HiTrip: Historical trajectory interpolation for trawlers via deep learning on multi-source data

Citation: Zhao, Zhongning, Chen, Jiaxuan, Shi, Yuqi, Hong, Feng, Jiang, Guiyuan et al. 2024. "HiTrip: Historical trajectory interpolation for trawlers via deep learning on multi-source data." Ocean Engineering, 292.

As Published: 10.1016/j.oceaneng.2023.116588

Publisher: Elsevier BV

Persistent URL: https://hdl.handle.net/1721.1/156438

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Ocean Engineering 292 (2024) 116588

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Research paper

HiTrip: Historical trajectory interpolation for trawlers via deep learning on multi-source data

Zhongning Zhao [a], Jiaxuan Chen [a], Yuqi Shi [a], Feng Hong [a,*], Guiyuan Jiang [a], Haiguang Huang [b], Jinhua Zhao [c]

[a] College of Information Science and Engineering, Ocean University of China, Qingdao, China
[b] College of Computer Science and Artificial Intelligence, Wenzhou University, Wenzhou, China
[c] Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, USA

ARTICLE INFO

Keywords:
Trajectory interpolation
Deep learning
VMS
Historical fishing effort distribution

ABSTRACT

The Vessel Monitoring System (VMS) on trawlers has revolutionized our understanding of spatiotemporal fishing activities. However, the low temporal resolution of historical VMS datasets complicates a precise analysis of fishing effort distribution. One inherent challenge for precise interpolation is the stark contrast between trawler movement patterns during steaming, characterized by straight lines, and fishing, which often involves consecutive turns. In this study, we introduce HiTrip, a deep learning approach that interpolates historical VMS data from two-hour intervals down to three minutes by harnessing both VMS and marine hydrological datasets. The proposed deep learning model, integrating ResNet, LSTM, and MLP, seamlessly synthesizes spatial features from coarse fishing effort distributions, sea surface factor fields, and current fields, while accounting for the temporal relationships within trajectory segments. Evaluated on 1855 East China Sea trawler VMS records and Copernicus Climate Data Store hydrological factor data, HiTrip achieves a 0.20 km interpolation error, meeting a finery 0.005° × 0.005° spatial resolution demand for fishing effort distribution analysis. Ablation study validates the efficacy of our deep learning model integrating multi-source datasets. Moreover, when evaluated on a diverse Global Fishing Watch dataset, which includes 45 trawlers spanning various global maritime regions, HiTrip maintains a 0.40 km error, emphasizing its broad generalization ability.

1. Introduction

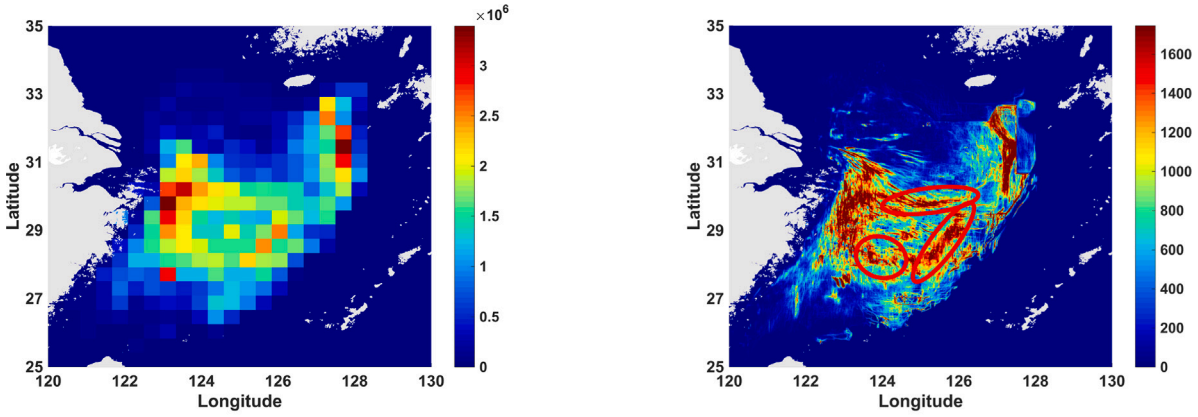
Introduced by the European Union (EU) in 2002, the Vessel Monitoring Systems (VMS) primarily aimed to bolster fishery control and enforcement (Campbell et al., 2014). Over time, the amassed VMS records have illuminated rich spatiotemporal trends of fishing activities (Lee et al., 2010; Zong et al., 2016; Behivoke et al., 2021; Bond et al., 2021). Today, these datasets serve a plethora of purposes. They are instrumental in predicting vessel collision risks to ensure safe navigation (Shaobo et al., 2020; Murray and Perera, 2021), gauging maritime traffic congestion (Zhang et al., 2020; Yang et al., 2022; Qu et al., 2023; Huang et al., 2023), identifying unusual fishing behaviors (Rong et al., 2020; Solano-Carrillo et al., 2021).

One of the most notable applications of VMS datasets is their utilization in evaluating historical fishing effort distribution (Zhao et al., 2021; Li et al., 2022). Such evaluations enable authorities to discern trends in fishery resource consumption and formulate sustainable fishing strategies. Yet, due to equipment constraints, the EU previously mandated a two-hour sampling interval for these datasets (Mills et al., 2007). As a consequence, most VMS datasets are available only in one-hour or two-hour sampling interval (Bertrand et al., 2005; Kourti et al., 2005; Katara and Silva, 2017). This prolonged interval obscures specific trajectory segments between sampled endpoints, thus diminishing the spatial precision of historical fishing effort distribution analysis.

Early VMS datasets typically have long sampling intervals and cannot be directly resampled, leading to an underestimation of historical fishing efforts (Skaar et al., 2011). This underestimation significantly impacts understanding the evolution of fishing effort distribution. Firstly, it may conceal historical environmental degradation or overfishing in specific regions, as these issues might have been undetected due to underreported fishing activities. Secondly, perceived increase in fishing efforts over time might be attributed to shorter sampling intervals rather than actual increases in fishing activities. These issues complicate the task of fisheries management in accurately assessing the historical progression of fishery resources and formulating effective fishing management strategies.

* Corresponding author.
E-mail address: hongfeng@ouc.edu.cn (F. Hong).

https://doi.org/10.1016/j.oceaneng.2023.116588
Received 23 October 2023; Received in revised form 3 December 2023; Accepted 16 December 2023
Available online 20 December 2023



(a) (b)

Fig. 1. Fishing effort distributions calculated from the same VMS dataset but with different spatial resolution and sampling interval. (a) 0.4° × 0.4° and two hours, (b) 0.01° × 0.01° and three minutes.

Table 1
Review of the sampling interval and region.

Source reference	Sampling interval (min)	Region
Salthaug and Johannessen (2006)	60	Norway
Mills et al. (2007)	120	UK
Walter et al. (2007)	120	USA
van et al. (2008)	120	Netherlands
Pedersen et al. (2008)	60	Germany
Mullowney and Dawe (2009)	60	Canada
Russo et al. (2011)	20	Italy
Zhang et al. (2016)	3	China
Lopes et al. (2018)	10	Portugal
Hong et al. (2019)	3	China
Azevedo and Silva (2020)	10	Portugal
Liu et al. (2023)	10	China

For instance, Katara and Silva (2017) revealed that using a two-hour sampling interval resulted in missing 42% of fishing activities when compared to a ten-minute interval. Similarly, Lopes et al. (2018) found that a two-hour interval failed to capture 89% of fishing hauls, leading to a 70% reduction in the estimated fishing effort compared to data collected every ten minutes.

Fig. 1 shows a case study of the fishing effort distributions derived from the East China Sea VMS dataset, albeit with varying spatial resolutions and sampling intervals. The total fishing efforts in Fig. 1(a) and 1(b) stand at 53,163 h and 61,929 h, respectively. A reduced spatial resolution from 0.01° to 0.4°, paired with an extended sampling interval shift from three minutes to two hours, results in a 8,766 h underestimation of the total fishing effort for the latter. Furthermore, certain highlighted fishing hotspots evident in Fig. 1(b) are conspicuously absent in Fig. 1(a). This discrepancy stems from potential fishing activities concealed between the two-hour interval endpoints, which consequently go unrecorded. Such problems are particularly pronounced for trawlers, as most individual fishing operations last less than three hours (Hong et al., 2019).

Table 1 shows the sampling intervals across different historical periods and maritime regions. It reveals that many early VMS datasets, particularly those from fishing vessels in the Atlantic Ocean, had sampling intervals exceeding one or two hours. This was primarily due to EU regulations mandating that VMS records be transmitted every two hours. The prolonged sampling intervals in those VMS datasets contribute to an underestimation of fishing effort, affecting the comprehension of historical fishery resource evolution and limiting the effectiveness of fishing planning. Hence, this study aims to interpolate the sampling interval from two hours to three minutes to achieve a more accurate estimation of historical fishing effort distribution. This upsampling will facilitate a more in-depth insight into historical fishing effort distributions and their implications on benthic ecosystems (Jones, 1992). Moreover, a VMS historical dataset with enhanced temporal granularity is advantageous for numerous applications, including trajectory forecasting and collision risk assessments.

To address the impacts of VMS datasets with long sampling interval on the estimation of fishing effort distribution, trajectory interpolation is often used to decrease the sampling interval. Previous research on interpolating fishing vessel trajectories typically falls into two main categories: mathematical methods and deep learning approaches.

Mathematical interpolation methods such as Linear Interpolation (LI) (Natale et al., 2015), cubic Hermite spline (cHs) (Hintzen et al., 2010), and the Catmull–Rom modification (CRm) (Russo et al., 2011) have been previously utilized for interpolating the trajectories of fishing vessels. These techniques rely on mathematical regression to create smooth interpolated curves. However, their inherent design, which focuses on producing smooth curves, limits their effectiveness, particularly in accurately representing multiple or sharp turns in trajectory segments. This limitation is particularly evident in the interpolation of complex trawler trajectories spanning two-hours segments. These segments often consist of successive turns, a common aspect of fishing activities. For example, Hintzen et al. (2010), Russo et al. (2011) reported that when trawler trajectory segments were interpolated from two-hour intervals to 6-minute and 20-minute intervals, the interpolation errors were 1.18 km and 0.65 km, respectively. Traditional methods lack the refinement needed to precisely capture the specifics of these turns, including their exact positions or angles. Consequently, they tend to introduce errors when interpolating the intricate movements of trawler fishing trajectories.

Deep learning methodologies, with their ability to harness vast amounts of data and their advanced learning capabilities, have been increasingly employed in various studies addressing trajectory issues related to fishing vessels (Zhao et al., 2020; Capobianco et al., 2021; Liu et al., 2021). These approaches have been primarily focused on addressing data loss issues (Liu et al., 2021) and improving the accuracy of trajectory predictions (Guo et al., 2018; Wang et al., 2020). However, there has been less emphasis on the interpolation of historical VMS datasets in existing research.

In our previous research (Teng et al., 2021), we created a deep learning model based on MLP (Multi-Layer Perceptron) and LSTM (Long Short-Term Memory) model to interpolate VMS datasets from a 30-minute to a three-minute interval. This was achieved without incorporating hydrological data. The effectiveness of the proposed model stems from the consistency of trawlers' operational states, either fishing or steaming, within a 30-minute interval. This is in contrast to the usual duration of typical continuous fishing activities, which often

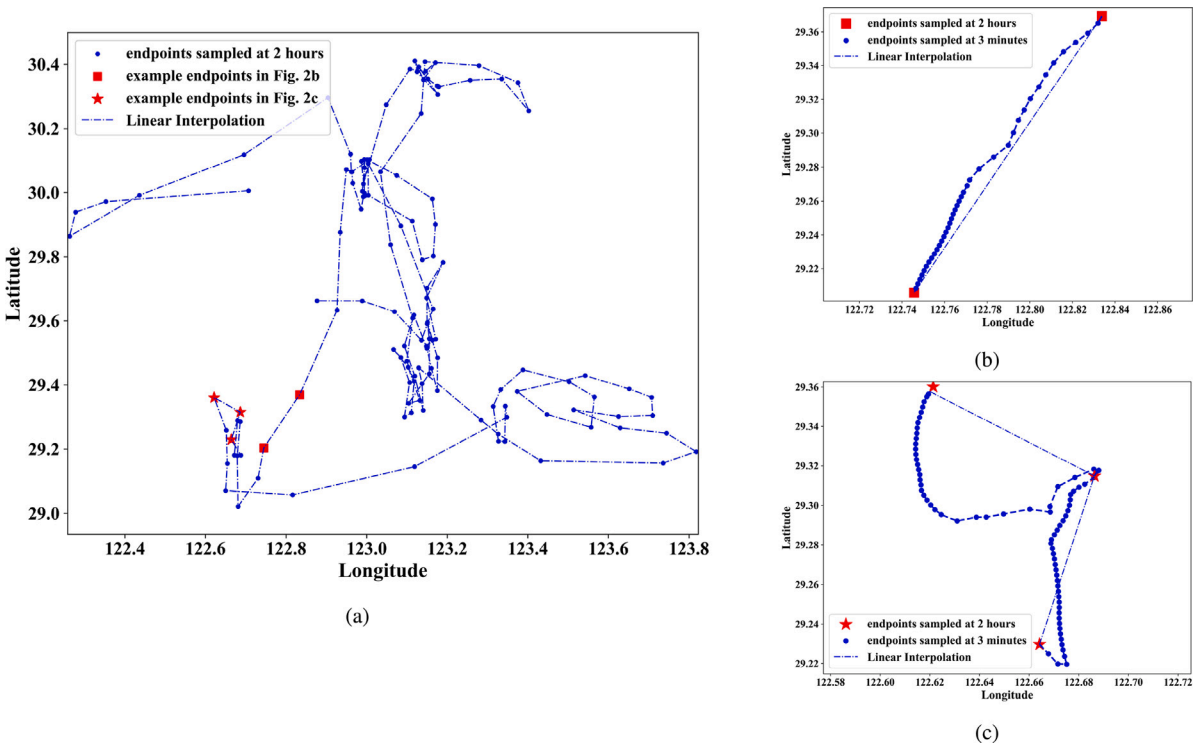


Fig. 2. Visualization of a single voyage with different sampling intervals. (a) single voyage with two hours sampling interval, dot line represents linear interpolation between two endpoints; (b) steaming trajectory with three minutes sampling interval; (c) fishing trajectory with three minutes sampling interval.

lasts around three hours. The proposed learning model is proficient in learning the movement patterns of trawlers within these 30-minute intervals, enabling it to produce interpolation results with a low average error of just 0.09 km.

