel hotspot areas. In terms of the weight coefficient set, the weight coefficient of the KDHSA method selected fixed distance.

### 2.2.7. Global Moran's I

Spatial autocorrelation is characterized by spatial correlations among nearby locations. Spatial autocorrelation is more complex than one-dimensional autocorrelation because spatial correlation is a multidimensional phenomenon. Global Moran's I is a measure of the overall clustering of spatial data. This index can be defined as formula 8:

$$I = \frac{N}{W} \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \omega_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{N} (x_i - \bar{x})^2} \tag{8}$$

where N is the number of spatial units indexed by I and j; x is the variable of interest; $\bar{x}$ is the mean of x; $\omega_{ij}$ denotes the elements of the matrix of spatial weights with zeroes on the diagonal; and W is the sum of all elements $\omega_{ij}$.

#### 2.2.8. Spatial similarity statistics

The fishing grounds calculated from fishery data were compared with the estimated commercial fishing data by computing the similarity statistic I [29,30]. This global metric accumulates the differences between two predictions considered in pairs, producing a single value that reflects how similar the two distributions are based on their variable intensity and spatial distribution. The similarity statistic I varies between 0 (when the two spatial distributions exhibit no overlap) and 1 (when the two spatial distributions are exactly the same).

## 3. Results

### 3.1. Speed characteristics of the fishing vessels

The speed distribution of all the Chinese mainland light purse seine vessels in the Northwest Pacific Ocean based on the GMM [Fig. 4(a-c)] exhibited an obvious bimodal feature. The first peak occurred between 0 and 2 knots, and the second peak occurred between 6 and 8 knots. The frequency of speeds between 9 and 15 knots was extremely low. According to the operation characteristics of the fishing vessels, the above two peaks potentially correspond to the low-speed fishing status and high-speed sailing status, respectively. The mean speed of the first peak is approximately 1.2 knots, and the standard deviation is approximately 0.8 knots.

The speed distributions of all fishing vessels during the day and at night based on the GMM (Fig. 4) also indicated bimodal distributions for both periods. During the nighttime, the two speed peaks occurred between 0 and 2 knots and between 6 and 7 knots, which is similar to the results obtained for all fishing vessels.

Based on the daytime and nighttime fishing vessel speed distribution characteristics, fishing vessels with speeds ranging from 0 to 1.6 (0, mean +standard deviation) knots were defined as engaged in fishing operations.

### 3.2. Analysis of the single fishing vessel fishing behaviour

Fig. 5(a and b) shows the speed and dwell time distribution of the light purse seine vessel with MMSI number 412420956 in September 2020. The speed distribution map (Fig. 5-a) reveals that the vessel track consists of several medium-speed points and a large number of low-speed points. The dwell time distribution map (Fig. 5-b) indicates that the vessel exhibited long dwell times in some spatial areas, while the dwell times at some track points were very short. After travelling at a medium speed to a given area, the vessel would stop for a long time, and would move slowly. These two findings showed that the vessel exhibited long dwell times at low-speed points, while the vessel had very short dwell times at continuous high-speed points.

### 3.3. Monthly variation in the fishing effort

Fig. 6 shows the trend in the fishing effort from 2020 to 2022. According to the statistical results, the fishing effort of light purse seine vessels sharply increased from May to June, and the fishing effort continued to increase from June to July. From July to

