

Article

VEPO-S2S: A Vessel Portrait Oriented Trajectory Prediction Model Based on S2S Framework

Xinyi Yang ^{1,2,3,†}, Zhonghe Han ^{1,3,4,5,†}, Yuanben Zhang ^{1,3,4,5}, Hu Liu ^{2,*}, Siye Liu ^{1,3,4,5}, Wanzheng Ai ² and Junyi Liu ^{1,3,4,5}

¹ Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China; yxy915200@163.com (X.Y.); hanzh@aircas.ac.cn (Z.H.); zhangyb@aircas.ac.cn (Y.Z.); 18010076513@163.com (S.L.); liujy004735@aircas.ac.cn (J.L.)

² School of Naval Architecture and Maritime, Zhejiang Ocean University, Zhoushou 361022, China; aiwanzheng@126.com

³ Key Laboratory of Network Information System Technology, Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China

⁴ Key Laboratory of Target Cognition and Application Technology(TCAT), Beijing 100190, China

⁵ School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100190, China

* Correspondence: liuhu@zjou.edu.cn

† These authors contributed equally to this work.

Abstract: The prediction of vessel trajectories plays a crucial role in ensuring maritime safety and reducing maritime accidents. Substantial progress has been made in trajectory prediction tasks by adopting sequence modeling methods, containing recurrent neural networks (RNNs) and sequence-to-sequence networks (Seq2Seq). However, (1) most of these studies focus on the application of trajectory information, such as the longitude, latitude, course, and speed, while neglecting the impact of differing vessel features and behavioral preferences on the trajectories. (2) Challenges remain in acquiring these features and preferences, as well as enabling the model to sensibly integrate and efficiently express them. To address the issue, we introduce a novel deep framework VEPO-S2S, consisting of a Multi-level Vessel Trajectory Representation Module (Multi-Rep) and a Feature Fusion and Decoding Module (FFDM). Apart from the trajectory information, we first defined the Multi-level Vessel Characteristics in Multi-Rep, encompassing Shallow-level Attributes (vessel length, width, draft, etc.) and Deep-level Features (Sailing Location Preference, Voyage Time Preference, etc.). Subsequently, Multi-Rep was designed to obtain trajectory information and Multi-level Vessel Characteristics, applying distinct encoders for encoding. Next, the FFDM selected and integrated the above features from Multi-Rep for prediction by employing both a priori and a posteriori mechanisms, a Feature Fusion Component, and an enhanced decoder. This allows the model to efficiently leverage them and enhance overall performance. Finally, we conducted comparative experiments with several baseline models. The experimental results demonstrate that VEPO-S2S is both quantitatively and qualitatively superior to the models.

Keywords: deep learning; vessel trajectory prediction; sequence-to-sequence network; maritime accidents; risk assessment; vessel portrait



Citation: Yang, X.; Han, Z.; Zhang, Y.; Liu, H.; Liu, S.; Ai, W.; Liu, J. VEPO-S2S: A Vessel Portrait Oriented Trajectory Prediction Model Based on S2S Framework. *Appl. Sci.* **2024**, *14*, 6344. <https://doi.org/10.3390/app14146344>

Academic Editor: Phivos Mylonas, Katia Lida Kermanidis and Manolis Maragoudakis

Received: 3 June 2024

Revised: 13 July 2024

Accepted: 16 July 2024

Published: 20 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The shipping industry has become more important in the global economy, accounting for over 90% of global freight in recent decades [1]. Consequently, ensuring maritime safety and enhancing sailing efficiency has become even more urgent. The use of an automatic identification system (AIS) for predicting ship trajectories can prevent collisions and provide risk assessment for regulators. Specifically, this task involves forecasting future paths based on historical trajectory points. Some algorithms [2,3], such as the Kalman filter and support vector machines, enable relatively accurate predictions. However, these

models are often constrained by simplifications and exhibit mediocre performance when confronted with more complex situations [4].

Today, deep learning has made significant progress and has found broad application across diverse domains. Recurrent neural networks (RNNs), as time series prediction models, have been widely applied in trajectory prediction, but they suffer from issues such as gradient vanishing and exploding. In recent years, researchers have consistently improved trajectory prediction approaches based on RNNs and achieved noteworthy results. The authors of [5,6] proposed a GRU-based model to capture the temporal dynamics of trajectory sequences. This model can learn the nonlinear and complex relationships between inputs and outputs, encoding the historical motion patterns of vessels. The authors of [7] proposed a trajectory model based on long-short-term memory (LSTM) that learns vessel movement patterns from the current environment and time. The authors of [8] proposed a trajectory prediction method combining bidirectional long short-term memory (BiLSTM) and density-based spatial clustering of applications with noise (DBSCAN). This method integrates vessel trajectory patterns detected using DBSCAN to further enhance the performance. The authors of [9,10] attempted to incorporate attention mechanisms to capture crucial information. However, these methods can only predict one point sequentially, resulting in rapid error accumulation in multi-step predictions. The emergence of Seq2Seq models has significantly alleviated this issue. Seq2Seq is a type of encoder–decoder neural network that is initially used in the field of machine translation, and it has been widely applied to trajectory prediction. It supports multi-point output in a single iteration, effectively reducing the error accumulation. The authors of [11] developed a model based on ConvLSTM and Seq2Seq, enhancing the ability to capture global temporal dependencies. The authors of [12] divided the sea area using a spatial grid based on the Seq2Seq model and achieved good results in long-term prediction. The authors of [13] proposed the METO-S2S model, which employs a multi-semantic decoder, taking into account the effects of various ship semantic data on trajectory forecasting. In addition to methods based on RNNs that utilize temporal information, another explored approach involves utilizing spatial information for modeling, with graph convolutional networks (GCNs) being the most representative. To address the issue of spatiotemporal dependencies, [14] combined a k-GCN with LSTM, using the GCN to capture spatial correlations between nodes and the LSTM to handle spatiotemporal correlations of nodes, enabling the prediction of vessel speeds. The authors of [15] introduced a DAA-SGCN model, utilizing an ST-GCN to extract spatial social interaction features and an RT-CNN to extract temporal features, fully considering the social interactions between vessels. The authors of [16] not only considered the vessel's own intentions but also took into account the impact of the static environment and surrounding dynamically interacting agents. This research largely focused on applying trajectory information for prediction and achieved noteworthy achievements. However, due to the intricate dependencies in historical information and the strong influence of spatial correlations, only relying on trajectory information makes it difficult to attain precise predictive outcomes. Moreover, Multi-level Vessel Characteristics, such as vessel attributes and Sailing Location Preferences, also play a crucial role in trajectory prediction. According to ship maneuverability standards [17], course stability and turning ability are crucial metrics for maneuverability, dependent on the block coefficient, which is determined by a vessel's attributes. Variations in a vessel's attributes significantly impact maneuverability, thereby affecting decisions regarding ports, fairways, and routes. Furthermore, Sailing Location Preferences reveal their tendencies toward specific maritime areas, which should receive more attention in predictions. As depicted in Figure 1, two types of vessels exhibit distinct motion trajectories. Compared to trawlers, cargo ships typically have larger volumes and higher block coefficients, resulting in a larger turning radius and poorer course stability. To mitigate the potential risks, cargo ships tend to select broader shipping lanes and fairways, strictly adhering to established schedules to ensure punctual cargo delivery and enhance overall logistical efficiency, resulting in smoother and more regular sailing trajectories. Conversely, trawlers operate within specific fishing areas, constrained by the distribution of fishery resources

and relevant regulatory policies, often resulting in irregular and concentrated navigation paths. Therefore, it is crucial to investigate the behavioral patterns of different vessels and conduct tailored predictive analyses based on vessel attributes and operational areas.

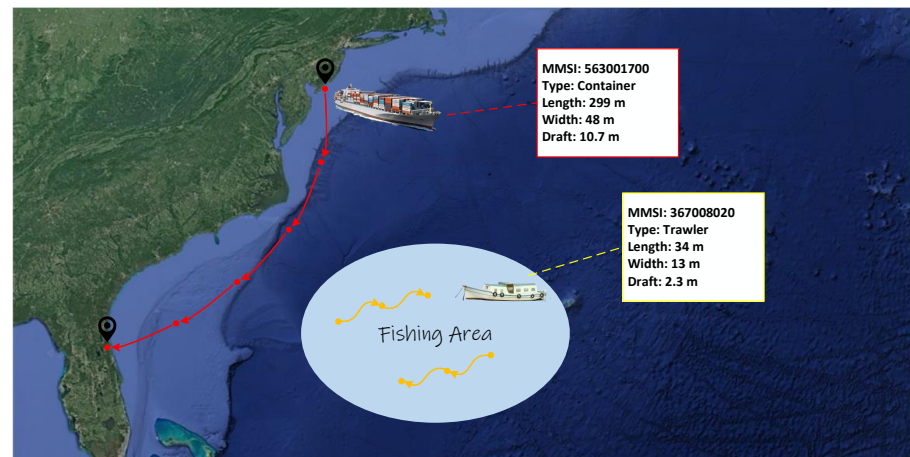


Figure 1. Trajectory examples for various vessel types, illustrate significant differences in navigation trajectories under the influence of various vessel attributes and types.

For this reason, challenges still persist in obtaining more comprehensive characteristics, along with their advisable selection and implementation. Inspired by user personas [18], we incorporated Shallow-level Attributes and Deep-level Features, defining Multi-level Vessel Characteristics to construct a comprehensive vessel portrait. Considering the aforementioned challenges, we propose a vessel trajectory prediction model VEPO-S2S based on the Seq2Seq architecture, comprising a Multi-level Vessel Trajectory Representation Module (Multi-Rep) and a Feature Fusion and Decoding Module (FFDM). Multi-Rep serves the function of acquiring and expressing features, consisting of two components: the Feature Acquisition Component and the Feature Expression Component. In the Feature Acquisition Component, we first specify the trajectory information that includes the longitude, latitude, speed, course, and sailing distance. Then, Multi-level Vessel Characteristics are defined, covering Shallow-level Attributes (such as the length, width, draft, etc.) as well as Deep-level Features (Sailing Location Preference, Voyage Time Preference, etc.). All of these are acquired through the Feature Acquisition Component and then encoded separately using three independent encoders within the Feature Expression Component.

Apart from that, despite the incorporation of trajectory information and vessel characteristics into the model, basic Seq2Seq models encounter challenges in discerning and leveraging them efficiently. Therefore, it is imperative to select and integrate the trajectory information and vessel characteristics before applying them. To achieve this purpose, we propose the FFDM module that consists of a Portrait Selection Component, Feature Fusion Component, and Multi-head Decoding Component. At first, the Portrait Selection Component discerns the most relevant vessel characteristics for the current prediction environment via analyzing the encoded characteristics. Then, the Feature Fusion Component is designed to merge trajectory information from the Multi-Rep module with relevant vessel characteristics. Finally, the output serves as the input for the Multi-head Decoding Component, which is designed based on the traditional Seq2Seq decoder. The Multi-head Decoding Component consists of two distinct GRU blocks, each controlling the proportion of trajectory information and vessel characteristics during prediction, providing more precise output results.

In summary, the main contributions of this paper can be summarized as follows:

- We propose a vessel trajectory prediction framework VEPO-S2S, which encompasses the multi-level vessel trajectory representation (Multi-Rep) module and Feature Fusion and Decoding Module (FFDM). This framework fully takes into account trajectory

information and vessel characteristics, ensuring their sensible integration and efficient expression to achieve more accurate results.

- We propose the Multi-Rep module, which integrates trajectory information with Multi-level Vessel Characteristics and employs multiple encoders for encoding. This module has the ability to capture temporal representations of the trajectories as well as the detailed portrait of the vessels.
- To address the challenge of effectively fusing and representing multiple characteristics within our model, we propose the FFDM. This module selects and integrates characteristics by employing a priori and a posteriori mechanisms, a Feature Fusion Component, and an enhanced decoder. The FFDM can better represent the spatiotemporal correlation among historical trajectories.
- We conducted comparative experiments on several baseline models. The experimental results demonstrated that the VEPO-S2S outperformed other baseline models in both quantitative and qualitative aspects, producing more robust and accurate prediction results <https://github.com/AIR-SkyForecast/AIR-SkyForecast-VEPO-S2S/new/main> (accessed on 15 July 2024).

2. Related Works

2.1. Vessel Trajectory Prediction

Traditional trajectory prediction methods have achieved favorable results in forecasting trajectories for vehicles, ships, and pedestrians. The authors of [19] proposed a dynamically assisted inertial navigation method for estimating observed values. The authors of [20] introduced a mathematical modeling-based Kalman filtering method for long-range surface tracking, enabling the direct prediction of the target position and heading without requiring coordinate system conversion. To improve ship motion prediction accuracy under environmental disturbances, ref. [21] proposed a ship motion recognition algorithm based on the least squares method. However, those methods exhibited limited predictive accuracy when faced with complex situations.

In recent years, the development of deep learning methods for vessel trajectory prediction has progressed rapidly, and significant advancements have been made in this task. Most research adopts the RNN structure. Some research is based on an LSTM [22] or a GRU [6]. Moreover, to investigate ship prediction under varying trajectory densities, ref. [23] proposed a model based on an LSTM and the K-nearest neighbor (KNN). The authors of [24] introduced the MP-LSTM model, which integrates the strengths of TPNet and LSTM, addressing the shortcomings of existing methods in terms of both the accuracy and model complexity.

