Global oceanic mesoscale eddies trajectories prediction with knowledge-fused neural network

Xinmin Zhang, Baoxiang Huang, *Member, IEEE,* Ge Chen, Linyao Ge, Milena Radenkovic, and Guojia Hou

***Abstract*—Efficient eddy trajectory prediction driven by multi-information fusion can facilitate the scientific research of oceanography, while the complicated dynamics mechanism makes this issue challenging. Benefiting from ocean observing technology, the eddy trajectory dataset can be qualified for data-intensive research paradigms. In this paper, the dynamics mechanism is used to inspire the design idea of the eddy trajectory prediction neural network (termed EddyTPNet) and is also transformed into prior knowledge to guide the learning process. This study is among the first to implement eddy trajectory prediction with physics informed neural network. First, an in-depth analysis of the kinematic characteristics indicates that the longitude and latitude of the trajectory should be decoupled; Second, the directional dispersion prior knowledge of global eddy propagation is embedded into the decoder of the EddyTPNet to improve the performance; Finally, EddyTPNet is implemented to predict the trajectories of global long-lived eddies. The extensive experimental results indicate that EddyTPNet can reliably forecast the eddy motion for the following 7 days and ensure a promising daily mean geodetic error of roughly 7.18 km. This exploratory study provides valuable insights into solving the prediction problem of ocean phenomena by using knowledge-based time series neural networks.**

***Index Terms*—Eddy trajectory prediction, Directional divergence physical information, Deep learning, Knowledge-fused neural network.**

I. Introduction

BIG data oceanography is a fertile domain for interdisciplinary research. Mesoscale eddies, the oceanic counterpart of atmospheric storms, are common and complex oceanic flow phenomena [1], [2], with the vast majority of them propagating westward at velocities comparable to those of long Rossby waves [3]. As the crucial link in the ocean's energy cascade, their movement propels the transmission of salt [4], [5], heat [6], carbon [7], [8], and other marine resources. Consequently, mesoscale eddy trajectories are of significant scientific interest [9], [10], climate change modeling [11], [12], and marine ecosystem improvement [13]. The patterns and dynamic mechanism of eddy motion are complicated [14], in this context, the efficient prediction of eddy has been a challenge [15].

Essentially, trajectory prediction is a complex time series prediction problem [16], [17]. In other words, the propagation trend of eddy can be predicted by using historical eddy motion data, but its uncertainty is higher than that of ordinary time series problems. The particularity of the detailed problem can be summarized in the following aspects. (1) Eddies vary in their rotation patterns, lifetimes, and survival intervals [18]. In addition, each form of eddies moves uniquely at various times and places, as depicted in Fig.1, increasing the difficulty of interpreting and forecasting movement; (2) Complex background fields, such as current, wind, and topography, strengthen the nonlinearity of mesoscale eddy and the complexity of the problem [19]; (3) Most mesoscale eddies propagate westward at a speed similar to that of Rossby waves, while only a few traveling eastward. The imbalance of data sets also makes track prediction more challenging.

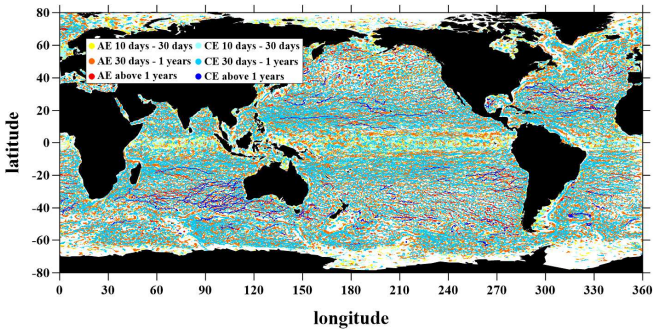

Fig. 1. The distribution of global mesoscale eddy motion trajectories for 2019-2020. The red and blue solid circles indicate the distribution of anticyclonic eddies and cyclonic eddies respectively. The color lightens as the life cycle decreases.

Theoretically, the assurance of ocean data products depends on the utilization of multiple observational data and synchronization techniques [20]–[22]. Furthermore, the prediction of oceanic phenomena has been enhanced through the use of numerical simulations [23]–[26], physical theories [27], and dynamic statistical analysis [28], [29], leading to substantial achievements and the enrichment of existing knowledge. Additionally, deep learning methods [30]–[32] have been successfully implemented in specific domains to improve effectiveness by integrating prior knowledge [33], [34].

This research was financially supported by the National Natural Science Foundation of China (No.42276203,42030406), the Laoshan Laboratory (No.LSKJ202204302), and the Natural Science Foundation of Shandong Province (No.ZR2021MD001). (Corresponding authors: Baoxiang Huang)

Xinmin zhang is with Qingdao University, Qingdao, 266071, China (e-mail: 2021023803@qdu.edu.cn).

Linyao Ge is with the Frontiers Science Center for Deep Ocean Multispheres and Earth System, School of Marine Technology, Ocean University of China, Qingdao, 266100, China (e-mail: linyaoge@stu.ouc.edu.cn).

Baoxiang Huang is with the College of Computer Science and Technology, Qingdao University, Qingdao 266071, China, and also with the Laboratory for Regional Oceanography and Numerical Modeling, Laoshan Laboratory, Qingdao, 266100, China. (e-mail: hbx3726@163.com).

Ge Chen is with the Frontiers Science Center for Deep Ocean Multispheres and Earth System, School of Marine Technology, Ocean University of China, Qingdao 266100, China, and also with the Laboratory for Regional Oceanography and Numerical Modeling, Laoshan Laboratory, Qingdao, 266100, China. (e-mail: gechen@ouc.edu.cn)

Benefiting from the intensive trajectory data of oceanic eddies, various learning architectures have been applied to the Earth observation [35], [36] with the rise of artificial intelligence [37], [38]. As for the prediction of eddy trajectories, multiple linear regression techniques were employed to create straightforward statistical prediction models [39] and correlate them with changes in eddy propagation position and other ocean parameters [40]. The method for predicting eddy properties and propagation trajectories using long short-term memory (LSTM) [41] and extra tree algorithms were proposed [27]. By combining data from different sources, the gate recurrent units (GRU)-based deep learning framework was applied to the prediction of eddy trajectories [42]. In addition, the new loss function called weighted mean square estimation was also proposed to improve model performance. The previous studies in artificial intelligence oceanography have demonstrated that deep learning methods yielded impressive performance [32], [43]. These studies focus mainly on data correlation analyses and do not consider the influence of time series on the prediction of eddy trajectories, nor the limitation of the prediction region and the lack of consideration of the influence of the external physical environment. Nevertheless, these results demonstrate the interpretability of the mesoscale eddy trajectory prediction problem and provide important insights for our research.

The objective of this paper is to implement oceanic mesoscale eddies trajectory prediction with the knowledge-fused neural network. Specifically, the proposed approach is to use historical data spanning 10 days to predict the trajectory of the next 7 days. The main contributions can be briefly summarized as follows.

1) The angular momentum propagation mechanism of eddies is embedded as the physical constraint into the proposed EddyTPNet to capture the dynamic characteristics, thereby improving prediction accuracy.
2) EddyTPNet is first trained using global data to predict the propagation process of long-lived eddies on a global scale, and then is transferred to complicated local regions with fine-tuning strategies.
3) Extensive experiments have been conducted to evaluate the proposed method for eddy trajectory prediction. The results demonstrate that the method can achieve promising performance both globally and locally.

The remainder of the paper is organized as follows. Section II formulates the problem definition and explains the idea of dynamics mechanism factorization. The proposed methodology is illustrated in section III, including trajectory data preprocessing, the architecture of the Knowledge-fused deep neural network, and the loss function. Section IV indicates the comprehensive experiments in detail. Finally, section V concludes and discusses the present work.

II. Preliminaries

A. Problem definition

As previously stated, the task of this study is to predict the trajectory of the next 7 days using 10 days of historical data. In this exploratory study, eddies with lifetimes of more than one year were chosen as the main research objects for the consideration of observation reliability, generalization, and the limitation of Graphic Processing Unit (GPU) memory. Specifically, the main reasons are as follows: (1) short-lived eddies are unstable, and the evolution process of short-lived eddies is easily affected by the external environment. Considering the merging and splitting of eddies, short-lived eddies may survive in another form. (2) There are certain errors in the observation of short-lived eddies. Usually, the scales of short-lived eddies are small, and there are some errors in satellite observation and tracking algorithms. (3) The related studies have demonstrated that the behaviors of short-lived and long-lived eddies are very different in terms of population, migration, and evolution [18]. (4) So using all eddies trajectories simultaneously to train a deep neural network will increase the computational cost but result in poor generalization.

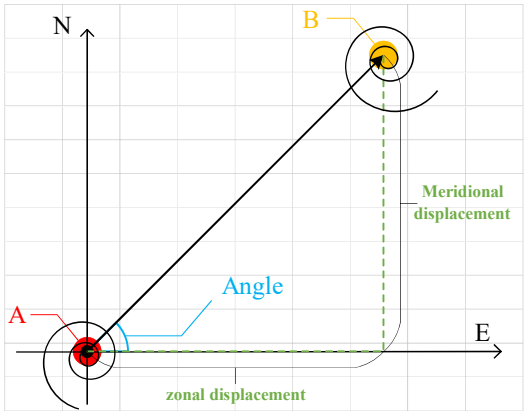

Fig. 2. Decomposition of mesoscale eddy motion trajectories. The red solid point A is the position of the eddy's starting point and the yellow solid point B is the position of the eddy's ending point. The angle formed by the Cartesian coordinate system with the east direction is the instantaneous eddy movement angle. The vertical and lateral movement distances are the meridional displacement and zonal displacement respectively.

EddyTPNet takes 10 days of historical trajectory $(X^n_{t-9}, X^n_{t-8}, ..., X^n_t)$ as input to predicts the next 7 days locations $(Y^n_{t+1}, Y^n_{t+2}, ..., Y^n_{t+7})$ of the eddy propagation, as described in following Eq.1.

$$(Y^n_{t+1}, Y^n_{t+2}, ..., Y^n_{t+7}) = EddyTPNet(X^n_{t-9}, X^n_{t-8}, ..., X^n_t) \quad (1)$$

Where n denotes the nth eddy and t denotes the tth time within that trajectory. The relevant data processing and the specific training test procedure within EddyTPNet will be described specifically in Section III.

B. Dynamics mechanism factorization towards network design

Through an in-depth analysis of the dynamics of eddy propagation, it is noted that the planetary gradient of the Coriolis parameter creates a mass imbalance in the eddy, causing it to move westward. Furthermore, in the radial motion, the eddy is subjected to the baroclinic shear of the flow field and the eddy diverges towards the poles and the equator, respectively [44]. The knowledge is factorized into two parts, one for the

longitude and latitude decoupling of the trajectory, the other is divergence priority.

The main idea of trajectory latitude and longitude decoupling is to predict zonal and meridional displacements rather than positions. The zonal and meridional displacements can determine the position of the eddy motion at the next step, as shown in Fig .2. Meanwhile, divergence knowledge plays an important role in weighing the dominance of the radial and latitudinal directions of the eddy. In addition, the directional dispersion can effectively guide the model in making more accurate predictions of propagation when the external factors are unknown.

III. Methodology

This section presents the overall pipeline for implementing eddy trajectories prediction. Specifically, section III-B introduces the materials and data processing method in detail. The structure of the EddyTPNet, the training process, and the testing process are detailed in section III-C. And the loss function is formulated in section III-D.

A. Overall pipeline

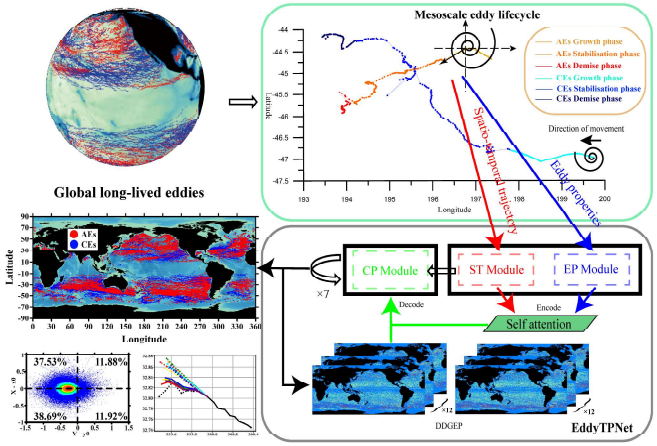

Fig. 3. Pipeline of the proposed methodology for global oceanic mesoscale eddies trajectories prediction.

The pipeline of the entire study is shown in Fig.3. First, based on the prior knowledge of the eddy dynamic mechanism and the divergence of eddy propagation direction, the directional divergence of global eddy propagation (DDGEP) grid data is constructed to simulate the physical background field and eddy propagation direction, which is used to guide and match the deep learning model [45]. Second, following the encoder-decoder structure, EddyTPNet is proposed for the mesoscale prediction of eddy propagation, which integrates LSTM cells [46] and GRU units [47] for the dynamic interaction of information. The decoder cyclic prediction (CP) module uses a combination of spatio-temporal trajectory data and directional divergence of global eddy propagation (DDGEP) data as input. The encoder consists of two modules: one focused on eddy current characteristics (EP) and the other focused on spatio-temporal trajectory characteristics (ST). A self-attention mechanism [48] is added between the encoder and the decoder to extract important information from the time series. Finally, the loss function of the optimized network is the mean absolute geodetic error (MAGE) loss, which combines earth distance and $L1$ regularized loss.

B. Data preprocessing

1) Materials and processing: The materials used in this study are obtained from the Satellite altimetric Mesoscale Eddy Trajectories Atlas, which cover the years from 1993 to 2022 and include amplitude, radius, speed-average, latitude, longitude, and time data. The data consists of Absolute Dynamic Topography (ADT) maps as input, which are first filtered by Lanczos, and then the eddy detection algorithm is used to outline the eddy closure contours. The center of the eddy is defined as the center of the circle which fits best with the contour of the maximum speed. Finally, the eddies are correlated by a tracking algorithm. This dataset is stored in the form of sequence points, and eddy information is recorded for each spatio-temporal state. Based on this dataset, the eddy characteristic dataset and spatio-temporal trajectory dataset are created.

Eddy properties dataset. This dataset comprises eddy properties because faster-moving and more energetic eddies typically have greater radius, amplitudes, and velocities [49], and the historical variance in these properties mostly indicates the stability of the eddy trajectory.