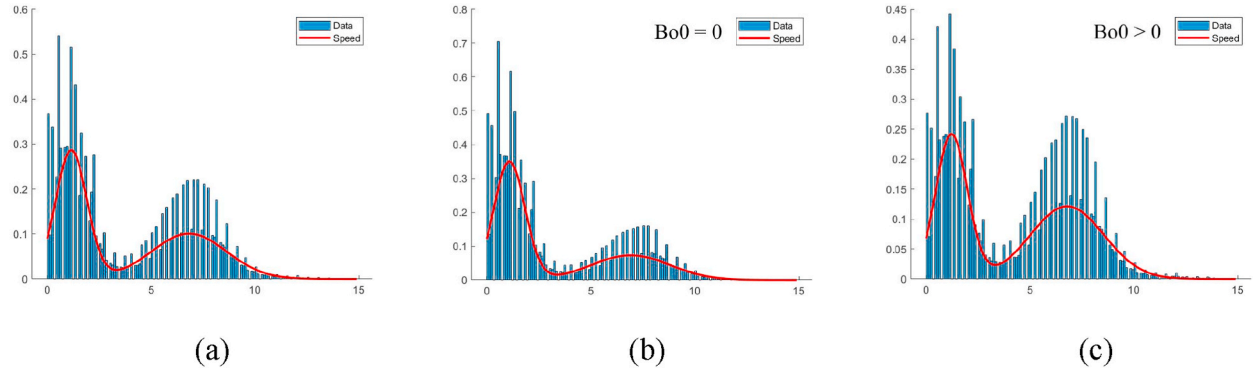


**Fig. 4.** (a) Speed distribution of the fishing vessels based on the GMM. (b) Speed distribution of the fishing vessels based on the GMM when the Bo0 = 0. (c) Speed distribution of the fishing vessels based on the GMM when the Bo0 > 0.

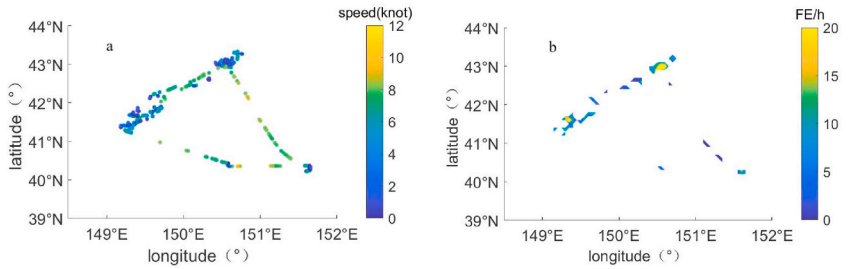


**Fig. 5.** (a) Speed distribution of a single fishing vessel. (b) FE distribution of a single fishing vessel.

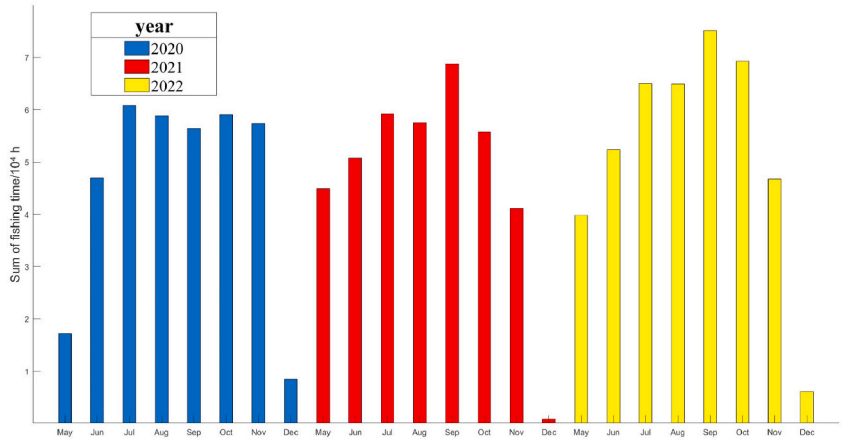


**Fig. 6.** Monthly variation in the fishing effort from May to December.

November, the fishing effort fluctuated, and in December, the fishing effort sharply decreased.

### *3.4. Comparison of the mapping results of the different methods*

Based on the fishing detection results, spatial information on the fishing intensity, kernel density and hotspots from June to November 2020–2022 was mapped to determine the fishing grounds [Fig. 7(a-c)]. The spatial distributions of fishing grounds obtained using the three different methods were similar. Fishing activity was mainly located at 149°E−154°E, 40°N-43°N, both outside and along the exclusive economic zone boundary, indicating an obvious northwestward distribution. There were several scattered fishing regions located outside of this area. The spatial distributions of light purse seine fishing activity mapped by the three methods in this paper also differed [Fig. 7(a-c)]. In terms of the spatial distribution, the high-intensity distribution areas were located at 149°E−153°E, 40°N-43°N [Fig. 7(a)(FE)]. The cumulative fishing time in this area reached up to 100–200 h, while in other areas, it was less than 50 h. The high-value area of the kernel density map was smaller than that of fishing intensity maps, but the hotspot map showed no information on cold spots, and only indicated hotspot regions.

In addition, Fig. 7(a–c) shows that the number of gathering blocks in the fishing intensity map was greater than that in the kernel density and hotspot maps, while the number of gathering blocks in the hotspot map was the smallest. The boundary of the fishing intensity map was sharp, whereas the boundaries of the kernel density and hotspot maps were very smooth. Several scattered regions observed in the fishing intensity map were smoothed out and disappeared in the hotspot map.