Meanwhile, some scholars attempted to use Seq2Seq architecture to address prediction problems. The authors of [25] introduced a neural network model based on an LSTM and Seq2Seq, utilized to capture long-term dependencies in historical data within trajectories. The authors of [26] proposed the ST-Seq2Seq model based on GRU architecture. The authors of [27] proposed a trajectory prediction model based on BiGRU and Seq2Seq, which fully considers the interactions among ships. Furthermore, several other trajectory prediction models have been proposed. The authors of [13] introduced the METO-S2S model based on a multi-semantic decoder, considering the influence of various ship semantic information on trajectory prediction. They also used semantic vectors (SLV) to guide model predictions in PESO [28], achieving outstanding results on the open-source AIS dataset in the United States. In addition to the Seq2Seq model, [29] applied the Transformer framework combined with LSTM to capture historical trajectories in time series and overcome issues related to distant information decay. To express the interdependence between ships, [30] proposed the spatiotemporal multi-graph convolutional neural network (STMGCN) model, which models both spatiotemporal data and ship types separately. The authors of [31] combined graph attention convolution (GAT) with an extended causal convolution structure and designed the GAGW model. The graph attention convolution network is responsible for extracting interaction information between different ships in space.

The majority of the aforementioned studies primarily focus on the utilization of shallow-level trajectory information. These studies typically use speed, course, and position as model inputs. However, this is insufficient for guiding ship avoidance in an intricate situation. Acquiring richer and deeper characteristics, as well as their sensible application, is crucial for guiding ship avoidance and overall route planning. Therefore, current research on vessel trajectory prediction will pay more attention to excavating the abundant characteristics and understanding the dynamics of real-world environments.

2.2. Seq2Seq Model

The Seq2Seq model has been widely applied in the field of machine translation [32], which consists of an encoder and a decoder, where the encoder embeds the input information and generates a high-dimensional semantic vector, while the decoder decodes it and outputs the result. We mainly present the related research in regression tasks based on Seq2Seq, including power prediction, runoff prediction, and stock prediction.

In power forecasting, ref. [33] proposed a Seq2Seq model based on an LSTM that takes into account the inherent correlation within the data, effectively capturing the sequential relationships in time series. To address the problem of low accuracy in short-term temperature predictions, a Seq2Seq-based model was proposed by [34]. In the domain of runoff prediction, ref. [35] made improvements to the Seq2Seq by replacing the RNN structure with a linear layer to handle historical data. In addition, the introduction of an attention mechanism led to a higher prediction accuracy. In [36], TEN-Seq2Seq was introduced for handling tabular data and well depths, which exhibited better robustness compared to LSTM and FCNN. The authors of [37] proposed a novel method to predict a reservoir level using LSTM and attention mechanism-based Seq2Seq modeling. The authors of [38] proposed a structure for stock price predictions based on Seq2Seq networks.

The Seq2Seq model has also made significant progress in the field of sea surface temperature (SST). The authors of [39] applied the Seq2Seq model with two-module attention (TMA-Seq2Seq) for long-term time series SST prediction, receiving superior performance compared to data-driven methods. In [40], a novel Seq2Seq network was proposed to achieve the k-step-ahead prediction based on the characteristics of sea clutter. The authors of [41] utilized the Seq2Seq model to provide a spatiotemporal forecast of the probability of sea ice, leading to higher accuracy.

2.3. User Personas

User personas are a product of internet development, which allows the discovery of differences among individuals within groups. The authors of [42] proposed an employee user persona model based on neural networks, which establishes personas according to employee skill levels and mental states, enabling personalized job recommendations for enterprise employees. The authors of [43] proposed a method for web service hybrid recommendations based on user personas to address the cold start problem for new users, improving both the accuracy and recommendation quality.

Recently, predicting future behavior based on user profiles has become a popular direction. The authors of [44] transformed users' emotional preference features into attention information and combined them with LSTM models to predict the personality traits of online users. The authors of [45] proposed the T-LSTM model for user occupation prediction, overcoming challenges in predictive performance which offers a novel and effective approach for accurate user occupation prediction. The authors of [46] introduced a method for predicting impulsive rewards in minors using user profiles, facilitating accurate forecasts of impulsive reward behaviors in underage users. The authors of [47] applied persona prediction in the field of academic warnings for university students. Constructing student personas to explore the relationship between student factors and academic performance provides strong guidance for teachers and administrators to adjust teaching plans.

In this work, we created a profile for each ship and introduced a novel Seq2Seq-based model, which proves to be more suitable in practice for collision detection and risk warning.

3. Proposed Method

We present the method in three parts. First, we provide definitions and the problem statement. Next, we provide a comprehensive overview of the data processing. Then, we describe the detailed process of constructing the vessel portrait. Finally, we provide a comprehensive description of our proposed model VEPO-S2S including the Multi-level Vessel Trajectory Representation Module and the Feature Fusion and Decoding Module.

4. Definitions and Problem Statement

The objective of VEPO-S2S is to predict the future trajectory of a vessel based on AIS data. To articulate our approach more clearly, we provide the following definitions:

[Vessel Trajectory] A trajectory point is defined as a tuple $x_t = (lon_t, lat_t, sog_t, cog_t, dist_t, l, w, d, t, \alpha, \beta, \gamma)$ at time t , in which x_t is composed of longitude lon_t , latitude lat_t , speed sog_t , course cog_t , sailing distance $dist_t$, length l , width w , draft d , type t , Sailing Location Preference α , Voyage Time Preference β , and Anchoring Time Preference γ , respectively. A vessel trajectory $X = (x_{t_0}, x_{t_1}, \dots, x_{t_n})$ is defined as a chronological sequence, where $\{t_i, i = 0, 1, 2, \dots, n\}$ is a set of timestamps.

[Position Sequence] The position of the ship at time t is defined as a tuple $y_t = (lon_t, lat_t)$, and the sequence of positions of the vessel at time $(1, 2, \dots, t)$ is defined as $Y = (y_1, y_2, \dots, y_t)$.

[Vessel Trajectory Prediction] Given an observed trajectory $X = (x_1, x_2, \dots, x_t)$ at timestamp $(1, 2, 3, \dots, t)$, the objective is to predict the trajectory $Y = (y_{t+1}, y_{t+2}, \dots, y_{t+k})$ at the following timestamps $(t + 1, t + 2, \dots, t + k)$.

4.1. Data Preprocessing

AIS data preprocessing is essential for training deep learning models, especially for models that require trajectory information and vessel characteristics. In VEPO-S2S, we selected AIS data from southwestern and southeastern coastal waters in the US for training, validation, and testing. The dataset includes static attributes such as the Maritime Mobile Service Identity (MMSI), vessel length, and width. Additionally, it encompasses dynamic information of vessel navigation such as the longitude, latitude, speed, and course. The original AIS data may experience adverse weather conditions during the reception process, leading to signal transmission delays and reception errors [48]. Moreover, the performance of deep learning models could be adversely affected by data loss resulting from technical issues and equipment maintenance. Therefore, we conducted comprehensive preprocessing of the AIS data before training (see Figure 2).

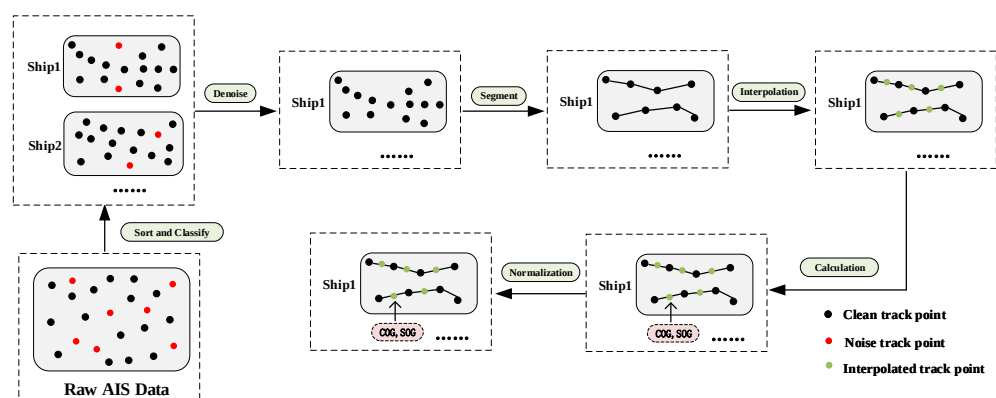


Figure 2. The process of data preprocessing.

The process is shown in the following steps:

(1) Sort and Classify: We filtered vessels with complete information on the length, width, draft, and type, then separated the trajectory data of each vessel based on the Maritime Mobile Service Identity (MMSI) number, and sorted them in ascending order of timestamps.

(2) Denoise: We removed points with duplicate timestamps and unreasonable longitude and latitude.

(3) Segment: We separated the trajectory into different segments when the time interval between two adjacent trajectory points exceeded 60 min or when the distance between three consecutive trajectory points was less than 100 m.

(4) Interpolate: We employed cubic spline interpolation to ensure a 10-min interval between consecutive trajectory points.

(5) Compute: We computed the course and speed for each trajectory point.

(6) Normalize: We normalized the longitude, latitude, speed, course, length, width, and draft using the min–max normalization method, as expressed in Equation (1)

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

where x is the original data, $x_{\min}$ and $x_{\max}$ represent the minimum and maximum value in the trajectory data, respectively. x_{norm} is the normalized data.

4.2. Vessel Portrait Construction

This chapter accomplishes the construction of vessel user portraits based on AIS data, including the establishment of a label system and the creation of vessel portraits.

4.2.1. Label System Construction

As shown in Figure 3, we established a label system based on Shallow-level Attributes and Deep-level Features. The Shallow-level Attributes included a series of fundamental attributes of a vessel (such as the length, width, draft, and type). These attributes significantly impact the maneuvering performance of vessels. According to ship maneuverability standards [17], both course stability and turning ability are pivotal indicators of maneuverability and are affected by the ship's block coefficient. The block coefficient is defined as the displacement of a ship divided by the product of its length, width, and draft. Moreover, different types of ships have different block coefficients due to variations in the shape of their underwater hulls. For vessels of the same displacement, ships with smaller block coefficients (such as container ships) exhibit better course stability but poorer turning ability than those with larger block coefficients (such as tankers). Therefore, they require wider navigational fairways to reduce the risk of collisions with other vessels. In addition, these attributes (length, width, draft, and type) also play a crucial role in the selection of fairways, ports, and routes. According to coastal engineering manuals [49], the fairway width is typically two to five times the ship's breadth. Vessels must consider both the width and depth when navigating to ensure safety and efficiency. In port selection, according to the PIANC [50], large vessels need to choose ports with sufficient berth and maneuvering space to ensure safe berthing. In route planning, vessels must consider their turning radius and draft, choosing suitable routes to avoid the risk of grounding or collision. Consequently, these attributes are crucial for feasibility and must be thoroughly considered to ensure more accurate predictions of different vessels.

To model trajectory information and Multi-level Vessel Characteristics more effectively, we take into account not only Shallow-level Attributes but also Deep-level Features. The Deep-level Features are defined as the Sailing Location Preference, the Voyage Time Preference, and the Anchoring Time Preference. The Sailing Location Preference reflects the behavioral pattern of the ship. For example, container ships engaged in liner shipping, typically operate on fixed routes and within port areas for cargo handling and transport [51]. The fixed routes and regular schedules of liner shipping ensure logistics timeliness, reducing losses and enhancing revenue. Meanwhile, trawlers primarily operate in specific fishing areas [52], where their Sailing Location Preferences are influenced by the distribution of fisheries resources. Unlike liner shipping, trawlers have a more flexible navigation pattern, often adjusting their fishing locations based on the season, to comply with regulatory constraints and to increase revenue. This preference provides a more comprehensive

understanding of vessel behavior and improves the accuracy of trajectory prediction. Regarding the Voyage Time Preference and Anchoring Time Preference, container ships tend to minimize the anchorage time [53], strictly adhering to schedules to optimize operational efficiency. This operational mode not only ensures the timely transportation of goods but also helps to reduce the operating costs. In contrast, trawlers' Voyage Time Preferences are more influenced by fisheries management regulations and market demands. This temporal information contributes to a deeper understanding of vessel behavior patterns and empowers prediction models to precisely capture fluctuations in vessel movements over time.

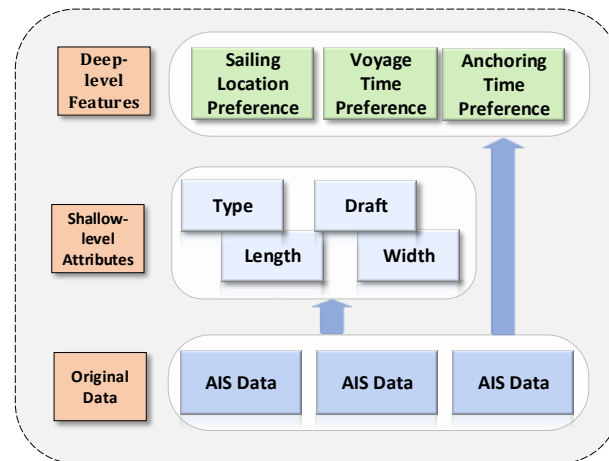


Figure 3. Label-level modeling and analysis process.