Spatio-temporal trajectory dataset. The eddy characteristics of the time, longitude, and latitude are crucial to estimating the eddy movement patterns. The input of longitude, latitude, and time can determine the precise place and time of eddy, as eddy motion is affected by geographical location, time, topography, and seasonality [50]. The motion of the eddies can be represented by the values of motion angle and displacement, as shown in Fig.2, so the meridional displacement, zonal displacement, and direction of motion are used to extend the spatial and temporal information. The detailed process can be formulated as Eq.2-4.

$$Z_t = lon_t - lon_{t-1}, \tag{2}$$

$$M_t = lat_t - lat_{t-1}, \tag{3}$$

$$A_t = \angle(\arctan\frac{|Z_t|}{|M_t|}). \tag{4}$$

Where t, lon, and lat are $t-step$, longitude, and latitude, respectively. The zonal displacement, meridional displacement, and motion direction are represented by Z, M, and A, respectively. Eq.2 and Eq.3 are the differences between the zonal and meridional displacements respectively, and instead of using longitude and latitude, the prediction of the meridional and zonal displacements is executed. The reason is that the input to the neural network is normalized data, and the amount of variation in longitude and latitude is negligible after normalization. Here $arctan$ function is used as the solution in Eq.4, and the relationship between meridional and zonal displacements can be more clearly reflected. The meridional and zonal displacements are utilized to determine the final

angle, which is then calculated using the Cartesian coordinate system as a base and used to match the newly generated DDGEP data. It is worth noting that the eddy polarity can always be guaranteed throughout the encoding and decoding process by incorporating polarity into the dataset along with state flag bits 0 and 1, respectively, as inputs to the neural network.

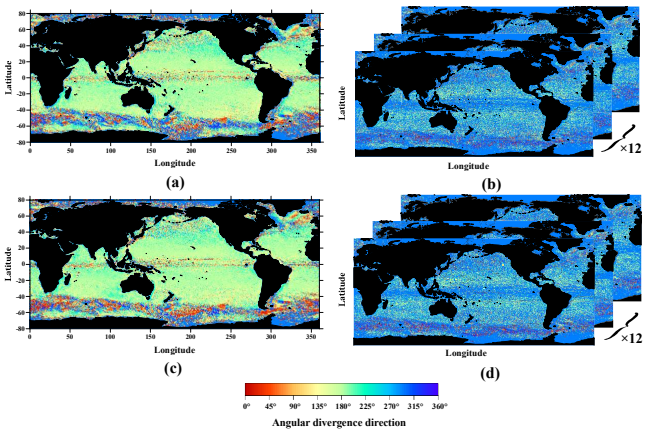

Fig. 4. Global cyclone and anticyclone eddy directional dispersion grid plots. (a) and (c) represent the statistical grid maps of the global 1993-2022 anticyclonic and cyclonic eddy directional dispersion grid maps, respectively. (b) and (d) record the anticyclonic and cyclonic eddy directional dispersion grid maps for January-December, respectively.

2) Directional divergence grid data: The idea of directional divergence comes from the physical mechanism of divergence in the direction of propagation of mesoscale eddy trajectories. First, the motion from one trajectory point to another can be determined by the amount of displacement and the angle, where the displacement in different directions is represented by the meridional displacement and the zonal displacement. The relationship between the meridional and the zonal motions is not isolated, they interact with each other, thus the usage of angles to model the dynamic connection between the relationship can guide the direction of eddy propagation. Second, mesoscale eddy trajectories are significantly influenced by wind, currents, and topography, leading to their irregular motion. Typically, the introduction of information from external background fields is necessary to enhance the accuracy of eddy trajectory predictions. Systematic analysis of the physical and data-driven dispersion of eddy propagation directions suggests that both physical and data-driven directional dispersion appropriately capture the influence of external forces [51]. Moreover, historical eddies tend to traverse regions with similar propagation directions, exhibiting minimal variability in external factors. Therefore, in the absence of incorporating multiple sources of data, such as external background fields, the creation of DDGEP data serves to reduce data usage while maintaining a high level of accuracy. It is known that cyclonic and anticyclonic eddies have different motion mechanisms and a strong correlation with seasons. Here DDGEP data with the size of 2×12 was synthesized based on the eddy polarity and different months. Each DDGEP data is calculated from eddy data of the same month and polarity. Specifically, the directional dispersion data within each grid is calculated by the following Eq.5-8.

$$\overline{lon}_t^i = \frac{(lon_{t+1} + lon_t) - (lon_t + lon_{t-1})}{2}, \tag{5}$$

$$\overline{lat}_t^i = \frac{(lat_{t+1} + lat_t) - (lat_t + lat_{t-1})}{2}, \tag{6}$$

$$A_k^i = \angle\left(\arctan\frac{\left|\overline{lat}_t^i\right|}{\left|\overline{lon}_t^i\right|}\right), \tag{7}$$

$$DDGEP_k = \frac{1}{N}\sum_{N=1}^{C} A_k^N. \tag{8}$$

Where A_k^i denotes the ith angular value within the kth grid, C denotes the number of angular values within each grid, and $DDGEP_k$ means the calculated angular value within the kth grid. The data is a matrix of size 2880×1440 with $1/8^\circ \times 1/8^\circ$ spatial resolution and the calculation process is slightly different from that of the angles, with the amount of data involved in each grid being determined by the number of eddies passing through. Specifically, Eq.5 and Eq.6 are averaged over the two consecutive days of each eddy's position to avoid transient noise. Eq.7 indicates that the angular value of each eddy is calculated. The grid region in which that eddy is located is determined, and then Eq.8 is calculated by averaging the angular values already within each grid so that the divergence angle within that grid matches the direction of propagation of most eddies, thus avoiding the effect of trajectory anomalous eddies. Due to the small amount of variation in latitude and longitude and the constant querying of data during the test, we eventually determined a grid accuracy of $1/8^\circ$ for the angular data, while ensuring accuracy and efficiency. We superimpose each DDGEP in Fig.4, it can reveal that most of the grid divergence tends to the west and a small portion to the east, which is consistent with the theory related to eddy propagation. The final directional dispersion will provide some guidance in the prediction process for the proportional distribution of the meridional and zonal displacements of the eddies and the prediction of the direction of the eddy propagation.

C. Knowledge-fused deep neural network

The overall architecture of the proposed EddyTPNet is presented in Fig.5. EddyTPNet is composed of three primary parts, namely the EP module, which focuses on eddy properties, the ST module, which can extract spatio-temporal trajectory information, and the CP module, which possesses mesoscale prediction capability.

1) Eddy properties module: LSTM has been utilized for capturing temporal information, and empirical findings suggest that the characteristics of eddies do exert a certain level of influence on their trajectory based on the analysis of variable importance. However, the magnitude of this influence is relatively minor [27]. The stable state of the eddy is constrained by the attributes of the eddy, and inputting it together with time-space information such as longitude and latitude may lose the

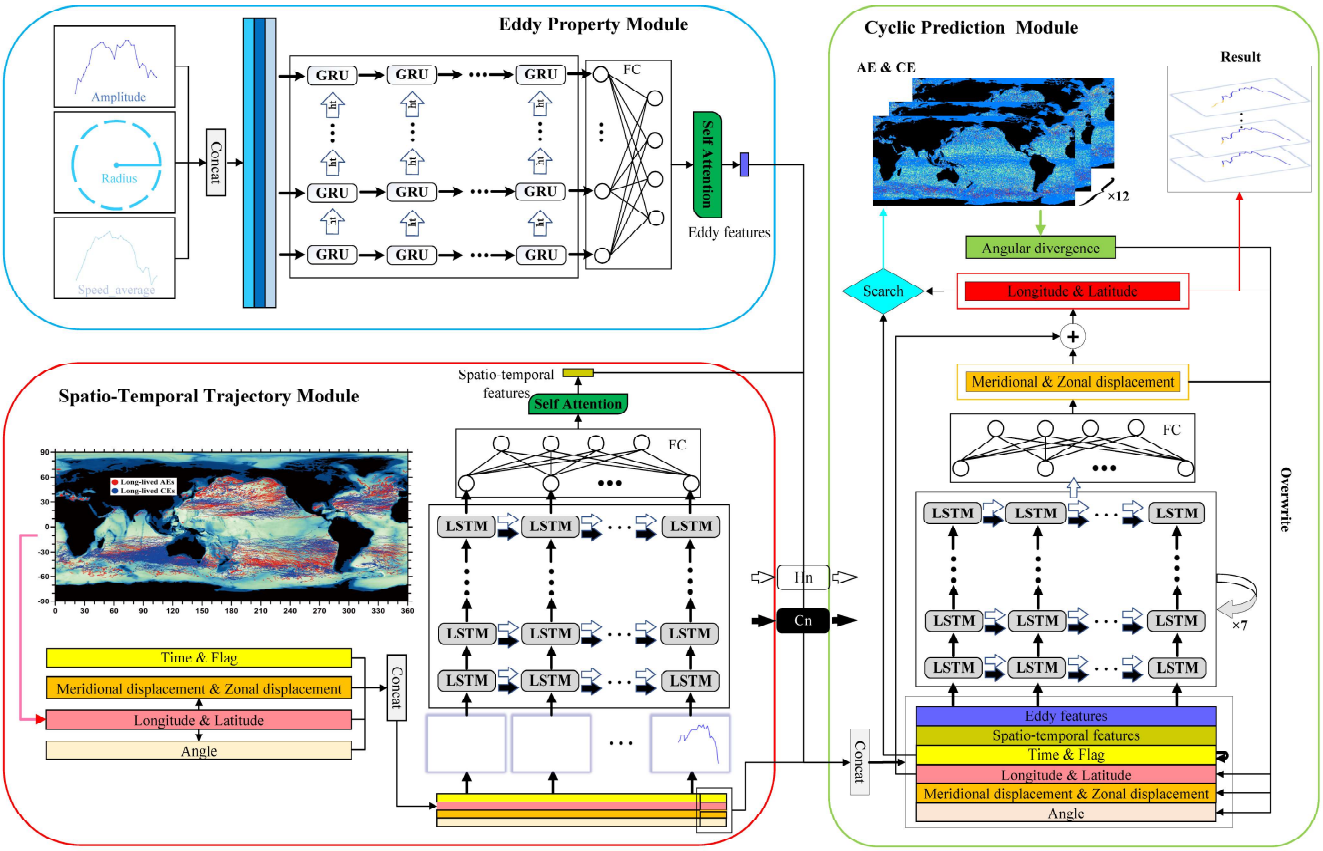

Fig. 5. The overall architecture of Knowledge-fused deep neural network for trajectory prediction of the global mesoscale eddy. The ST module, EP module, and CP module are each represented by a red, blue, and green box. Self-attention connects the modules while GRU extracts the eddy properties and LSTM derives the spatio-temporal information. DDGEP data embeds the CP module for continuous indexing.

state constraint information. The GRU unit, which also has an excellent performance in capturing timing information, has fewer parameters and is faster to train. Therefore, the GRU unit is employed in the EP module to capture the eddy energy changes of the historical sequence to constrain the state trend of the eddy in the future. Additionally, the future state of eddies is unknown, so it needs to be separated. The internal structure is shown in Fig.6(a), where the input is e_t, containing the radius, amplitude, and speed-average at step t, and the hidden state at step $t-1$ is represented by h_{t-1}. The processing of the GRU unit can be expressed as Eq.9-11.

$$Z_t = \sigma(W_z \cdot [e_t, h_{t-1}] + b_z), \tag{9}$$

$$R_t = \sigma\left(W_r \cdot [e_t, h_{t-1}] + b_r\right), \tag{10}$$

$$h_t = (1 - Z_t) * h_{t-1} + Z_t * \tan h\left(W \cdot [R_t * h_{t-1}, e_t] + b\right). \tag{11}$$

The reset gate R and the update gate Z determine the degree of retention of historical and current attributes. W and b denote the weights and biases of the corresponding gates respectively. σ is the sigmoid activation function and $*$ is the Hadamard product. h_t is used as an output to represent the current state and provide the basis for the predictions of the CP module. However, the spatio-temporal information is more important in influencing the eddy trajectory.

2) Spatio-temporal information module: The ST module uses LSTM units to memorize and encode historical spatio-temporal trajectory features. Time and eddy polarity are embedded to capture seasonal information for different eddies. We use the spatio-temporal trajectory information at time $t-1$ as the input to the ST module at time t, denoted by s_t. The output of the module can be formulated as Eq.12-16.

$$F_t = \sigma\left(W_f \cdot [h_{t-1}, s_t] + b_f\right), \tag{12}$$

$$I_t = \sigma\left(W_i \cdot [h_{t-1}, s_t] + b_i\right), \tag{13}$$

$$C_t = I_t * \tan h\left(W \cdot [h_{t-1}, s_t] + b_c\right) + F_t * C_{t-1}, \tag{14}$$

$$O_t = \sigma\left(W_o \cdot [h_{t-1}, s_t] + b_o\right), \tag{15}$$

$$h_t = O_t * \tan h\left(C_t\right). \tag{16}$$

LSTM generally performs better than GRU on the finer tasks such as latitude and longitude prediction because of the large number of parameters. Compared to GRU units, the LSTM unit is divided into three gates: input gate I_t, which determines the amount of input information at time t, forgetting gate F_t, which determines the amount of information forgotten at time $t-1$, and output gate O_t, which determines the amount of output, as shown in Fig.6(b). A separate memory unit C_t is used to save the current LSTM state information and pass it to the LSTM the next time.

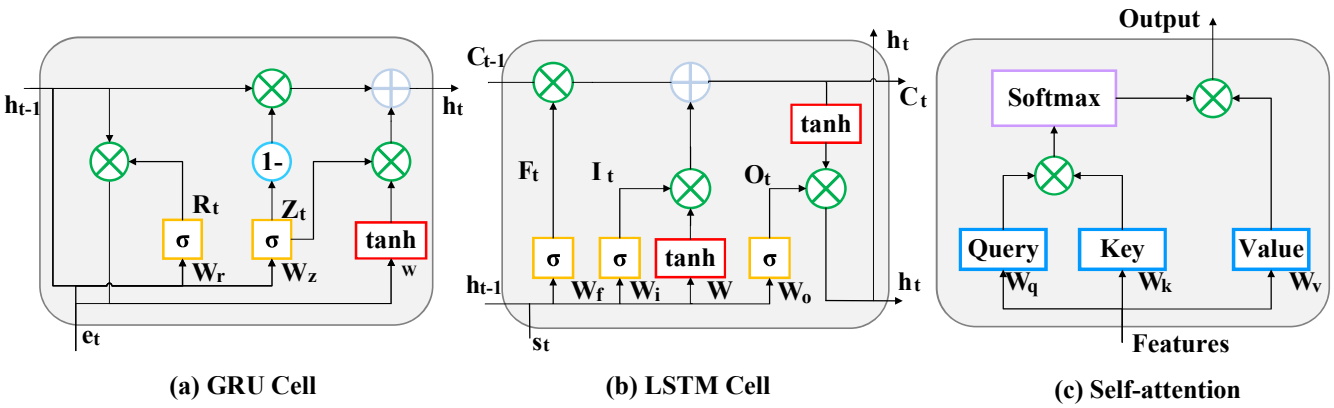

Fig. 6. (a) GRU unit structure diagram. (b) LSTM unit structure diagram. (c) Self-attention structure diagram.