The information on fishing grounds extracted using the three methods mentioned above is provided in Table S1. Overall, the distribution patterns of the fishing grounds were consistent, but there were differences in the specific characteristics of the hotspot regions. The data mining method yielded the largest number of hotspots in fishing activity blocks. The number of blocks extracted by the KDE method was similar to that extracted by the KDHSA method, but the block areas were much larger than those extracted by the latter method. Blocks with an area smaller than 10 $km^2$, extracted by all three methods, constituted only a small proportion of the total block area, while the majority of the blocks were larger than 10 $km^2$, indicating concentrated and continuous fishing activity aggregation areas.

Therefore, the fishing intensity map could better show the high-low value distribution, indicating more gathering blocks, although the boundary is sharp.

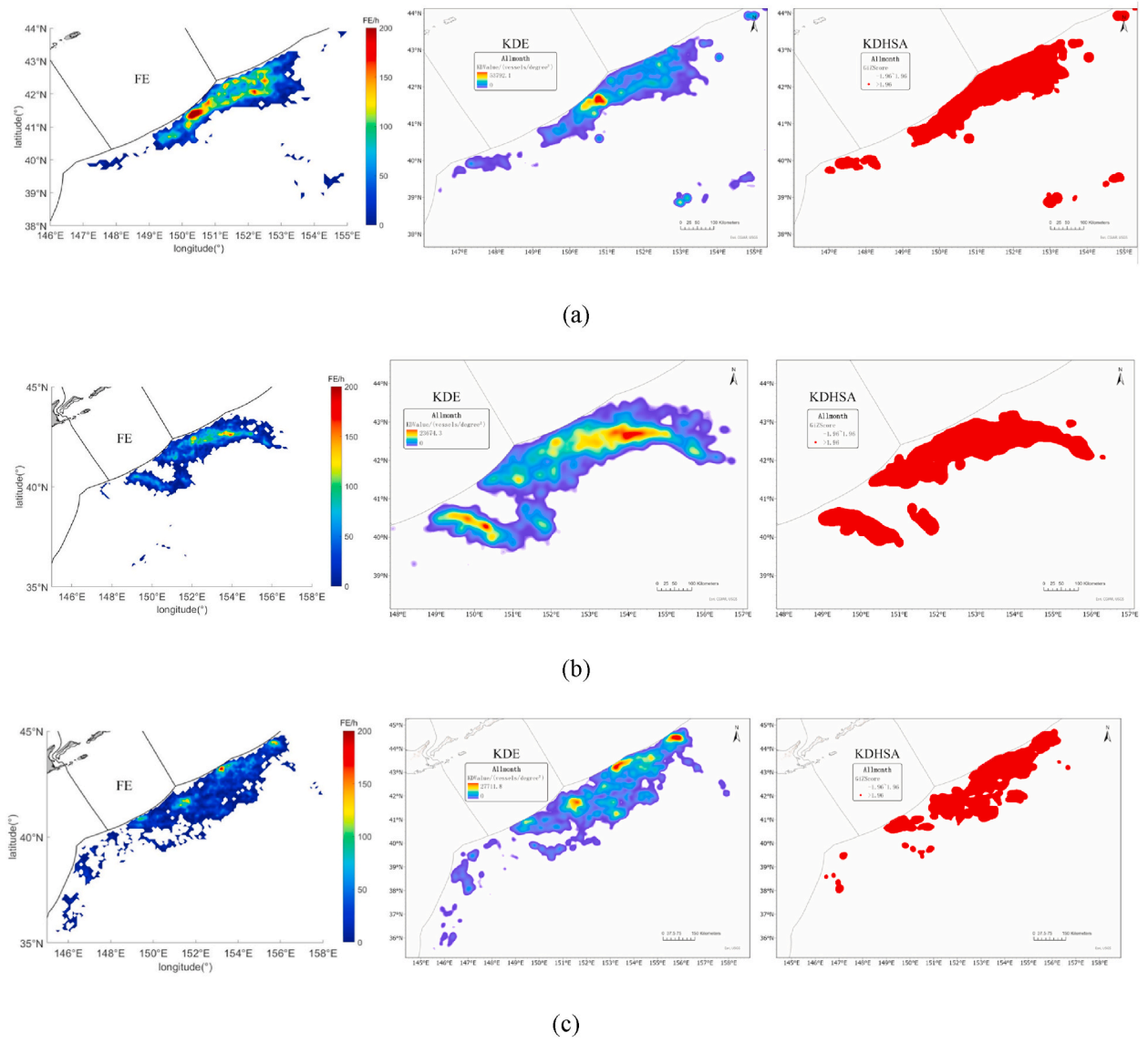


**Fig. 7.** (a) Light purse seine vessel distribution mapping results by FE, KDE and KDHSA during 2020. (b) Light purse seine vessel distribution mapping results by FE, KDE and KDHSA during 2021. (c) Light purse seine vessel distribution mapping results by FE, KDE and KDHSA during 2022.