As for other type of ship, most of the interpolation studies deal with simple trajectories, such as the trajectory of a ferry (Gao et al., 2021) or a cargo (Xue et al., 2017) in an inland river, which does not contain many turns due to the limitation of the river channel or gears. It is important to note that these studies often focus on simpler trajectory patterns of specific vessels, influenced by river channel constraints or the vessels' operational nature. However, trawlers have a large operation range in the ocean, and will conduct multiple turns during fishing. Hence, applying these simpler interpolation methods to trawlers is unlikely to yield accurate reconstructions of their distinct movement patterns, differentiating them from other ship types.

Previous interpolation methods face significant limitations with longer sampling intervals like two hours. Over these extended periods, a trawler trajectory segment might include multiple turns. Solely depending on VMS records in these cases fails to provide a comprehensive insight into the different operational states and turns within the trawler trajectory segment. Consequently, for such longer intervals, these interpolation methods may fall short in accurately representing the intricate movement patterns of the trawler trajectory.

Attaining accurate interpolation from a two-hour interval down to three minutes meets two major challenges. Firstly, there is a pronounced difference in the trajectory patterns observed during steaming versus fishing. The former predominantly manifests as straight lines, whereas the latter often displays a series of consecutive turns. This highlights the importance of distinguishing between these operational states for accurate interpolation. However, the broad sampling interval frequently conceals nuanced movement patterns, making it challenging to identify the operation states of intermediate trajectory segments using only two consecutive VMS records. For instance, Fig. 2(a) showcases a single trawler voyage sampled at a two-hour interval. In contrast, Fig. 2(b) and 2(c) display two trajectory segments extracted from Fig. 2(a), but sampled at a three-minute interval. Notably, even though they are located in close geographical proximity, the segments distinctly represent the diverse movement patterns of trawling and fishing activities.

Secondly, even if the target trajectory segment is correctly recognized as in the fishing activities, it is still a challenge to accurate reconstruct the turn morphology inside the target segment. For instance, the two consecutive trajectory segments in Fig. 2(c) are both within one fishing activities, while their turns are different in direction and curvature.

To address these challenges, we delve into the pivotal factors that play a role in distinguishing operation states and influencing turns: (1) The four continuous VMS records sketch the basic shape of the target trajectory segment. Meanwhile, the velocity and course of these records pinpoint potential turn locations. (2) Coarse fishing effort distribution, even when derived from VMS records with a two-hour sampling rate, sheds light on fishing hotspots. This serves as a direct clue to differentiate the fishing state. (3) Sea surface factors fields can act as another discerning markers. For instance, certain areas become probable fishing hotspots when they exhibit significant variability in the sea surface temperature field compared to their surroundings. Such variability can lead to fish stock aggregation, making these regions more attractive for trawlers' fishing activities (Zhang et al., 2018; Iiyama et al., 2018). (4) Trawlers usually move orthogonal or in opposition to the currents while trawling their nets. So current fields can offer insights into the trawlers' course adjustments during fishing.

Drawing from these insights, we introduce a novel historical trawler trajectory interpolation system, called HiTrip, using deep learning on both VMS and hydrological factor datasets. HiTrip ambitiously seeks to refine the sampling intervals of historical VMS datasets from two-hour spans down to just three minutes. Beyond just harnessing VMS records, our approach integrates coarse fishing effort distributions, sea surface factors fields, and current fields, offering a comprehensive understanding of operation states and inferring possible turns.

HiTrip operates on an encoder–decoder structure. The encoder is tailored to extract features from all crucial factors, whereas the decoder

is designed to produce VMS records at a finer sampling interval. Within the encoder, the Multi-Layer Perceptron (MLP) is leveraged to capture the overall trajectory shape, as represented by the sparsely sampled set of four input VMS records. To assess the contributions of coarse fishing effort distributions, sea surface factors fields, and current fields, we employ three parallel Residual Networks (ResNet) (He et al., 2016), with each network dedicated to a specific data source. Transitioning to the decoder phase, the sophisticated features drawn by the encoder are streamlined and channeled into the Long Short-Term Memory (LSTM) network (Hochreiter and Schmidhuber, 1997) to generate the features for each finery time step corresponding to the output. Finally, these features are honed by an MLP to create a finely up-sampled trajectory segment.

In the case study, we leveraged a VMS dataset acquired from the BeiDou Satellite, covering records from 1855 otter trawlers operating in the East China Sea. This dataset possesses a sampling interval of three minutes. We sourced marine hydrological factor fields from the Copernicus Climate Data Store,[1] offering a spatial–temporal resolution of $0.25° \times 0.25°$, updated on a daily basis. HiTrip showcases a commendable average interpolation error of 0.20 km, meeting the requirements for analyzing fishing effort distribution with a spatial resolution of $0.005° \times 0.005°$. Through meticulous ablation studies, we underscore the benefits of our deep learning model that seamlessly integrates multi-source data for accurate interpolation. Additionally, we evaluated HiTrip's efficacy using a distinct dataset from Global Fishing Watch,[2] covering 45 trawlers across multiple global maritime areas. The evaluations highlight an average distance error of 0.40 km, reinforcing HiTrip's adaptability and effectiveness across varied marine contexts.

2. Materials

This section begins by introducing the three datasets: the VMS dataset sourced from the BeiDou Satellite, the VMS dataset from Global Fishing Watch, and the marine hydrological factor fields. Following this, we delve into the correlation between trajectories and coarse fishing effort distribution, as well as the relationship between trajectories and marine hydrological factors fields.

2.1. VMS dataset in the East China Sea

We use the VMS dataset captured via the China BeiDou Satellite System with a short sampling interval of three minutes. The primary fishing regions of this dataset cover Zhoushan and Yushan fishing grounds, with their external fishing grounds situated within the geographical bounds of $120°E - 130°E$ and $25°N - 35°N$. This dataset contains VMS records from September 2015 to May 2017 for a total of 1855 trawlers.

Each VMS record of a trawler includes five properties: *t*imestamp, *lon*gitude, *lat*itude, *v*elocity, and *c*ourse. To provide a calculable representation for the course of trawlers, we transform the course to its sine and cosine value. Thus, each record can be expressed as $r_t = \{lon_t, lat_t, v_t, c_t^s, c_t^c\}$.

Because the task of HiTrip is to interpolate the sampling interval from two hours to three minutes, we down-sample the original VMS dataset to two hours as the dataset to be interpolated, called low-sampled interval dataset (L). At the same time, the original VMS dataset is employed as the ground-truth dataset for assessing the interpolation accuracy, called high-sampled interval dataset (H).

The fishing cessation period for trawlers in the East China Sea is from June to September. Consequently, both L and H are partitioned into two distinct parts: 2015.09–2016.05 and 2016.09–2017.05. To realize the target of interpolating the historical VMS dataset, we use the VMS dataset from Sep. 2016 to May. 2017 for analysis and interpolation model training while employing the VMS dataset from Sep. 2015 to May. 2016 as the historical VMS dataset to be interpolated.

2.2. Trawler dataset in global maritime regions

To evaluate the generalization of HiTrip, we employ the trawler dataset downloaded from Global Fishing Watch. This dataset comprises 45 trawlers, primarily operating within the Norwegian Sea, Pacific Ocean, and Indian Ocean. The records cover the period from Sep. 2015 to Nov. 2016.

To represent the target of interpolating the historical VMS dataset, we divide the trawler dataset into two equal parts chronologically. The latter half of this dataset is used for transfer learning based on our well-trained interpolation model, while the former half is utilized to assess the generalization ability of HiTrip.

2.3. Marine hydrological factor fields

We source marine hydrological factor fields from the Copernicus Climate Data Store, encompassing both sea surface factors and current fields. These fields have a spatial–temporal resolution of $0.25° \times 0.25°$ and are updated daily. The sea surface factors include sea surface height (SSH), temperature (SST), and salinity (SSS). Current fields, on the other hand, are depicted by a two-dimensional vector representing eastward and northward sea water velocities.

To enhance the model's understanding of the influence of current fields on trawler trajectories, we modify the representation in line with the trawler's course. We transform each record in the current fields for a given date d into three components: sine, cosine, and magnitude, denoted respectively as C_d^s, C_d^c, C_d^m.

Based on the above description, given a date d, the sea surface factors fields and current fields can be represented as $S_d = \{SSH_d, SST_d, SSS_d\}$ and $C_d = \{C_d^s, C_d^c, C_d^m\}$, respectively.

2.4. Coarse fishing effort distribution

From the low-sampled dataset L, we derive the coarse fishing effort distributions. We divide the fishing grounds into spatial cells matching the resolution of the hydrological factor fields, set at $0.25° \times 0.25°$. The distribution is aggregated daily. Leveraging a velocity threshold of 5.5 knots, as established in our prior research (Huang et al., 2019), we filter out specific fishing records. These records, indicative of a two-hour fishing effort, are subsequently aligned with their respective grid and date. Consequently, for any given date d, the coarse fishing effort distribution is denoted as D_d.

2.5. Relationship analysis

For precise interpolation, it is essential to identify the operational state and deduce the turns inherent to fishing behaviors in interpolated trajectory segments. We first examine the correlation between coarse fishing effort distributions and trawler activities. Then, we explore the indirect influence of sea surface factor fields on trawler behaviors. Lastly, we assess how current fields affect the course changes with the trajectory segment, especially during fishing activities.

Firstly, regions with high concentrations in the coarse fishing effort distributions signify intense fishing activities, identifying them as fishing hotspots. Such hotspots can hint at the potential fishing states of individual trajectory segments. When a trawler enters one of these hotspots, it is highly likely to be engaged in fishing. Similarly, if a hotspot lies ahead, the trawler might be inclined to head towards it. Thus, the coarse fishing effort distribution can assist in discerning the probable operational state of a specific trajectory segment.

[1] https://marine.copernicus.eu/

[2] https://globalfishingwatch.org/data-download/datasets/public-training-data-v1

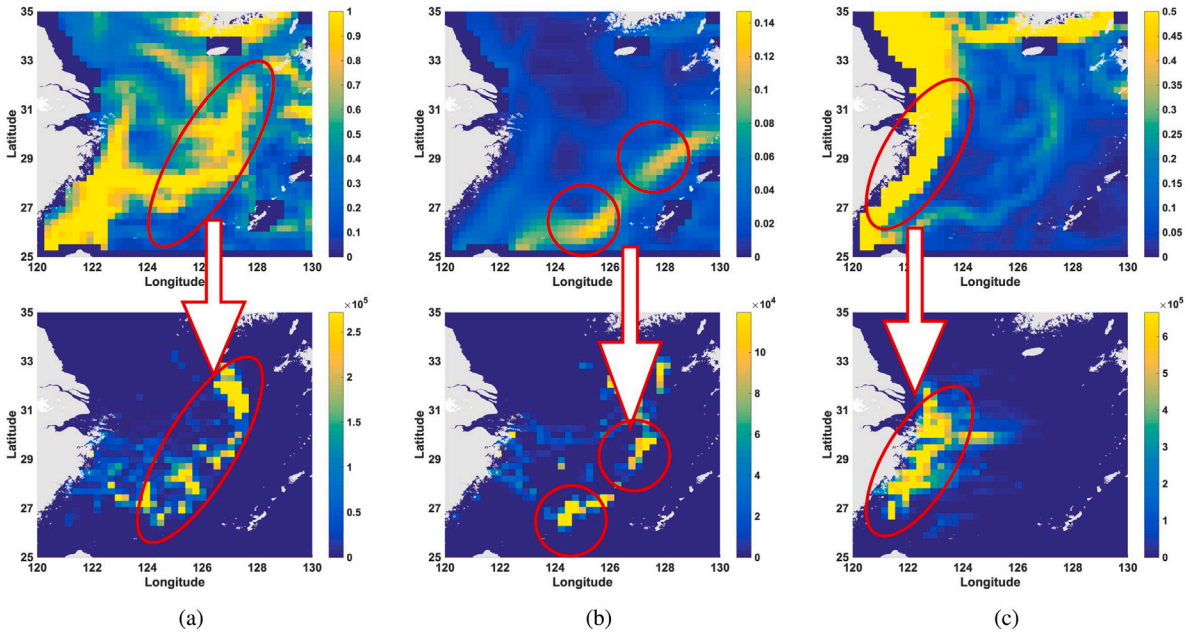


Fig. 3. Spatial variability distributions of SST, SSH, SSS, and the corresponding coarse fishing effort distributions at same date. (a) Mar. 12, 2016, (b) Apr. 30, 2016, (c) Nov. 11, 2015.