4.2.2. Vessel Portrait Construction

A vessel portrait consists of Shallow-level Attributes and Deep-level Features. Regarding the processing of Shallow-level Attributes, we employ the following approach: firstly, select AIS data with non-empty attributes (such as length, width, draft, and type). And then we randomly select 100 data points based on the Maritime Mobile Service Identity (MMSI). For each attribute, consider the value with the highest frequency as the current vessel's attribute to construct the vessel's shallow profile. This process can be expressed in the formula as Equation (2)

$$Y = \{y_i | y_i = \{x_1^*, x_2^*, \dots, x_m^*\}, D_k \subset X, \forall x_j^*, P(x_j^* | D_k) \rightarrow \max\} \quad (2)$$

where Y represents the Shallow-level Attributes of all vessels, and each element y_i denotes those of the i -th vessel. D_k represents the AIS data collection with complete attributes, X denotes the attribute set of all AIS data, and each element x_j^* represents the highest frequency value for D_k .

After obtaining Shallow-level Attributes, we focused on the process of acquiring Deep-level Features, which include the Sailing Location Preference, the Voyage Time Preference, and the Anchoring Time Preference. For the Sailing Location Preference, due to the difference in the quantity and distribution of ship trajectory points, we employed HDBSCAN (hierarchical density-based spatial clustering of applications with noise) [54] for cluster analysis. The clustering results are shown in Figure 4. Different colors represent different clusters, and the black labels denote the clustering centers. Meanwhile, to capture the Voyage Time Preference and the Anchoring Time Preference, we divided a day into 24 segments and assigned each vessel's trajectory points to the corresponding periods. The distribution of trajectory points in each period reflects the temporal preferences of the vessels. After processing, each vessel's profile can be expressed as Equation (3)

$$S_{mmsi} = \{l_{mmsi}, w_{mmsi}, d_{mmsi}, t_{mmsi}, \alpha_{mmsi}, \beta_{mmsi}, \gamma_{mmsi}\} \quad (3)$$

where l_{mmsi} , w_{mmsi} , d_{mmsi} , and t_{mmsi} , respectively, represent the Shallow-level Attributes of the length, width, draft, and type, α_{mmsi} represents the Voyage Time Preference, β_{mmsi} stands for the Anchoring Time Preference, and γ_{mmsi} is the Sailing Location Preference. Whereas α_{mmsi} and β_{mmsi} are transformed into two 24-dimensional features, γ_{mmsi} is converted into a 114-dimensional feature. The utilization of the vessel portrait is elaborated in Section 4.3.1.

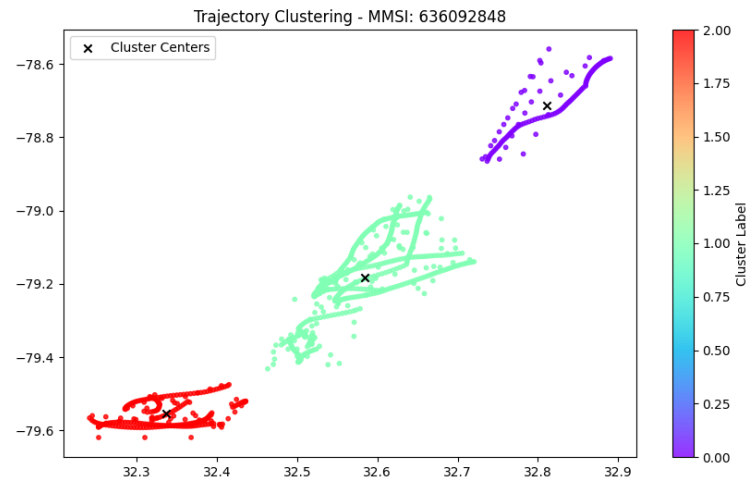


Figure 4. The trajectory cluster result. Different colors represent different clusters, and the black labels indicate cluster centers.

4.3. VEPO-S2S Model

We propose a novel trajectory prediction model VEPO-S2S based on the Seq2Seq model, and the structure is shown in Figure 5. As the figure shows, VEPO-S2S consists of the Multi-level Vessel Trajectory Representation Module (Multi-Rep) and the Feature Fusion and Decoding Module (FFDM). The Multi-Rep is targeted to acquire trajectory information and Multi-level Vessel Characteristics, encoding them with several encoders. The FFDM is designed to select and merge the above information and characteristics from Multi-Rep for prediction.

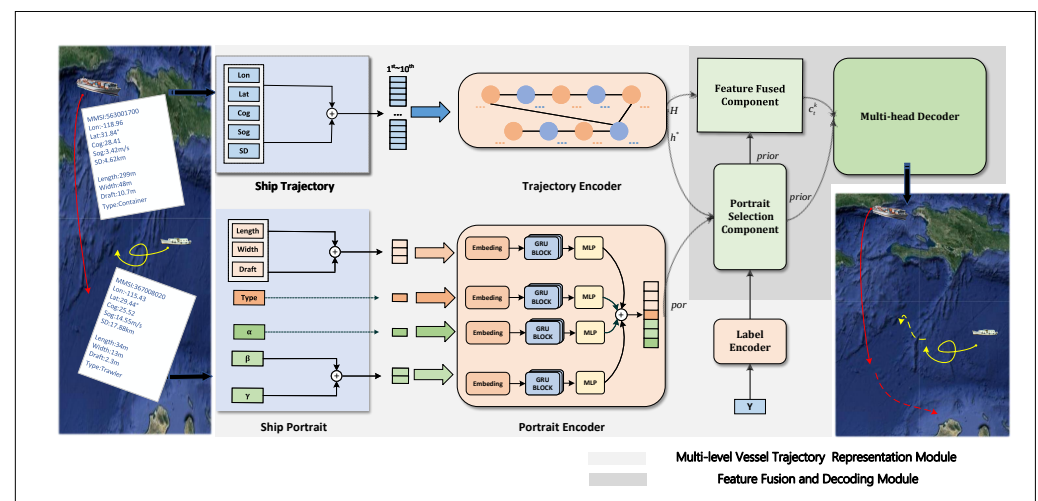


Figure 5. The structure of VEPO-S2S comprises the Multi-level Vessel Trajectory Representation Module (Multi-Rep) and the Feature Fusion and Decoding Module (FFDM). The Multi-Rep is designed to obtain trajectory information and Multi-level Vessel Characteristics, applying distinct encoders for encoding. The FFDM is targeted to select and integrate the above characteristics from Multi-Rep for prediction.

4.3.1. Multi-Level Vessel Trajectory Representation Module

The Multi-level Vessel Trajectory Representation Module is designed to acquire trajectory information and Multi-level Vessel Characteristics and apply distinct encoders for encoding. In this subsection, we introduce the Multi-level Vessel Trajectory Representation Module, which consists of the Feature Acquisition Component and the Feature Representation Component. For the Feature Acquisition Component, we obtained trajectory information and Multi-level Vessel Characteristics through data preprocessing, as described in Section 4.1, and vessel portrait construction, as described in Section 4.2. Simultaneously, building on the RNN and Seq2Seq models, we introduced the Feature Representation Component, consisting of three distinct encoders. Those encoders are designed to separately handle different input characteristics from the Feature Acquisition Component. The trajectory encoder is responsible for encoding the trajectory information (including the longitude, latitude, speed, course, and navigation distance). This process can be expressed by the following Equation (4)

$$\begin{aligned} X_{traj} &= (x_1, x_2, \dots, x_{10}) \\ x_n &= (lon_n, lat_n, sog_n, cog_n, dis_n) \\ H, h, h^* &= Enc_{traj}(X_{traj}) \end{aligned} \quad (4)$$

where X_{traj} represents the trajectory information, including the normalized longitude, latitude, speed, course, and navigation distance for ten trajectory points. $H = [h_1, h_2, \dots, h_{10}]$ signifies the hidden state at each time step, h^* denotes the hidden state at the final time step, and h represents the ultimate hidden state. Similar to the trajectory encoder, the task of the label encoder Enc_{label} is to encode the gold trajectory and output the encoded state H_y , where Y_{label} represents five trajectory points containing the longitude and latitude.

$$\begin{aligned} Y_{label} &= (x_{11}, x_{12}, \dots, x_{15}) \\ y_n &= (lon_n, lat_n) \\ H_y &= Enc_{traj}(Y_{label}) \end{aligned} \quad (5)$$

The task of the Portrait Feature Encoder is to embed Multi-level Vessel Characteristics into a high-dimensional vector. First, the normalized continuous numerical values (including the length, width, and draft) were concatenated and embedded into an eight-dimensional semantic vector. Second, the discrete vessel type was transformed into a continuous value for model input and individually embedded into another semantic vector. Third, we encoded two 24-dimensional Deep-level Features to capture the Voyage Time Preference and the Anchoring Time Preference (as mentioned in Section 4.2.2), while the Sailing Location Preference was encoded separately. Finally, they were concatenated to form a seven-dimensional feature vector, which was input into the Portrait Feature Encoder for encoding. This process can be expressed by the following Equation (6)

$$\begin{aligned} sp &= \text{concat}(l, w, d) \\ \hat{sp} &= \text{embed}_{sp}(sp) \\ \widehat{type} &= \text{embed}_{type}(type) \\ tim &= \text{concat}(\alpha, \beta) \\ \widehat{tim} &= \text{embed}_{tim}(tim) \\ \hat{\gamma} &= \text{embed}_{\gamma}(\gamma) \\ por &= \text{concat}(Enc_{por}(\hat{sp}, \widehat{type}, \widehat{tim}, \hat{\gamma})) \end{aligned} \quad (6)$$

where por represents the portrait feature, and embed_{sp} , embed_{type} , embed_{tim} , and embed_{α} are the embedding layers. Enc_{por} is the Portrait Feature Encoder. l , w , and d represent the length, width, and draft, respectively, t is the type, α represents the Voyage Time Preference,

β represents the Anchoring Time Preference, and γ is the Sailing Location Preference (as mentioned in Equation (3)).

4.3.2. Feature Fusion and Decoding Module

Despite incorporating trajectory information and vessel characteristics into the model, the basic Seq2Seq models still have difficulty efficiently discerning and leveraging. Therefore, the Feature Fusion and Decoding Module was designed to select and integrate the trajectory information with Multi-level Vessel Characteristics, applying a priori and a posteriori mechanisms in the Portrait Selection Component, a Feature Fusion Component, and a Multi-head Decoder Component. The goal of the Portrait Selection Component is to identify and select suitable Multi-level Vessel Characteristics for prediction; hence, we use a prior distribution and a posterior distribution together in the vessel characteristic selection, and the framework is shown in Figure 6. The prior distribution selects the characteristics based on the similarity between the vector h^* from the trajectory encoder and the portrait vector por , which helps to filter out the more relevant characteristics in the early stages of the model, reducing the computational overhead. This process can be expressed as Equation (7)

$$Z_{prior}(por = por_i|h^*) = \frac{\exp(por_i * h^*)}{\sum_{i=1}^7 \exp(por_i * h^*)} \quad (7)$$

where $por_i \in (por_1, por_2, \dots, por_7)$, $*$ is the dot product, h^* is the trajectory coding vector, and por is the portrait vector. Specifically, the model assigns higher weights to vectors with greater similarity by comparing the dot product results of different characteristics, reducing the interference of redundant information and increasing the computational efficiency.

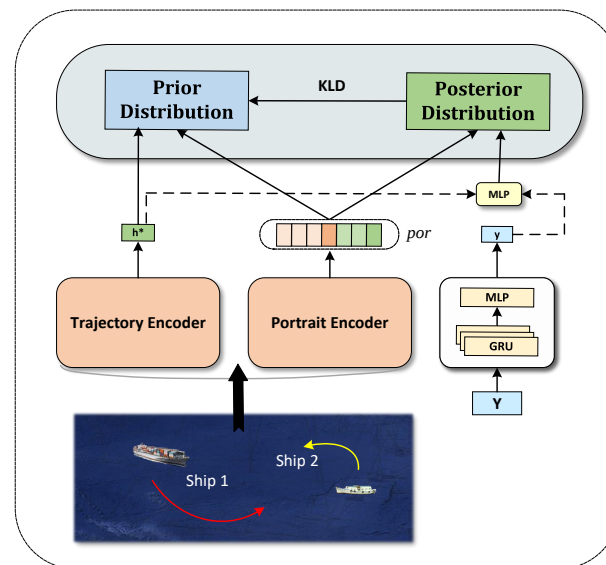


Figure 6. The Portrait Selection Component consists of prior distribution and posterior distribution. Prior distribution expresses the trajectory coding vector h^* and portrait feature por , and the posterior distribution incorporates the label y to improve the accuracy of selection. Meanwhile, the KLD is designed to bridge the gap between the prior distribution and the posterior distribution, allowing the prior distribution to benefit from the posterior distribution and generate more accurate results.

However, only relying on the prior distribution can not enable obtaining accurate results, as it is typically based on assumptions or historical data, which do not fully reflect the real situation; hence, it is impossible to select the appropriate characteristics to guide the generation. In contrast, the characteristics used in label y can be obtained through posterior distribution. Therefore, the posterior distribution, derived by combining the trajectory

vector h^* and label y , can more effectively guide the selection of the profile, which can be expressed as

$$Z_{post}(por = por_i | h^*, y) = \frac{\exp(por_i * MLP([h^*; y]))}{\sum_{i=1}^7 \exp(por_i * MLP([h^*; y]))} \quad (8)$$

where MLP is a linear layer, $*$ is the dot product, and $;$ represents the vector splicing.