**Table 1**
Spatial similarity indices of the three methods.

| | | Jun | Jul | Aug | Sep | Oct | Nov |
|---|---|---|---|---|---|---|---|
| 2020 | FE | 0.78 | 0.79 | 0.82 | 0.75 | 0.82 | 0.80 |
| | KDE | 0.78 | 0.78 | 0.80 | 0.66 | 0.79 | 0.79 |
| | KDHSA | 0.77 | 0.77 | 0.79 | 0.66 | 0.78 | 0.79 |
| 2021 | FE | 0.75 | 0.77 | 0.79 | 0.78 | 0.81 | 0.78 |
| | KDE | 0.75 | 0.77 | 0.79 | 0.78 | 0.82 | 0.78 |
| | KDHSA | 0.73 | 0.73 | 0.75 | 0.71 | 0.79 | 0.76 |
| 2022 | FE | 0.70 | 0.79 | 0.81 | 0.73 | 0.66 | 0.70 |
| | KDE | 0.71 | 0.80 | 0.81 | 0.78 | 0.68 | 0.71 |
| | KDHSA | 0.69 | 0.77 | 0.80 | 0.79 | 0.67 | 0.71 |

### 3.5. Spatial similarity index

The spatial distribution of the light purse seine fishery data from 2020 to 2022 is shown in Fig. S1 [31]. Catch data were mainly collected outside and along the periphery of the Japanese exclusive economic zone boundary. Higher catch data occurred closer to the Japanese exclusive economic zone boundary. The farther away from the Japanese exclusive economic zone boundary, the lower the production was. High catch data were located in the area of 149°–153°E, 40.5°-42.5°N.

Table 1 provides the spatial similarity index between the fishing ground information (0.25° × 0.25°) extracted by the three methods and the catch data for each month. Except for September, the spatial similarity index reached approximately 0.8. The spatial distribution of the fishing ground information extracted by the three methods was highly similar to that of the catch data. Among the three methods, the spatial similarity index between the fishing effort and catch data was the highest. The spatial similarity index between the fishing ground information extracted by the KDHSA method and catch data was the lowest.

### 3.6. Spatial distribution characteristics of the fishing activity

Based on the comparisons in Sections 2.4 and 2.5, the fishing intensity method was adopted to map the fishing grounds. The monthly distribution of the fishing effort from 2020 to 2022 is shown in Fig. S2. Overall, the spatial distribution of the fishing activity of light purse seine vessels showed longitudinal and latitudinal variations. The fishing fleets moved northeastward outside and along the exclusive economic zone boundary from June to August, and then returned from September to November.

Although the operational locations were similar, the spatial distribution of fishing vessel operations exhibited different shapes in the same months of the different years reflecting the spatial dynamic changes in fishing vessel operations.

Fig. 8(a–c) shows the distribution of the centre of gravity from June to November during 2020–2022. The centre of gravity shifted northeastward beyond and along the boundary of the Japanese exclusive economic zone from low longitudes and latitudes, reaching its peak in August 2020–2021, and in September 2022. Then, the vessels moved southwestward outside and along the boundary of the Japanese exclusive economic zone. During this period, the centre of gravity occurred closer to the Japanese exclusive economic zone.

### 3.7. Global spatial pattern

The results of spatial autocorrelation analysis of the fishing effort in each year are listed in Tables S2, S3 and S4. There were significant differences in fishing effort during the various months. The least and greatest fishing efforts differed each year. All months exhibited a positive skewness with a positively skewed frequency distribution. A kurtosis value greater than 3, indicates a low and narrow peak, suggesting that high fishing effort values dominated the study area with fewer low-value regions. The coefficient of variation for each month was low, indicating low variation in the fishing effort across the study area. The global Moran's index results showed values greater than 0 for all months, with high Z scores and P values less than 0.01, indicating significant spatial autocorrelation and notably clustered distribution pattern of the fishing effort in each month.