3) Cyclic prediction module: The encoder-decoder with the self-attention mechanism often achieves better results. Attention as shown in Fig.6(c) is used as part of the input to the CP module, extracting key information from the last-moment hidden states of the EP and ST modules respectively. The calculation of the self-attention mechanism is as the following Eq.17.

$$Attention = W_v x * softmax\left(W_q x * [W_k x]^T\right) \quad (17)$$

Where x represents the output of the Eddy Properties Module and the Spatio-temporal Information Module, W_q, W_k and W_v correspond to weight vectors associated with value, query, and key, respectively. These weight vectors are used to calculate the weighted sum of the value vectors based on attention weights, thereby assigning varying importance to different positions and enhancing the model's capability to model long-range dependencies.

The CP module acts as a decoder with a built-in unit of LSTM and the initial states h and c are the last layers of hidden states of the ST module. The inputs are longitude, latitude, meridional displacement, zonal displacement, angle, time, and the self-attention outputs of the other two models. After preliminary experiments, it is not ideal to directly predict the eddy trajectory for the next 7 days, so we split the multi-step prediction problem [52] into multiple single-step prediction problems. The decoder is designed to predict a 1-day trajectory using 5 days of historical data, and the implicit state of this time is used as the initial state for the next prediction, and the prediction process is carried out iteratively through a sliding window technique, resulting in a total of 7 consecutive predictions. In the prediction process that follows, where the input data remains unknown, the 5-day data undergoes a continuous update utilizing predicted values, alongside calculations and lookup operations. This iterative approach ultimately ensures data integrity and fosters accurate prediction outcomes. After introducing the divergence of physical information, the physical embedding process of the training procedure is derived from the given angular scattering features. On the contrary, the testing procedure is comparatively complex. The specific algorithmic flow of the CP module, taking the testing process as an example, is shown in Algorithm. 1.

Algorithm 1 CP Module with the Directional Divergence Algorithm

Input:
Trajectory characteristics:
$(T_n)_{n=0}^{12} = (lat_n, lon_n, z_n, m_n, t_n, a_n, f_n)_{n=0}^{12}$;
Encoder information: $(s_n, x_n)_{n=0}^{4}$;
Encoder status: h_n, c_n;
Output:
for $t = 0; t < 7; t++$ **do**
//5 days of data assignment;
$Input = (T_n)_{n=j}^{j+4}, (s_n, x_n)_{n=0}^{4}$;
//Forecast;
$Output, h_n, c_n = DecoderNet(Input, h_n, c_n)$;
//Storing predicted values;
$(z_n, m_n)_{n=j+5}^{j+6} = Output$;
//Inverse normalization;
$z, m = InvNorm((z_n, m_n)_{n=j+5}^{j+6})$;
$lat, lon, f, t = InvNorm((lat_n, lon_n, f_n, t_n)_{n=j+4}^{j+5})$;
//Summation;
$lat, lon = z, m + lat, lon$;
//Normalized for next prediction;
$(lat_n, lon_n)_{n=j+5}^{j+6} = Norm(lat, lon)$;
//Search DDGEP data;
$a = Search(lat, lon, f, t)$;
//Normalized for next prediction;
$(a_n)_{n=j+5}^{j+6} = Norm(a)$;
end for
return $(lat_n, lon_n)_{n=5}^{12}$

u, v, z, m, t, f, and a represent latitude, longitude, meridional displacement, zonal displacement, time, angle, and flag respectively. In this process, the predicted meridional and zonal displacements are added to the previous day's longitude and latitude to generate the new predicted longitude and latitude, and then the eddy polarity, month, and predicted longitude and latitude are used to find the corresponding value in the DDGEP data for the next prediction input, which is used

to guide the model prediction.

D. Mean absolute geodetic error loss

The accuracy of the model prediction depends heavily on the loss function, and since the geodetic distance [40] enables a practical test of the accuracy of the eddy trajectory prediction, the MAGE loss function [53] adopted in this paper can be expressed as Eq.18.

$$MAGE = L1Loss + MGDLoss \tag{18}$$

MAGE consists of two components, the $L1$ and mean geodetic distance (MGD) losses, respectively. Specifically, the $L1$ loss as Eq.19 is used to reduce the Euclidean distance between the predicted and true values, and the MGD loss as Eq.20 is adopted to model the physical environment in which the predicted and true values are located.

$$L1 = \frac{1}{n}\sum_{i=0}^{n}(\frac{1}{m}\sum_{t=1}^{m}|x_i^t - y_i^t|) \tag{19}$$

$$MGDLoss = \frac{1}{n}\sum_{i=0}^{n}(\frac{1}{m}\sum_{t=1}^{m}\Delta\alpha_i^t) * R \tag{20}$$

Where m denotes the mth day, n is the amount of data in a batch, and x_i^t,y_i^t, and $\Delta\alpha_i^t$ indicate the predicted and true values and the central angle between them at time t, and the ith time, respectively. This central angle can be specified as Eq.21 and Eq.22, which are derivable and continuous in the domain of definition, thus the whole function is differentiable.

$$\Delta\alpha = \arctan x \tag{21}$$

$$x = \frac{\sqrt{(\cos\varphi_2 \sin\Delta\gamma)^2 + \cos\varphi_2 \sin\varphi_1 - \sin\varphi_1 \cos\varphi_2 \cos\Delta\gamma}}{\sin\varphi_1 \sin\varphi_2 + \cos\varphi_1 \cos\varphi_2 \cos\Delta\gamma} \tag{22}$$

Where γ and φ denote longitude and latitude, respectively. $\Delta\gamma$ denotes the difference in longitude between the predicted and true values.

IV. Experimental results and analysis

A. Implementation details

1) Datasets partitioning: The datasets employed in this study consist of the eddy properties dataset and the spatio-temporal trajectory dataset. To explore the influence of historical sequence length on prediction, each entry in the eddy properties dataset includes 30 days of historical feature data. In contrast, the spatio-temporal trajectory dataset comprises 30 days of historical trajectory data and 7 days of future actual data. A total of 400,648 data sequences were randomly divided (with a random seed set to 42), with 280,453 sequences allocated for training, 60,098 sequences for validation, and 60,097 sequences for testing. The data was subsequently normalized and fed into the model for training.

TABLE I
THE PREDICTION RESULTS FOR EDDYTPNET VARIANTS.

Method	MGD(km)	SGD(km)	Precision(%)
EddyTPNet-CP	7.30	51.08	84.58
EddyTPNet-ET	7.28	50.96	84.66
EddyTPNet-ST	7.20	50.37	84.90
EddyTPNet-10-7	7.18	50.28	84.99
EddyTPNet-20-7	7.19	50.36	84.91
EddyTPNet-30-7	**7.17**	**50.16**	**85.10**

2) Experimental setup: The proposed EddyTPNet was implemented on the PyTorch deep learning library. The experiments were performed under Windows 11 (Nvidia GeForce RTX 3090GPU). The hidden states of the EP module, ST module, and CP module were set to 64, 128, and 128, respectively, and the network layers were set to 4 layers. Due to the abundance of data, the direction of model descent is more accurate when more training data is obtained in each batch, and the processing speed can be accelerated, we finally set the batch size to 2048. The initial value of the learning rate was set to 0.0005 to ensure accurate convergence of the network and was adjusted using the ReduceLRPlateau strategy, with the patience value set to 5. The optimizer was Adam, and the number of training rounds was 50.

The final model predicts the trajectory for the next 7 days by learning 10 days of historical data. We will demonstrate the feasibility of the proposed method by comparing with different models in terms of the length of the history sequence, the presence or absence of physical information, and the ability to generalize the local region.

3) Baselines: To evaluate the performance of the proposed method, Extra Trees [27], Multiple Linear Regression [40], Random Forest [27], Gradient Boosting [27], MesoGRU [42], MesoLSTM [42], Seq2Seq [54], EddyTPNet-ST, EddyTPNet-EP, and EddyTPNet-CP models were selected as the state-of-the-art baselines.

Extra Trees, Random Forest, Gradient Boosting and Multiple Linear Regression. Traditional machine learning models. One model was built for each prediction day, for a total of seven models.

MesoLSTM and MesoGRU. The MesoGRU model is a variant of the GRU model, consisting of 4 layers of GRU. It utilized historical data from the previous three days to predict future data for the next day. To achieve multi-step prediction, this model employed a sliding window approach for training and testing. For this experiment, the hidden state was set to 128. Additionally, the experimental setup for the MesoLSTM model was the same as the MesoGRU model, except for the substitution of GRU units with LSTM units.

Seq2Seq. The Seq2Seq model employs an encoder-decoder architecture composed of 4 recurrent neural networks (RNNs). The encoder encoded historical data of the past 10 days of eddies, while the decoder performed daily predictions using a sliding window approach, making a total of 7 predictions. The network was configured with 128 hidden states.

EddyTPNet-CP. EddyTPNet-CP, as a variant of EddyTPNet, removed the encoding part of the Eddy attribute module and the spatio-temporal module. However, as the decoder of

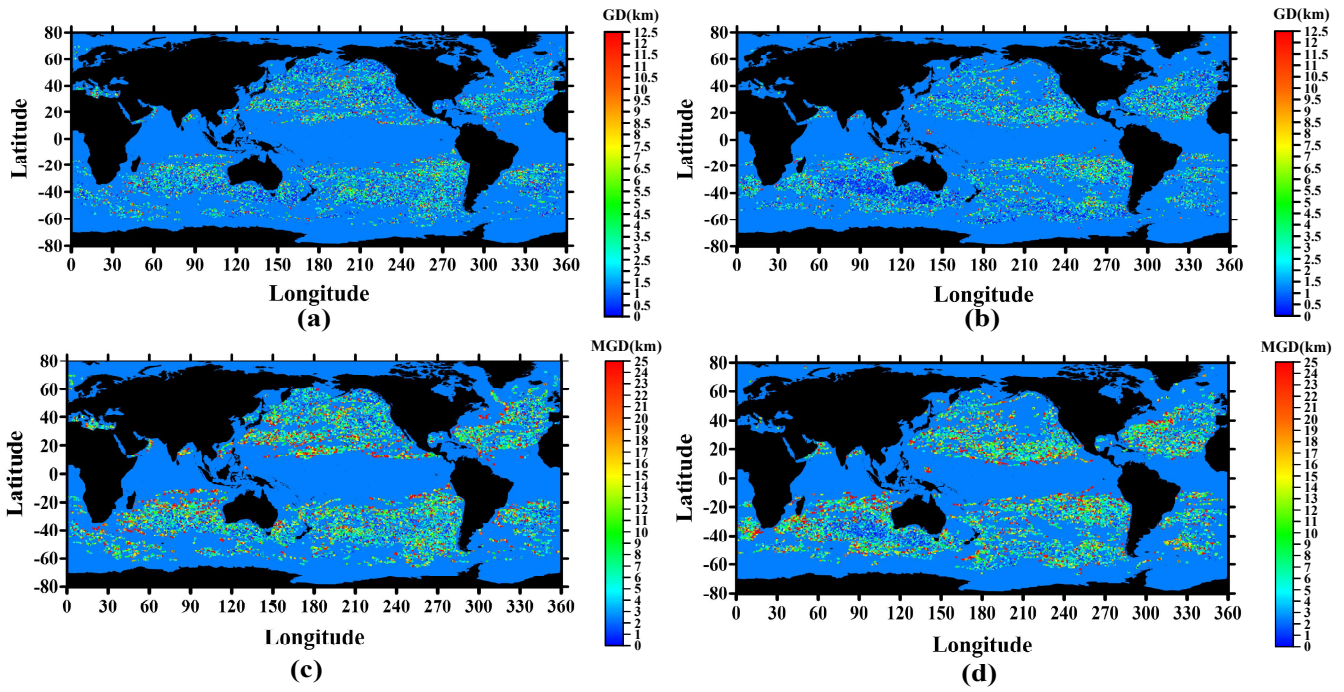

Fig. 7. Global geodetic distance distribution of test sets. (a) and (b) represent the Geodetic Distance (GD) distribution of the next day prediction results for Anticyclonic eddies and Cyclonic eddies respectively. (c) and (d) denote the Mean Geodetic Distance (MGD) distribution of the 7-day forecast results for Anticyclonic eddies and Cyclonic eddies respectively.

TABLE II
PREDICTION ERRORS OF MODELS UNDER DIRECTIONAL DIVERGENCE PHYSICAL INFORMATION.

Method	MGD(km)		SGD(km)		Precision(%)	
	Angular	Without angular	Angular	Without angular	Angular	Without angular
MesoGRU	8.02	8.87	56.12	62.05	82.20	79.29
MesoLSTM	8.02	8.86	56.13	62.00	82.20	79.37
Seq2Seq	7.44	7.67	52.04	53.70	84.19	83.80
EddyTPNet-CP	7.30	7.77	51.08	54.37	84.58	83.47
EddyTPNet-ET	7.28	7.75	50.96	54.22	84.66	83.54
EddyTPNet-ST	7.20	7.67	50.37	53.60	84.90	83.84
EddyTPNet	**7.18**	7.63	**50.28**	53.38	**84.99**	83.97

EddyTPNet, it still possesses the ability of multi-step prediction. The training and prediction processes remain the same as the CP module while maintaining the same hyperparameters as EddyTPNet.

EddyTPNet-ST. The model conformed to the basic architecture of the Seq2Seq and was based on EddyTPNet with the EP module removed and other parameters set in the same way as EddyTPNet.

EddyTPNet-EP. This model was a model where only the EP module and the CP module are retained, the encoder-decoder structure is split, and the parameter settings remain unchanged. The eddy properties were used as input to the EP module, with self-attention as the connection between the EP module and the CP module.

4) Evaluation metrics: Applying the above model, the evaluation indicators used for the assessment of the results were accuracy, geographical MGD, and summed geodetic distance (SGD) respectively.

MGD and SGD. MGD and SGD are the mean geodetic distance and the sum of 7-day geodetic distances, respectively. They can reflect the error between the predicted and true values by the true Earth distance.

Precision. As grid data with 12.5km or 25km resolution can be applied to real scenarios, we define trajectories with MGD errors below 12.5km as prediction accuracy.

B. Comparison and analysis of prediction results

To facilitate marine exploration tasks, it is customary to deploy equipment in advance, hence the establishment of a 7-day prediction window. The decoding module utilizes a recurrent encoder to generate the next predicted value, ensuring that long-term predictions do not compromise short-term accuracy. The trajectory is only predicted for the following day, and the distribution of geographic distances (GD) is depicted in Fig. 7(a-b), showcasing highly accurate prediction results. Furthermore, Fig. 7(c-d) presents the distribution of the average geographic distance (MGD) within the 7-day

prediction window, highlighting the strong predictive efficacy of the proposed EddyTPNet.

Here, we label the error as MGD $>$ 12.5km, and the poor forecast area mainly exists near the shore, the equator, and poles, as shown in Fig.8. Nevertheless, it can be seen that the error distributions are relatively uniform and there are no problems with regional failures. To further demonstrate the effectiveness and complexity of the method, several comparative experiments were conducted.