Simultaneously, there is a significant gap between the prior distribution and the posterior distribution. To address this issue, the Kullback–Leibler Divergence (KLD) loss is employed to compel their proximity. It can effectively correct errors in the prior distribution and guide the profile selection to benefit the model. The stability of the KLD divergence lies in its mathematical properties, ensuring convergence and reliability during training. By minimizing the KLD loss, the system can strike a suitable balance between the prior and posterior distributions. The formula for the KLD divergence is expressed as follows:

$$D_{KL}(P||Q) = \sum_i P(i) \log\left(\frac{P(i)}{Q(i)}\right) \quad (9)$$

where P represents the posterior distribution, which comprises the characteristics required under the guidance of real labels.

In general, a straightforward approach to leverage the selected characteristics for result generation is to directly append these characteristics to the encoder's input. However, this approach usually fails to yield satisfactory results due to the lack of flexibility in controlling the degree of characteristic involvement introduced. Therefore, we introduce the Feature Fusion Component to optimize the utilization of the characteristics. Compared with directly connecting those characteristics, we used a more flexible way to integrate them. Here, we applied an LSTM to fuse the prior distribution $prior$ and the historical trajectory H , and the results of the LSTM took into account the continuity and correlation between characteristics over time. Additionally, $prior$ served as the initial hidden state of the LSTM, and the trajectory representation H obtained by the trajectory encoder was used as the input at each step. Finally, we obtained the fused semantic vector c_t^k , and the process can be expressed through the following following Equation (10)

$$c_t^k = LSTM(prior, H) \quad (10)$$

To regulate the involvement of Multi-level Vessel Characteristics in the prediction, we introduced the Multi-head Decoding Component; the framework diagram is depicted in Figure 7. This component comprises two GRU blocks and a fusion unit that efficiently synthesizes the hidden states generated by the two GRU blocks to predict future trajectories. The design is formulated to adjust the weighting between the trajectory information and the vessel characteristics during the prediction process. The orange region is a standard GRU module that receives the trajectory information h . Additionally, it takes the preceding prediction value $y - 1$ as input, producing its hidden state T_i^n . Another GRU is dedicated to adding the prior distribution por into the predictions; it also takes the fused semantic vector c_t^k , trajectory information h , and the preceding prediction value $y - 1$ as inputs, generating the feature representation T_i^p . Ultimately, T_i^n and T_i^p are fused through the fusion gate to produce the final trajectory. This process can be expressed by Equation (11):

$$\begin{aligned} T_i^n &= GRU^n(y - 1, h) \\ inp_p &= concat(y - 1, c_t^k, prior) \\ T_i^p &= GRU_p(inp_p, h) \\ O_t &= \sigma(W_z * [\tanh(w_y * T_i^n); \tanh(W_k * T_i^p)]) \\ T_i &= O_t * T_i^n + (1 - O_t) * T_i^p \end{aligned} \quad (11)$$

where $i \in (11, 12, \dots, 15)$, and W_z , W_y , and W_k correspond to weight matrices with different coefficients. σ and $\tanh$ are the sigmoid activation function and tanh activation function, respectively.

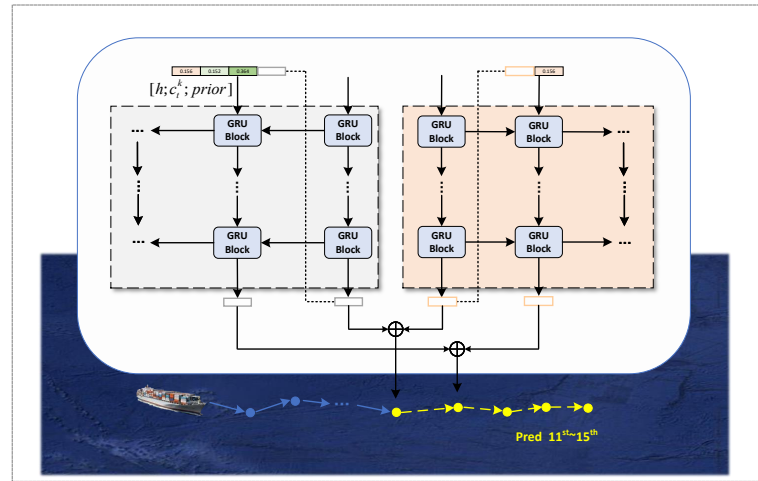


Figure 7. The Multi-head Decoder Component consists of two GRU blocks and a fusion unit. It can flexibly adjust the weighting between the trajectory information and the vessel characteristics during the prediction process.

4.3.3. Loss Function

The VEPO adopts the root mean square error (RMSE) and the Kullback–Leibler Divergence (KLD) (as mentioned in Equation (9)) as the loss function. The objective of this work is to utilize the first m points denoted by $X_k = (x_1^k, x_2^k, \dots, x_m^k)$ to predict the subsequent n track points, where m and n are the hyperparameters. The prediction sequence and target sequence are represented by $Y_k = (y_{m+1}^k, y_{m+2}^k, \dots, y_{m+n}^k)$, and $\hat{Y}_k = (\hat{y}_{m+1}^k, \hat{y}_{m+2}^k, \dots, \hat{y}_{m+n}^k)$, respectively. The primary aim is to minimize the loss function during training, ensuring greater accuracy in predicting the last n trajectory points. The expression for the loss function is shown in the following Equation (12).

$$\mathcal{L} = \sqrt{\frac{1}{n} \sum_{k=1}^n (Y_k - \hat{Y}_k)^2} + D_{\text{KL}} \quad (12)$$

5. Experiments

To validate the effectiveness of the VEPO model, we conducted a series of quantitative and qualitative experiments. Specifically, we first introduce the experiment settings, including the hyperparameter settings, experimental environments, datasets, baseline models, and evaluation metrics. Subsequently, we present the quantitative comparison results of our proposed method and other baseline models. Following this, we describe the ablation experiments conducted to substantiate the effectiveness of different components of the model. Finally, we depict the prediction results of VEPO-S2S through qualitative analysis.

5.1. Experiment Settings

5.1.1. Dataset

We adopted AIS data (<https://marinecadastre.gov/accessais/>) (accessed on 1 June 2024) from the coastal waters of the southeastern and southwestern United States for training, validation, and testing [13]. As shown in Table 1, there are 68 types of vessels and 28,645 vessels in this dataset. At the same time, due to the particularity of the portrait construction (as mentioned in Section 4.2.2), we selected the vessel attributes (including the length, width, draft, and type). After processing, we obtained 45 types with a total of 6194

vessels. The detailed distribution of the vessel types is shown in Figure 8. In our dataset, the maximum number of ship types was 45, consisting of passenger ships, pleasure crafts, sailing ships, etc. The distribution of ship types was uneven, and they were mainly classified into five categories: passenger ships, pleasure craft, sailing ships, tug tow, and fishing. Passenger ships and pleasure craft together accounted for over 23% of the total. The remaining ships, mainly consisting of cargo ships, container ships, and tankers, collectively constitute 21% of the dataset. To enhance the richness and comprehensiveness of the data, we employed a sliding window approach, dividing the normalized data into groups of 15 trajectory points with a sliding step of 1 (see Figure 9). Subsequently, each vessel was allocated to training, validation, and testing sets in a ratio of 8:1:1. This approach increased the diversity of the data and ensured that each vessel received sufficient validation.

Table 1. The detail of our dataset.

Dataset	Region	Track Points Count	Type Count	Vessel Count
Total	Coast of the United States	144,445,580	68	28,645
Ours	Coast of the United States	4,930,061	45	6194

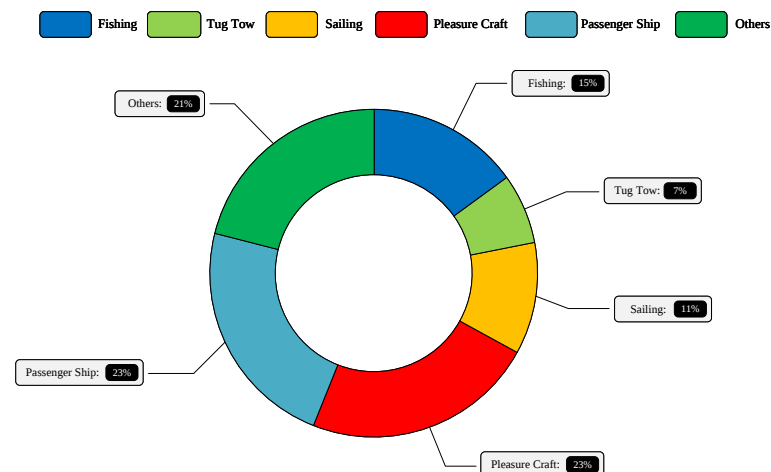


Figure 8. The distribution of vessel type in our processed dataset.

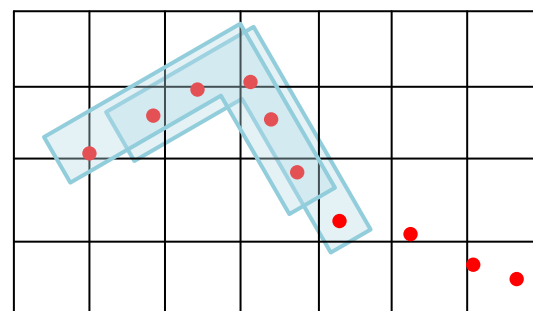


Figure 9. An example of dividing a dataset using the sliding window method, and the red dots represent trajectory points.

5.1.2. Hyperparameters Settings and Experimental Environment

We utilized ten trajectory points as inputs to predict the last five trajectory points, and those two hyperparameters can be flexibly adjusted to adapt to different tasks. Meanwhile, the epochs were set to 40 for model training, and the best valuation performance was saved. The learning rate was 0.001 with a weight decay of 0.0, the optimizer was set to Adam, and the batch size was set to 128. Moreover, the hidden sizes in the GRU were

set to 64, and the hidden layer number was 2. In the above hyperparameters, we selected two typical cases for visualization, which were the number of layers and the hidden sizes, the details of which are shown in Figures 10 and 11. The experiments were all based on Python 3.8 using the PyTorch framework. We trained the model using the Ubuntu operating system and GTX 3090Ti on the server for experiments.

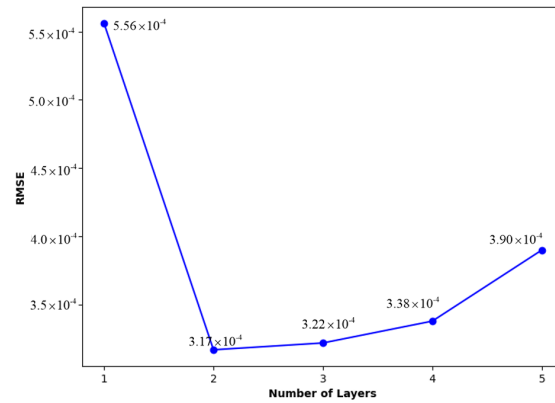


Figure 10. The influence of different GRU layers on the prediction of the VEPO-S2S model, according to the RMSE. The X-axis represents the number of layers, and the Y-axis represents the RMSE loss value.

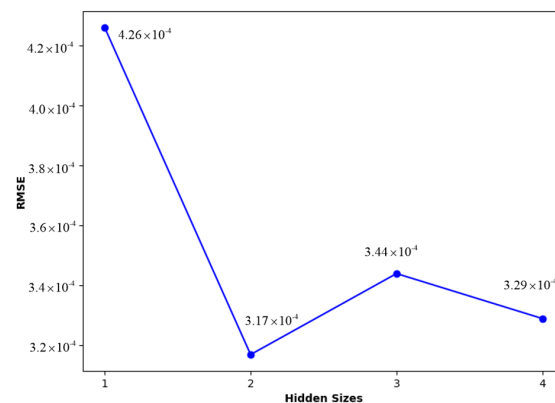


Figure 11. The influence of different hidden sizes on the prediction of the VEPO-S2S model, according to the RMSE. The X-axis represents the number of hidden sizes, and the Y-axis represents the RMSE loss value.

5.1.3. Baselines

To better evaluate the performance of VEPO-S2S, we compared it with several baseline models using the same dataset. In contrast to Seq2Seq, recurrent neural networks are limited to predicting multiple consecutive track points. In our experiments, we continuously predicted five trajectory points using an RNN, which was achieved through the sliding window method. The baselines were as follows:

- (1) Kalman: A linear optimal estimation model;
- (2) VAR: A statistical model for multivariate time-series prediction;
- (3) ARIMA: A statistical time-series forecasting model;
- (4) LSTM: A type of recurrent neural network, which consists of two layers;
- (5) BiLSTM: Similarly to LSTM, the BiLSTM is composed of two bidirectional LSTM layers;
- (6) GRU: Similar to an LSTM;
- (7) BiGRU: Similar to a GRU but with two bidirectional GRU layers;
- (8) LSTM-LSTM: A Seq2Seq model with a two-layer LSTM as the encoder and decoder;

- (9) BiLSTM–LSTM: A Seq2Seq model with a two-layer BiLSTM as the encoder and a four-layer LSTM as the decoder;
- (10) GRU–GRU: Similar to an LSTM–LSTM;
- (11) BiGRU–GRU: Similar to a BiLSTM–LSTM;
- (12) Transformer: A Seq2Seq model based on attention mechanism;
- (13) METO-S2S: An S2S-based vessel trajectory prediction method with a multiple-semantic encoder and a type-oriented decoder.