## 4. Discussion

### 4.1. Reliability analysis of the AIS data

We conducted queries regarding the study area using three GFW identifiers for light purse seine vessels. The search criteria included 'purse_seines', 'other_seines', and 'other_purse_seines' within the range of 146°E to 160°E and 35°N to 50°N from May to December. Eight vessels did not originate from the Chinese mainland in 2020, but very few records were available. There were no vessels in the fishing areas in 2021 and 2022. The fishing vessels of the Chinese mainland dominated in terms of absolute numbers in the high seas.

The transmission rate of the AIS is critical for real-time and accurate data delivery. The AIS comprises two main categories of devices, A and B. Class B devices are given lower priority in networks and broadcast weaker signals (2–5 W instead of 12.5 W) than

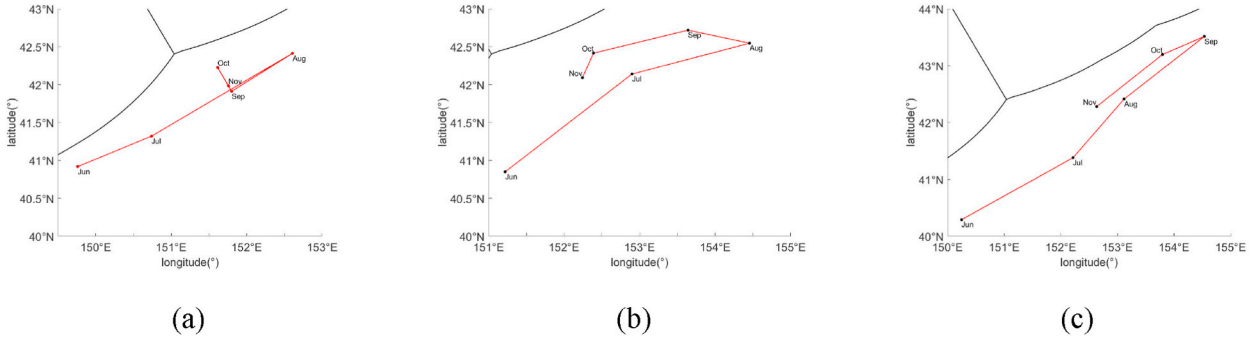


**Fig. 8.** (a) Change in the gravity centre of fishing effort from June to November in 2020. (b) Change in the gravity centre of fishing effort from June to November in 2021. (c) Change in the gravity centre of fishing effort from June to November in 2022.

class A devices. In the fleet of the Chinese mainland, Class B devices are the most common (>99% of the MMSI).

At the data preprocessing stage, the AIS records of six fishing vessels with fewer than 100 records were excluded. Possible reasons for the missing data for these vessels could be AIS device malfunctions or the decision by captains to deactivate the devices, leading to underestimation of the fishing effort [1,11]. Recently, the fishery management authorities of the Chinese mainland have noted that fishing vessels operating in the high seas ensure that their AIS devices are always activated. In the future, AIS information can provide more accurate and comprehensive data for fishery analysis and management.

The spatial distribution of fishing grounds closely aligns with that of Japan's EEZ, indicating the need for increased attention to the management of fishing vessels in this area. To ensure effective fishery management, it is necessary to separately manage fishing vessels engaged in fishing operations in this area, establish a tracking and supervision system, and ensure that fishing vessels conducting fishing operations within this area maintain their AIS devices continuously activated. Furthermore, due to the limitations of the AIS data, such as temporary deactivation of AIS devices by human intervention leading to brief information gaps, it is challenging to distinguish this situation from data loss due to transmission issues. It is necessary to integrate remote sensing imagery data and build a comprehensive supervision system to monitor the operations of fishing vessels more effectively.

### 4.2. Detection of light purse seine fishing behaviour

In existing research, the Gaussian mixture model (GMM) has been used to extract threshold values for identifying fishing vessel operational states, without separating daytime and nighttime information components [13,23]. Compared to traditional methods that rely on the experience of fishermen and histogram-based approaches to determine speed thresholds, the GMM is more systematic and objective. A more accurate speed threshold can be used to extract more precise fishing track points. Typically, fishing vessels engage in either daytime or nighttime operations for specific fishing methods, while very few operate continuously for 24 h. Therefore, neglecting to differentiate between daytime and nighttime track data when identifying fishing behaviour could lead to overestimation of the fishing effort, which could cause ineffective fishery research and management.