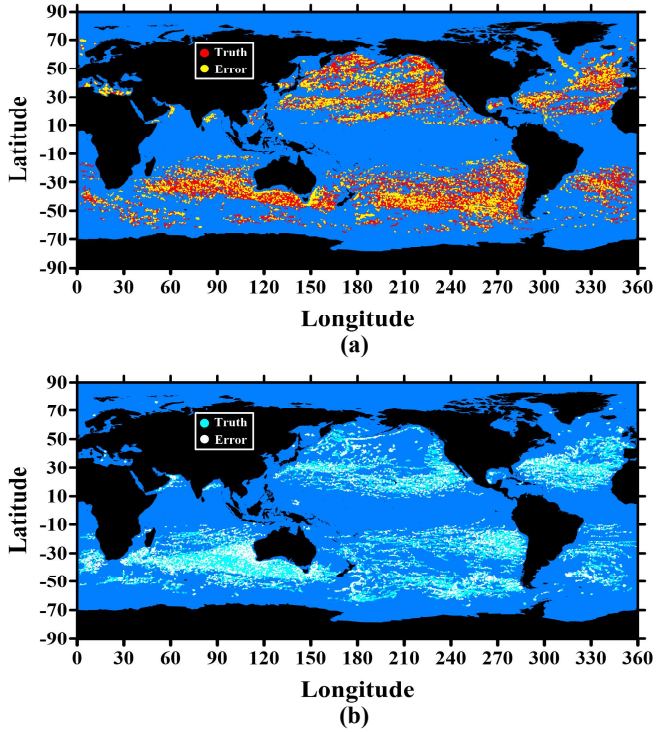

Fig. 8. Global distribution of test sets and errors. (a) denotes Anticyclonic eddies and (b) denotes Cyclonic eddies. Among them, the trajectory of the Mean Geodetic Distance (MGD) $>$ 12.5km is marked as an error.

1) Global Comparative Experiment: The EddyTPNet variants are first compared. On the one hand, EddyTPNet-10-7, EddyTPNet-20-7, and EddyTPNet-30-7 are used to represent the prediction ability of EddyTPNet for different lengths of historical sequences, which provide the encoder with 10, 20, and 30 days of historical spatio-temporal and eddy state information, respectively. After introducing the directional divergence physics information, Table I shows that 30 days has the best prediction effect on 7 days, but in general, different lengths of historical sequences do not have much effect on EddyTPNet. The reason is that the orientation information embedded by the neural network largely replaces the encoding process. Therefore, the subsequent experiments are still based on 10 days predicting 7 days. On the other hand, EddyTPNet-CP, EddyTPNet-ST, and EddyTPNet-EP are used to demonstrate the importance of each module. Table I also indicates that each module is effective in improving the network accuracy when directional dispersion is included, and the ST module is more important than the ET module, as illustrated in Table II.

The experimentation now shifts towards comparing the MesoGRU, MesoLSTM, and Seq2Seq models. Combined with Fig.9 and Table II, it can be found that MesoLSTM and MesoGRU are less expressive and not sensitive enough to the motion distance of eddy currents. In contrast, the prediction accuracy of the Seq2Seq model, the basic architecture of EddyTPNet, is much higher than that of LSTM and GRU, indicating that the encoder part can remember and extract the historical propagation distance of eddy motion.

The most expressive EddyTPNet has good prediction performance in predicting both motion distance and motion trajectory direction. The physical information of the direction dispersion is crucial for the accuracy of the eddy trajectory prediction. Table II presents the statistical results for different models with and without the introduction of directionally dispersive physical information. The results imply that the introduction of directionally dispersive physical information can improve the effectiveness of the deep neural network, and the angular dispersion information shown in Fig.9 can greatly fit the dispersion ratios of the true values, which solves the problem of the imbalance between the westward and eastward eddy data.

Further, some examples of single trajectories are plotted in Fig.10. A closer look reveals that the prediction results without the introduction of the directional dispersion physical information are more inclined to smooth and less varied trajectories. whereas, with the introduction of the directional dispersion physical information, the network can accurately capture the instantaneous variations of the eddies. The results indicate that EddyTPNet has the highest prediction accuracy and the smallest prediction error of 84.99%, while MGD and SGD are 7.18 km and 50.28 km, respectively.

The daily prediction errors and accuracies of each model were calculated separately to fully demonstrate the advantages of each model in the time series task, as shown in Table III. The prediction accuracy of each model decreases over time due to the accumulation of errors. Nevertheless, the results show that EddyTPNet still outperforms the other models in different periods, and still achieves a trajectory prediction rate of 71.26% on the seventh day. Furthermore, in contrast to other models, EddyTPNet exhibits a consistent increase in daily prediction errors, with a difference of approximately 1.4 km. This observation suggests that the model demonstrates greater stability in forecasting long sequences.

2) Local Regional Prediction Results: To examine the generalizability of the proposed method, we opted to focus our research on the South China Sea and the North Atlantic region [55]. The South China Sea stands out due to its distinctive geographical location and intricate topography, which has led to the development of prediction methods specifically tailored to this area's eddy trajectories. Moreover, the North Atlantic also serves as a significant source region for eddies of varying lifespan scales. For our study, we gathered data from the South China Sea region ($5^\circ - 25^\circ N, 105^\circ - 125^\circ E$) and the North Pacific region ($20^\circ - 40^\circ N, 20^\circ - 40^\circ W$) spanning a period of 30 years (1993-2022). This data was divided into 31,050 and 80,751 trajectory data, respectively. Subsequently, these datasets were partitioned into training, validation, and testing

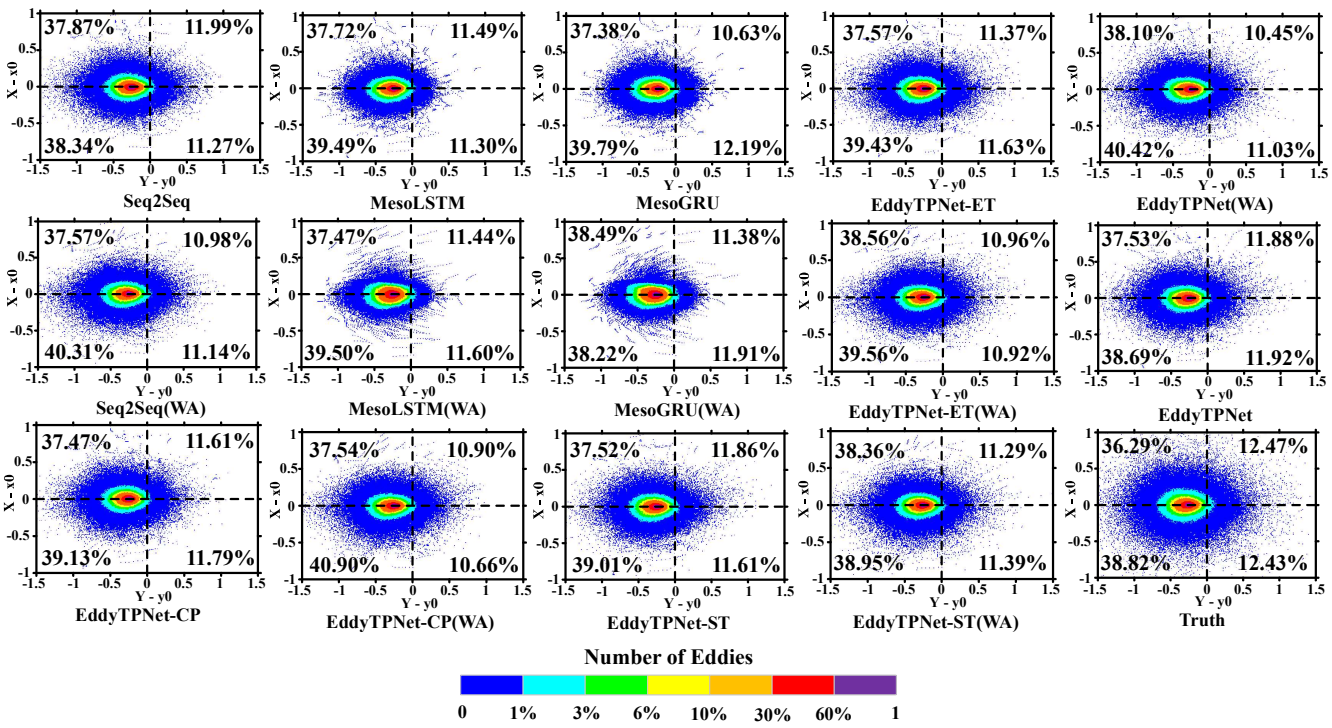

Fig. 9. Distribution of eddy trajectories predicted by different methods. Each grid is populated by the number of coordinates consisting of the meridional and zonal displacements of the eddy. Where the color indicates the number of eddies per unit grid, decreasing from purple to blue, respectively. The percentages in the figure indicate the proportion of the number of eddies accounted for by the four quadrants in the Cartesian coordinate system. (WA) denotes no angle dispersion information.

TABLE III
DAILY PREDICTION ERRORS UNDER DIFFERENT MODELS.

Days	MesoLSTM		MesoGRU		Seq2Seq		EddyTPNet-CP		EddyTPNet-ET		EddyTPNet-ST		EddyTPNet	
	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)
1	3.24	96.30	3.25	96.28	3.07	96.44	3.02	96.51	3.02	96.50	3.00	96.50	**3.00**	**96.51**
2	4.89	92.44	4.90	92.41	4.53	93.26	4.44	93.38	4.45	93.22	4.41	93.33	**4.41**	**93.34**
3	6.51	87.74	6.52	87.77	5.98	89.39	5.87	89.44	5.88	89.45	5.80	89.63	**5.80**	**89.69**
4	8.04	82.74	8.05	82.79	7.39	84.97	7.26	85.24	7.25	85.22	7.17	85.49	**7.16**	**85.50**
5	9.60	77.16	9.60	77.38	8.86	80.12	8.71	80.44	8.69	80.54	8.58	80.87	**8.56**	**80.89**
6	11.15	71.73	11.15	71.92	10.36	74.87	10.17	75.54	10.13	75.70	10.00	76.15	**9.97**	**76.19**
7	12.68	66.43	12.67	66.72	11.86	69.72	11.62	70.58	11.56	71.00	11.40	71.21	**11.37**	**71.37**

sets with a ratio of 7:1.5:1.5. In both regions, we employed various methods, including Extra Tree [27], Multiple Linear Regression [40], Random Forest [27], Gradient Boosting [27], and MesoGRU [42], to predict eddies trajectories. However, there are complex background fields in different oceanic regions, which result in significant differences in local areas. To accommodate these regional differences, EddyTPNet migrates the model trained on a global scale to local regions for fine-tuning experiments, with an initial learning rate set to 0.00001 [56]. Ultimately, results with prediction errors of less than 25 kilometers are considered accurate predictions. A summary of the 7-day forecast outcomes can be found in Table IV.

Meanwhile, the daily evaluations for prediction results of the South China Sea and the North Atlantic region are summarized in Table V and Table VI for more efficient comparisons. It can be seen that EddyTPNet's prediction accuracy exceeds

TABLE IV
PREDICTION RESULTS OF DIFFERENT MODELS IN THE SOUTH CHINA SEA AND THE NORTH ATLANTIC.

The South China Sea	MGD(km)	SGD(km)	Preicision(%)
Extra Trees	34.75	243.25	28.98
MesoGRU	32.09	224.63	42.98
Multiple Linear Regression	24.05	168.35	63.03
Random Forest	23.17	162.20	65.67
Gradient Boosting	23.07	161.49	65.52
EddyTPNet	**21.35**	**149.45**	**69.32**
The North Atlantic	**MGD(km)**	**SGD(km)**	**Preicision(%)**
Extra Trees	15.47	108.29	88.74
MesoGRU	14.92	104.44	87.41
Multiple Linear Regression	10.42	72.94	93.49
Random Forest	10.21	71.49	93.76
Gradient Boosting	11.44	80.11	93.73
EddyTPNet	**8.77**	**61.39**	**94.72**

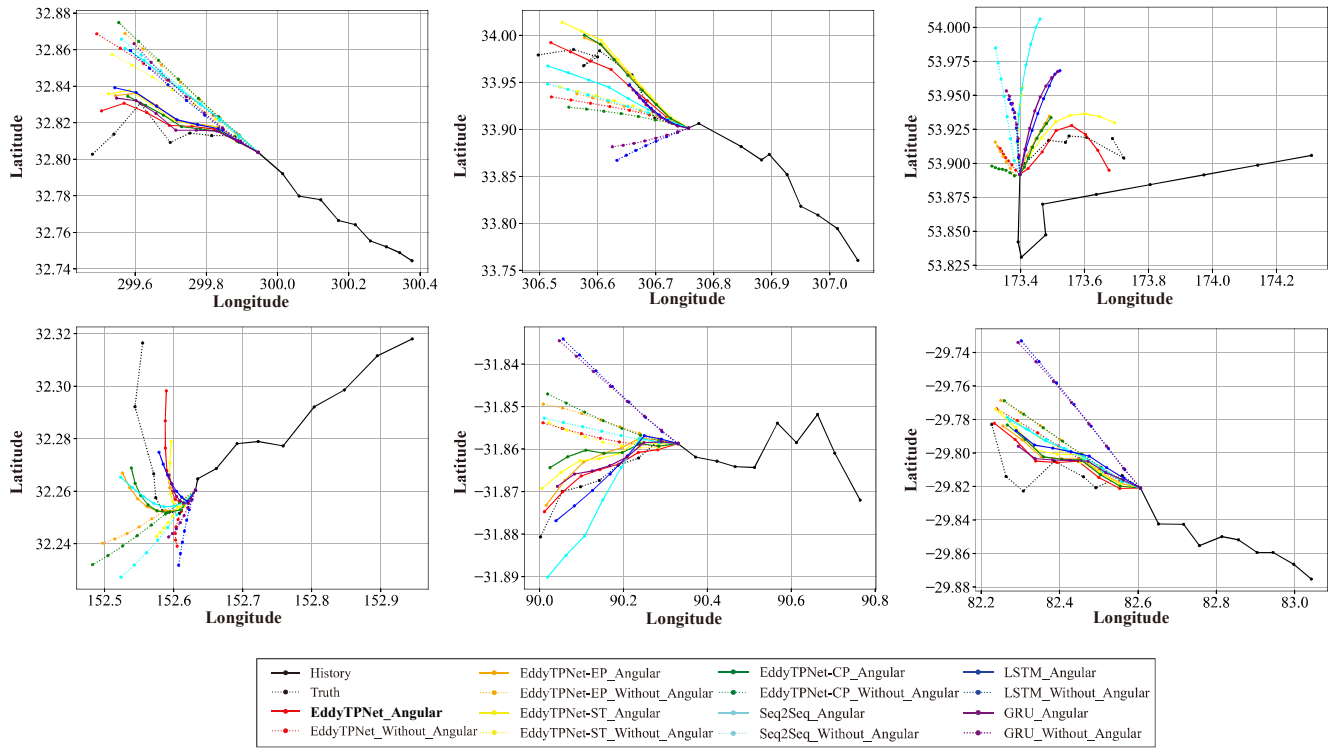

Fig. 10. Examples of trajectory prediction by different methods under directional divergence and without directional divergence.

that of the competing models. Nevertheless, there is still much room for improvement in detailed positioning accuracy. On the one hand, there are abundant types of eddy tracks in the South China Sea and the North Atlantic region, and the mechanism of each eddy is different. The insufficient number of eddy track data of each type makes it difficult to fit the eddy tracks, and it is unscientific to embed all the data into a lightweight model. On the other hand, there is a large discrepancy between the historical data and the predicted data, which reflects the complexity of the evolution of eddy motion and its dynamical mechanisms.