5.1.4. Evaluation Metrics

To evaluate the prediction performance of the proposed model, we used four evaluation metrics, including the root mean square error (RMSE), the mean absolute error (MAE), the average displacement error (ADE), and the final displacement error (FDE). The RMSE focuses on measuring the stability of the result, while the MAE evaluates the prediction ability of a model. The ADE stands for the average Euclidean distance error between the predicted position and the actual position. Additionally, the FDE focuses on the final accuracy of the predictions.

$$\begin{aligned}
 RMSE &= \sqrt{\frac{1}{n} \sum_{k=1}^n (y_k - \hat{y}_k)^2} \\
 MAE &= \frac{1}{n} \sum_{k=1}^n \|y_k - \hat{y}_k\|_1 \\
 ADE &= \frac{1}{n} \sum_{k=1}^n \|y_k - \hat{y}_k\|_2 \\
 FDE &= \|y_{final} - \widehat{y}_{final}\|_2
 \end{aligned} \tag{13}$$

where y_k and $\hat{y}_k$ represent the true position and the predicted position, n is the total number of predicted track points, and $\|\cdot\|_1$ and $\|\cdot\|_2$ denote one norm and the Euclidean distance, respectively. y_{final} and $\widehat{y}_{final}$ represent the final position of the actual trajectory and the predicted trajectory, respectively. It is worth noting that the lower the evaluation index, the better the generalization ability of the model.

5.2. Model Performance Comparison

5.2.1. Comparison Results

Through comparisons with the baselines using the evaluation metrics of the RMSE, MAE, and ADE, our model demonstrated strong robustness, as shown in Tables 2–4. The LSTM, BiLSTM, GRU, and BiGRU employ a sliding window approach for the prediction of multiple points, while the other six Seq2Seq models do not. Our model outperformed the baselines in the third to fifth trajectory points. However, the GRU achieved better results in the MAE in Table 3 and the ADE in Table 4, regarding the first and second points. This is mainly because the GRU has a structural advantage in predicting short-term sequences due to its simplicity. However, VEPO-S2S experiences a slight decrease when handling short-term predictions, as the portrait of the vessel may not be easily discernible. When the prediction length increases, our model performs better than the baseline models, according to the RMSE, MAE, and ADE.

Simultaneously, considering that trajectory prediction is also a time series forecasting problem, we compared VEPO-S2S with several time series forecasting models, including the ARIMA, Kalman Filter, and VAR. As Tables 2–4 show, our model was superior to others; the three-time series prediction models performed well in short-term forecasting, but the accuracy rapidly declined as the forecast horizon increased. Specifically, the VAR model became ineffective after the third time step because it failed to capture long-term complex nonlinear relationships. Different from time series models that only utilize position sequences, VEPO-S2S benefits from additional prior information, such as the ship length

and width, as well as its powerful ability to construct spatiotemporal correlations from historical trajectory points, making it more advantageous in forecasting tasks.

Table 2. Comparison results of VEPO-S2S with various baselines under RMSE evaluation metric. Here, 10->5 represents the RMSE value of the last 5 trajectories predicted by 10 historical trajectories.

Model Name	10->1	10->2	10->3	10->4	10->5
Kalman	2.42×10^{-4}	3.00×10^{-4}	4.22×10^{-4}	5.51×10^{-4}	6.71×10^{-4}
VAR	0.98×10^{-4}	2.80×10^{-4}	50.43×10^{-4}	-	-
ARIMA	3.42×10^{-4}	5.30×10^{-4}	7.21×10^{-4}	9.15×10^{-4}	11.12×10^{-4}
LSTM	0.88×10^{-4}	1.16×10^{-4}	2.48×10^{-4}	3.46×10^{-4}	4.60×10^{-4}
BiLSTM	2.64×10^{-4}	2.65×10^{-4}	3.51×10^{-4}	4.41×10^{-4}	5.45×10^{-4}
GRU	0.86×10^{-4}	1.58×10^{-4}	2.41×10^{-4}	3.32×10^{-4}	4.30×10^{-4}
BiGRU	3.16×10^{-4}	3.24×10^{-4}	3.96×10^{-4}	4.71×10^{-4}	5.21×10^{-4}
LSTM-LSTM	1.22×10^{-4}	1.94×10^{-4}	2.76×10^{-4}	3.64×10^{-4}	4.58×10^{-4}
BiLSTM-LSTM	1.44×10^{-4}	2.21×10^{-4}	3.07×10^{-4}	3.99×10^{-4}	4.96×10^{-4}
GRU-GRU	0.98×10^{-4}	1.69×10^{-4}	2.49×10^{-4}	3.36×10^{-4}	4.29×10^{-4}
BiGRU-GRU	1.34×10^{-4}	2.21×10^{-4}	3.14×10^{-4}	4.11×10^{-4}	5.12×10^{-4}
Transformer	1.31×10^{-4}	1.74×10^{-4}	2.35×10^{-4}	3.07×10^{-4}	3.86×10^{-4}
METO-S2S	2.02×10^{-4}	2.03×10^{-4}	2.28×10^{-4}	2.73×10^{-4}	3.35×10^{-4}
Ours	0.74×10^{-4}	1.11×10^{-4}	1.71×10^{-4}	2.40×10^{-4}	3.17×10^{-4}

Table 3. Comparison results of VEPO-S2S with various baselines under MAE evaluation metric. Here, 10->5 means predicted through 10 historical trajectories.

Model Name	10->1	10->2	10->3	10->4	10->5
Kalman	1.73×10^{-4}	2.34×10^{-4}	3.27×10^{-4}	4.24×10^{-4}	5.16×10^{-4}
VAR	0.87×10^{-4}	2.12×10^{-4}	28.6×10^{-4}	-	-
ARIMA	2.69×10^{-4}	3.99×10^{-4}	5.33×10^{-4}	6.69×10^{-4}	8.08×10^{-4}
LSTM	0.43×10^{-4}	0.77×10^{-4}	1.18×10^{-4}	1.66×10^{-4}	2.21×10^{-4}
BiLSTM	1.50×10^{-4}	1.52×10^{-4}	1.87×10^{-4}	2.35×10^{-4}	2.87×10^{-4}
GRU	0.36×10^{-4}	0.66×10^{-4}	1.01×10^{-4}	1.42×10^{-4}	1.87×10^{-4}
BiGRU	1.53×10^{-4}	1.65×10^{-4}	1.84×10^{-4}	2.15×10^{-4}	2.38×10^{-4}
LSTM-LSTM	0.70×10^{-4}	1.00×10^{-4}	1.34×10^{-4}	1.72×10^{-4}	2.13×10^{-4}
BiLSTM-LSTM	0.88×10^{-4}	1.20×10^{-4}	1.56×10^{-4}	1.94×10^{-4}	2.35×10^{-4}
GRU-GRU	0.51×10^{-4}	0.78×10^{-4}	1.10×10^{-4}	1.46×10^{-4}	1.85×10^{-4}
BiGRU-GRU	0.74×10^{-4}	1.10×10^{-4}	1.50×10^{-4}	1.92×10^{-4}	2.37×10^{-4}
Transformer	0.81×10^{-4}	0.95×10^{-4}	1.16×10^{-4}	1.42×10^{-4}	1.72×10^{-4}
METO-S2S	1.43×10^{-4}	1.46×10^{-4}	1.53×10^{-4}	1.70×10^{-4}	1.94×10^{-4}
Ours	0.53×10^{-4}	0.74×10^{-4}	0.99×10^{-4}	1.28×10^{-4}	1.60×10^{-4}

Table 4. Comparison results of VEPO-S2S with various baselines under ADE evaluation metric. Here, 10->5 means predicted through 10 historical trajectories.

Model Name	10->1	10->2	10->3	10->4	10->5
Kalman	3.42×10^{-4}	4.62×10^{-4}	8.22×10^{-4}	11.52×10^{-4}	14.29×10^{-4}
VAR	1.38×10^{-4}	5.60×10^{-4}	123.53×10^{-4}	-	-
ARIMA	4.83×10^{-4}	7.11×10^{-4}	9.42×10^{-4}	11.76×10^{-4}	14.14×10^{-4}
LSTM	0.68×10^{-4}	1.22×10^{-4}	1.87×10^{-4}	2.64×10^{-4}	3.51×10^{-4}
BiLSTM	2.37×10^{-4}	2.38×10^{-4}	2.99×10^{-4}	3.74×10^{-4}	4.56×10^{-4}
GRU	0.59×10^{-4}	1.05×10^{-4}	1.62×10^{-4}	2.27×10^{-4}	2.99×10^{-4}
BiGRU	2.42×10^{-4}	2.73×10^{-4}	4.00×10^{-4}	4.91×10^{-4}	5.20×10^{-4}
LSTM-LSTM	1.11×10^{-4}	1.59×10^{-4}	2.13×10^{-4}	2.73×10^{-4}	3.38×10^{-4}
BiLSTM-LSTM	1.39×10^{-4}	1.90×10^{-4}	2.47×10^{-4}	3.07×10^{-4}	3.72×10^{-4}
GRU-GRU	0.80×10^{-4}	1.23×10^{-4}	1.74×10^{-4}	2.31×10^{-4}	2.93×10^{-4}
BiGRU-GRU	1.16×10^{-4}	1.75×10^{-4}	2.38×10^{-4}	3.05×10^{-4}	3.76×10^{-4}
Transformer	1.28×10^{-4}	1.50×10^{-4}	1.83×10^{-4}	2.25×10^{-4}	2.73×10^{-4}
METO-S2S	2.26×10^{-4}	2.30×10^{-4}	2.41×10^{-4}	2.68×10^{-4}	3.07×10^{-4}
Ours	0.83×10^{-4}	1.17×10^{-4}	1.52×10^{-4}	2.01×10^{-4}	2.52×10^{-4}

5.2.2. Exploration on Seq2Seq Structure of VEPO-S2S

We conducted several experiments on the structure of the encoder and decoder in VEPO-S2S to achieve optimal performance, including a VEPO-BiGRU-GRU, a VEPO-LSTM-LSTM, a VEPO-BiLSTM-LSTM, and a VEPO-GRU-GRU. As shown in Table 5, the best result was obtained with the VEPO-GRU-GRU. Therefore, in the subsequent experiments, we utilized the VEPO-S2S with a GRU-GRU structure for further experimentation.

Table 5. Exploration results on different Seq2Seq structure.

Model Name	Encoder	Decoder	RMSE	MAE	ADE	FDE
VEPO-S2S	BiGRU	GRU	3.60×10^{-4}	1.99×10^{-4}	3.15×10^{-4}	5.28×10^{-4}
VEPO-S2S	LSTM	LSTM	3.46×10^{-4}	1.72×10^{-4}	2.72×10^{-4}	4.94×10^{-4}
VEPO-S2S	BiLSTM	LSTM	4.19×10^{-4}	2.48×10^{-4}	3.39×10^{-4}	5.97×10^{-4}
VEPO-S2S	GRU	GRU	3.17×10^{-4}	1.60×10^{-4}	2.52×10^{-4}	4.57×10^{-4}

5.2.3. Further Analysis

In contrast to the evaluation of continuous trajectory prediction tasks, we conducted individual points evaluation using the RMSE, MAE, and FDE. As we can see from Tables 6–8, the prediction errors of all the models exhibited a noticeable increase from the first to the fifth prediction point. This is attributed to a significant reduction in the available information for each prediction, moving from the first to the last prediction. RNN baseline models employ a sliding window approach for prediction and gradually accumulate inaccuracies with each prediction. Conversely, the Seq2Seq baseline models have the ability to simultaneously predict multiple points, which can reduce the tendency for error escalation compared with the RNN baseline models. Additionally, time series models rely on linear relationships between multiple time series; however, in long-term prediction tasks, nonlinear features become more prominent, causing inaccuracies to increase rapidly over time. Furthermore, our model almost outperforms all baselines. We also observe a similar performance between METO-S2S and VEPO-S2S at the fourth and fifth points in Figure 6, indicating that both VEPO and METO are excellent models for long sequence prediction tasks.

Table 6. The quantitative analysis results of each trajectory point under the RMSE evaluation index.

Model Name	First	Second	Third	Fourth	Fifth
Kalman	2.42×10^{-4}	3.27×10^{-4}	5.81×10^{-4}	8.15×10^{-4}	10.11×10^{-4}
VAR	0.98×10^{-4}	3.74×10^{-4}	86.61×10^{-4}	-	-
ARIMA	3.42×10^{-4}	6.63×10^{-4}	9.93×10^{-4}	13.28×10^{-4}	16.72×10^{-4}
LSTM	0.88×10^{-4}	2.10×10^{-4}	3.63×10^{-4}	5.43×10^{-4}	7.60×10^{-4}
BiLSTM	2.64×10^{-4}	2.65×10^{-4}	4.77×10^{-4}	6.39×10^{-4}	8.38×10^{-4}
GRU	0.86×10^{-4}	2.06×10^{-4}	3.52×10^{-4}	5.16×10^{-4}	6.96×10^{-4}
BiGRU	3.16×10^{-4}	2.42×10^{-4}	5.24×10^{-4}	6.44×10^{-4}	6.83×10^{-4}
LSTM-LSTM	1.22×10^{-4}	2.46×10^{-4}	3.90×10^{-4}	5.50×10^{-4}	7.20×10^{-4}
BiLSTM-LSTM	1.44×10^{-4}	2.77×10^{-4}	4.30×10^{-4}	5.95×10^{-4}	7.68×10^{-4}
GRU-GRU	0.98×10^{-4}	2.18×10^{-4}	3.59×10^{-4}	5.14×10^{-4}	6.84×10^{-4}
BiGRU-GRU	1.34×10^{-4}	2.82×10^{-4}	4.45×10^{-4}	6.17×10^{-4}	7.96×10^{-4}
Transformer	1.31×10^{-4}	2.08×10^{-4}	3.24×10^{-4}	4.59×10^{-4}	6.07×10^{-4}
METO-S2S	2.02×10^{-4}	2.03×10^{-4}	2.71×10^{-4}	3.76×10^{-4}	5.23×10^{-4}
Ours	0.74×10^{-4}	1.44×10^{-4}	2.47×10^{-4}	3.76×10^{-4}	5.22×10^{-4}

Table 7. The quantitative analysis results of each trajectory point under the MAE evaluation index.