### 4.3. Comparison of the spatial information extracted by the three methods

The spatial resolution of the kernel density and hotspot analysis maps in this paper is $0.01^\circ \times 0.01^\circ$, while that of the fishing intensity map is $0.1^\circ \times 0.1^\circ$. At a resolution of $0.01^\circ \times 0.01^\circ$, the fishing intensity map exhibited discrete and small patchy areas as well as blank spaces (Fig. S3). Overly fragmented and incomplete fishery information is indeed detrimental to fishery management and research. Therefore, in this study, a fishing intensity map with a resolution of $0.1^\circ \times 0.1^\circ$ was created. Both the kernel density and hotspot methods, since they produce raster outputs, could yield in smoother representations.

The results are influenced by various methodology-related factors in mapping fishing grounds and issues related to the data [12]. The distribution and movement of the extracted fishing grounds were approximately similar across the three methods. Notably, the spatial similarity index between the fishing ground information extracted by the FE methods and catch data from the actual production operations was the highest, while the KDEHSA method exhibited the lowest similarity. This suggests that the KDHSA method introduces some divergence in spatial characterization relative to the actual production operations, highlighting a slight difference. The spatial distributions of the fishing intensity and production data for each month demonstrated the highest similarity.

Chen et al., (2022) noted that the KDHSA method maximizes the benefits of both the KDE and HSA techniques. This involves combining the distance attenuation effect of the KDE method [13] with the statistical index of the HSA technique [14]. However, in this paper, the hotspot map showed no information on cold spots, and small scattered regions were excessively smoothed and even disappeared. Additionally, the associated spatial similarity index was inferior to that of the other methods. In particular, hotspot maps cannot display the spatial variation in the value of fishing grounds. Therefore, KDHSA method is unsuitable for evaluating fishing ground pressures.

Although the boundary of the fishing intensity map was sharp, it contained the largest gathering blocks, with the best spatial similarity index, and it exhibited spatial distribution characteristics well. The kernel density map also showed the spatial variation in fishing grounds with smooth boundaries. However, overall, its performance was inferior to that of the fishing intensity map.

### 4.4. Spatiotemporal distribution of fishing vessels activity

The results showed that the main fishing areas of the Chinese mainland light purse seine vessels occur near the exclusive economic zone of Japan. This spatial distribution is highly similar to that of the fishery catch data of the Chinese mainland light purse seine vessels. The fishing vessels of the Chinese mainland constitute the primary fishing force in the region, indicating that the calculated spatial information of fishing activity reflects the spatial distribution of fishery resources for major fish species to some extent. The main target species for the Chinese mainland light purse seine vessels in this region are chub mackerel and Pacific sardine [32], with high catch yields concentrated from May to November. The monthly variation in fishing effort (Fig. 6) also indicated that the fishing season extends from May to November, corresponding to the monthly distribution of catch yields.

Chub mackerel is a warm-water marine fish species known for its rapid growth and large individual size. Its notable swimming ability enables it to achieve long-distance migrations each year. In August, chub mackerel gradually migrates northeastward to the area at approximately 170°E, and then begins to migrate southwestward to the area east of 150°E in September [33,34]. The distribution of Pacific sardines was similar to chub mackerel. This finding with the results of the monthly spatial variation in the fishing effort in this study. The spatiotemporal variation in fishing activity is highly consistent with the changes in the fishing grounds of target

species [19]. The above analysis indicates that the extracted spatiotemporal distribution information of fishing vessel activity in this study could indirectly reveal the migratory patterns and spatial distribution characteristics of the major target species in the sea area.

## 5. Conclusion

In this study, AIS data from the Chinese mainland light purse seine fishing vessels in the Northwest Pacific Ocean from May to December 2020 to 2022 were used to extract spatial information on fishing activity. All three methods could effectively extract spatial information on fishing vessel operations. (1) the fishing intensity exhibited the highest spatial consistency with the fishing data and was well-suited for spatial analysis of fishing pressures in light purse seine fishing in the Northwest Pacific Ocean. (2) The Chinese mainland light purse seine fishing vessels operate and travel along the exclusive economic zone boundary. The spatiotemporal variation in the fishing activity of the Chinese mainland light purse seine fishing vessels is similar to the migratory patterns of the target fish species. (3) AIS technology must be integrated with other methods and technologies to enhance fishing vessel operational management and prevent fishing vessels from operating beyond the defined boundaries.

## Funding

This research was funded by the Laoshan Laboratory (LSKJ202201804,LSKJ202201801), the National Key R&D Program of China ( 2023YFD2401303), the Special Funds of Basic Research of Central Public Welfare Institute (2019T09).

## Institutional review board statement

Not applicable.