Consequently, we try to determine the inconsistent eddy trajectories and conduct elementary statistics on eddy trajectories, and preliminary classification of the phenomenon of complex eddy trajectories. The inconsistent trajectory cases that can be used to prove our hypothesis are shown in Fig.11, here we temporarily define them as anomalous eddy trajectories [57].

It is necessary to trace the complex ocean dynamics mechanism [58] for the cause of the anomalous eddy trajectory. The conclusion can be drawn that the introduction of prior knowledge has indeed solved the influence caused by uncertain factors. It is undeniable that the exploration of anomalous eddy trajectories still requires further research. In general, the proposed EddyTPNet with the introduction of directional divergence physical information can achieve promising prediction results.

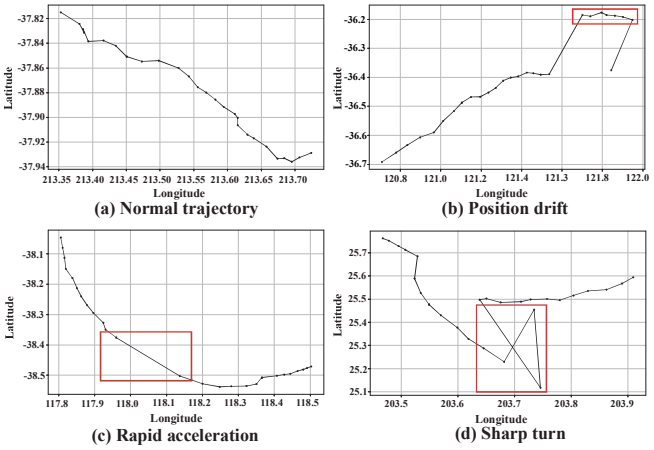

Fig. 11. Schematic diagram of the classification of eddy trajectories.

V. Conclusions and Future Work

The trajectory prediction of the oceanic eddy is a scientific issue that has not been well tackled in oceanography. With the idea of artificial intelligence for science in mind, this paper proposes a deep learning model for the estimation of eddy trajectories in oceans. It combines different modules using recurrent neural networks (GRU, LSTM) and self-attention modules to forecast the trajectory location of eddies up to 7 days in the future, based on 10 days of data (the data is mainly the amplitude, the radius, the speed average, the current displacement, the current location, the time). Moreover, a map created by averaging the past eddy trajectories (angular divergences) at every grid location was used to help the network. The results are promising and the comparison with state-of-the-art are performed.

TABLE V
DAILY PREDICTION RESULTS OF DIFFERENT MODELS IN THE SOUTH CHINA SEA.

Days	Extra Trees		MesoGRU		Mutiple Linear Regression		Random Forest		Gradient Boosting		EddyTPNet	
	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)
1	14.81	84.89	21.4	70.74	9.61	93.80	9.21	93.82	10.59	93.71	**8.28**	**94.74**
2	21.83	69.29	21.23	70.50	14.43	85.10	14.50	85.32	14.71	85.77	**12.74**	**87.27**
3	29.44	51.93	21.53	69.90	19.29	75.10	19.62	74.67	19.08	96.00	**17.25**	**78.34**
4	35.09	40.34	36.64	38.45	23.39	64.80	24.21	63.31	23.10	65.89	**21.32**	**69.15**
5	41.33	33.21	36.79	38.04	27.76	56.68	29.14	53.11	27.34	56.85	**25.87**	**60.82**
6	47.59	25.89	36.82	37.81	31.76	48.30	33.61	45.28	31.41	48.60	**29.90**	**52.92**
7	53.13	20.46	50.25	23.08	35.96	40.79	38.09	37.81	35.34	41.76	**34.12**	**46.46**

TABLE VI
DAILY PREDICTION RESULTS OF DIFFERENT MODELS IN THE NORTH ATLANTIC.

Days	Extra Trees		MesoGRU		Mutiple Linear Regression		Random Forest		Gradient Boosting		EddyTPNet	
	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)
1	7.29	95.72	9.92	93.91	4.57	98.06	4.86	97.76	7.23	97.96	**4.06**	**98.07**
2	10.36	92.19	9.90	93.84	6.65	96.28	6.80	96.24	8.59	96.29	**5.72**	**96.73**
3	13.08	87.72	9.95	93.87	8.61	94.45	8.58	94.37	10.05	94.39	**7.29**	**95.08**
4	15.59	82.90	16.97	81.85	10.47	91.97	10.25	92.19	11.34	92.17	**8.74**	**93.23**
5	18.32	77.30	16.93	81.44	12.35	88.97	11.93	89.88	12.83	89.52	**10.26**	**90.91**
6	20.57	71.20	17.10	81.30	14.23	85.64	13.67	86.86	14.32	86.61	**11.86**	**88.50**
7	23.09	64.96	23.71	62.79	16.07	81.38	15.40	83.49	15.75	83.25	**13.43**	**85.26**

The limitation of this study is that despite directional divergence information being used to fit the eddy movement trend, there are still eastward moving trajectories with low accuracy. As shown in Fig.9, the number of predicted eastward trajectories is significantly reduced. Possible reasons and future research may exist in:

(1) Since most trajectories are westward, the neural network neutralizes the predictions of east-west trajectories to ensure a high fit. Due to the difference in dynamic mechanism, the prediction distance of the westward trajectory is shortened and the ability to represent the eastward trajectory is decreased;

(2) The DDGEP data can guarantee the motion trend of most eddies while can not accurately represent the few eastward trajectories as the data are averaged. The more detailed dynamic mechanism and motion rules should be embedded to enhance the prediction ability;

(3) The discovery of anomalous eddy trajectories will lead to a deeper understanding of the movement of eddies. The causes of the abrupt changes and the environmental impact are also worthy of further study.

Notwithstanding these limitations, this exploratory study offers some insight into the feasibility of solving the prediction problem of ocean phenomena based on knowledge-fused time series deep neural networks. Within this knowledge-fused deep learning framework, there are various potentials to obtain a better outlook for oceanic eddy trajectory prediction and explore the model's generalization capability in other oceanic phenomena prediction.

VI. ACKNOWLEDGMENTS

The altimetric Mesoscale Eddy Trajectories Atlas (META3.2 DT) was produced by SSALTO/DUACS and distributed by AVISO+ with support from CNES, in collaboration with Mediterranean Institute for Advanced Studies (IMEDEA). The dataset includes allsat version and twosat version and can be downloaded with the linage META3.2 DT. In the spirit of reproducibility, the Python code of the proposed prediction methodology can be accessed with the linkage EddyTPNet.

REFERENCES

[1] X. Wang, X. Wang, C. Li, Y. Zhao, and P. Ren, "Data-attention-yolo (day): A comprehensive framework for mesoscale eddy identification," *Pattern Recognition*, vol. 131, p. 108870, 2022.

[2] Q. Liu, Y. Liu, and X. Li, "A deep learning model for eddy tracking based on multi-source remote sensing imagery," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, 2021, pp. 7576–7579.

[3] D. B. Chelton, M. G. Schlax, and R. M. Samelson, "Global observations of nonlinear mesoscale eddies," *Progress in Oceanography*, vol. 91, no. 2, pp. 167–216, 2011.

[4] Q. Wang, C. Pang, and C. Dong, "Role of submesoscale processes in the isopycnal mixing associated with subthermocline eddies in the philippine sea," *Deep Sea Research Part II: Topical Studies in Oceanography*, vol. 202, p. 105148, 2022.

[5] D. Dong, P. Brandt, P. Chang, F. Schütte, X. Yang, J. Yan, and J. Zeng, "Mesoscale eddies in the northwestern pacific ocean: Three-dimensional eddy structures and heat/salt transports," *Journal of Geophysical Research: Oceans*, vol. 122, no. 12, pp. 9795–9813, 2017.

[6] Y. Du, X. Yang, J. Yang, S. Tan, W. Ma, Z. Li, and X. Li, "Effects of temperature on sea surface radar backscattering under neutral and nonneutral atmospheric conditions for wind retrieval applications: a numerical study," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 4, pp. 2727–2743, 2020.

[7] Q. Xia, G. Li, and C. Dong, "Global oceanic mass transport by coherent eddies," *Journal of Physical Oceanography*, vol. 52, no. 6, pp. 1111–1132, 2022.

[8] J. Martínez-Moreno, A. M. Hogg, M. H. England, N. C. Constantinou, A. E. Kiss, and A. K. Morrison, "Global changes in oceanic mesoscale currents over the satellite altimetry record," *Nature Climate Change*, vol. 11, no. 5, pp. 397–403, 2021.

[9] D. Zhao, Y. Xu, X. Zhang, and C. Huang, "Global chlorophyll distribution induced by mesoscale eddies," *Remote Sensing of Environment*, vol. 254, p. 112245, 2021.
[10] M. Naeije, K. Wakker, R. Scharroo, and B. Ambrosius, "Observation of mesoscale ocean currents from geosat altimeter data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 47, no. 5, pp. 347–368, 1992.
[11] J. Ji, C. Dong, X. Liu, T. Liu, Y. Yu, K. T. L. K. Sian, B. J. Bethel, and H. Zhao, "Influence of oceanic mesoscale eddy on the atmospheric boundary layer based on an idealized model," *Deep Sea Research Part II: Topical Studies in Oceanography*, vol. 202, p. 105146, 2022.
[12] C. Dong, L. Liu, F. Nencioli, B. J. Bethel, Y. Liu, G. Xu, J. Ma, J. Ji, W. Sun, H. Shan *et al.*, "The near-global ocean mesoscale eddy atmospheric-oceanic-biological interaction observational dataset," *Scientific Data*, vol. 9, no. 1, p. 436, 2022.
[13] A. S. J. Wyatt, J. J. Leichter, L. Washburn, L. Kui, P. J. Edmunds, and S. C. Burgess, "Hidden heatwaves and severe coral bleaching linked to mesoscale eddies and thermocline dynamics," *Nature Communications*, vol. 14, no. 1, p. 25, 2023.
[14] Q. Wang, C. Dong, J. Dong, H. Zhang, and J. Yang, "Submesoscale processes-induced vertical heat transport modulated by oceanic mesoscale eddies," *Deep Sea Research Part II: Topical Studies in Oceanography*, vol. 202, p. 105138, 2022.
[15] Y. Liu, Q. Zheng, and X. Li, "Characteristics of global ocean abnormal mesoscale eddies derived from the fusion of sea surface height and temperature data by deep learning," *Geophysical Research Letters*, vol. 48, no. 17, p. e2021GL094772, SEP 16 2021.
[16] B. Lim and S. Zohren, "Time-series forecasting with deep learning: a survey," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2194, p. 20200209, 2021.
[17] Y. Zhou, C. Lu, K. Chen, and X. Li, "Multilayer fusion recurrent neural network for sea surface height anomaly field prediction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–11, 2021.
[18] G. Chen and G. Han, "Contrasting short-lived with long-lived mesoscale eddies in the global ocean," *Journal of Geophysical Research: Oceans*, vol. 124, no. 5, pp. 3149–3167, 2019.
[19] F. Teng, C. Dong, L. K. Sian, K. T. Choy, J. Ji, and W. Zhu, "Wind work on oceanic mesoscale eddies in the northeast tropical pacific ocean," *Frontiers in Marine Science*, 2023.
[20] J. Shi, D. Lü, Y. Wang, Y. Du, Y. Pang, D. Yang, X. Wang, X. Dong, and X. Yang, "Recent progress of earth science satellite missions in china," *Chin. J. Space Sci*, vol. 42, pp. 712–723, 2022.
[21] J. Souza, C. de Boyer Montégut, C. Cabanes, and P. Klein, "Estimation of the agulhas ring impacts on meridional heat fluxes and transport using argo floats and satellite data," *Geophysical Research Letters*, vol. 38, no. 21, 2011.
[22] F. Nencioli, C. Dong, T. Dickey, L. Washburn, and J. C. McWilliams, "A vector geometry–based eddy detection algorithm and its application to a high-resolution numerical model product and high-frequency radar surface velocities in the southern california bight," *Journal of atmospheric and oceanic technology*, vol. 27, no. 3, pp. 564–579, 2010.
[23] Y. Du, X. Yang, J. Yang, S. Tan, W. Ma, Z. Li, and X. Li, "Effects of temperature on sea surface radar backscattering under neutral and nonneutral atmospheric conditions for wind retrieval applications: A numerical study," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 4, pp. 2727–2743, 2020.
[24] Y. Du, X. Yang, J. Yang, and X. Li, "A numerical study of sst effects on ocean radar backscattering," in *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2020, pp. 5725–5728.
[25] A. R. Robinson and W. G. Leslie, "Estimation and prediction of oceanic eddy fields," *Progress in Oceanography*, vol. 14, pp. 485–510, 1985.
[26] X. Jiang, C. Dong, Y. Ji, C. Wang, Y. Shu, L. Liu, and J. Ji, "Influences of deep-water seamounts on the hydrodynamic environment in the northwestern pacific ocean," *Journal of Geophysical Research: Oceans*, vol. 126, no. 12, p. e2021JC017396, 2021.
[27] X. Wang, H. Wang, D. Liu, and W. Wang, "The prediction of oceanic mesoscale eddy properties and propagation trajectories based on machine learning," *Water*, vol. 12, no. 9, 2020.
[28] J. Wang, "A nowcast/forecast system for coastal ocean circulation using simple nudging data assimilation," *Journal of Atmospheric and Oceanic Technology*, vol. 18, no. 6, pp. 1037–1047, 2001.
[29] C. Dong, G. Xu, G. Han, B. J. Bethel, W. Xie, and S. Zhou, "Recent developments in artificial intelligence in oceanography," *Ocean-Land-Atmosphere Research*, vol. 2022, 2022.
[30] H. Wang and X. Li, "Deepblue: Advanced convolutional neural network applications for ocean remote sensing," *IEEE Geoscience and Remote Sensing Magazine*, 2023.
[31] Y. Ren, X. Li, and H. Xu, "A deep learning model to extract ship size from sentinel-1 sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2021.
[32] Y. Ren, X. Li, X. Yang, and H. Xu, "Development of a dual-attention u-net model for sea ice and open water classification on sar images," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2021.
[33] Z. Liao, Q. Gu, S. Li, and Y. Sun, "A knowledge transfer-based adaptive differential evolution for solving nonlinear equation systems," *Knowledge-Based Systems*, vol. 261, p. 110214, 2023.
[34] X. Li and R. Fablet, "Foreword to the special issue on the remote sensing of the world oceans," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 11, pp. 4895–4897, 2016.
[35] Y. Liu, X. Li, and Y. Ren, "A deep learning model for oceanic mesoscale eddy detection based on multi-source remote sensing imagery," in *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2020, pp. 6762–6765.
[36] Y. Liu, L. Gao, Q. Liu, and X. Li, "Dual-branch neural network for mesoscale eddy identification based on multi-variables remote sensing data," in *2023 Photonics & Electromagnetics Research Symposium (PIERS)*. IEEE, 2023, pp. 275–279.
[37] Z.-H. Zhou, *Machine learning*. Springer Nature, 2021.
[38] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
[39] Z. Zhang, X. Yang, L. Shi, B. Wang, Z. Du, F. Zhang, and R. Liu, "A neural network framework for fine-grained tropical cyclone intensity prediction," *Knowledge-Based Systems*, vol. 241, p. 108195, 2022.
[40] J. Li, G. Wang, H. Xue, and H. Wang, "A simple predictive model for the eddy propagation trajectory in the northern south china sea," *Ocean Science*, vol. 15, no. 2, pp. 401–412, 2019.
[41] J. Wen, J. Yang, Y. Li, and L. Gao, "Harmful algal bloom warning based on machine learning in maritime site monitoring," *Knowledge-Based Systems*, vol. 245, p. 108569, 2022.
[42] X. Wang, X. Wang, M. Yu, C. Li, D. Song, P. Ren, and J. Wu, "Mesogru: Deep learning framework for mesoscale eddy trajectory prediction," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2021.
[43] Y.-R. Wang and X.-M. Li, "Arctic sea ice cover data from spaceborne synthetic aperture radar by deep learning," *Earth System Science Data*, vol. 13, no. 6, pp. 2723–2742, 2021.
[44] J. Dong, B. Fox-Kemper, H. Zhang, and C. Dong, "The scale of submesoscale baroclinic instability globally," *Journal of Physical Oceanography*, vol. 50, no. 9, pp. 2649–2667, 2020.
[45] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," *Nature Reviews Physics*, vol. 3, no. 6, pp. 422–440, JUN 2021.
[46] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[47] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, 1994.
[48] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 30, 2017.
[49] J. Ji, C. Dong, B. Zhang, Y. Liu, B. Zou, G. P. King, G. Xu, and D. Chen, "Oceanic eddy characteristics and generation mechanisms in the kuroshio extension region," *Journal of Geophysical Research: Oceans*, vol. 123, no. 11, pp. 8548–8567, 2018.
[50] T. Rohr, C. Harrison, M. C. Long, P. Gaube, and S. C. Doney, "The simulated biological response to southern ocean eddies via biological rate modification and physical transport," *Global Biogeochemical Cycles*, vol. 34, no. 6, 2020.
[51] G. Chen, X. Chen, and C. Cao, "Divergence and dispersion of global eddy propagation from satellite altimetry," *Journal of Physical Oceanography*, vol. 52, no. 4, pp. 705–722, 2022.
[52] N. Mohajerin and S. L. Waslander, "Multistep prediction of dynamic systems with recurrent neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3370–3383, 2019.
[53] L. Ge, B. Huang, X. Chen, and G. Chen, "Medium-range trajectory prediction network compliant to physical constraint for oceanic eddy," *IEEE Transactions on Geoscience and Remote Sensing*, 2023.
[54] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," *Advances in neural information processing systems*, vol. 27, 2014.