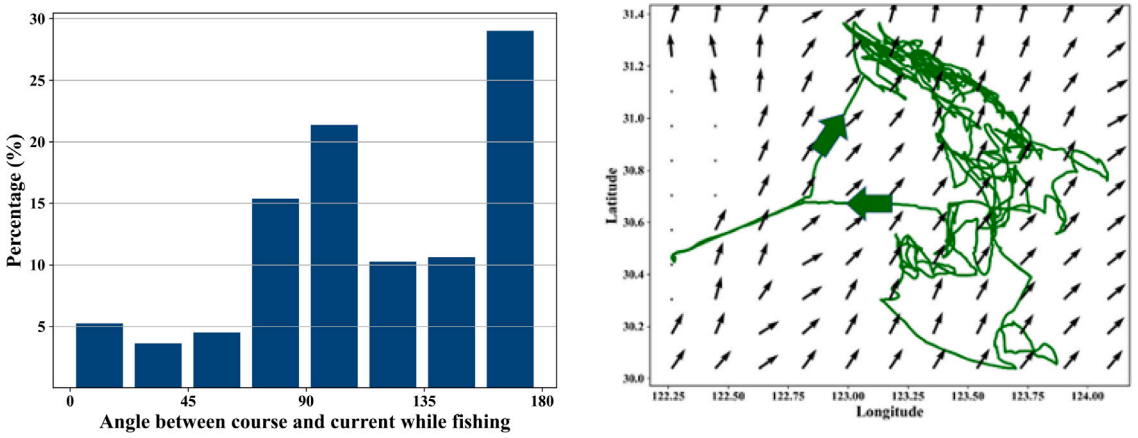


Fig. 4. Effect of current on trajectory. (a) distribution of angles between course and current while fishing. (b) voyage visualization with the background of current, green arrow represents the starting and ending course of the voyage.

Regions with pronounced variability in sea surface factors fields, such as *SST*, relative to their adjacent areas, typically suggest nutrient-rich conditions. Such regions can induce fish aggregation and, subsequently, influence trawler fishing behaviors. By measuring the variability within each grid and juxtaposing it with the fishing hotspots from the coarse fishing effort distributions, we can ascertain the relationship between variability in sea surface factor fields and fishing effort distribution.

For each grid denoted as (x, y), we compute the variability (Var) as the average spatial difference in relation to its adjacent 3×3 grids. Using *SST* as a representative example, this variability is determined according to Eq. (1). Here, the notation $|*|$ stands for the absolute value calculation.

$$Var^{SST}_{(x,y)} = \frac{1}{8} \sum_{x',y'} |SST_t(x, y) - SST_t(x', y')|, \quad where \quad x' \in [x-1, x+1], \; y' \in [y-1, y+1] \tag{1}$$

Fig. 3 depicts the examples of the variability distributions for *SSH*, *SST*, and *SSS*, juxtaposed against their corresponding coarse fishing effort distributions for the same date. Significantly, the areas highlighted in red in Fig. 3 emphasize that fishing efforts intensify in regions exhibiting pronounced variability in sea surface factors fields. Statistical findings reveal that of the grids ranking in the top 20% for the highest fishing efforts, a vast majority – 99% for *SSH*, 78% for *SST*, and 86% for *SSS* – exhibit variability exceeding their respective median values. This observation accentuates the indirect influence of sea surface factor fields on trawler fishing activities, suggesting heightened fishing activities in areas with marked variability.

Thirdly, we then assess how currents influence the course of fine-grained trajectory segments. By calculating the angle between the current and the course of fishing records from *H*, we find that, as depicted in Fig. 4(a), most trawlers during fishing operations maintain a substantial angle (ranging from 90° to 180°) relative to the current. A representative voyage, visualized alongside the current field and shown in Fig. 4(b), reinforces this observation, with most fishing trajectories deviating significantly from the direction of the currents. This suggests that current fields can enhance HiTrip's ability to infer the turn pattern for the target trajectory segments.

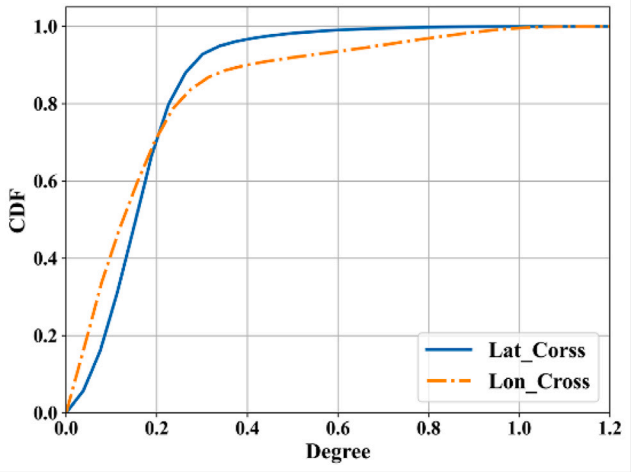


Fig. 5. CDF of latitude and longitude span of each single sample.

From our in-depth analysis of various influencing factors and their relation to trajectory segments, we discern that both coarse fishing effort distributions and sea surface factors fields play pivotal roles in identifying the fishing states of trawlers. Moreover, current fields prove instrumental in inferring turn patterns. Consequently, HiTrip integrates multi-source data, encompassing VMS records, coarse fishing effort distributions, sea surface factors fields, and current fields.

3. Methodology

This section initiates by defining the interpolation problem. Following that, we introduce the encoder–decoder architecture of the interpolation model. Furthermore, we introduce two optimizations for trajectory interpolation. Finally, we provide the training specifics for HiTrip.

3.1. Problem formulation

Using inputs that include two consecutive VMS records r_t^L and r_{t+1}^L, a coarse fishing effort distribution D_t, sea surface factors fields S_t, and a current field C_t, HiTrip will interpolate n records on the trajectory segment of r_t^L and r_{t+1}^L according to the interpolation target. Subsequently, we begin by explaining the construction of each input, followed by the articulation of HiTrip's interpolation problem.

Given that trawlers in the East China Sea typically engage in a single fishing activity lasting less than three hours (Huang et al., 2019), the trajectory segment to be interpolated between r_t^L and r_{t+1}^L may encompass steaming, fishing, or a combination of both operation states. To comprehensively account for the potential entire fishing activity within the target trajectory segment, we enhance the input by including two additional records: r_{t-1}^L and r_{t+2}^L. This expanded input can be represented as $R_t^L = \{r_{t-1}^L, r_t^L, r_{t+1}^L, r_{t+2}^L\}$.

When considering marine factors, which encompass D_t, S_t, and C_t, it is crucial to recognize that within a six-hour duration (the time span of R_t^L), the operational area for trawlers is notably smaller than the entire fishing ground. To supply the model with highly correlated D_t, S_t, and C_t aligned with the trajectory segment to be interpolated, we opt for a subregion from the marine factor fields. This approach empowers the interpolation model to extract precise features influencing the morphology of the target trajectory segment, as opposed to information from unrelated marine factor fields far away.

The computation of the subregion entails two steps: center grid calculation and range determination. Firstly, we calculate the center grid (x, y) of R_t^L using Eq. (2). Here, $r_{t,lon}^L$ and $r_{t,lat}^L$ are the longitude and latitude values of r_t^L, Lon_{min} and Lat_{min} denote the minimum longitude and latitude values within the dataset L, which are $120°E$ and $25°N$, respectively. Additionally, res indicates the resolution of D_t, S_t, and C_t, which is 0.25°. The function $avg(*)$ signifies the calculation of the average value.

$$x = \lfloor \frac{avg(r_{t-1,lon}^L, r_{t,lon}^L, r_{t+1,lon}^L, r_{t+2,lon}^L,) - Lon_{min}}{res} \rfloor$$
$$y = \lfloor \frac{avg(r_{t-1,lat}^L, r_{t,lat}^L, r_{t+1,lat}^L, r_{t+2,lat}^L,) - Lon_{min}}{res} \rfloor \quad (2)$$

Secondly, we calculate the range of movement that trawlers can traverse within six hours to determine the size of subregion. Fig. 5 shows that most trawlers nearly cover an area with a latitude and longitude span of less than 1° within six hours. Taking into account that trawlers may move in various directions, we choose a (2°×2°) subregion to cover the possible movement area, i.e., 9 × 9 grids centered around the coordinates (x, y).

The interpolation problem can be formalized as Eq. (3). Here, θ is the set of all learnable parameters. $\hat{R}_t^H$ is the predicted records from the interpolation model, denoted as $\hat{R}_t^H = \{\hat{r}_{t+1}^H, \hat{r}_{t+2}^H, \ldots, \hat{r}_{t+n-1}^H, \hat{r}_{t+n}^H\}$. The corresponding ground-truth records of $\hat{R}_t^H$ is R_t^H, denoted as $R_t^H = \{r_{t+1}^H, r_{t+2}^H, \ldots, r_{t+n-1}^H, r_{t+n}^H\}$. In this study, the sampling intervals for L and H are two hours and three minutes, respectively. Consequently, the value of n is determined to be 39.

$$\hat{R}_t^H = G_\theta(R_t^L, D_t^{sub}, S_t^{sub}, C_t^{sub}) \quad (3)$$

3.2. Model overview

HiTrip's objective is to derive spatiotemporal patterns from multiple sources and interpolate VMS records from two-hour intervals down to three minutes. We employ an encoder–decoder architecture, where the decoder, equipped with LSTM, guides the encoder in feature extraction, as illustrated in Fig. 6. The decoder's role is to produce the interpolated n records using LSTM, while the encoder concentrates on extracting and organizing spatial features from multiple sources into n time steps for the decoder.

The encoder tackles the challenge of having four input VMS records compared to the number of interpolated records ($n = 39$) by extracting as many relevant features as it can from four different types of input sources: low-sampled VMS records, coarse fishing effort distribution, sea surface factor fields, and current fields. Each of these sources goes through a specific feature extraction network within the encoder, tailored to pull out spatial features for the corresponding target time steps. The decoder, powered by LSTM, then instructs each feature extraction network in the encoder on how to create spatial features and generate interpolated records by learning from temporal patterns. Thus, HiTrip effectively learns the spatiotemporal patterns embedded within the four types of input sources and is capable of performing the interpolation from two-hour intervals to three-minute intervals.

3.3. Encoder

The encoder is tasked with bridging the information gap by pulling out as many pertinent spatial features as possible from the input. This input consists of four successive VMS records taken at two-hour intervals, a distribution of coarse fishing effort, sea surface factors fields, and current fields, all showcasing distinct spatial traits. The encoder efficiently compiles these features to build a basis for the decoder to generate interpolation results. Subsequently, we will detail each component of the encoder, emphasizing how they individually extract features from each data source.

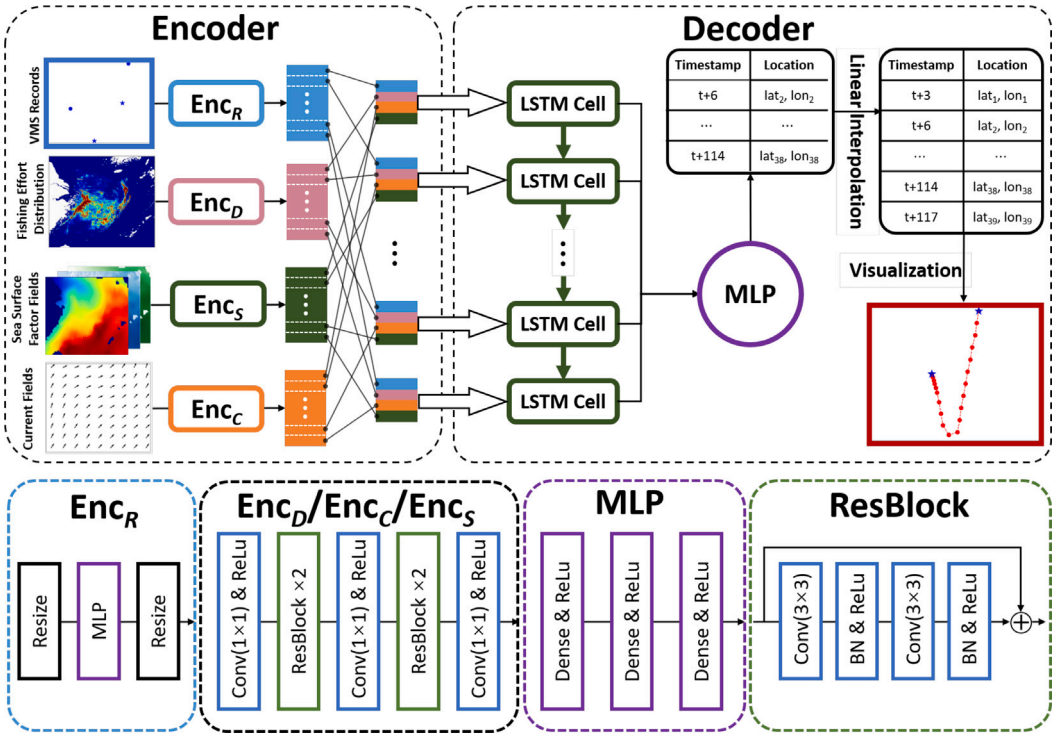


Fig. 6. HiTrip framework. The top two blocks are the encoder–decoder architecture of HiTrip, while bottom four blocks is the module of HiTrip. Enc_R is tasked with feature extraction for VMS records, $Enc_D/Enc_C/Enc_S$ are responsible for capturing spatial features from coarse fishing effort distribution, sea surface factors fields, and current fields, respectively. After extract features from multi-source inputs, Encoder organizes features according to the time steps of the trajectory points to be interpolated. Decoder applies LSTM as its core to integrate temporal features and generates interpolated points by MLP.

3.3.1. VMS records

Given a coarsely-sampled trajectory segment, represented as R_t^L, the objective of HiTrip is to develop an interpolation model that deduces the finer-grained trajectory segment R_t^H, consisting of n records. Considering that R_t^L contains only four records, the spatial features extracted from it should be a multiple of n.