## Informed consent statement

Not applicable.

## Data availability statement

Publicly available datasets of vessels number were analysed in this study.

## CRediT authorship contribution statement

**Lijun Wan:** Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Tianfei Cheng:** Visualization. **Wei Fan:** Supervision, Project administration. **Yongchuang Shi:** Validation. **Heng Zhang:** Data curation. **Shengmao Zhang:** Validation. **Linlin Yu:** Software, Project administration, Methodology, Formal analysis. **Yang Dai:** Visualization. **Shenglong Yang:** Writing – review & editing, Software, Project administration, Methodology, Formal analysis, Data curation, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors of this research would like to thank all the researchers who contributed to the research process during our study.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e28953.

## References

[1] M. Taconet, D. Kroodsma, J.A. Fernandes, Global Atlas of AIS-based fishing activity — challenges and opportunities, in: Security Research Hub Reports, 2019. https://digitalcommons.fiu.edu/srhreports/iuufishing/iuufishing/84/. (Accessed 6 February 2024).
[2] R. Chen, X. Wu, B. Liu, Y. Wang, Z. Gao, Mapping coastal fishing grounds and assessing the effectiveness of fishery regulation measures with AIS data: a case study of the sea area around the Bohai Strait, China, Ocean Coast Manag. 223 (2022) 106136.
[3] D. Le Guyader, C. Ray, F. Gourmelon, D. Brosset, Defining high-resolution dredge fishing grounds with Automatic Identification System (AIS) data, Aquat. Living Resour. 30 (2017) 39.
[4] F. Natale, M. Gibin, A. Alessandrini, M. Vespe, A. Paulrud, Mapping fishing effort through AIS data, PLoS One 10 (2015) e0130746.