[55] A. Ioannou, S. Speich, and R. Laxenaire, "Characterizing mesoscale eddies of eastern upwelling origins in the atlantic ocean and their role in offshore transport," *Frontiers in Marine Science*, vol. 9, p. 835260, 2022.
[56] X. Zhang, H. Wang, S. Wang, Y. Liu, W. Yu, J. Wang, Q. Xu, and X. Li, "Oceanic internal wave amplitude retrieval from satellite images based on a data-driven transfer learning model," *Remote Sensing of Environment*, vol. 272, p. 112940, 2022.
[57] Y. Liu, Q. Zheng, and X. Li, "Characteristics of global ocean abnormal mesoscale eddies derived from the fusion of sea surface height and temperature data by deep learning," *Geophysical Research Letters*, vol. 48, no. 17, p. e2021GL094772, 2021.
[58] E. J. Drenkard and K. B. Karnauskas, "Strengthening of the pacific equatorial undercurrent in the soda reanalysis: Mechanisms, ocean dynamics, and implications," *Journal of Climate*, vol. 27, no. 6, pp. 2405–2416, 2014.

Xinmin Zhang is pursuing the M.S. degree in computer application technology from the Qingdao University, Qingdao.His research interests include satellite remote sensing of the ocean, big data oceanography, and deep learning.

Baoxiang Huang (Member, IEEE) received the B.S. degree in traffic engineering from the Shandong University of Technology, China, in 2002, the M.S. degree in mechatronic engineering from Shandong University, China, in 2005, and the Ph.D. degree in computer engineering from Ocean University of China, in 2011, China. She was an academic visitor of Nottingham University. Currently, she is a Professor at the College of Computer Science and Technology, Qingdao University, China. Her research interests include remote sensing image processing and analysis, big data oceanography, and artificial intelligence.

Ge Chen received the B.S. degree in marine physics, the M.S. degree in satellite oceanography, and the Ph.D. degree in physical oceanography from the Ocean University of China (OUC), Qingdao, China, in 1988 and 1990, respectively. After graduation, he worked as a Post-Doctoral Fellow at The French Research Institute for the Exploitation of the Sea (IFREMER), Brest, France, from 1994 to 1996. Since 1997, he has been a Professor of satellite oceanography and meteorology with the OUC. In 2001, he received the National Science Fund for Outstanding Young Scientists awarded by the Natural Science Foundation of China and became the Chair Professor of the Cheung Kong Scholars Program nominated by the Chinese Ministry of Education. He is also the Chief Scientist for Ocean Science Satellite Missions at the National Laboratory for Ocean Science and Technology, Qingdao. He is the author of more than 180 peer-reviewed scientific articles published in internationally recognized journals. His research interests include satellite remote sensing of the ocean and big data oceanography.

Linyao Ge received the B.S. degree in computer application technology and the M.S. degree in computer technology from Qingdao University,Qingdao, China, in 2018 and 2021, respectively.He is pursuing the Ph.D. degree in computer application technology from the Ocean University of China, Qingdao.His research interests include satellite remote sensing of the ocean, big data oceanography, and deep learning.

Milena Radenkovic received the Dipl.−Ing. degree in electric and electronic engineering from the University of Nis, Nis, Serbia, in 1998, and the Ph.D. degree in computer science from the University of Nottingham, Nottingham, U.K., in 2002. She has authored more than 80 papers in premium conferences and journal venues. Her research interests include intelligent mobile and disconnection-tolerant networking, complex temporal graphs, self-organized security, distributed predictive analytics with applications to autonomous vehicles, mobile social networks, smart manufacturing, and predictive telemetry. Dr. Radenkovic was a recipient of multiple EPSRC and EU grants for her research. She has organized and chaired multiple ACM and IEEE conferences and served on many program committees. She is an Editor for premium journals such as the Ad Hoc Networks (Elsevier), the IEEE Transactions on Parallel and Distributed Computing, and ACM Multimedia.

Guojia Hou received his B.S. degree in computer science and technology from Linyi University in 2010, and Ph.D. degree in computer application technology from the Ocean University of China in 2015, respectively. He is currently an Associate Professor at the School of Computer Science and Technology, Qingdao University. His current research interests include underwater vision, image/video processing, and image quality assessment.

Global oceanic mesoscale eddies trajectories prediction with knowledge-fused neural network

Xinmin Zhang, Baoxiang Huang, *Member, IEEE,* Ge Chen, Linyao Ge, Milena Radenkovic, and Guojia Hou

***Abstract*—Efficient eddy trajectory prediction driven by multi-information fusion can facilitate the scientific research of oceanography, while the complicated dynamics mechanism makes this issue challenging. Benefiting from ocean observing technology, the eddy trajectory dataset can be qualified for data-intensive research paradigms. In this paper, the dynamics mechanism is used to inspire the design idea of the eddy trajectory prediction neural network (termed EddyTPNet) and is also transformed into prior knowledge to guide the learning process. This study is among the first to implement eddy trajectory prediction with physics informed neural network. First, an in-depth analysis of the kinematic characteristics indicates that the longitude and latitude of the trajectory should be decoupled; Second, the directional dispersion prior knowledge of global eddy propagation is embedded into the decoder of the EddyTPNet to improve the performance; Finally, EddyTPNet is implemented to predict the trajectories of global long-lived eddies. The extensive experimental results indicate that EddyTPNet can reliably forecast the eddy motion for the following 7 days and ensure a promising daily mean geodetic error of roughly 7.18 km. This exploratory study provides valuable insights into solving the prediction problem of ocean phenomena by using knowledge-based time series neural networks.**

***Index Terms*—Eddy trajectory prediction, Directional divergence physical information, Deep learning, Knowledge-fused neural network.**

I. Introduction

BIG data oceanography is a fertile domain for interdisciplinary research. Mesoscale eddies, the oceanic counterpart of atmospheric storms, are common and complex oceanic flow phenomena [1], [2], with the vast majority of them propagating westward at velocities comparable to those of long Rossby waves [3]. As the crucial link in the ocean's energy cascade, their movement propels the transmission of salt [4], [5], heat [6], carbon [7], [8], and other marine resources. Consequently, mesoscale eddy trajectories are of significant scientific interest [9], [10], climate change modeling [11], [12], and marine ecosystem improvement [13]. The patterns and dynamic mechanism of eddy motion are complicated [14], in this context, the efficient prediction of eddy has been a challenge [15].

Essentially, trajectory prediction is a complex time series prediction problem [16], [17]. In other words, the propagation trend of eddy can be predicted by using historical eddy motion data, but its uncertainty is higher than that of ordinary time series problems. The particularity of the detailed problem can be summarized in the following aspects. (1) Eddies vary in their rotation patterns, lifetimes, and survival intervals [18]. In addition, each form of eddies moves uniquely at various times and places, as depicted in Fig.1, increasing the difficulty of interpreting and forecasting movement; (2) Complex background fields, such as current, wind, and topography, strengthen the nonlinearity of mesoscale eddy and the complexity of the problem [19]; (3) Most mesoscale eddies propagate westward at a speed similar to that of Rossby waves, while only a few traveling eastward. The imbalance of data sets also makes track prediction more challenging.

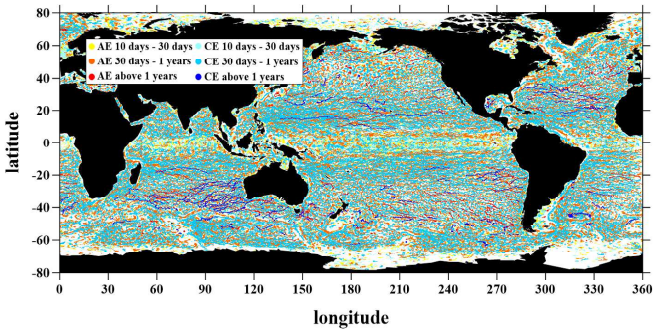

Fig. 1. The distribution of global mesoscale eddy motion trajectories for 2019-2020. The red and blue solid circles indicate the distribution of anticyclonic eddies and cyclonic eddies respectively. The color lightens as the life cycle decreases.

Theoretically, the assurance of ocean data products depends on the utilization of multiple observational data and synchronization techniques [20]–[22]. Furthermore, the prediction of oceanic phenomena has been enhanced through the use of numerical simulations [23]–[26], physical theories [27], and dynamic statistical analysis [28], [29], leading to substantial achievements and the enrichment of existing knowledge. Additionally, deep learning methods [30]–[32] have been successfully implemented in specific domains to improve effectiveness by integrating prior knowledge [33], [34].

This research was financially supported by the National Natural Science Foundation of China (No.42276203,42030406), the Laoshan Laboratory (No.LSKJ202204302), and the Natural Science Foundation of Shandong Province (No.ZR2021MD001). (Corresponding authors: Baoxiang Huang)

Xinmin zhang is with Qingdao University, Qingdao, 266071, China (e-mail: 2021023803@qdu.edu.cn).

Linyao Ge is with the Frontiers Science Center for Deep Ocean Multispheres and Earth System, School of Marine Technology, Ocean University of China, Qingdao, 266100, China (e-mail: linyaoge@stu.ouc.edu.cn).

Baoxiang Huang is with the College of Computer Science and Technology, Qingdao University, Qingdao 266071, China, and also with the Laboratory for Regional Oceanography and Numerical Modeling, Laoshan Laboratory, Qingdao, 266100, China. (e-mail: hbx3726@163.com).

Ge Chen is with the Frontiers Science Center for Deep Ocean Multispheres and Earth System, School of Marine Technology, Ocean University of China, Qingdao 266100, China, and also with the Laboratory for Regional Oceanography and Numerical Modeling, Laoshan Laboratory, Qingdao, 266100, China. (e-mail: gechen@ouc.edu.cn)

Benefiting from the intensive trajectory data of oceanic eddies, various learning architectures have been applied to the Earth observation [35], [36] with the rise of artificial intelligence [37], [38]. As for the prediction of eddy trajectories, multiple linear regression techniques were employed to create straightforward statistical prediction models [39] and correlate them with changes in eddy propagation position and other ocean parameters [40]. The method for predicting eddy properties and propagation trajectories using long short-term memory (LSTM) [41] and extra tree algorithms were proposed [27]. By combining data from different sources, the gate recurrent units (GRU)-based deep learning framework was applied to the prediction of eddy trajectories [42]. In addition, the new loss function called weighted mean square estimation was also proposed to improve model performance. The previous studies in artificial intelligence oceanography have demonstrated that deep learning methods yielded impressive performance [32], [43]. These studies focus mainly on data correlation analyses and do not consider the influence of time series on the prediction of eddy trajectories, nor the limitation of the prediction region and the lack of consideration of the influence of the external physical environment. Nevertheless, these results demonstrate the interpretability of the mesoscale eddy trajectory prediction problem and provide important insights for our research.

The objective of this paper is to implement oceanic mesoscale eddies trajectory prediction with the knowledge-fused neural network. Specifically, the proposed approach is to use historical data spanning 10 days to predict the trajectory of the next 7 days. The main contributions can be briefly summarized as follows.

1) The angular momentum propagation mechanism of eddies is embedded as the physical constraint into the proposed EddyTPNet to capture the dynamic characteristics, thereby improving prediction accuracy.
2) EddyTPNet is first trained using global data to predict the propagation process of long-lived eddies on a global scale, and then is transferred to complicated local regions with fine-tuning strategies.
3) Extensive experiments have been conducted to evaluate the proposed method for eddy trajectory prediction. The results demonstrate that the method can achieve promising performance both globally and locally.

The remainder of the paper is organized as follows. Section II formulates the problem definition and explains the idea of dynamics mechanism factorization. The proposed methodology is illustrated in section III, including trajectory data preprocessing, the architecture of the Knowledge-fused deep neural network, and the loss function. Section IV indicates the comprehensive experiments in detail. Finally, section V concludes and discusses the present work.