Model Name	First	Second	Third	Fourth	Fifth
Kalman	1.73×10^{-4}	2.94×10^{-4}	5.13×10^{-4}	7.13×10^{-4}	8.83×10^{-4}
VAR	0.87×10^{-4}	3.38×10^{-4}	81.62×10^{-4}	-	-
ARIMA	2.69×10^{-4}	5.26×10^{-4}	7.99×10^{-4}	10.77×10^{-4}	13.62×10^{-4}
LSTM	0.43×10^{-4}	1.11×10^{-4}	2.01×10^{-4}	3.11×10^{-4}	4.40×10^{-4}
BiLSTM	1.50×10^{-4}	1.52×10^{-4}	2.60×10^{-4}	3.79×10^{-4}	4.94×10^{-4}
GRU	0.36×10^{-4}	0.95×10^{-4}	1.72×10^{-4}	2.63×10^{-4}	3.67×10^{-4}
BiGRU	1.53×10^{-4}	1.77×10^{-4}	2.53×10^{-4}	3.09×10^{-4}	3.26×10^{-4}
LSTM-LSTM	0.70×10^{-4}	1.30×10^{-4}	2.02×10^{-4}	2.85×10^{-4}	3.77×10^{-4}
BiLSTM-LSTM	0.88×10^{-4}	1.53×10^{-4}	2.28×10^{-4}	3.08×10^{-4}	3.98×10^{-4}
GRU-GRU	0.51×10^{-4}	1.04×10^{-4}	3.59×10^{-4}	5.14×10^{-4}	6.84×10^{-4}
BiGRU-GRU	1.34×10^{-4}	2.82×10^{-4}	4.45×10^{-4}	6.17×10^{-4}	7.96×10^{-4}
Transformer	0.81×10^{-4}	1.09×10^{-4}	1.58×10^{-4}	2.20×10^{-4}	2.93×10^{-4}
METO-S2S	2.02×10^{-4}	2.03×10^{-4}	2.21×10^{-4}	2.71×10^{-4}	2.92×10^{-4}
Ours	0.53×10^{-4}	0.96×10^{-4}	1.48×10^{-4}	2.13×10^{-4}	2.88×10^{-4}

Table 8. The quantitative analysis results of each trajectory point under the FDE evaluation index.

Model Name	First	Second	Third	Fourth	Fifth
Kalman	3.42×10^{-4}	4.62×10^{-4}	8.22×10^{-4}	11.52×10^{-4}	14.29×10^{-4}
VAR	1.38×10^{-4}	5.30×10^{-4}	122.49×10^{-4}	-	-
ARIMA	4.83×10^{-4}	9.38×10^{-4}	14.05×10^{-4}	18.79×10^{-4}	23.65×10^{-4}
LSTM	0.68×10^{-4}	1.76×10^{-4}	3.19×10^{-4}	4.93×10^{-4}	6.99×10^{-4}
BiLSTM	2.37×10^{-4}	2.38×10^{-4}	4.22×10^{-4}	6.00×10^{-4}	7.84×10^{-4}
GRU	0.59×10^{-4}	1.48×10^{-4}	2.75×10^{-4}	4.20×10^{-4}	5.86×10^{-4}
BiGRU	2.42×10^{-4}	2.84×10^{-4}	4.00×10^{-4}	4.91×10^{-4}	5.20×10^{-4}
LSTM-LSTM	1.11×10^{-4}	2.06×10^{-4}	3.21×10^{-4}	4.52×10^{-4}	5.98×10^{-4}
BiLSTM-LSTM	1.39×10^{-4}	2.41×10^{-4}	3.61×10^{-4}	4.89×10^{-4}	6.32×10^{-4}
GRU-GRU	0.80×10^{-4}	1.66×10^{-4}	2.76×10^{-4}	4.02×10^{-4}	5.42×10^{-4}
BiGRU-GRU	1.16×10^{-4}	2.23×10^{-4}	3.65×10^{-4}	5.06×10^{-4}	6.60×10^{-4}
Transformer	1.28×10^{-4}	1.72×10^{-4}	2.50×10^{-4}	3.50×10^{-4}	4.67×10^{-4}
METO-S2S	2.22×10^{-4}	2.30×10^{-4}	2.71×10^{-4}	3.49×10^{-4}	4.62×10^{-4}
Ours	0.83×10^{-4}	1.51×10^{-4}	2.34×10^{-4}	3.37×10^{-4}	4.57×10^{-4}

The proposed VEPO-S2S model consists of the Multi-level Vessel Trajectory Representation Module (Multi-Rep) and the Feature Fusion and Decoding Module (FFDM). The Multi-Rep is specifically designed not only to integrate trajectory information but also to fully consider the vessel features and behavioral preferences, encoding them with distinct

encoders to enrich the representation of features. Additionally, the FFDM selects and integrates the above information and features based on the current prediction environment, which allows the model to leverage them efficiently. These two advantages make VEPO-S2S more accurate than the other baselines.

5.3. Ablation Study

To investigate the function of the Multi-level Vessel Characteristics in VEPO-S2S and the Feature Fusion and Decoding Module, we designed several ablation experiments, which are introduced in the following part:

1. *w/o sa*. Delete the input of the Shallow-level Attributes in the Multi-level Vessel Trajectory Representation Module, including the vessel length, width, draft, and type (see Section 4.3.1);
2. *w/o k1*. Delete the Sailing Location Preference in the Multi-level Vessel Trajectory Representation Module (see Section 4.3.1);
3. *w/o k2*. Delete the Voyage Time Preference from the Multi-level Vessel Trajectory Representation Module (see Section 4.3.1);
4. *w/o k3*. Delete the Anchoring Time Preference from the Multi-level Vessel Trajectory Representation Module (see Section 4.3.1);
5. *w/o p_s*. Delete the Portrait Selection Component from the Feature Fusion and Decoding Module (see Section 4.3.2);
6. *w/o f_f*. Delete the Feature Fusion Component from the Feature Fusion and Decoding Module (see Section 4.3.2);
7. *w/o muti_d*. Delete the Multi-head Decoder Component from the Feature Fusion and Decoding Module, using a single GRU for decoding instead, which does not receive trajectory information separately (see Section 4.3.2).

The results are shown in Table 9. We evaluated the performance using the RMSE, MAE, ADE, and FDE. The results indicate that the deletion of the Portrait Selection Component had the most significant impact on the RMSE metric, decreasing it from 3.17×10^{-4} to 4.48×10^{-4} . This indicates that the Portrait Selection Component plays a crucial role in the model. This is because the Portrait Selection Component is responsible for selecting the vessel characteristics that are most suitable for the current environment. When the Portrait Selection Component is removed, the model's performance significantly declines. Additionally, it is observed that the deletion of the Multi-head Decoder Component had the least impact, with a decrease from 3.17×10^{-4} to 3.32×10^{-4} . This is because enhancing the decoder does not affect the overall structure. The model still has the ability to select and learn how to use the corresponding characteristics to generate accurate predictions. However, reinforcing the decoder leads to a slight improvement in model performance. Consistent with the above analysis, the MAE, ADE, and FDE evaluation metrics exhibited similar trends.

Table 9. Quantitative results on different ablation studies.

Ablation	RMSE	MAE	ADE	FDE
VEPO-S2S	3.17×10^{-4}	1.60×10^{-4}	2.52×10^{-4}	4.57×10^{-4}
<i>w/o sa</i>	3.53×10^{-4}	1.88×10^{-4}	2.96×10^{-4}	5.27×10^{-4}
<i>w/o k1</i>	3.50×10^{-4}	1.82×10^{-4}	2.88×10^{-4}	5.07×10^{-4}
<i>w/o k2</i>	3.50×10^{-4}	1.88×10^{-4}	2.96×10^{-4}	5.19×10^{-4}
<i>w/o k3</i>	3.55×10^{-4}	1.95×10^{-4}	3.07×10^{-4}	5.38×10^{-4}
<i>w/o p_s</i>	4.48×10^{-4}	1.99×10^{-4}	3.15×10^{-4}	5.84×10^{-4}
<i>w/o f_f</i>	3.42×10^{-4}	1.81×10^{-4}	2.88×10^{-4}	5.13×10^{-4}
<i>w/o muti_d</i>	3.32×10^{-4}	1.76×10^{-4}	2.77×10^{-4}	4.87×10^{-4}

After removing various innovations, the performance of the model decreased. Hence, the model is equipped with all the characteristics and components to achieve optimal results, demonstrating the effectiveness of our innovative points.

5.4. Qualitative Analysis

In order to better analyze the performance of VEPO-S2S, we selected several comparable models for several qualitative analyses, as described in this subsection.

5.4.1. Baselines Comparing

As shown in Figure 12, our model achieved accurate predictions in both scenarios compared to other baselines. The difficulty of the prediction increased from (a) to (b). In (a), the cargo ship had a straight route. Most of the models produced satisfactory predictions, especially VEPO and METO. However, the GRU deviated from the true trajectory. (b) shows a container ship turning, where our model performs the best. Additionally, the METO-S2S model accurately predicted the first four points, but it deviated from the actual trajectory in the last point. The GRU-GRU model performs better than the GRU model; however, it still struggles to achieve satisfactory prediction results. In practice, incorrect predictions can easily lead to accidents. It can be observed that the results predicted by VEPO-S2S were superior to others, which can avoid safety issues.

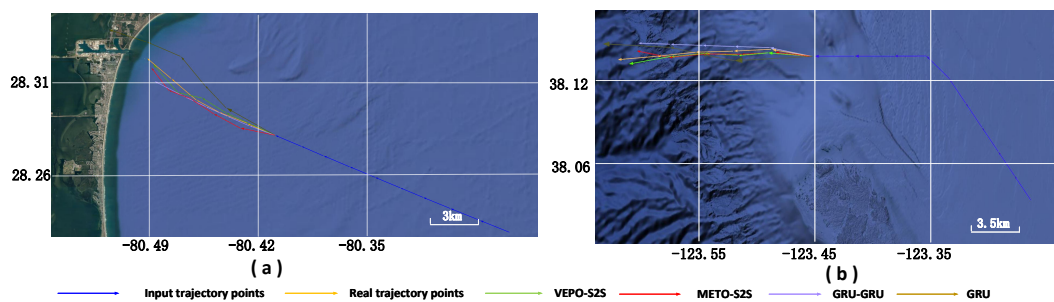


Figure 12. The predictions of cargo ships and container ships under various models, with the difficulty of predictions increasing from (a) to (b). Our model performs the best in both scenarios. In (a), which involves straight-line navigation of cargo ships, all models except GRU achieve decent prediction results. In (b), which involves container ship turning. Additionally, the METO-S2S model is also able to accomplish the prediction tasks to some extent. However, other models struggle to achieve satisfactory prediction performance.

5.4.2. Visual Result of the Seq2Seq Structure

As we can see from Figure 13, VEPO-GRU-GRU is more robust compared with VEPO-BiGRU-GRU, VEPO-BiLSTM-LSTM, and VEPO-LSTM-LSTM. (a) depicts a smooth trajectory, indicating the normal navigation of a cargo ship. (b) shows a curved trajectory, possibly suggesting avoidance maneuvers by an oil tanker. Across various vessel types and motion states, all structures of the VEPO-S2S model consistently exhibit satisfactory performance. Notably, the VEPO-S2S model with the GRU-GRU structure shows better robustness.

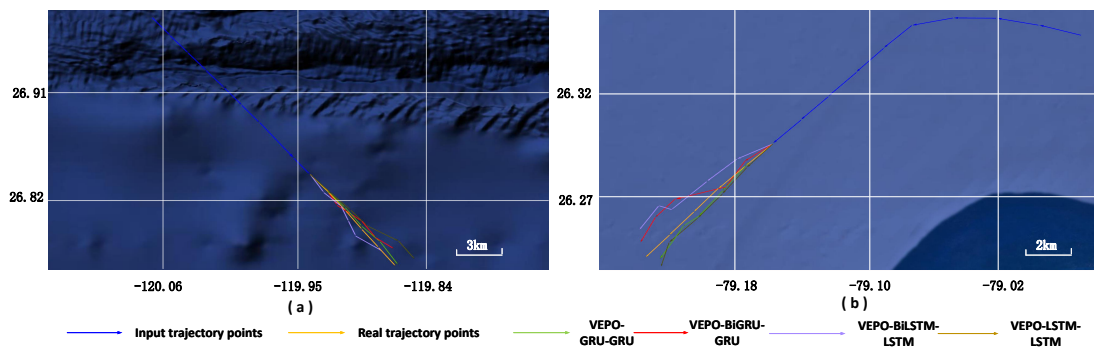


Figure 13. The trajectory predictions of the cargo ship and oil tanker using different structures of the VEPO-S2S model, with the predictive difficulty increasing gradually from (a) to (b). As depicted in the graph, VEPO-GRU-GRU achieves the best predictive performance.