The aim of this module is to extract spatial patterns from the input VMS record sequence and arrange them according to targeted time steps. Each record in the input sequence comprises multi-dimensional parameters, such as latitude, longitude, velocity, and course. The complex and nonlinear relationships among these parameters call for the use of Multi-Layer Perceptron (MLP). MLP is highly effective in discerning and learning these intricate patterns, primarily due to its nonlinear activation functions and the presence of multiple hidden layers. Consequently, we employ a three-layer MLP to extract spatial patterns from the VMS record sequence, resulting in a feature vector $X^R \in \mathbb{R}^{n\times M}$. Here M signifies the number of features in the feature vector for each time step. This process is crucial for bridging the informational gap of the interpolation, ensuring that the spatial patterns of the VMS records are represented and utilized in the subsequent steps of the deep learning model. The network architecture used for extracting features from VMS records is represented in Eq. (4).

$$X^R = Enc_R(R_t^L) \tag{4}$$

3.3.2. Coarse fishing effort distribution

Derived from the coarse-grained VMS dataset R^L, the hotspots in the coarse fishing effort distribution pinpoint areas where many trawlers engage in fishing. The hotspots within the coarse fishing effort distribution are key indicators, which signify areas of intense fishing activities, often indicated by a high density of trawlers' VMS records. An area is identified as a fishing hotspot when its fishing effort values substantially deviate from those of the surrounding regions. These hotspots are crucial as they can indicate potential operational states of the trajectory segments being interpolated. Thus, extracting these hotspots as primary spatial features from the coarse fishing effort distribution becomes essential, providing vital insights into the spatial dynamics of fishing activities.

Convolutional operations in Convolutional Neural Networks (CNN) excel at detecting variations among adjacent grids, and by integrating multiple levels of CNNs, the network gains the ability to recognize variations across grids that are further apart. However, a significant challenge arises when multiple CNN layers are stacked: the gradient vanishing problem. This issue can impede the network's learning process. To overcome this, Residual Networks (ResNet) are employed. ResNet effectively solves the gradient vanishing problem by introducing shortcut connections that link the input directly to the output. These shortcuts facilitate a direct flow of gradients, mitigating the risk of gradient vanishing. Consequently, we utilize ResNet as the foundation of our spatial feature extraction module, capitalizing on its robustness in maintaining gradient flow.

As illustrated in Fig. 6, we initiate by applying a CNN layer (with 1×1 filter size, F filters) in Enc_D to enhance the channel of the provided input D_t^{sub} from $\mathbb{R}^{1\times 9\times 9}$ to $\mathbb{R}^{F\times 9\times 9}$. We then incorporate two Residual Blocks to extract high-level features, maintaining a uniform structure of the features. The configuration of the Residual Block, detailed in the bottom right of Fig. 6, aligns with the recommendations in He et al. (2016). It comprises two convolutional layers, each with a 3×3 filter size and F filters, and incorporates a ReLU function in between to introduce non-linearity. Following the initial pair of Residual Blocks, another CNN layer (with 1×1 filter size, $2F$ filters) is applied, amplifying the output channel of the prior layers from $\mathbb{R}^{F\times 9\times 9}$ to $\mathbb{R}^{2F\times 9\times 9}$.

Then an additional pair of Residual Blocks is utilized, further honing the high-level features while retaining the feature dimensions at $\mathbb{R}^{2F\times 9\times 9}$. Given that the decoder processes the features extracted by the encoder in sequential time steps, the concluding CNN layer (with 1×1 filter size, n filters) is employed to consolidate the feature set across n

time steps. The output of the feature extraction network is represented by $X^D \in \mathbb{R}^{n\times 9\times 9}$. This output is vital for the precise interpolation, ensuring that the spatial dynamics are accurately represented. The whole procedure can be summarized as:

$$X^D = Enc_D(D_t^{sub}) \tag{5}$$

3.3.3. Sea surface factors fields

Besides the coarse fishing effort distribution, both Fig. 3 and the statistical findings in Section 2.5 reveal that sea surface factors fields can also work as an indicator of fishing hotspots. Given that certain areas within the sea surface factors fields become probable fishing hotspots due to their pronounced variability compared to neighboring regions, it is imperative to extract the spatial features from the sea surface factor fields.

Given the need to capture spatial distribution variations, we utilize a network structure analogous to that applied for the coarse fishing effort distribution to extract the spatial features of the sea surface factors field. This feature extraction network Enc_S for sea surface factors fields can be summarized as:

$$X^S = Enc_S(S_t^{sub}) \tag{6}$$

3.3.4. Current fields

The observations from Fig. 4 indicate that trawlers, when engaged in fishing activities, tend to maintain a significant angle relative to the current. This suggests that the current fields can be instrumental in aiding the interpolation model to predict course changes of the trawler. Since R_t^L provides only a rough sketch of the target trajectory segment, it is essential to extract spatial features of the current fields across grids to depict potential course of each interpolated trajectory segment. For this extraction, we employ a network structure mirroring that of Enc_S, resulting in spatial features represented as $X^C \in \mathbb{R}^{n\times 9\times 9}$. The feature extraction process for current fields using the network Enc_C is outlined as:

$$X^C = Enc_C(C_t^{sub}) \tag{7}$$

In summary, from the spatial feature extractions described above, we derive four distinct features: VMS records feature X^R, coarse fishing effort distribution feature X^D, sea surface factors fields feature X^S, and current fields feature X^C. As depicted in Fig. 6, these four feature vectors are utilized by the decoder to produce the fine-grained trajectory segment R_t^H. We merge these four features from the encoder, resulting in the consolidated feature vector $V \in \mathbb{R}^{n\times(M+243)}$, shown in Eq. (8), where *Concat* is the tensor concatenate operation.

$$V = Concat(X^R, X^D, X^S, X^C) \tag{8}$$

3.4. Decoder

The interpolated trajectory in our study is characterized by a sequential pattern, where each point in the trajectory is influenced by its preceding points. The decoder's primary objective is to encapsulate this temporal pattern present in the interpolated trajectory. Long Short-Term Memory (LSTM) Neural Network, a specialized variant of the Recurrent Neural Network, is particularly effective in capturing these temporal dependencies. Therefore, to accurately represent this inherent sequential nature and to effectively organize the spatial features provided by the encoder, we have chosen LSTM as the core component of our decoder. By employing LSTM, we ensure that the temporal dynamics of the trajectory are captured and represented in a manner that reflects the movement and behavior patterns of trawlers.

Furthermore, Bidirectional LSTM (Bi-LSTM) can further enhance this capability by analyzing both past and future context in the data. However, the utility of Bi-LSTM is constrained in scenarios where interpolated points are generated sequentially, and the future context crucial for Bi-LSTM's functionality is unavailable for the points generated earlier. This limitation is particularly pertinent in our case, where interpolated points are produced in a sequential manner, meaning the information needed for subsequent points is not available at the time of generating previous points. Therefore, we choose LSTM as the main framework for our decoder. Its ability to effectively handle the sequential nature of the data and organize spatial features makes it more apt for our specific interpolation challenge than Bi-LSTM.

Table 2
Average distance errors of linear interpolation from different sampling interval to three minutes.

Sampling interval (minutes)	6	12	18	24	30	45	60
Distance error (km)	0.01	0.09	0.32	0.42	0.67	0.89	1.01

The LSTM's time step is configured to n, enabling it to sequentially generate feature vectors of each time step in alignment with the records to be interpolated. Moreover, the LSTM input, represented as $V \in \mathbb{R}^{n\times(M+243)}$, is structured according to these time steps. This arrangement ensures that in every time step, the feature $V_i \in \mathbb{R}^{(M+243)}$ can be fed into the LSTM. Upon generating these feature vectors for each time step, a MLP is further applied to map each feature vector to its final record, encompassing both longitude and latitude coordinates. The process of decoder can be summarized as:

$$\hat{r}_{t+i}^H = MLP(\ LSTM\ (V_i)\ ), \quad i \in [1, n] \tag{9}$$

3.5. Optimization

Upon constructing features into $V \in \mathbb{R}^{n\times(M+243)}$ through the encoder, a pronounced information gap persists between the four sparsely sampled VMS records and the desired 39 VMS records for interpolation. This discrepancy makes it challenging for the deep learning model to yield accurate interpolation outcomes. Moreover, while the LSTM captures the sequential nature of the outputs, the first output record r_{t+1}^H is dependent exclusively on the confined feature vector V_1. This limitation might introduce initial error, which could subsequently cascade and impact the generation of following records. To counter these challenges, we introduce two optimizations, labeled as opt_1 and opt_2.

To mitigate the evident information gap, we lessen the number of records predicted in R_t^H, thereby reducing the burden on deep learning model. Nonetheless, this strategy requires additional interpolation step to reach the intended target sampling interval. The primary focus of opt_1 is to find a suitable target sampling interval for the deep learning model that minimizes the information gap while keeping the interpolation error minimal.

To pinpoint the optimal sampling interval, we assess linear interpolation's capability across various down-sampled intervals transitioning to three minutes. Table 2 shows the interpolation errors encountered when using linear interpolation to convert the VMS data from various sampling intervals to a three-minute interval. Notably, when upsampling from intervals exceeding twelve minutes, the error hovers around the hundred-meter range. However, comparing the errors from twelve-minute and six-minute intervals, there is a nine-fold difference in interpolation error, with the latter showing a mere distance error of approximately 10 m. This significant reduction is attributed to the typical behavior of trawlers; they are unlikely to undertake complex maneuvers within a short six-minute duration, like executing a sharp turn. Hence, we adjust our deep learning model's target sampling rate from three minutes to six minutes and subsequently produce the three-minute record through linear interpolation as our first optimization. The prediction target of the interpolation model changes to $\hat{R}_t^{H'} = \{\hat{r}_{t+1}^{H'}, \hat{r}_{t+2}^{H'}, \ldots, \hat{r}_{t+m}^{H'}\}$, where $m = \frac{n-1}{2}$. After obtaining $\hat{R}_t^{H'}$, we generate the fine-grained $\hat{R}_t^H$ by Eq. (10).

$$\hat{r}_{t+i}^H = \begin{cases} \dfrac{\hat{r}_{t+(i-1)/2}^{H'} + \hat{r}_{t+(i+1)/2}^{H'}}{2} & if\ i\ \textbf{is}\ odd \\ \hat{r}_{t+i/2}^{H'} & if\ i\ \textbf{is}\ even \end{cases} \tag{10}$$

The second optimization is targeted at mitigating the accumulated error resulting from the limited features available for record $\hat{r}^H_{t+1}$. Typically, record $\hat{r}^H_{t+k}$ generated at time step k can harness both the feature V_k and the temporal features extracted by the LSTM from the previous k-1 steps. However, record $\hat{r}^H_{t+1}$ relies solely on V_1. Consequently, record $\hat{r}^H_{t+1}$ has fewer constraints and is susceptible to introducing errors that can ripple through the generation of subsequent records.

To mitigate this issue, we integrate both $\hat{r}^H_t$ and $\hat{r}^H_{t+m+1}$ into the interpolation target, which align with the known starting and ending points of the trajectory segment to be interpolated, specifically r^L_t and r^L_{t+1}. Therefore, the first predicted record $\hat{r}^H_t$ is limited by r^L_t, and the subsequent record $\hat{r}^H_{t+1}$ benefits from the feature V_1 derived from the encoder and the temporal features harnessed by the LSTM from its preceding time step. This strategy signifies that the generation of the fine-grained trajectory segment is anchored by the known endpoints, providing a reliable boundary for the entire trajectory's position. After opt_2, the prediction target of the deep learning model can be denoted as $\hat{R}^{H'}_t = \{\hat{r}^H_t, \hat{r}^{H'}_{t+1}, \hat{r}^{H'}_{t+2}, \ldots, \hat{r}^{H'}_{t+m-1}, \hat{r}^{H'}_{t+m}, \hat{r}^H_{t+m+1}\}$. The corresponding ground-truth is $R^{H'}_t = \{r^L_t, r^H_{t+1}, r^H_{t+3}, \ldots, r^H_{t+n-3}, r^H_{t+n-1}, r^L_{t+1}\}$.

3.6. Loss function

In deep learning, the loss function serves to quantify the difference between the prediction results and the ground truth data, thereby guiding the optimization of model parameters. In this study, we utilize Mean Square Error (*MSE*) as the loss function, as presented in Eq. (11). Throughout the training process, the objective of the model is to diminish the *MSE* between the predicted and the true records.

$$\mathcal{L} = \frac{1}{n}\sum_{t=1}^{n} \|r^H_t - \hat{r}^H_t\|^2 \tag{11}$$

4. Results

This section begins by introducing six baseline methods against which HiTrip's performance will be compared. Then we use the geographical distance as evaluation metric to compare HiTrip and six baseline methods. We also provide visualization of several interpolation examples to highlight the improvements made by spatial feature extraction module from multi-source input. Subsequently, we compare the fishing effort distribution estimated by HiTrip and six baseline methods to show the effectiveness of HiTrip. Several ablation studies are conducted to assess the contributions of each component of HiTrip. Finally, we employ other trawler dataset from Global Fishing Watch to demonstrate HiTrip's generalizability.

4.1. Baselines

We evaluate HiTrip against six baselines, categorized into two groups: mathematical and deep learning-based methods.

4.1.1. Mathematical methods

We use the following mathematical methods to directly interpolate the sampling interval of R^L_t from two hours to three minutes.

- **Linear Interpolation (LI)**: We link the consecutive VMS records, r^L_t and r^L_{t+1}, using a line segment and then sample the interpolated positions along this segment.
- **cubic Hermite spline (cHs)** (Hintzen et al., 2010): cHs formulates a spline interpolation function using position, course, and velocity. The interpolated records can be derived by inputting the target timestamps into this equation.
- **Catmull–Rom modification (CRm)** (Russo et al., 2011): CRm employs all accessible VMS records to determine the course deviation induced by currents. It then utilizes this course deviation to refine the interpolated records derived from the Catmull–Rom model. In this study, we leverage current fields from the Copernicus Climate Data Store, rather than computing currents from VMS records, which can enhance the accuracy of CRm.