II. Preliminaries

A. Problem definition

As previously stated, the task of this study is to predict the trajectory of the next 7 days using 10 days of historical data. In this exploratory study, eddies with lifetimes of more than one year were chosen as the main research objects for the consideration of observation reliability, generalization, and the limitation of Graphic Processing Unit (GPU) memory. Specifically, the main reasons are as follows: (1) short-lived eddies are unstable, and the evolution process of short-lived eddies is easily affected by the external environment. Considering the merging and splitting of eddies, short-lived eddies may survive in another form. (2) There are certain errors in the observation of short-lived eddies. Usually, the scales of short-lived eddies are small, and there are some errors in satellite observation and tracking algorithms. (3) The related studies have demonstrated that the behaviors of short-lived and long-lived eddies are very different in terms of population, migration, and evolution [18]. (4) So using all eddies trajectories simultaneously to train a deep neural network will increase the computational cost but result in poor generalization.

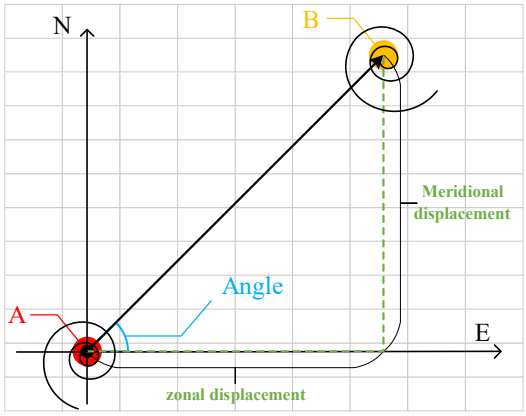

Fig. 2. Decomposition of mesoscale eddy motion trajectories. The red solid point A is the position of the eddy's starting point and the yellow solid point B is the position of the eddy's ending point. The angle formed by the Cartesian coordinate system with the east direction is the instantaneous eddy movement angle. The vertical and lateral movement distances are the meridional displacement and zonal displacement respectively.

EddyTPNet takes 10 days of historical trajectory $(X^n_{t-9}, X^n_{t-8}, ..., X^n_t)$ as input to predicts the next 7 days locations $(Y^n_{t+1}, Y^n_{t+2}, ..., Y^n_{t+7})$ of the eddy propagation, as described in following Eq.1.

$$(Y^n_{t+1}, Y^n_{t+2}, ..., Y^n_{t+7}) = EddyTPNet(X^n_{t-9}, X^n_{t-8}, ..., X^n_t) \quad (1)$$

Where n denotes the nth eddy and t denotes the tth time within that trajectory. The relevant data processing and the specific training test procedure within EddyTPNet will be described specifically in Section III.

B. Dynamics mechanism factorization towards network design

Through an in-depth analysis of the dynamics of eddy propagation, it is noted that the planetary gradient of the Coriolis parameter creates a mass imbalance in the eddy, causing it to move westward. Furthermore, in the radial motion, the eddy is subjected to the baroclinic shear of the flow field and the eddy diverges towards the poles and the equator, respectively [44]. The knowledge is factorized into two parts, one for the

longitude and latitude decoupling of the trajectory, the other is divergence priority.

The main idea of trajectory latitude and longitude decoupling is to predict zonal and meridional displacements rather than positions. The zonal and meridional displacements can determine the position of the eddy motion at the next step, as shown in Fig .2. Meanwhile, divergence knowledge plays an important role in weighing the dominance of the radial and latitudinal directions of the eddy. In addition, the directional dispersion can effectively guide the model in making more accurate predictions of propagation when the external factors are unknown.

III. Methodology

This section presents the overall pipeline for implementing eddy trajectories prediction. Specifically, section III-B introduces the materials and data processing method in detail. The structure of the EddyTPNet, the training process, and the testing process are detailed in section III-C. And the loss function is formulated in section III-D.

A. Overall pipeline

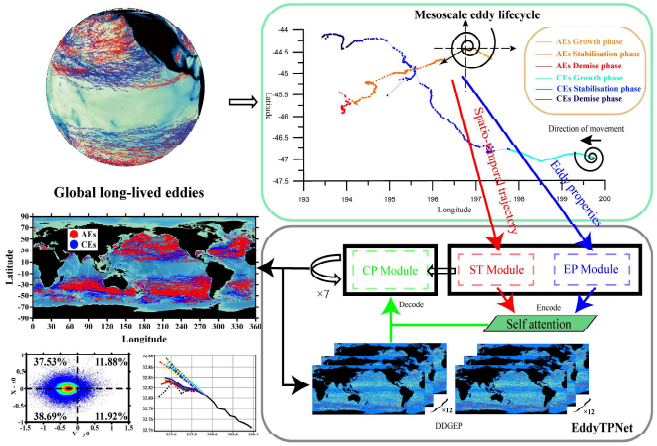

Fig. 3. Pipeline of the proposed methodology for global oceanic mesoscale eddies trajectories prediction.

The pipeline of the entire study is shown in Fig.3. First, based on the prior knowledge of the eddy dynamic mechanism and the divergence of eddy propagation direction, the directional divergence of global eddy propagation (DDGEP) grid data is constructed to simulate the physical background field and eddy propagation direction, which is used to guide and match the deep learning model [45]. Second, following the encoder-decoder structure, EddyTPNet is proposed for the mesoscale prediction of eddy propagation, which integrates LSTM cells [46] and GRU units [47] for the dynamic interaction of information. The decoder cyclic prediction (CP) module uses a combination of spatio-temporal trajectory data and directional divergence of global eddy propagation (DDGEP) data as input. The encoder consists of two modules: one focused on eddy current characteristics (EP) and the other focused on spatio-temporal trajectory characteristics (ST). A self-attention mechanism [48] is added between the encoder and the decoder to extract important information from the time series. Finally, the loss function of the optimized network is the mean absolute geodetic error (MAGE) loss, which combines earth distance and $L1$ regularized loss.

B. Data preprocessing

1) Materials and processing: The materials used in this study are obtained from the Satellite altimetric Mesoscale Eddy Trajectories Atlas, which cover the years from 1993 to 2022 and include amplitude, radius, speed-average, latitude, longitude, and time data. The data consists of Absolute Dynamic Topography (ADT) maps as input, which are first filtered by Lanczos, and then the eddy detection algorithm is used to outline the eddy closure contours. The center of the eddy is defined as the center of the circle which fits best with the contour of the maximum speed. Finally, the eddies are correlated by a tracking algorithm. This dataset is stored in the form of sequence points, and eddy information is recorded for each spatio-temporal state. Based on this dataset, the eddy characteristic dataset and spatio-temporal trajectory dataset are created.

Eddy properties dataset. This dataset comprises eddy properties because faster-moving and more energetic eddies typically have greater radius, amplitudes, and velocities [49], and the historical variance in these properties mostly indicates the stability of the eddy trajectory.

Spatio-temporal trajectory dataset. The eddy characteristics of the time, longitude, and latitude are crucial to estimating the eddy movement patterns. The input of longitude, latitude, and time can determine the precise place and time of eddy, as eddy motion is affected by geographical location, time, topography, and seasonality [50]. The motion of the eddies can be represented by the values of motion angle and displacement, as shown in Fig.2, so the meridional displacement, zonal displacement, and direction of motion are used to extend the spatial and temporal information. The detailed process can be formulated as Eq.2-4.

$$Z_t = lon_t - lon_{t-1}, \tag{2}$$

$$M_t = lat_t - lat_{t-1}, \tag{3}$$

$$A_t = \angle(\arctan\frac{|Z_t|}{|M_t|}). \tag{4}$$

Where t, lon, and lat are $t-step$, longitude, and latitude, respectively. The zonal displacement, meridional displacement, and motion direction are represented by Z, M, and A, respectively. Eq.2 and Eq.3 are the differences between the zonal and meridional displacements respectively, and instead of using longitude and latitude, the prediction of the meridional and zonal displacements is executed. The reason is that the input to the neural network is normalized data, and the amount of variation in longitude and latitude is negligible after normalization. Here $arctan$ function is used as the solution in Eq.4, and the relationship between meridional and zonal displacements can be more clearly reflected. The meridional and zonal displacements are utilized to determine the final

angle, which is then calculated using the Cartesian coordinate system as a base and used to match the newly generated DDGEP data. It is worth noting that the eddy polarity can always be guaranteed throughout the encoding and decoding process by incorporating polarity into the dataset along with state flag bits 0 and 1, respectively, as inputs to the neural network.

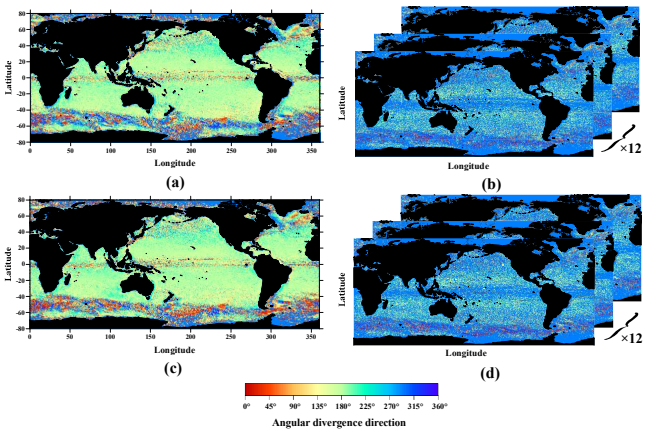

Fig. 4. Global cyclone and anticyclone eddy directional dispersion grid plots. (a) and (c) represent the statistical grid maps of the global 1993-2022 anticyclonic and cyclonic eddy directional dispersion grid maps, respectively. (b) and (d) record the anticyclonic and cyclonic eddy directional dispersion grid maps for January-December, respectively.

2) Directional divergence grid data: The idea of directional divergence comes from the physical mechanism of divergence in the direction of propagation of mesoscale eddy trajectories. First, the motion from one trajectory point to another can be determined by the amount of displacement and the angle, where the displacement in different directions is represented by the meridional displacement and the zonal displacement. The relationship between the meridional and the zonal motions is not isolated, they interact with each other, thus the usage of angles to model the dynamic connection between the relationship can guide the direction of eddy propagation. Second, mesoscale eddy trajectories are significantly influenced by wind, currents, and topography, leading to their irregular motion. Typically, the introduction of information from external background fields is necessary to enhance the accuracy of eddy trajectory predictions. Systematic analysis of the physical and data-driven dispersion of eddy propagation directions suggests that both physical and data-driven directional dispersion appropriately capture the influence of external forces [51]. Moreover, historical eddies tend to traverse regions with similar propagation directions, exhibiting minimal variability in external factors. Therefore, in the absence of incorporating multiple sources of data, such as external background fields, the creation of DDGEP data serves to reduce data usage while maintaining a high level of accuracy. It is known that cyclonic and anticyclonic eddies have different motion mechanisms and a strong correlation with seasons. Here DDGEP data with the size of 2×12 was synthesized based on the eddy polarity and different months. Each DDGEP data is calculated from eddy data of the same month and polarity. Specifically, the directional dispersion data within each grid is calculated by the following Eq.5-8.

$$\overline{lon}_t^i = \frac{(lon_{t+1} + lon_t) - (lon_t + lon_{t-1})}{2}, \tag{5}$$

$$\overline{lat}_t^i = \frac{(lat_{t+1} + lat_t) - (lat_t + lat_{t-1})}{2}, \tag{6}$$

$$A_k^i = \angle \left(\arctan \frac{\left|\overline{lat}_t^i\right|}{\left|\overline{lon}_t^i\right|} \right), \tag{7}$$

$$DDGEP_k = \frac{1}{N} \sum_{N=1}^{C} A_k^N. \tag{8}$$

Where A_k^i denotes the ith angular value within the kth grid, C denotes the number of angular values within each grid, and $DDGEP_k$ means the calculated angular value within the kth grid. The data is a matrix of size 2880×1440 with $1/8^\circ \times 1/8^\circ$ spatial resolution and the calculation process is slightly different from that of the angles, with the amount of data involved in each grid being determined by the number of eddies passing through. Specifically, Eq.5 and Eq.6 are averaged over the two consecutive days of each eddy's position to avoid transient noise. Eq.7 indicates that the angular value of each eddy is calculated. The grid region in which that eddy is located is determined, and then Eq.8 is calculated by averaging the angular values already within each grid so that the divergence angle within that grid matches the direction of propagation of most eddies, thus avoiding the effect of trajectory anomalous eddies. Due to the small amount of variation in latitude and longitude and the constant querying of data during the test, we eventually determined a grid accuracy of $1/8^\circ$ for the angular data, while ensuring accuracy and efficiency. We superimpose each DDGEP in Fig.4, it can reveal that most of the grid divergence tends to the west and a small portion to the east, which is consistent with the theory related to eddy propagation. The final directional dispersion will provide some guidance in the prediction process for the proportional distribution of the meridional and zonal displacements of the eddies and the prediction of the direction of the eddy propagation.

C. Knowledge-fused deep neural network

The overall architecture of the proposed EddyTPNet is presented in Fig.5. EddyTPNet is composed of three primary parts, namely the EP module, which focuses on eddy properties, the ST module, which can extract spatio-temporal trajectory information, and the CP module, which possesses mesoscale prediction capability.