5.4.3. Qualitative Ablation Results

We conducted several detailed studies to better illustrate the changes in the experimental results before and after ablation. Seven figures display the results of the experiments on the Shallow-level Attributes, the Sailing Location Preference, the Voyage Time Preference, the Anchoring Time Preference, the Portrait Selection Component, the Feature Fusion Component, and the Multi-head Decoder Component for a tugboat. More precisely, we present four trajectories in each image, including the input track points, labels, and the predictive results before and after ablation.

For each figure, it is evident that the green curve closely aligns with the actual trajectory. We observe that if a vessel lacks any one of the Shallow-level Attributes, the Sailing Location Preference, the Voyage Time Preference, or the Anchoring Time Preference, there are significant biases in both the direction and distance in the predicted results.

Figure 14 illustrates the results with and without Shallow-level Attributes. In the absence of Shallow-level Attributes, VEPO-S2S predictions deviate from the actual trajectory at the fourth and fifth points. This is because Shallow-level Attributes determine the vessel's inertia and turning capabilities. Typically, larger vessels have more difficulty in altering their current motion states. When there is a lack of Shallow-level Attributes, the model struggles to accurately assess these abilities of the vessel. Therefore, in long-term predictions, the model fails to provide effective guidance and leads to deviations from the correct trajectory in later stages.

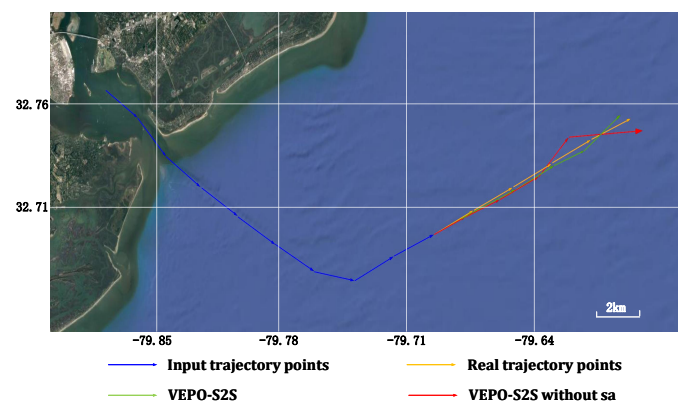


Figure 14. The predicted trajectories of a tugboat, where the green and red lines are the prediction results using VEPO-S2S with and without Shallow-level Attributes, respectively. The Shallow-level Attributes are associated with the vessel's inertia and turning capabilities. The model without Shallow-level Attributes is not able to grasp this ability well, which may cause errors.

Figure 15 demonstrates the results with and without the Sailing Location Preference. The Sailing Location Preference helps the model recognize a vessel's adaptability to terrain.

Vessels with varying degrees of adaptability to terrains choose different collision avoidance routes. Without the Sailing Location Preference, the model struggles to capture details of routes, which makes it difficult to generate corresponding choices and leads to oscillations in the predicted trajectories.

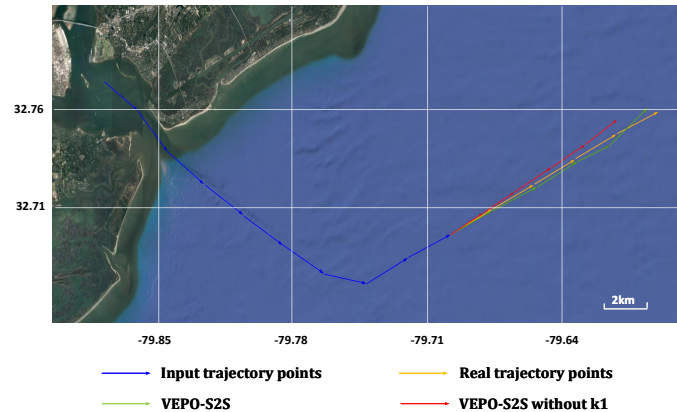


Figure 15. The predicted trajectories of a tugboat, where the green and red lines represent the prediction results of VEPO-S2S with and without considering the Sailing Location Preference, respectively. The Sailing Location Preference helps the model identify the adaptability of the vessel to the geographical environment. When the Sailing Location Preference is not considered, the VEPO-S2S model is not correct.

Figure 16 displays the comparison results with and without the Voyage Time Preference. The Voyage Time Preference reflects the habits of the crews and vessel sailing at various times. For instance, the collision avoidance maneuvers are different during high and low vessel traffic periods. Without guidance from the Voyage Time Preference, the predicted results exhibit considerable fluctuations.

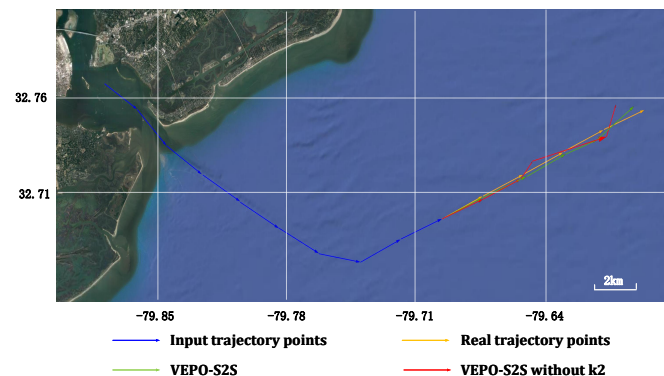


Figure 16. The predicted trajectories of a tugboat, where the green and red lines represent the prediction results of the VEPO-S2S model with and without considering the Voyage Time Preference, respectively. The Voyage Time Preference is related to the habits of the crews. When the Voyage Time Preference is not considered, the VEPO-S2S model may produce inaccurate predictions.

Figure 17 shows the results with and without the Anchoring Time Preference, which is related to working and resting habits. The trajectory of the tugboat changes more significantly when it is in working condition. Therefore, it is difficult for the model to accurately determine the current movement of the tugboat without the Anchoring Time Preference, which leads to deviations from the true trajectory points.

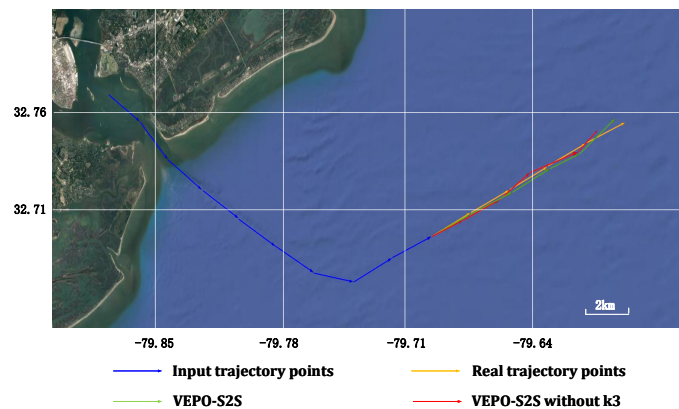


Figure 17. The predicted trajectories of a tugboat, where the green and red lines represent the prediction results of the VEPO-S2S model with and without considering the Anchoring Time Preference, respectively. The Anchoring Time Preference helps the model identify the working and resting habits of the vessel. In the absence of the Anchoring Time Preference, the model produces an incorrect estimation for each timestamp.

Figure 18 demonstrates the visual comparison results of the Portrait Selection Component. Without the Portrait Selection Component, the result deviates from the real track. This is because the Portrait Selection Component can select the most relevant characteristics for prediction and without it, irrelevant characteristics may be introduced into the prediction process, which results in significant deviations in the predicted outcomes.

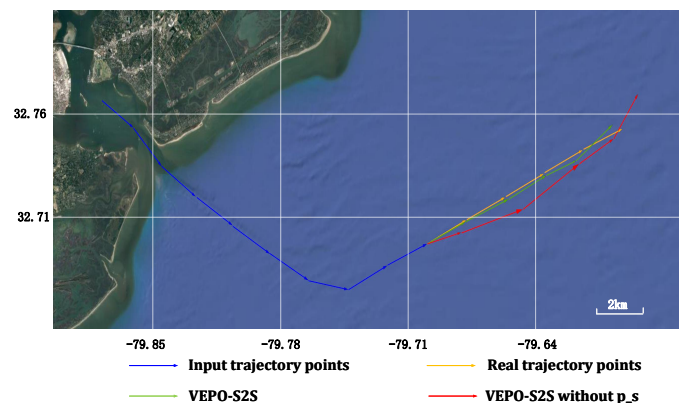


Figure 18. The predicted trajectories of a tugboat, where the green and red lines are the prediction results using VEPO-S2S with and without the Portrait Selection Component, respectively. The Portrait Selection Component is responsible for selecting the most relevant characteristics for prediction. Without this component, the model cannot select appropriate characteristics to assist in prediction, leading to a decrease in model robustness.

Figure 19 shows the visual comparison results of the Feature Fusion Component, which effectively integrates the trajectory information with the vessel features and enhances the correlation between them. Without the Feature Fusion Component, the vessel characteristics struggle to be fully expressed, which leads to incorrect predictions.

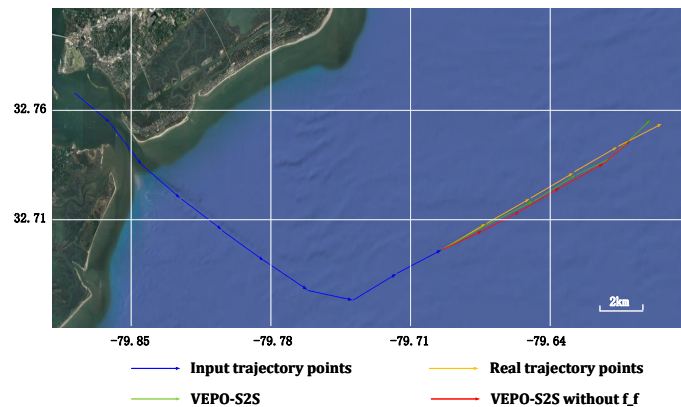


Figure 19. The predicted trajectories of a tugboat, where green and red lines are the prediction results by VEPO-S2S with and without Feature Fused Component, respectively. The Feature Fused Component effectively integrates trajectory information and vessel characteristics and increases the correlation between both. Without the Feature Fused Component, vessel characteristics are difficult to express adequately in the VEPO-S2S model, leading to a decrease in model accuracy.

Figure 20 denotes the comparison results of the Multi-head Decoder Component, which can adjust the level of engagement in the trajectory information and vessel features. The model's adaptability decreases when removing the Multi-head Decoder Component. Based on these ablation studies, the VEPO-S2S model demonstrates satisfactory accuracy and robustness.

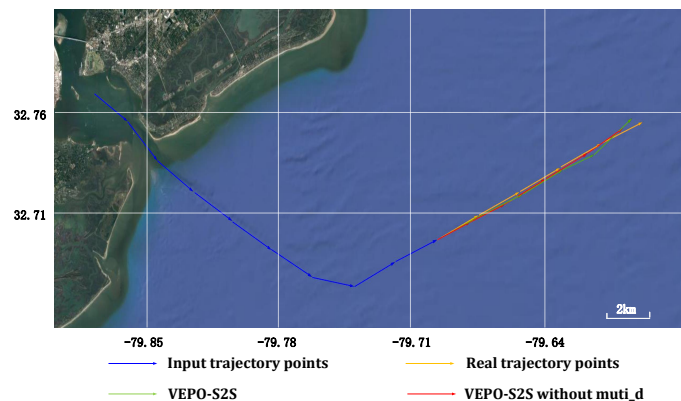


Figure 20. The predicted trajectories of a tugboat, where the green and red lines are the prediction results using VEPO-S2S with and without the Multi-head Decoder Component, respectively. The Multi-head Decoder Component regulates the involvement of the trajectory information and vessel characteristics in the prediction process. Without it, the adaptability of the VEPO-S2S model decreases.

6. Conclusions

Through a study of the relevant literature, we found that vessel features and behavioral preferences have a significant impact on trajectories. Therefore, this study proposes a new trajectory prediction model, VEPO-S2S, which fully considers the trajectory information, vessel features, and behavioral preferences. VEPO-S2S consists of two parts: the Multi-level Vessel Trajectory Representation Module and the Feature Fusion and Decoding Module. The Multi-level Vessel Trajectory Representation Module obtains trajectory information (such as the longitude, latitude, course, speed, and sailing distance) along with Multi-level Vessel Characteristics, encompassing Shallow-level Attributes (vessel length, type, and draft) and Deep-level Features (Sailing Location Preference, Voyage Time Preference, and anchoring time preference). These are encoded using multiple encoders. The Feature Fusion and Decoding Module aims to select the most relevant vessel characteristics for the current prediction environment and integrate them with the trajectory information before

decoding with an enhanced decoder. The experimental results demonstrate that this model outperforms other baseline models qualitatively and exhibits excellent performance on grid-based maps.

7. Future Works

In this work, we took into account the impact of features and preferences beyond trajectory information on trajectory prediction. In the future, we aim to optimize our model to enhance the efficiency and prediction accuracy, validated across more global navigation datasets. Moreover, we will explore other prominent models, such as the time-series large model. Additionally, there are other factors influencing vessel movements, including weather, sea conditions, ocean currents, and reefs. Therefore, we aim to incorporate more influencing factors into the modeling process and undertake further investigations in more complex scenarios.

Author Contributions: Conceptualization, Z.H.; methodology, Z.H. and X.Y.; software, S.L.; validation, H.L. and J.L.; formal analysis, Y.Z.; investigation, Y.Z. and H.L.; resources, H.L.; data curation, W.A.; writing—original draft preparation, X.Y.; writing—review and editing, Z.H.; visualization, W.A. and X.Y.; supervision, S.L. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Zhejiang Province, China under Grant LY21E090005, and the Bureau of Science and Technology Project of Zhoushan (2021C21010).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data used in this paper can be downloaded from public website: marinecadastre.gov/accessais (accessed on 31 December 2021).