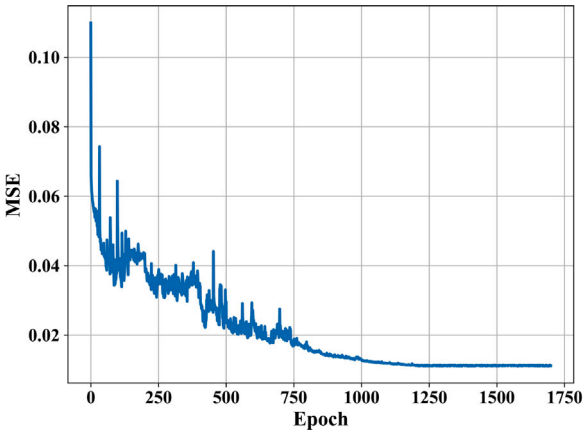


Fig. 7. Training loss.

4.1.2. Deep learning methods

For the baseline deep learning methods, we apply the same two optimizations as used in HiTrip.

- **Bidirectional Gated Recurrent Unit (Bi-GRU)** (Wang et al., 2020): This research designs a Bi-GRU network to extract the spatiotemporal features from longitude, latitude, velocity, course, and time.
- **Spatio-Temporal Sequence to Sequence model (ST-Seq2Seq)** (You et al., 2020): This study adopts the encoder–decoder framework, leveraging two GRU networks for the encoder and decoder roles. The encoder first translates the VMS records into a context vector, which is subsequently used by the decoder to produce the target trajectory.
- **Bidirectional Long Short-Term Memory Neural Network (Bi-LSTM)** (Liu et al., 2021): This study employs a Bi-LSTM architecture to generate the interpolations between two successive records.

4.2. Training details

The East China Sea dataset serves as the benchmark for evaluating the interpolation accuracy of HiTrip and other baselines. This dataset is further leveraged in our ablation study. We partition the VMS dataset based on the local cessation period, specifically from September 2015 to May 2016 and from September 2016 to May 2017. To simulate the interpolation process for historical VMS datasets, the VMS records from September 2016 to May 2017 is employed to train the interpolation model of HiTrip and the three deep learning baseline models. The trainset contains 526,736 samples, and the validset includes 131,688 samples. We evaluate the interpolation accuracy using the VMS records from September 2015 to May 2016, which contains 354,591 samples. Fig. 7 shows the training loss, which demonstrate the network convergence.

4.3. Interpolation accuracy

To assess the accuracy of our interpolation in terms of spatial distance, we utilize Geographical Distance Error (DE), which measures the geographical distance between the interpolated records and the actual ground truth within the Geographic Coordinate System, calculated with Eq. (12). Here, R indicates the Earth's radius. We have chosen not to use traditional metrics like RMSE (Root Mean Square Error), MSE (Mean Squared Error), and MAE (Mean Absolute Error), despite being direct outputs of the deep learning model. The reason is that these metrics do not provide an absolute measure of the distance error,

Table 3
Average distance error comparison among different models, the errors are presented in "mean ± standard deviation".

Model	LI	cHs	CRm	Bi-GRU	Bi-LSTM	ST-Seq2Seq	**HiTrip**
Distance error (km)	1.95 ± 1.05	1.40 ± 0.73	1.08 ± 0.67	1.07 ± 0.69	0.99 ± 0.62	0.94 ± 0.60	**0.20 ± 0.10**

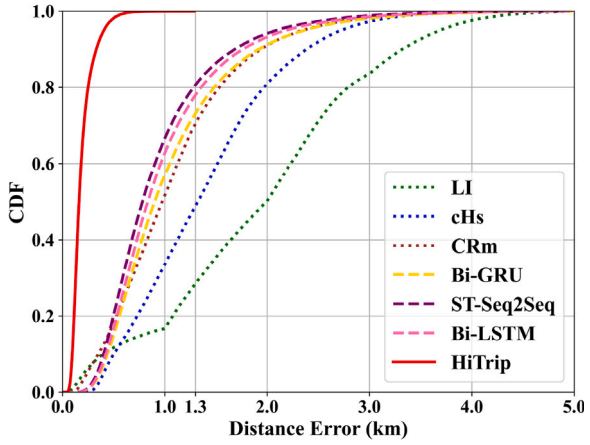


Fig. 8. CDF of *DE* for LI, cHs, CRm, Bi-GRU, ST-Seq2Seq, Bi-LSTM, and HiTrip.

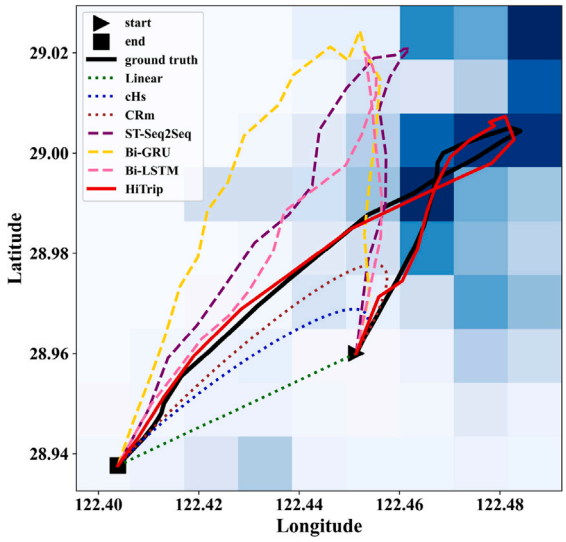


Fig. 9. Interpolation results comparison among different models with the backgrounds of coarse fishing effort distribution.

especially important in oceanic contexts where calculations should consider the Earth's curvature. Therefore, DE is implemented as the primary measure to evaluate the accuracy of our interpolation model, ensuring a more geographically precise assessment.

$$DE = \frac{R}{n}\left(\sum_{i=1}^{n}[acos(sin(r^H_{t,lat}) \times sin(\hat{r}^H_{t,lat}) + cos(r^H_{t,lat}) \times cos(\hat{r}^H_{t,lat}) \times cos(r^H_{t,lon} - \hat{r}^H_{t,lon}))]\right) \quad (12)$$

Using the first part of the VMS dataset, spanning September 2015 to May 2016, we evaluate the interpolation outcomes of HiTrip against the baselines, shown in Fig. 8. Given the tendency of trawlers to execute numerous turns during a two-hour fishing window, LI records the largest *DE*. Both cHs and CRm, considering they incorporate velocity and course, register lower *DE*s than LI. Notably, the inclusion of current fields in CRm, which is absent in both LI and cHs, renders it the most accurate among mathematical methods, bringing it on par with the deep learning model of Bi-GRU.

Deep learning methods, leveraging their robust learning capabilities, have the potential to replicate trajectory turns, thereby offering lower *DE*s than their mathematical counterparts. HiTrip outperforms all methods, with the highest *DE* capped at a mere 1.30 km. This highlights the efficacy of using deep learning models to harness spatial features from diverse data sources, leading to accurate interpolation of trajectory segments.

To evaluate the consistency and precision of HiTrip and other baselines, we compute the mean and standard deviation of *DE* across all interpolated trajectory segments, as presented in Table 3. HiTrip achieves the minimal *DE* at 0.20 km, accompanied by the smallest standard deviation of 0.10 km. When compared with CRm and ST-Seq2Seq, the best ones in mathematical and deep learning models respectively, HiTrip demonstrates reductions in average *DE* by 81% and 79%. Moreover, HiTrip's standard deviation is notably lower than both CRm and ST-Seq2Seq, suggesting that HiTrip's predictions are more consistent. In line with the Six Sigma guidelines (Yadav et al., 2018), the majority of HiTrip's *DE*s fall below 0.5 km. Given that a single degree in the Geographic Coordinate System roughly equates to 100 km, HiTrip proves adept at analyzing fishing effort distribution with a spatial resolution close to $0.005° \times 0.005°$ and a temporal resolution of three minutes.

Fig. 9 provides a visual representation of an interpolation example, shedding light on the superior accuracy of HiTrip compared to other models, especially in terms of coarse fishing effort distributions. LI's straight segment results in the highest *DE* when the interpolated section contains a turn. While cHs and CRm depict the turn with smoother curves, they miss the sharpness of the actual turn. Deep learning techniques do recognize the segment's turn pattern, but misplace the turning positions. On the other hand, HiTrip not only identifies the sharp turns but also pinpoints the precise turning location, typically occurring in fishing hotspots. This highlights that the Enc_D network captures the core spatial features from the coarse fishing effort distributions, allowing HiTrip to accurately recreate the turns in the interpolated trajectory segment.

To showcase the influence of implementing sea surface factors fields, Fig. 10 displays three interpolation outcomes against the background of the variability in *SSH*, *SSS*, and *SST* fields, as determined by Eq. (1). In Fig. 10(a), mathematical methods produce either straight lines or smooth curves. While deep learning methods do introduce turns in the trajectory, they misplace the turn positions, resulting in substantial *DE*s. Extracting the spatial features from *SSH* variability via the Enc_S network, HiTrip accurately regenerate turns in regions with marked variability, closely mirroring the ground truth. Fig. 10(b) and 10(c) also demonstrate the effects of applying Enc_S on the variability fields of *SSS* and *SST*, respectively.

Fig. 10 showcases the efficacy of our custom-designed feature extraction network, Enc_S, in pinpointing the spatial features inherent within the sea surface factors fields. This insight enhances HiTrip's capability to detect prospective fishing regions, thereby facilitating a more precise reconstruction of turns within the target trajectory segments.

To showcase the impact of exploiting the current fields through the network Enc_C, we present an interpolation example with the background of the current fields in Fig. 11. All methods, except for LI, generate interpolations with turns. HiTrip's outcomes indicate that during fishing activities, such as setting, towing, or retrieving nets, trawlers often navigate at significant angles to the current or even against it.

Consequently, with both Enc_D and Enc_S identifying potential fishing areas and the associated turning positions, and Enc_C determining the course at each interpolation point, HiTrip attains remarkable interpolation precision.

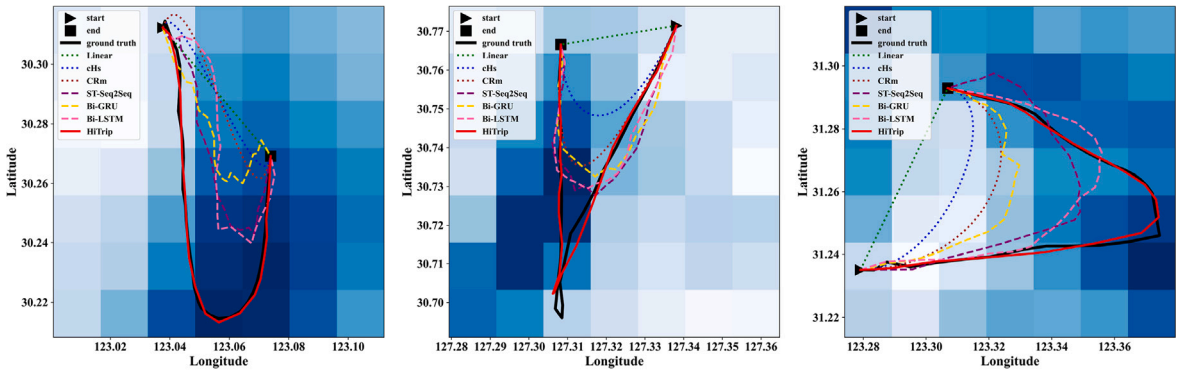


Fig. 10. Interpolation results comparison among different models with the backgrounds of the spatial variability of (a) SSH, (b) SSS, and (c) SST.

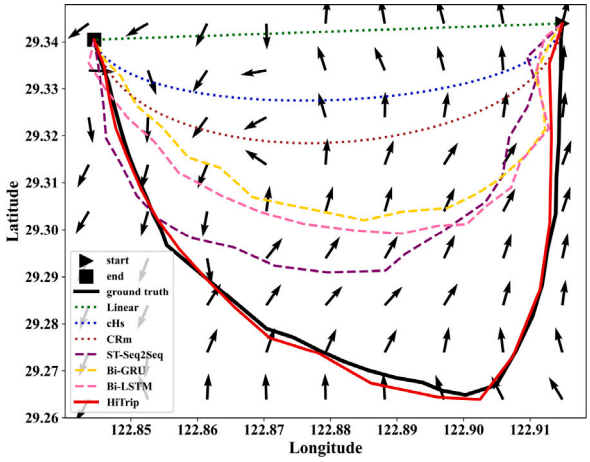


Fig. 11. Interpolation results comparison among different models with the background of current fields.

4.4. Estimation of fishing effort distribution

The main objective of analyzing the historical VMS dataset is to track changes in fishing effort distribution over time. So the second metric we used is the difference between the estimated fishing effort distribution based on the interpolated results and the actual distribution. Considering the maximum spatial resolution of HiTrip is approximately $0.005° \times 0.005°$, we juxtapose the fishing effort distributions of HiTrip and the baseline methods at spatial resolutions of $0.005° \times 0.005°$, $0.01° \times 0.01°$, and $0.1° \times 0.1°$. The time resolution for these distributions is set to one month, a standard duration for such analyses.