[5] G.O. Crespo, D.C. Dunn, G. Reygondeau, K. Boerder, B. Worm, W. Cheung, D.P. Tittensor, P.N. Halpin, The environmental niche of the global high seas pelagic longline fleet, Sci. Adv. 4 (2018) eaat3681.
[6] M.A. Cimino, M. Anderson, T. Schramek, S. Merrifield, E.J. Terrill, Towards a fishing pressure prediction system for a western pacific EEZ, Sci. Rep. 9 (2019) 461.
[7] Y. Wenhao, A. Tinghua, The visualization and analysis of POI features under network space supported by kernel density estimation, Acta Geod. Cartogr. Sinica 44 (2015) 82.
[8] Y.B. Wang, Y. Wang, Estimating catches with automatic identification system (AIS) data: a case study of single otter trawl in Zhoushan fishing ground, China, Iran, J. Fish. Sci. 15 (2016) 75–90.
[9] N.T. Hintzen, F. Bastardie, D. Beare, G.J. Piet, C. Ulrich, N. Deporte, J. Egekvist, H. Degel, VMStools: open-source software for the processing, analysis and visualisation of fisheries logbook and VMS data, Fish. Res. 115–116 (2012) 31–43.
[10] S. Yang, H. Zhang, W. Fan, H. Shi, Y. Fei, S. Yuan, Behaviour impact analysis of tuna purse seiners in the western and central pacific based on the BRT and GAM models, Front. Mar. Sci. 9 (2022), https://doi.org/10.3389/fmars.2022.881036.
[11] S. Yang, L. Yu, F. Wang, T. Chen, Y. Fei, S. Zhang, W. Fan, The environmental niche of the tuna purse seine fleet in the western and central Pacific Ocean based on different fisheries data, Fishes Sahul 8 (2023) 78.
[12] G.I. Lambert, S. Jennings, J.G. Hiddink, N.T. Hintzen, H. Hinz, M.J. Kaiser, L.G. Murray, Implications of using alternative methods of vessel monitoring system (VMS) data analysis to describe fishing activities and impacts, ICES J. Mar. Sci. 69 (2012) 682–693.
[13] G. Borruso, Network density estimation: a GIS approach for analysing point patterns in a network space, Trans. GIS 12 (2008) 377–402.
[14] Y. Feng, X. Chen, F. Gao, Y. Liu, Impacts of changing scale on Getis-Ord Gi* hotspots of CPUE: a case study of the neon flying squid (Ommastrephes bartramii) in the northwest Pacific Ocean, Acta Oceanol. Sin. 37 (2018) 67–76.
[15] T. Russo, A. Parisi, S. Cataudella, Spatial indicators of fishing pressure: preliminary analyses and possible developments, Ecol. Indicat. 26 (2013) 141–153.
[16] R.A. Campbell, A new spatial framework incorporating uncertain stock and fleet dynamics for estimating fish abundance, Fish Fish. 17 (2016) 56–77.
[17] R.A. Campbell, Constructing stock abundance indices from catch and effort data: some nuts and bolts, Fish. Res. 161 (2015) 109–130.
[18] S.W. Dai, et al., Distribution of resource and environment characteristics of fishing ground of Scomber japonicas in the North Pacific high seas, Marine Fisheries 39 (4) (2017) 372–382.
[19] Z. Guoqing, S.H.I. Yongchuang, F.A.N. Wei, C.U.I. Xuesen, T. Fenghua, Study on main catch composition and fishing ground change of light purse seine in Northwest Pacific, South China Fisheries Science 18 (2022) 33–42.
[20] J. VanDerWal, L. Falconi, S. Januchowski, L. Shoo, SDMTools: Tools for Processing Data Associated with Species Distribution Modelling Exercises, Version, 2012.
[21] D.L. Warren, R.E. Glor, M. Turelli, Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution, Evolution 62 (2008) 2868–2883.
[22] D.A. Kroodsma, J. Mayorga, T. Hochberg, N.A. Miller, K. Boerder, F. Ferretti, A. Wilson, B. Bergman, T.D. White, B.A. Block, P. Woods, B. Sullivan, C. Costello, B. Worm, Tracking the global footprint of fisheries, Science 359 (2018) 904–908.
[23] S.H.I. Huimin, F.A.N. Wei, Z. Han, Y. Shenglong, Spatial Analysis of Fishing Intensity for Illex Argentinus Based on Fishing Vessel Tracks, 南方水产科学. (n.d.) https://doi.org/10.12131/20210010..
[24] Z. Yan, R. He, X. Ruan, H. Yang, Footprints of fishing vessels in Chinese waters based on automatic identification system data, J. Sea Res. 187 (2022) 102255.
[25] P. Lehodey, M. Bertignac, J. Hampton, A. Lewis, J. Picaut, El niño southern oscillation and tuna in the western pacific, Nature 389 (1997) 715–718.
[26] H. Jin, K.Z. Liu, J. Ma, et al., Research on vessel arrival rules based on Gaussian Mixture Model[J], J. Wuhan Univ. Technol. (Transp. Sci. Eng.) 44 (1) (2020) 162–166.
[27] T.D. White, F. Ferretti, D.A. Kroodsma, E.L. Hazen, A.B. Carlisle, K.L. Scales, S.J. Bograd, B.A. Block, Predicted hotspots of overlap between highly migratory fishes and industrial fishing fleets in the northeast Pacific, Sci. Adv. 5 (2019) eaau3761.
[28] W. Yu, T. Ai, M. Yang, J. Liu, Detecting “Hot Spots” of facility POIs based on kernel density estimation and spatial autocorrelation technique, Geomatics Inf. Sci. Wuhan Univ. 41 (2016) 221–227.
[29] J. Lee, A.B. South, S. Jennings, Developing reliable, repeatable, and accessible methods to provide high-resolution estimates of fishing-effort distributions from vessel monitoring system (VMS) data, ICES J. Mar. Sci. 67 (2010) 1260–1271.
[30] S. Jennings, J. Lee, Defining fishing grounds with vessel monitoring system data, ICES J. Mar. Sci. 69 (2011) 51–63.
[31] M. Hinton, M. Maunder, Inter-American Tropical, Methods for standardizing CPUE and how to select among them. https://www.spc.int/DigitalLibrary/Doc/FAME/Meetings/SCTB/16/MWG_7.pdf, 2003.
[32] T.-Y. Hsu, Y. Chang, M.-A. Lee, R.-F. Wu, S.-C. Hsiao, Predicting skipjack tuna fishing grounds in the western and central Pacific Ocean based on high-spatial-temporal-resolution satellite data, Rem. Sens. 13 (2021) 861.
[33] A. Yatsu, T. Watanabe, M. Ishida, H. Sugisaki, L.D. Jacobson, Environmental effects on recruitment and productivity of Japanese sardine Sardinops melanostictus and chub mackerel Scomber japonicus with recommendations for management, Fish. Oceanogr. 14 (2005) 263–278.
[34] W. Yu, A. Guo, Y. Zhang, X. Chen, W. Qian, Y. Li, Climate-induced habitat suitability variations of chub mackerel Scomber japonicus in the East China Sea, Fish. Res. 207 (2018) 63–73.