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

Abbreviations

The following abbreviations are used in this manuscript:

AIS	Automatic Identification System
Seq2Seq	Sequence-to-Sequence Network
VEPO	the VEssel Portrait Oriented model
Multi-Rep	Multi-level Vessel Trajectory Representation Module
FFDM	the Feature Fusion and Decoding Module

References

1. Yang, C.-H.; Wu, C.-H.; Shao, J.-C.; Wang, Y.-C.; Hsieh, C.-M. Ais-based intelligent vessel trajectory prediction using bi-lstm. *IEEE Access* **2022**, *10*, 24302–24315.
2. Fossen, S.; Fossen, T.I. Extended kalman filter design and motion prediction of ships using live automatic identification system (ais) data. In Proceedings of the 2018 2nd European Conference on Electrical Engineering and Computer Science (EECS), Bern, Switzerland, 20–22 December 2018; pp. 464–470.
3. Kawan, B.; Wang, H.; Li, G.; Chhantyal, K. Data-driven modeling of ship motion prediction based on support vector regression. In Proceedings of the 58th SIMS, Reykjavik, Iceland, 25–27 September 2017.
4. Li, H.; Jiao, H.; Yang, Z. Ship trajectory prediction based on machine learning and deep learning: A systematic review and methods analysis. *Eng. Appl. Artif. Intell.* **2023**, *126*, 107062.
5. Suo, Y.; Chen, W.; Claramunt, C.; Yang, S. A ship trajectory prediction framework based on a recurrent neural network. *Sensors* **2020**, *20*, 5133.
6. Wang, C.; Ren, H.; Li, H. Vessel trajectory prediction based on ais data and bidirectional gru. In Proceedings of the 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL), IEEE, Nanchang, China, 15–17 May 2020; pp. 260–264.
7. Tang, H.; Yin, Y.; Shen, H. A model for vessel trajectory prediction based on long short-term memory neural network. *J. Mar. Eng. Technol.* **2022**, *21*, 136–145.

8. Park, J.; Jeong, J.; Park, Y. Ship trajectory prediction based on bi-LSTM using spectral-clustered AIS data. *J. Mar. Sci. Eng.* **2021**, *9*, 1037.
9. Zhang, S.; Wang, L.; Zhu, M.; Chen, S.; Zhang, H.; Zeng, Z. A bi-directional lstm ship trajectory prediction method based on attention mechanism. In Proceedings of the 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), IEEE, Chongqing, China, 12–14 March 2021; pp. 1987–1993.
10. Bao, K.; Bi, J.; Gao, M.; Sun, Y.; Zhang, X.; Zhang, W. An improved ship trajectory prediction based on ais data using mha-bigru. *J. Mar. Sci. Eng.* **2022**, *10*, 804.
11. Wu, W.; Chen, P.; Chen, L.; Mou, J. Ship trajectory prediction: An integrated approach using convlstm-based sequence-to-sequence model. *J. Mar. Sci. Eng.* **2023**, *11*, 1484.
12. Nguyen, D.-D.; Van, C.L.; Ali, M.I. Vessel trajectory prediction using sequence-to-sequence models over spatial grid. In Proceedings of the 12th ACM International Conference on Distributed and Event-Based Systems, Lyon, France, 25–28 June 2018; pp. 258–261.
13. Zhang, Y.; Han, Z.; Zhou, X.; Li, B.; Zhang, L.; Zhen, E.; Wang, S.; Zhao, Z.; Guo, Z. Meto-s2s: A s2s based vessel trajectory prediction method with multiple-semantic encoder and type-oriented decoder. *Ocean. Eng.* **2023**, *277*, 114248.
14. Zhao, J.; Yan, Z.; Chen, X.; Han, B.; Wu, S.; Ke, R. k-gcn-lstm: A k-hop graph convolutional network and long-short-term memory for ship speed prediction. *Phys. Stat. Mech. Its Appl.* **2022**, *606*, 128107.
15. Wang, S.; Li, Y.; Xing, H.; Zhang, Z. Vessel trajectory prediction based on spatio-temporal graph convolutional network for complex and crowded sea areas. *Ocean. Eng.* **2024**, *298*, 117232.
16. Cao, D.; Li, J.; Ma, H.; Tomizuka, M. Spectral temporal graph neural network for trajectory prediction. In Proceedings of the 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, 30 May–5 June 2021; pp. 1839–1845.
17. Yasukawa, H.; Yoshimura, Y. *Ship Maneuverability*; Seizando: Tokyo, Japan, 2018.
18. Cooper, A. *Why High-Tech Products Drive Us Crazy and How to Restore the Sanity*; Sams Publishing: Carnel, IN, USA, 2004.
19. Lu, J.; Xie, L. Ocean vehicle inertial navigation method based on dynamic constraints. *J. Navig.* **2018**, *71*, 1553–1566.
20. Cole, B.; Schamberg, G. Unscented kalman filter for long-distance vessel tracking in geodetic coordinates. *Appl. Ocean. Res.* **2022**, *124*, 103205.
21. Zhang, D.; Chu, X.; Wu, W.; He, Z.; Wang, Z.; Liu, C. Model identification of ship turning maneuver and extreme short-term trajectory prediction under the influence of sea currents. *Ocean. Eng.* **2023**, *278*, 114367.
22. Sun, Q.; Tang, Z.; Gao, J.; Zhang, G. Short-term ship motion attitude prediction based on lstm and gpr. *Appl. Ocean. Res.* **2022**, *118*, 102927.
23. Zhang, L.; Zhu, Y.; Su, J.; Lu, W.; Li, J.; Yao, Y. A hybrid prediction model based on knn-lstm for vessel trajectory. *Mathematics* **2022**, *10*, 4493.
24. Gao, D.-W.; Zhu, Y.-S.; Zhang, J.-F.; He, Y.-K.; Yan, K.; Yan, B.-R. A novel mp-lstm method for ship trajectory prediction based on ais data. *Ocean. Eng.* **2021**, *228*, 108956.
25. Forti, N.; Millefiori, L.M.; Braca, P.; Willett, P. Prediction oof vessel trajectories from ais data via sequence-to-sequence recurrent neural networks. In Proceedings of the ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Virtual, 4–9 May 2020; pp. 8936–8940.
26. You, L.; Xiao, S.; Peng, Q.; Claramunt, C.; Han, X.; Guan, Z.; Zhang, J. St-seq2seq: A spatio-temporal feature-optimized seq2seq model for short-term vessel trajectory prediction. *IEEE Access* **2020**, *8*, 218565–218574.
27. Chen, P.; Yang, F.; Mou, J.; Chen, L.; Li, M. Regional ship behavior and trajectory prediction for maritime traffic management: A social generative adversarial network approach. *Ocean. Eng.* **2024**, *299*, 117186.
28. Zhang, Y.; Han, Z.; Zhou, X.; Zhang, L.; Wang, L.; Zhen, E.; Wang, S.; Zhao, Z.; Guo, Z. Peso: A seq2seq-based vessel trajectory prediction method with parallel encoders and ship-oriented decoder. *Appl. Sci.* **2023**, *13*, 4307.
29. Jiang, D.; Shi, G.; Li, N.; Ma, L.; Li, W.; Shi, J. Trfm-ls: Transformer-based deep learning method for vessel trajectory prediction. *J. Mar. Sci. Eng.* **2023**, *11*, 880.
30. Liu, R.W.; Liang, M.; Nie, J.; Yuan, Y.; Xiong, Z.; Yu, H.; Guizani, N. Stmgcn: Mobile edge computing-empowered vessel trajectory prediction using spatio-temporal multigraph convolutional network. *IEEE Trans. Ind. Inform.* **2022**, *18*, 7977–7987.
31. Li, Y.; Li, Z.; Mei, Q.; Wang, P.; Hu, W.; Wang, Z.; Xie, W.; Yang, Y.; Chen, Y. Research on multi-port ship traffic prediction method based on spatiotemporal graph neural networks. *J. Mar. Sci. Eng.* **2023**, *11*, 1379.
32. Cho, K.; Merriënboer, B.V.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.
33. Gong, G.; An, X.; Mahato, N.K.; Sun, S.; Chen, S.; Wen, Y. Research on short-term load prediction based on seq2seq model. *Energies* **2019**, *12*, 3199.
34. Hwang, S.; Jeon, G.; Jeong, J.; Lee, J. A novel time series based seq2seq model for temperature prediction in firing furnace process. *Procedia Comput. Sci.* **2019**, *155*, 19–26.
35. Gao, S.; Zhang, S.; Huang, Y.; Han, J.; Luo, H.; Zhang, Y.; Wang, G. A new seq2seq architecture for hourly runoff prediction using historical rainfall and runoff as input. *J. Hydrol.* **2022**, *612*, 128099.
36. Wang, H.; Wang, S.; Chen, S.; Hui, G. Predicting long-term production dynamics in tight/shale gas reservoirs with dual-stage attention-based ten-seq2seq model: A case study in duvernay formation. *Geoenergy Sci. Eng.* **2023**, *223*, 211495.

37. Stefenon, S.F.; Seman, L.O.; Aquino, L.S.; Coelho, L.d.S. Wavelet-seq2seq-lstm with attention for time series forecasting of level of dams in hydroelectric power plants. *Energy* **2023**, *274*, 127350.
38. Mootha, S.; Sridhar, S.; Seetharaman, R.; Chitrakala, S. Stock price prediction using bi-directional lstm based sequence to sequence modeling and multitask learning. In Proceedings of the 2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 28–31 October 2020; pp. 0078–0086.
39. He, Q.; Li, W.; Hao, Z.; Liu, G.; Huang, D.; Song, W.; Xu, H.; Alqahtani, F.; Kim, J.-U. A tma-seq2seq network for multi-factor time series sea surface temperature prediction. *Comput. Mater. Contin.* **2022**, *73*, 2022.
40. Qu, Q.; Chen, H.; Lei, Z.; Li, B.; Du, Q.; Wang, Y. Sea clutter amplitude prediction via an attention-enhanced seq2seq network. *Remote Sens.* **2023**, *15*, 3234.
41. Asadi, N.; Lamontagne, P.; King, M.; Richard, M.; Scott, K.A. Seasonal Sea Ice Presence Forecasting of Hudson Bay Using seq2seq Learning. Available online: <https://s3.us-east-1.amazonaws.com/climate-change-ai/papers/icml2021/6/paper.pdf> (accessed on 1 June 2024)
42. Zhu, W.; Zhou, R. Employee portrait construction based on deep neural networks. In Proceedings of the 2022 9th International Conference on Dependable Systems and Their Applications (DSA), IEEE, Wulumuqi, China, 4–5 August 2022; pp. 139–142.
43. Ma, C.; Wang, Z.; Cheng, P.; Liu, X.; Yu, H.; Zhang, H. User portrait-based hybrid recommendation method of web services. *J. Comput.* **2021**, *32*, 197–212.
44. Zhao, J.; Zeng, D.; Xiao, Y.; Che, L.; Wang, M. User personality prediction based on topic preference and sentiment analysis using lstm model. *Pattern Recognit. Lett.* **2020**, *138*, 397–402.
45. Yan, S.; Zhao, T.; Deng, J. Predicting social media user occupation with content-aware hierarchical neural networks. In Proceedings of the 2022 8th International Conference on Big Data and Information Analytics (BigDIA), Guiyang, China, 24–25 August 2022; pp. 388–395.
46. Wang, N. Computer intelligent prediction model of internet impulsive reward based on user portrait algorithm. In Proceedings of the 2021 IEEE 4th International Conference on Automation, Electronics and Electrical Engineering (AUTEEE), Shenyang, China, 19–21 November 2021; pp. 545–548.
47. Mingyu, Z.; Sutong, W.; Yanzhang, W.; Dujuan, W. An interpretable prediction method for university student academic crisis warning. *Complex Intell. Syst.* **2022**, *8*, 323–336.
48. Liu, T.; Ma, J. Ship navigation behavior prediction based on ais data. *IEEE Access* **2022**, *10*, 47997–48008.
49. Silvester, R. *Coastal Engineering*; Elsevier Scientific Publishing Company: Amsterdam, The Netherlands, 1974.
50. Brolsma, J.U. Permanent International Association of Navigation Congresses. *PIANC, the World Association for Waterborne Transport Infrastructure: An Association in a Changing World, 1885–2010*; PIANC: Brussels, Belgium, 2010. Available online: <https://books.google.co.jp/books?id=ZaSwYgEACAAJ> (accessed on 1 June 2024).
51. Notteboom, T.E. Container shipping and ports: An overview. *Rev. Netw. Econ.* **2004**, *3*, 86–106.
52. Gillis, D.M. Ideal Free Distributions in Fleet Dynamics: A Behavioral Perspective on Vessel Movement in Fisheries Analysis. *Can. J. Zool.* **2003**, *81*, 177–187.
53. Mahpour, A.; Nazifi, A.; Amiri, A.M. Development of optimization model to reduce unloading and loading time at berth in container ports. *Iran. J. Sci. Technol. Trans. Civ. Eng.* **2021**, *45*, 2831–2840.
54. McInnes, L.; Healy, J.; Astels, S. hdbscan: Hierarchical density based clustering. *J. Open Source Softw.* **2017**, *2*, 205.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.