We specifically calculate the average fishing effort error (AFE) in areas where actual fishing activities were recorded. This approach allows us to assess the accuracy of various interpolation models in estimating fishing effort distribution. The average fishing effort error (AFE) and the average fishing effort value for month m, denoted as AE_m and AFE_m respectively, are derived using Eq. (13) and (14). Within these equations, $\hat{FE}_m$ labels the fishing effort distributions deduced from the interpolation outcomes of various methods, while FE_m represents the actual fishing effort distributions for month m. To ensure accuracy, both Eq. (13) and (14) consider only those grids that have recorded fishing activity, preventing the dilution of fishing effort errors into grids devoid of any fishing action. Table 4 presents a comparison of the monthly average fishing effort errors across grids during the evaluation period. It also offers the monthly average fishing effort values across these grids as references.

$$AE_m = avg(\sum_{x,y} |FE_m^{x,y} - \hat{FE}_m^{x,y}|,\ if\ FE_m^{x,y} > 0) \tag{13}$$

$$AFE_m = avg(\sum_{x,y} FE_m^{x,y},\ if\ FE_m^{x,y} > 0) \tag{14}$$

In Table 4, as spatial resolution increases, leading to a rise in grid numbers, the AFE per grid subsequently decreases. Across all spatial resolutions and months, LI consistently registers the highest error, while cHs and CRm follow with lessor errors. The interpolations derived from deep learning methods outshine those produced by traditional mathematical approaches in terms of precision. HiTrip stands out by consistently recording the lowest AE across every spatial resolution and month, surpassing all alternative methods. These results align with observations from Table 3, underscoring that superior interpolation techniques yield more accurate fishing effort distributions.

$$EER_m = \frac{AE_m}{avg(\sum_{x,y} |FE_m^{x,y}|,\ if\ FE_m^{x,y} > 0)} \tag{15}$$

To present a direct measure of estimation accuracy for the fishing effort distribution across different interpolation models, we further introduce the fishing effort error ratio (EER) in Eq. (15). EER calculates the ratio of AFE to the average fishing effort per grid each month. This metric offers a comparative assessment of the relative precision of different interpolation models in estimating fishing effort distribution.

Table 4 shows that the EER consistently rise with the increase of spatial resolution across all methods. For instance, considering the EER in September, mathematical methods have a minimum value of 10.4% for CRm at a spatial resolution of $0.1° \times 0.1°$. In comparison, prior deep learning methods exhibit a minimum error ratio of 8.1% for ST-Seq2Seq. When we enhance the spatial resolution, the error ratios for mathematical methods exceeds 70%, whereas those for the earlier deep learning methods surpasses 50%. In contrast, HiTrip's EER in September stand at 2.1%, 24.9%, and 18.6% for spatial resolutions of $0.1° \times 0.1°$, $0.01° \times 0.01°$, and $0.005° \times 0.005°$, respectively. This indicates that HiTrip effectively supports fishing effort distribution even at the finer spatial resolution of $0.005° \times 0.005°$.

In October, all methods except LI exhibit the highest average fishing effort error. We visualize the error distributions for October's fishing efforts, $\|FE_{10} - \hat{FE}_{10}\|$, comparing HiTrip with CRm and ST-Seq2Seq, as depicted in Fig. 12. The same figure also presents the ground truth for context. Notably, HiTrip displays fewer global errors, which are evenly distributed across the fishing grounds. To delve deeper into distribution errors, three prominent fishing zones are labeled in Fig. 12(a). As the fishing season kicks off in September, numerous trawlers conduct fishing offshore (zones A and B), while some target more distant areas (zone C). The heatmap of the three fishing hot zones reveals HiTrip's superior accuracy in error distribution compared to its counterparts, aligning with the quantitative findings in Table 4.

Table 4
Estimation error of fishing effort over grids comparison among different models. We report the average fishing effort error and corresponding error ratio in Table.

Spatial resolution	Month	LI	cHs	CRm	Bi-GRU	Bi-LSTM	ST-Seq2Seq	HiTrip	Average fishing effort
	Sep	43.0(17.6%)	31.2(12.7%)	25.5(10.4%)	25.7(10.5%)	20.7(8.5%)	19.9(8.1%)	**5.0(2.1%)**	245.1
	Oct	63.1(14.5%)	44.2(10.1%)	35.6(8.2%)	34.8(8.0%)	29.6(6.8%)	27.3(6.3%)	**6.7(1.5%)**	436.8
	Nov	59.5(15.4%)	41.2(10.7%)	33.5(8.7%)	33.3(8.6%)	27.6(7.1%)	25.6(6.6%)	**5.8(1.5%)**	385.6
	Dec	57.2(15.4%)	38.7(10.4%)	31.3(8.4%)	29.8(8.0%)	24.9(6.7%)	24.5(6.6%)	**6.0(1.6%)**	370.7
0.1^o	Jan	48.8(16.5%)	34.7(11.7%)	28.1(9.5%)	27.0(9.1%)	24.0(8.1%)	21.9(7.4%)	**5.2(1.8%)**	295.9
	Feb	44.5(18.5%)	31.1(12.9%)	24.9(10.4%)	24.1(10.0%)	20.3(8.5%)	19.0(7.9%)	**4.5(1.9%)**	240.6
	Mar	50.1(17.3%)	34.2(11.8%)	27.9(9.6%)	27.5(9.5%)	22.2(7.7%)	22.0(7.6%)	**5.2(1.8%)**	289.0
	Apr	49.0(18.0%)	34.2(12.5%)	27.6(10.1%)	26.7(9.8%)	21.2(7.8%)	21.0(7.7%)	**5.0(1.8%)**	272.6
	May	63.1(27.1%)	45.6(19.6%)	37.9(16.3%)	34.8(14.9%)	28.9(12.4%)	26.8(11.5%)	**6.5(2.8%)**	232.9
	Sep	8.6(82.1%)	7.6(73.1%)	6.8(65.2%)	6.6(63.4%)	6.1(58.7%)	6.0(57.6%)	**2.6(24.9%)**	10.5
	Oct	10.2(80.4%)	8.9(70.6%)	7.9(62.6%)	7.6(59.8%)	7.1(56.4%)	6.9(54.7%)	**2.8(22.4%)**	12.7
	Nov	10.2(82.2%)	9.0(72.1%)	7.9(63.9%)	7.5(60.2%)	7.0(55.9%)	6.8(54.6%)	**2.8(22.3%)**	12.4
	Dec	9.7(81.0%)	8.4(70.6%)	7.5(62.5%)	7.1(59.8%)	6.7(56.0%)	6.5(54.6%)	**2.7(22.7%)**	11.9
0.01^o	Jan	9.4(80.7%)	8.2(70.4%)	7.2(62.1%)	6.9(59.4%)	6.5(55.9%)	6.3(54.4%)	**2.7(23.1%)**	11.6
	Feb	9.0(81.3%)	7.9(71.8%)	7.1(64.3%)	6.7(60.7%)	6.3(57.0%)	6.1(55.3%)	**2.6(23.4%)**	11.1
	Mar	9.2(81.0%)	8.0(70.8%)	7.1(62.7%)	6.9(60.4%)	6.5(56.9%)	6.3(55.5%)	**2.7(23.4%)**	11.3
	Apr	9.5(82.7%)	8.3(72.7%)	7.3(64.0%)	7.0(60.8%)	6.5(57.0%)	6.3(55.1%)	**2.6(23.0%)**	11.4
	May	10.7(95.0%)	9.3(83.2%)	8.4(74.5%)	7.8(69.8%)	7.4(65.6%)	7.2(63.7%)	**3.0(26.8%)**	11.2
	Sep	4.7(87.1%)	4.4(81.8%)	4.0(74.2%)	4.0(74.3%)	3.8(71.2%)	3.8(70.2%)	**1.1(18.6%)**	5.4
	Oct	5.2(86.9%)	4.8(81.1%)	4.4(73.4%)	4.3(72.1%)	4.2(69.6%)	4.1(68.4%)	**1.1(18.6%)**	6.0
	Nov	5.2(88.3%)	4.9(82.6%)	4.4(74.6%)	4.3(72.6%)	4.1(69.7%)	4.1(68.7%)	**1.1(18.6%)**	5.9
	Dec	5.0(86.6%)	4.7(80.9%)	4.2(73.0%)	4.2(71.9%)	4.0(69.3%)	4.0(68.2%)	**1.1(18.6%)**	5.8
0.005^o	Jan	5.0(86.7%)	4.6(80.7%)	4.2(72.9%)	4.1(72.1%)	4.0(69.0%)	3.9(67.9%)	**1.1(18.6%)**	5.7
	Feb	4.8(87.0%)	4.5(81.5%)	4.1(74.2%)	4.1(73.0%)	3.9(69.8%)	3.8(68.5%)	**1.0(18.8%)**	5.6
	Mar	4.9(86.5%)	4.5(80.8%)	4.1(73.2%)	4.1(72.5%)	3.9(70.0%)	3.9(68.9%)	**1.1(18.8%)**	5.6
	Apr	5.0(88.1%)	4.7(82.3%)	4.2(74.2%)	4.1(72.8%)	4.0(70.3%)	3.9(68.4%)	**1.1(18.7%)**	5.7
	May	5.3(105.4%)	5.0(98.1%)	4.5(89.2%)	4.3(85.9%)	4.2(83.0%)	4.1(81.8%)	**1.2(22.9%)**	5.0

Table 5
Ablation studies. We report the average distance errors (mean ± standard deviation) when applying the network components with their corresponding inputs in the interpolation model.

Module	VMS Records	Coarse Fishing Effort Distributions	Sea Surface Factor Fields	Current Fields	LSTM	opt_1	opt_2	Distance error (km)
Exp_1	✓				✓	✓	✓	0.93 ± 0.59
Exp_2	✓	✓			✓	✓	✓	**0.76 ± 0.53**
Exp_3	✓		✓		✓	✓	✓	0.87 ± 0.53
Exp_4	✓			✓	✓	✓	✓	0.88 ± 0.54
Exp_5	✓	✓	✓		✓	✓	✓	0.48 ± 0.28
Exp_6	✓	✓		✓	✓	✓	✓	**0.39 ± 0.23**
Exp_7	✓		✓	✓	✓	✓	✓	0.44 ± 0.26
Exp_8	✓	✓	✓	✓		✓	✓	**0.32 ± 0.19**
Exp_9	✓	✓	✓	✓	✓			0.54 ± 0.31
Exp_{10}	✓	✓	✓	✓	✓	✓		**0.38 ± 0.23**
Exp_{11}	✓	✓	✓	✓	✓		✓	0.45 ± 0.26
HiTrip	✓	✓	✓	✓	✓	✓	✓	**0.20 ± 0.10**

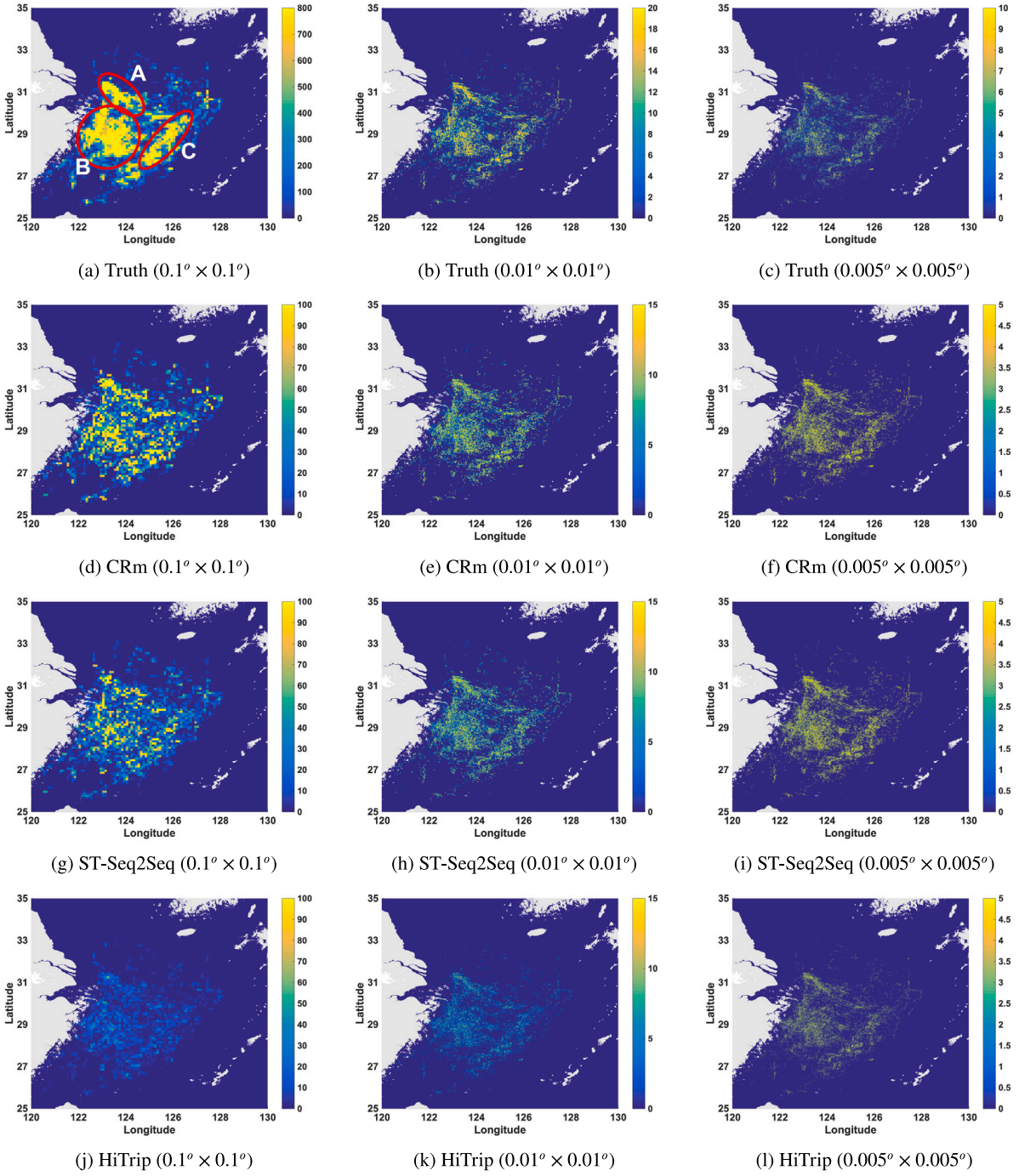
4.5. Ablation study

To analyze the contributions of each component within HiTrip, we carry out the ablation studies. All the settings and the results are shown in Table 5.

We first evaluate the interpolation accuracy by solely utilizing a Multi-Layer Perceptron (MLP) to extract spatial features from the input VMS records, under the directive guidance of an LSTM-based decoder. This methodology resulted in a DE of 0.93 km, which is impressively lower than the outcomes achieved by three other deep learning baselines. This finding emphasizes the superiority of our proposed MLP-LSTM encoder–decoder combination, demonstrating its effectiveness over the other deep learning models in accurately interpreting and processing the spatial features of VMS data.

Subsequently, we apply ResNet to extract spatial features from coarse fishing effort distribution, sea surface factors fields, and current fields, respectively, followed by their integration with the features extracted by MLP from VMS records. Adding spatial features extracted from the coarse fishing effort distribution resulted in the most significant improvement, achieving a low DE of 0.76 km. This finding underscores the importance of differentiating between fishing and steaming states to enhance the accuracy of trajectory interpolation. The notable improvement in interpolation precision is mainly due to ResNet's effective extraction of spatial features from the coarse fishing effort distribution. This approach allows for a more precise representation of the operational states in the trajectory, leading to a more accurate and reliable interpolation of the VMS data.

When we applied ResNet to extract spatial features from sea surface factors fields and merged these with the spatial features extracted from VMS records by MLP, the DE is larger compared to utilizing spatial features from coarse fishing effort distribution. This rise in DE is likely due to the indirect way that sea surface factors fields indicate fishing hotspots, resulting in a less accurate identification of fishing activities. Furthermore, the DE increases even more when using the combination of spatial features from current fields and VMS records. This rise highlights the challenges in interpolating trajectories when there is no clear distinction between fishing and steaming activities. It points to the necessity of having direct and distinct indicators of fishing activity to enhance the accuracy of interpolations in such models.



(a) Truth ($0.1^o \times 0.1^o$) (b) Truth ($0.01^o \times 0.01^o$) (c) Truth ($0.005^o \times 0.005^o$)

(d) CRm ($0.1^o \times 0.1^o$) (e) CRm ($0.01^o \times 0.01^o$) (f) CRm ($0.005^o \times 0.005^o$)

(g) ST-Seq2Seq ($0.1^o \times 0.1^o$) (h) ST-Seq2Seq ($0.01^o \times 0.01^o$) (i) ST-Seq2Seq ($0.005^o \times 0.005^o$)

(j) HiTrip ($0.1^o \times 0.1^o$) (k) HiTrip ($0.01^o \times 0.01^o$) (l) HiTrip ($0.005^o \times 0.005^o$)

Fig. 12. Visualization of distribution errors of fishing effort in October among three methods at different spatial resolutions. Best view in color. (a)–(c) are the ground-truth fishing effort distribution, (d)–(l) are the distribution errors between the fishing effort distribution calculated after applying different interpolation methods and ground-truth.

Another important finding from our ablation study is the enhanced interpolation accuracy achieved by combining spatial features extracted from current fields using ResNet with those from VMS records, and either coarse fishing effort distributions or sea surface factors fields. Such a combination outperforms the accuracy achieved when these features are individually paired with VMS record features. The synergistic effect of merging current field features with either coarse fishing effort distributions or sea surface factors fields considerably betters the prediction of course changes, thereby leading to more precise trajectory interpolations.

However, it is important to recognize that a combination of spatial features from both coarse fishing effort distribution and sea surface factors fields results in a higher DE than the other combinations mentioned. This outcome suggests that while it is vital to differentiate between fishing and steaming states, failing to thoroughly consider course changes can result in increased interpolation errors. This observation underscores the intricate nature of accurately interpolating trawler trajectories and the necessity for a comprehensive approach that takes into account the various dimensions of trawler movements and behaviors.

Moreover, to demonstrate HiTrip's proficiency in capturing the temporal sequence characteristics intrinsic to the target trajectory segment, we conducted a test by removing the LSTM component from the decoder and relying exclusively on an additional MLP to directly

Table 6
Interpolation error (mean ± standard deviation) comparison among different models in Global Fishing Watch dataset.

Model	LI	cHs	CRm	Bi-GRU	Bi-LSTM	ST-Seq2Seq	HiTrip^{GFW}	HiTrip$^{GFW}_{TL}$
Distance error (km)	1.86 ± 1.33	1.56 ± 0.97	1.30 ± 0.90	1.24 ± 0.85	1.09 ± 0.80	1.07 ± 0.80	0.89 ± 0.66	**0.40 ± 0.29**

produce the interpolated records. The results of this experiment show that omitting LSTM results in a 60% increase in *DE* compared to the original HiTrip configuration. This outcome clearly indicates that the inclusion of LSTM, with its ability to represent temporal sequencing, plays a crucial role in substantially improving the accuracy of the interpolation.

Lastly, to gauge the potency of the two proposed optimizations, we conduct tests excluding opt_1, opt_2, or both. The individual removal of opt_1 resulted in a *DE* rise by 125%, illuminating the pitfalls of not bridging the information gap between input and target records. Exclusion of opt_2 led to a 90% *DE* surge, underscoring the value of having known anchor points at the beginning and end of trajectory segments on limiting potential accumulated errors. When both optimizations are absent, the *DE* soars by 170%, reaffirming the indispensable nature of these optimizations in our model.

4.6. Generalization of HiTrip

In this section, we delve into HiTrip's generalization potential across diverse maritime areas. We utilize a trawler dataset sourced from Global Fishing Watch, which predominantly captures trawling activities in the Norwegian Sea, Pacific Ocean, and Indian Ocean. We split this dataset chronologically into two halves: the latter portion is allocated for transfer learning using the well-tuned HiTrip (referred to as HiTrip$^{GFW}_{TL}$) or for training a fresh HiTrip instance (designated as HiTrip^{GFW}). The initial half serves to evaluate the *DE*. Moreover, the three deep learning baseline models also undergo transfer learning based on their pre-trained models on the East China Sea dataset.

The sampling frequency of this dataset ranges from minutes to hours, suggesting it is not directly suitable for training or evaluation. To address this, we first employ linear interpolation to narrow down the sampling intervals. Given that trawlers typically do not execute intricate maneuvers in short time spans, we only apply this interpolation when the gap between two successive records is under twelve minutes. This tactic not only shortens the intervals as much as feasible but also keeps the added error to a minimum. As a result, the majority of the intervals between adjacent records are refined to three minutes. The trainset, validset, and testset contain 1372, 344, and 1716 samples, respectively. This diverse dataset distribution enables a comprehensive evaluation of the interpolation models across different maritime regions, temporal periods and data volumes, providing robust insights into their performance and applicability.

Table 6 showcases the *DE*s of the baseline models, HiTrip$^{GFW}_{TL}$, and HiTrip^{GFW}. Initially, the LI yields the highest *DE* among all models, yet it remains smaller than the *DE* presented in Table 3. This difference largely arises from our application of linear interpolation on the dataset to narrow down the sampling interval, causing some generated records to match the linearly interpolated ground-truth records. However, as certain ground-truth records are products of linear interpolation, the *DE*s of cHs and CRm are marginally higher than those in Table 3.

Besides, the three deep learning models solely depend on VMS records for inputs and do not extract the spatial features inherent in coarse fishing effort distributions, sea surface factors fields, and current fields. Consequently, transfer learning does not empower their well-trained models to assimilate enough features pertinent to the respective maritime regions, especially considering that the dataset volume is considerably smaller than the East China Sea dataset. This limitation is evident, as all of their *DE*s surpass those in Table 3.

Additionally, HiTrip^{GFW} registers a higher error than HiTrip$^{GFW}_{TL}$. This discrepancy mainly stems from constraints tied to the training data volume and the inherent difference when computing coarse fishing effort distribution across different marine regions. Nevertheless, the *DE* for HiTrip^{GFW} still outperforms that of ST-Seq2Seq, attesting to the merit of extracting spatial features from coarse fishing effort distributions, sea surface factors fields, and current fields.

Remarkably, HiTrip$^{GFW}_{TL}$ posts the most commendable *DE*. This outcome is attributed to HiTrip's proficiency in encapsulating the influence of coarse fishing effort distributions, sea surface factors fields, and current fields for the target trajectory segments. The spatial features of the new dataset then enhance the interpolation accuracy of HiTrip$^{GFW}_{TL}$. Furthermore, the slightly elevated *DE* of HiTrip$^{GFW}_{TL}$ in comparison to HiTrip in Table 3 can be linked to biases in the coarse fishing effort distribution calculation.

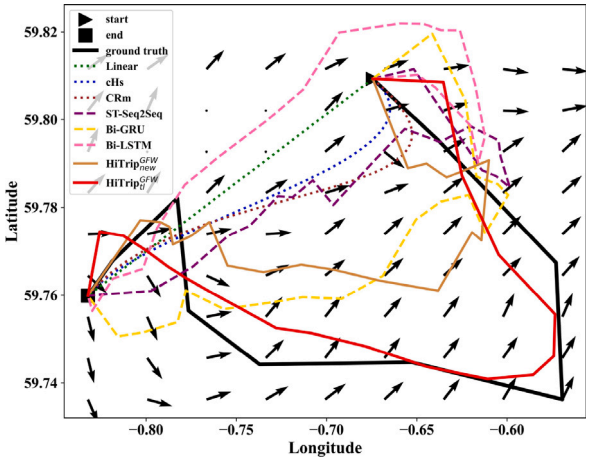


Fig. 13. Interpolation comparison of Global Fishing Watch dataset among different models with the background of current fields.

Fig. 13 illustrates an interpolation example in the Norwegian Sea from the Global Fishing Watch dataset. The interpolations by the three mathematical methods still yield either straight lines or gentle curves. The Bi-LSTM's results do suggest a turn but misdirect it, rendering it less accurate than even the basic mathematical methods. Bi-GRU and ST-Seq2Seq, on the other hand, manage to detect the trajectory turns but falter in pinpointing the exact turn locations.

Due to the intrinsic bias in estimating the coarse fishing effort distribution across different maritime regions, HiTrip^{GFW} faces challenges in adequately extracting spatial features using the limited training data available. This leads its interpolated results to diverge considerably from the actual trajectory segment. Nevertheless, when utilizing transfer learning and leveraging the well-trained HiTrip model in conjunction with this dataset, HiTrip$^{GFW}_{TL}$ can successfully extract spatial features on the corresponding coarse fishing effort distributions, sea surface factors fields, and current fields. As a result, it can more accurately reconstruct the target trajectory segments, producing outcomes that outperform other models.

5. Discussion

This study proposes HiTrip, a historical trawler trajectory interpolation system via deep learning on VMS and hydrological factor datasets. HiTrip is able to interpolate the sampling interval of the historical dataset from two hours to three minutes with a low interpolation error. The precision of HiTrip's interpolation is rooted in carefully

crafting deep learning model to encapsulate the influential factors. The deep learning model serves three primary functions: extracting features from multi-source data, articulating sequential characteristics, and implementing optimizations.

(1) Interpolation challenges mainly arise within trajectory segments that showcase turns, particularly evident during fishing activities. Current research typically differentiates the fishing state of target trajectory segments by scrutinizing the characteristics of two sequential records from low-sampled VMS datasets, often setting velocity thresholds (Lee et al., 2010). Nonetheless, trawlers might wrap up a fishing activity within a mere two-hour window, making it possible to overlook these events when reliant on infrequent VMS records. Thus, using only two adjacent VMS records for interpolation can lead to inaccuracies.

HiTrip goes beyond merely interpolating from VMS records; it enhances the feature set by integrating information from coarse fishing effort distributions, sea surface factors fields' variability, and current fields. These additional features address the insufficiency of relying on VMS records alone to identify fishing activities. The ablation study underscores the importance of the encoder in gleaning features from these diverse data sources. It further establishes that, even though the coarse fishing effort distributions are derived from low-sampled VMS datasets, they offer a direct insight into fishing hotspots. In comparison, the variability of the sea surface factors fields offers an indirect indicator of fishing activities. Furthermore, the current fields augment interpolation precision by signaling nuanced course changes.

(2) The outcomes within a target trajectory segment display a clear sequential pattern. Each record's position, velocity, and course are dictated by the previous one. HiTrip employs the LSTM network to capture this inherent sequential nature and further guides the encoder to extract fused features from multiple data sources for each time step. The ablation study emphasizes that overlooking the LSTM's ability to integrate these sequential features results in a significant decline in interpolation accuracy.

(3) Introducing two pivotal optimization strategies is essential for the accurate interpolation of historical trajectories, especially when narrowing the gap from a two-hour to a mere three-minute sampling interval. The first optimization tactically addresses the pronounced information discrepancy between the readily available four VMS records and the targeted 39 VMS records. Besides, the second strategy curtails the accumulation of interpolation errors. The efficacy and importance of these optimizations are further substantiated by the results of our ablation study.

Owing to its superior interpolation accuracy, HiTrip can facilitate analysis for fishing effort distributions at an enhanced spatial resolution of $0.005° \times 0.005°$. This refined granularity paves the way for advanced digital tools that can significantly aid in the planning and management of marine fishery resources. Furthermore, the tests on the Global Fishing Watch dataset attest to HiTrip's generalization. It can seamlessly adapt to different maritime regions while consistently delivering high-precision results.

Additionally, we observe from Table 1 that numerous early datasets from the Atlantic region feature sampling intervals longer than one or two hours. For historical datasets with a two-hour sampling interval, HiTrip can be effectively applied through transfer learning, as demonstrated in our interpolation of the GFW dataset. Regarding historical datasets with a one-hour sampling interval, the HiTrip model can be adapted and retrained to meet the specific interpolation needs, transitioning from a one-hour to a three-minute interval.

The primary limitation of HiTrip is its design specificity for trawlers, making it less suitable for direct application to other types of fishing vessels. Our study focused on the impact of marine hydrological factors on trawler trajectories, leading to the development of a model specifically tailored for these effects. Trawlers are characterized by frequent turns while fishing, typically in areas identified as fishing hotspots, marked by dense clusters of VMS records. The coarse fishing effort distribution and marine hydrological factor fields can help in indicating the operational states of trawlers, however, it is crucial to recognize that different fishing vessels display distinct trajectory patterns. For instance, gillnets exhibit straighter trajectory segments and sharp turns during fishing activities. These variations in movement patterns limits HiTrip's applicability across diverse types of fishing vessels.

The second significant limitation of HiTrip pertains to the amount of historical VMS dataset required for its training. Our comparison of interpolation results using historical VMS datasets from Global Fishing Watch and the East China Sea shows that the accuracy of HiTrip's interpolation is dependent on the volume of the training dataset. In situations where the available historical VMS dataset is limited, as with the Global Fishing Watch dataset, employing transfer learning techniques is more effective for achieving better interpolation accuracy than starting the learning process from scratch

6. Conclusion and future direction

This study propose HiTrip, a historical trawler trajectory interpolation system powered by deep learning. Leveraging VMS and hydrological factor datasets, HiTrip refines the sampling interval of historical VMS datasets from two hours to just three minutes.

HiTrip proposes a novel deep learning network that melds ResNet, LSTM, and MLP, which seamlessly integrates spatial features derived from coarse fishing effort distributions, sea surface factors fields, and current fields, while also accounting for the temporal associations among VMS records within the target trajectory segment. Validated by our case study on the East China Sea's dataset, HiTrip achieves an impressive interpolation error of just 0.20 km. This level of precision facilitates a detailed examination of historical fishing effort distribution at a refined spatial resolution, approaching $0.005° \times 0.005°$, as corroborated by the historical dataset from the East China Sea. Furthermore, when tested on a trawler dataset from Global Fishing Watch, spanning global maritime regions, HiTrip consistently demonstrates its vast adaptability and generalizability.

Although HiTrip focuses on the gear of trawlers, the proposed method also has the potential to be applied to other gears. Since coarse fishing effort distributions, sea surface factors fields, and current fields will also affect other gears of fishing vessels, our analysis and model can be adapted to those gears. Additionally, HiTrip can be utilized for other important fishery applications, such as recovering trawler trajectories that have been manually terminated due to illegal fishing.

CRediT authorship contribution statement

Zhongning Zhao: Methodology, Writing – original draft. **Jiaxuan Chen:** Methodology, Writing – original draft. **Yuqi Shi:** Methodology, Writing – original draft. **Feng Hong:** Supervision, Writing – review & editing. **Guiyuan Jiang:** Writing – review & editing. **Haiguang Huang:** Writing – review & editing. **Jinhua Zhao:** Writing – review & editing.

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.

Data availability

Data will be made available on request.

Acknowledgments

This work has been supported by National Natural Science Foundation of China (Grant No. 41976185).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.oceaneng.2023.116588.

References

Azevedo, M., Silva, C., 2020. A framework to investigate fishery dynamics and species size and age spatio-temporal distribution patterns based on daily resolution data: A case study using Northeast Atlantic horse mackerel. ICES J. Mar. Sci. 77 (7–8), 2933–2944.

Behivoke, F., Etienne, M.-P., Guitton, J., Randriatsara, R.M., Ranaivoson, E., Léopold, M., 2021. Estimating fishing effort in small-scale fisheries using GPS tracking data and random forests. Ecol. Indic. 123, 107321.

Bertrand, S., Burgos, J.M., Gerlotto, F., Atiquipa, J., 2005. Lévy trajectories of Peruvian purse-seiners as an indicator of the spatial distribution of anchovy (Engraulis ringens). ICES J. Mar. Sci. 62, 477–482.

Bond, T., McLean, D.L., Wakefield, C.B., Partridge, J.C., Prince, J., White, D., Boddington, D.K., Newman, S.J., 2021. Quantifying fishing activity targeting subsea pipelines by commercial trap fishers. Rev. Fish Biol. Fish. 31, 1009–1023.

Campbell, M.S., Stehfest, K.M., Votier, S.C., Hall-Spencer, J.M., 2014. Mapping fisheries for marine spatial planning: Gear-specific vessel monitoring system (VMS), marine conservation and offshore renewable energy. Mar. Policy 45, 293–300.

Capobianco, S., Millefiori, L.M., Forti, N., Braca, P., Willett, P., 2021. Deep learning methods for vessel trajectory prediction based on recurrent neural networks. IEEE Trans. Aerosp. Electron. Syst. 57, 4329–4346.

Gao, D., Zhu, Y., Zhang, J., He, Y., Yan, K., Boran, Y., 2021. A novel MP-LSTM method for ship trajectory prediction based on ais data. Ocean Eng. 228, 108956.

Guo, S., Liu, C., Guo, Z., Feng, Y., Hong, F., Huang, H., 2018. Trajectory prediction for ocean vessels base on K-order multivariate Markov chain. In: Wireless Algorithms Systems and Applications. Tianjin, China, pp. 140–150.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA, pp. 770–778.

Hintzen, N.T., Piet, G.J., Brunel, T., 2010. Improved estimation of trawling tracks using cubic Hermite spline interpolation of position registration data. Fish. Res. 101, 108–115.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9, 1735–1780.

Hong, F., Zhao, Z., Huang, H., Feng, Y., 2019. Discovering early birds in trawlers through VMS data analysis. In: OCEANS 2019 - Marseille. Marseille, France, pp. 1–4.

Huang, H., Hong, F., Liu, J., Liu, C., Feng, Y., Guo, Z., 2019. FVID: Fishing vessel type identification based on VMS trajectories. J. Ocean Univ. China 18, 403–412. http://dx.doi.org/10.1007/s11802-019-3717-9.

Huang, C., Qi, X., Zheng, J., Zhu, R., Shen, J., 2023. A maritime traffic route extraction method based on density-based spatial clustering of applications with noise for multi-dimensional data. Ocean Eng. 268, 113036.

Iiyama, M., Zhao, K., Hashimoto, A., Kasahara, H., Minoh, M., 2018. Fishing spot prediction by sea temperature pattern learning. In: 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans. OTO, Kobe, Japan, pp. 1–4.

Jones, J.B., 1992. Environmental impact of trawling on the seabed: A review. New Zealand J. Mar. Freshw. Res. 26, 59–67.

Katara, I., Silva, A., 2017. Mismatch between VMS data temporal resolution and fishing activity time scales. Fish. Res. 188, 1–5.

Kourti, N., Shepherd, I., Greidanus, H., Alvarez, M., Aresu, E., Bauna, T., Chesworth, J., Lemoine, G., Schwartz, G., 2005. Integrating remote sensing in fisheries control. Fish. Manag. Ecol. 12, 295–307.

Lee, J., South, A.B., Jennings, S., 2010. 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, 1260–1271.

Li, G., Lu, Z., Cao, Y., Zou, L., Chen, X., 2022. CPUE estimation and standardization based on VMS: A case study for squid-jigging Fishery in the equatorial of Eastern Pacific ocean. Fishes 8, 2.

Liu, Q., Chen, Y., Wang, J., Miao, H., Wang, Y., 2023. An example of fishery yield predictions from VMS-based navigational characteristics applied to double trawlers in China. Fish. Res. 261, 106614.

Liu, R.W., Nie, J., Garg, S., Xiong, Z., Zhang, Y., Hossain, M.S., 2021. Data-driven trajectory quality improvement for promoting intelligent vessel traffic services in 6G-enabled maritime IoT systems. IEEE Internet Things J. 8, 5374–5385.

Lopes, P., Campos, A., Fonseca, P., Parente, J., Antunes, N., 2018. B. In: Maritime Transportation and Harvesting of Sea Resources, Vol. 2. pp. 1263–1266.

Mills, C.M., Townsend, S.E., Jennings, S., Eastwood, P.D., Houghton, C.A., 2007. Estimating high resolution trawl fishing effort from satellite-based vessel monitoring system data. ICES J. Mar. Sci. 64, 248–255.

Mullowney, D., Dawe, E., 2009. Development of performance indices for the newfoundland and labrador snow crab (*Chionoecetes opilio*) fishery using data from a vessel monitoring system. Fisheries Research 100 (3), 248–254.

Murray, B., Perera, L.P., 2021. An AIS-based deep learning framework for regional ship behavior prediction. Reliab. Eng. Syst. Saf. 215, 107819.

Natale, F., Gibin, M., Alessandrini, A., Vespe, M., Paulrud, A., 2015. Mapping fishing effort through AIS data. PLoS One 10, e0130746.

Pedersen, S.A., Fock, H., Krause, J., Pusch, C., Sell, A.L., Böttcher, U., Rogers, S.I., Sköld, M., Skov, H., Podolska, M., Piet, G.J., Rice, J.C., 2008. Natura 2000 sites and fisheries in German offshore waters. ICES J. Mar. Sci. 66 (1), 155–169.

Qu, J., Liu, R.W., Guo, Y., Lu, Y., Su, J., Li, P., 2023. Improving maritime traffic surveillance in inland waterways using the robust fusion of AIS and visual data. Ocean Eng. 275, 114198.

Rong, H., Teixeira, A., Guedes Soares, C., 2020. Data mining approach to shipping route characterization and anomaly detection based on AIS data. Ocean Eng. 198, 106936.

Russo, T., Parisi, A., Cataudella, S., 2011. New insights in interpolating fishing tracks from VMS data for different métiers. Fish Res 108, 184–194.

Salthaug, A., Johannessen, T., 2006. The norwegian in-year monitoring fishery for sandeel in the north sea using satellite-based VMS data and landings information. ICES CM documents 2006/N:07.

Shaobo, W., Yingjun, Z., Lianbo, L., 2020. A collision avoidance decision-making system for autonomous ship based on modified velocity obstacle method. Ocean Eng. 215, 107910.

Skaar, K.L., Jørgensen, T., Ulvestad, B.K.H., Engås, A., 2011. Accuracy of VMS data from Norwegian demersal stern trawlers for estimating trawled areas in the Barents Sea. ICES J. Mar. Sci. 68 (8), 1615–1620.

Solano-Carrillo, E., Carrillo-Perez, B., Flenker, T., Steiniger, Y., Stoppe, J., 2021. Detection and geovisualization of abnormal vessel behavior from video. In: 2021 IEEE International Intelligent Transportation Systems Conference. ITSC, Indianapolis, IN, USA, pp. 2193–2199.

Teng, G., Zhao, Z., Hong, F., Wang, B., Huang, H., 2021. HIT: Hybrid interpolation for trawlers based on deep learning. In: OCEANS 2021: San Diego – Porto. San Diego, CA, USA, pp. 1–4.

van, F., Piet, G., Hintzen, T., 2008. Spatial fishing effort distribution based on interpolated vessel tracks from vessel monitoring system (VMS) data. ICES Document CM 2008/E: 35.

Walter, I., Hoenig, J.M., Gedamke, T., 2007. Correcting for effective area fished in fishery-dependent depletion estimates of abundance and capture efficiency. ICES J. Mar. Sci. 64 (9), 1760–1771.

Wang, C., Ren, H., Li, H., 2020. Vessel trajectory prediction based on AIS data and bidirectional GRU. In: 2020 International Conference on Computer Vision, Image and Deep Learning. CVIDL, Chongqing, China, pp. 260–264.

Xue, J., Wu, C., Chen, Z., Chen, X., 2017. A novel estimation algorithm for interpolating ship motion. In: 2017 4th International Conference on Transportation Information and Safety. ICTIS, pp. 557–562.

Yadav, G., Seth, D., Desai, T.N., 2018. Application of hybrid framework to facilitate lean six sigma implementation: A manufacturing company case experience. Prod. Plan. Control 29, 185–201.

Yang, Y., Shao, Z., Hu, Y., Mei, Q., Pan, J., Song, R., Wang, P., 2022. Geographical spatial analysis and risk prediction based on machine learning for maritime traffic accidents: A case study of Fujian sea area. Ocean Eng. 266, 113106.

You, L., Xiao, S., Peng, Q., Claramunt, C., Han, X., Guan, Z., Zhang, J., 2020. ST-Seq2Seq: A spatio-temporal feature-optimized Seq2Seq model for short-term vessel trajectory prediction. IEEE Access 8, 218565–218574.

Zhang, W., Feng, X., Goerlandt, F., Liu, Q., 2020. Towards a Convolutional Neural Network model for classifying regional ship collision risk levels for waterway risk analysis. Reliab. Eng. Syst. Saf. 204, 107127.

Zhang, F., Reid, K.B., Nudds, T.D., 2018. Effects of walleye predation on variation in the stock-recruitment relationship of Lake Erie yellow perch. J. Gt. Lakes Res. 44, 805–812.

Zhang, S., Yu, B., Zheng, Q., Zhou, W., 2016. Algorithm of trawler fishing effort extraction based on BeiDou vessel monitoring system data. In: Geo-Informatics in Resource Management and Sustainable Ecosystem. Wuhan, China, pp. 159–168.

Zhao, Z., Hong, F., Huang, H., Liu, C., Feng, Y., Guo, Z., 2021. Short-term prediction of fishing effort distributions by discovering fishing chronology among trawlers based on VMS dataset. Expert Syst. Appl. 184, 115512.

Zhao, Z., Tian, Y., Hong, F., Huang, H., Zhou, S., 2020. Trawler fishing track interpolation using LSTM for satellite-based VMS traces. In: Global Oceans 2020: Singapore – U.S. Gulf Coast. Biloxi, MS, USA, pp. 1–4.

Zong, Y., Huang, H., Hong, F., Zhen, Y., Guo, Z., 2016. Recognizing fishing activities via VMS trace analysis based on Mathematical Morphology. In: 2016 Techno-Ocean (Techno-Ocean). Kobe, Japan, pp. 465–470.