1) Eddy properties module: LSTM has been utilized for capturing temporal information, and empirical findings suggest that the characteristics of eddies do exert a certain level of influence on their trajectory based on the analysis of variable importance. However, the magnitude of this influence is relatively minor [27]. The stable state of the eddy is constrained by the attributes of the eddy, and inputting it together with time-space information such as longitude and latitude may lose the

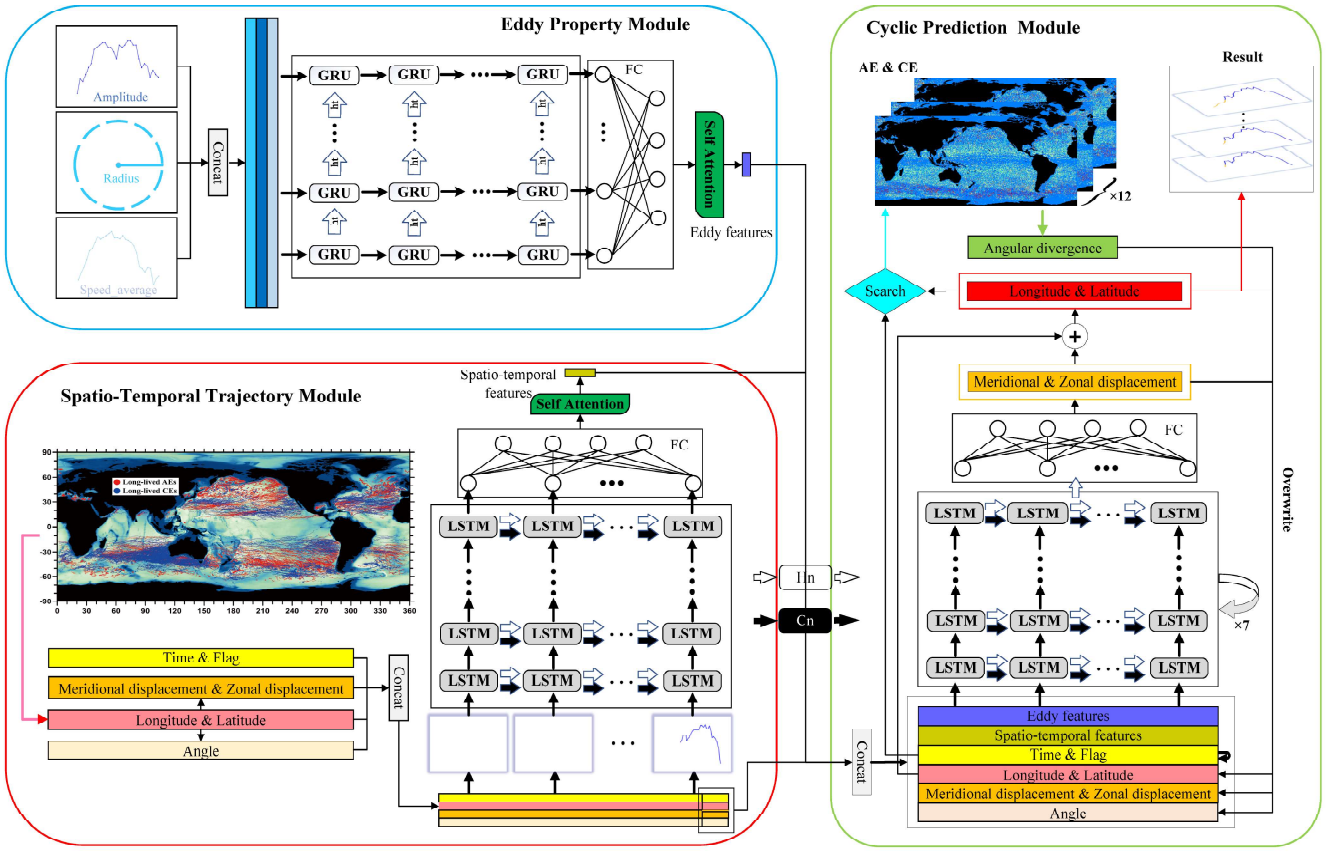

Fig. 5. The overall architecture of Knowledge-fused deep neural network for trajectory prediction of the global mesoscale eddy. The ST module, EP module, and CP module are each represented by a red, blue, and green box. Self-attention connects the modules while GRU extracts the eddy properties and LSTM derives the spatio-temporal information. DDGEP data embeds the CP module for continuous indexing.

state constraint information. The GRU unit, which also has an excellent performance in capturing timing information, has fewer parameters and is faster to train. Therefore, the GRU unit is employed in the EP module to capture the eddy energy changes of the historical sequence to constrain the state trend of the eddy in the future. Additionally, the future state of eddies is unknown, so it needs to be separated. The internal structure is shown in Fig.6(a), where the input is e_t, containing the radius, amplitude, and speed-average at step t, and the hidden state at step $t-1$ is represented by h_{t-1}. The processing of the GRU unit can be expressed as Eq.9-11.

$$Z_t = \sigma(W_z \cdot [e_t, h_{t-1}] + b_z), \tag{9}$$

$$R_t = \sigma\left(W_r \cdot [e_t, h_{t-1}] + b_r\right), \tag{10}$$

$$h_t = (1 - Z_t) * h_{t-1} + Z_t * \tan h\left(W \cdot [R_t * h_{t-1}, e_t] + b\right). \tag{11}$$

The reset gate R and the update gate Z determine the degree of retention of historical and current attributes. W and b denote the weights and biases of the corresponding gates respectively. σ is the sigmoid activation function and $*$ is the Hadamard product. h_t is used as an output to represent the current state and provide the basis for the predictions of the CP module. However, the spatio-temporal information is more important in influencing the eddy trajectory.

2) Spatio-temporal information module: The ST module uses LSTM units to memorize and encode historical spatio-temporal trajectory features. Time and eddy polarity are embedded to capture seasonal information for different eddies. We use the spatio-temporal trajectory information at time $t-1$ as the input to the ST module at time t, denoted by s_t. The output of the module can be formulated as Eq.12-16.

$$F_t = \sigma\left(W_f \cdot [h_{t-1}, s_t] + b_f\right), \tag{12}$$

$$I_t = \sigma\left(W_i \cdot [h_{t-1}, s_t] + b_i\right), \tag{13}$$

$$C_t = I_t * \tan h\left(W \cdot [h_{t-1}, s_t] + b_c\right) + F_t * C_{t-1}, \tag{14}$$

$$O_t = \sigma\left(W_o \cdot [h_{t-1}, s_t] + b_o\right), \tag{15}$$

$$h_t = O_t * \tan h\left(C_t\right). \tag{16}$$

LSTM generally performs better than GRU on the finer tasks such as latitude and longitude prediction because of the large number of parameters. Compared to GRU units, the LSTM unit is divided into three gates: input gate I_t, which determines the amount of input information at time t, forgetting gate F_t, which determines the amount of information forgotten at time $t-1$, and output gate O_t, which determines the amount of output, as shown in Fig.6(b). A separate memory unit C_t is used to save the current LSTM state information and pass it to the LSTM the next time.

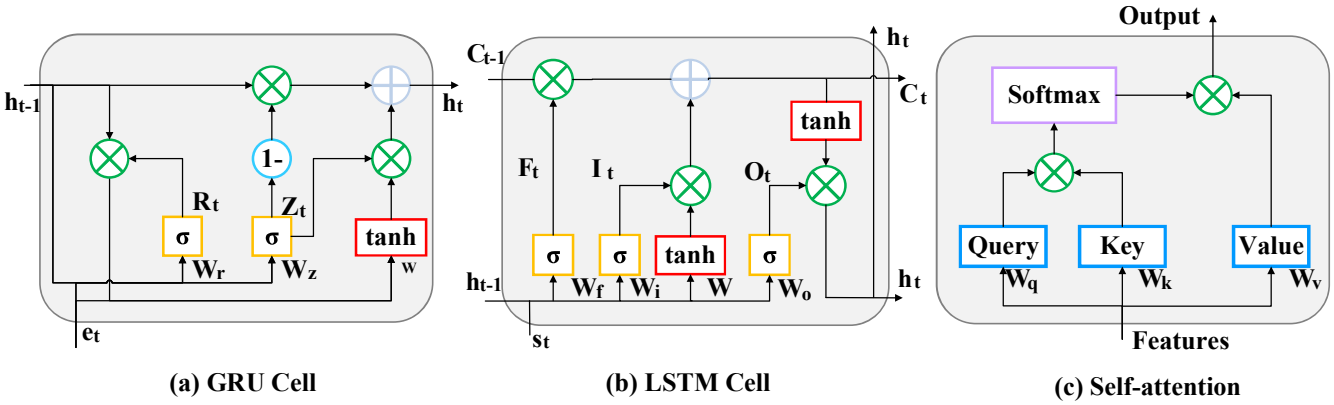

Fig. 6. (a) GRU unit structure diagram. (b) LSTM unit structure diagram. (c) Self-attention structure diagram.

3) Cyclic prediction module: The encoder-decoder with the self-attention mechanism often achieves better results. Attention as shown in Fig.6(c) is used as part of the input to the CP module, extracting key information from the last-moment hidden states of the EP and ST modules respectively. The calculation of the self-attention mechanism is as the following Eq.17.

$$Attention = W_v x * softmax\left(W_q x * [W_k x]^T\right) \quad (17)$$

Where x represents the output of the Eddy Properties Module and the Spatio-temporal Information Module, W_q, W_k and W_v correspond to weight vectors associated with value, query, and key, respectively. These weight vectors are used to calculate the weighted sum of the value vectors based on attention weights, thereby assigning varying importance to different positions and enhancing the model's capability to model long-range dependencies.

The CP module acts as a decoder with a built-in unit of LSTM and the initial states h and c are the last layers of hidden states of the ST module. The inputs are longitude, latitude, meridional displacement, zonal displacement, angle, time, and the self-attention outputs of the other two models. After preliminary experiments, it is not ideal to directly predict the eddy trajectory for the next 7 days, so we split the multi-step prediction problem [52] into multiple single-step prediction problems. The decoder is designed to predict a 1-day trajectory using 5 days of historical data, and the implicit state of this time is used as the initial state for the next prediction, and the prediction process is carried out iteratively through a sliding window technique, resulting in a total of 7 consecutive predictions. In the prediction process that follows, where the input data remains unknown, the 5-day data undergoes a continuous update utilizing predicted values, alongside calculations and lookup operations. This iterative approach ultimately ensures data integrity and fosters accurate prediction outcomes. After introducing the divergence of physical information, the physical embedding process of the training procedure is derived from the given angular scattering features. On the contrary, the testing procedure is comparatively complex. The specific algorithmic flow of the CP module, taking the testing process as an example, is shown in Algorithm. 1.

Algorithm 1 CP Module with the Directional Divergence Algorithm

Input:
Trajectory characteristics:
$(T_n)_{n=0}^{12} = (lat_n, lon_n, z_n, m_n, t_n, a_n, f_n)_{n=0}^{12}$;
Encoder information: $(s_n, x_n)_{n=0}^{4}$;
Encoder status: h_n, c_n;
Output:
for $t = 0; t < 7; t++$ **do**
//5 days of data assignment;
$Input = (T_n)_{n=j}^{j+4}, (s_n, x_n)_{n=0}^{4}$;
//Forecast;
$Output, h_n, c_n = DecoderNet(Input, h_n, c_n)$;
//Storing predicted values;
$(z_n, m_n)_{n=j+5}^{j+6} = Output$;
//Inverse normalization;
$z, m = InvNorm((z_n, m_n)_{n=j+5}^{j+6})$;
$lat, lon, f, t = InvNorm((lat_n, lon_n, f_n, t_n)_{n=j+4}^{j+5})$;
//Summation;
$lat, lon = z, m + lat, lon$;
//Normalized for next prediction;
$(lat_n, lon_n)_{n=j+5}^{j+6} = Norm(lat, lon)$;
//Search DDGEP data;
$a = Search(lat, lon, f, t)$;
//Normalized for next prediction;
$(a_n)_{n=j+5}^{j+6} = Norm(a)$;
end for
return $(lat_n, lon_n)_{n=5}^{12}$

u, v, z, m, t, f, and a represent latitude, longitude, meridional displacement, zonal displacement, time, angle, and flag respectively. In this process, the predicted meridional and zonal displacements are added to the previous day's longitude and latitude to generate the new predicted longitude and latitude, and then the eddy polarity, month, and predicted longitude and latitude are used to find the corresponding value in the DDGEP data for the next prediction input, which is used

to guide the model prediction.

D. Mean absolute geodetic error loss

The accuracy of the model prediction depends heavily on the loss function, and since the geodetic distance [40] enables a practical test of the accuracy of the eddy trajectory prediction, the MAGE loss function [53] adopted in this paper can be expressed as Eq.18.

$$MAGE = L1Loss + MGDLoss \tag{18}$$

MAGE consists of two components, the $L1$ and mean geodetic distance (MGD) losses, respectively. Specifically, the $L1$ loss as Eq.19 is used to reduce the Euclidean distance between the predicted and true values, and the MGD loss as Eq.20 is adopted to model the physical environment in which the predicted and true values are located.

$$L1 = \frac{1}{n}\sum_{i=0}^{n}(\frac{1}{m}\sum_{t=1}^{m}|x_i^t - y_i^t|) \tag{19}$$

$$MGDLoss = \frac{1}{n}\sum_{i=0}^{n}(\frac{1}{m}\sum_{t=1}^{m}\Delta\alpha_i^t) * R \tag{20}$$

Where m denotes the mth day, n is the amount of data in a batch, and x_i^t,y_i^t, and $\Delta\alpha_i^t$ indicate the predicted and true values and the central angle between them at time t, and the ith time, respectively. This central angle can be specified as Eq.21 and Eq.22, which are derivable and continuous in the domain of definition, thus the whole function is differentiable.

$$\Delta\alpha = \arctan x \tag{21}$$

$$x = \frac{\sqrt{\left(\cos\varphi_2 \sin\Delta\gamma\right)^2 + \cos\varphi_2 \sin\varphi_1 - \sin\varphi_1 \cos\varphi_2 \cos\Delta\gamma}}{\sin\varphi_1 \sin\varphi_2 + \cos\varphi_1 \cos\varphi_2 \cos\Delta\gamma} \tag{22}$$

Where γ and φ denote longitude and latitude, respectively. $\Delta\gamma$ denotes the difference in longitude between the predicted and true values.

IV. Experimental results and analysis

A. Implementation details

1) Datasets partitioning: The datasets employed in this study consist of the eddy properties dataset and the spatio-temporal trajectory dataset. To explore the influence of historical sequence length on prediction, each entry in the eddy properties dataset includes 30 days of historical feature data. In contrast, the spatio-temporal trajectory dataset comprises 30 days of historical trajectory data and 7 days of future actual data. A total of 400,648 data sequences were randomly divided (with a random seed set to 42), with 280,453 sequences allocated for training, 60,098 sequences for validation, and 60,097 sequences for testing. The data was subsequently normalized and fed into the model for training.

TABLE I
THE PREDICTION RESULTS FOR EDDYTPNET VARIANTS.

Method	MGD(km)	SGD(km)	Precision(%)
EddyTPNet-CP	7.30	51.08	84.58
EddyTPNet-ET	7.28	50.96	84.66
EddyTPNet-ST	7.20	50.37	84.90
EddyTPNet-10-7	7.18	50.28	84.99
EddyTPNet-20-7	7.19	50.36	84.91
EddyTPNet-30-7	**7.17**	**50.16**	**85.10**

2) Experimental setup: The proposed EddyTPNet was implemented on the PyTorch deep learning library. The experiments were performed under Windows 11 (Nvidia GeForce RTX 3090GPU). The hidden states of the EP module, ST module, and CP module were set to 64, 128, and 128, respectively, and the network layers were set to 4 layers. Due to the abundance of data, the direction of model descent is more accurate when more training data is obtained in each batch, and the processing speed can be accelerated, we finally set the batch size to 2048. The initial value of the learning rate was set to 0.0005 to ensure accurate convergence of the network and was adjusted using the ReduceLRPlateau strategy, with the patience value set to 5. The optimizer was Adam, and the number of training rounds was 50.

The final model predicts the trajectory for the next 7 days by learning 10 days of historical data. We will demonstrate the feasibility of the proposed method by comparing with different models in terms of the length of the history sequence, the presence or absence of physical information, and the ability to generalize the local region.

3) Baselines: To evaluate the performance of the proposed method, Extra Trees [27], Multiple Linear Regression [40], Random Forest [27], Gradient Boosting [27], MesoGRU [42], MesoLSTM [42], Seq2Seq [54], EddyTPNet-ST, EddyTPNet-EP, and EddyTPNet-CP models were selected as the state-of-the-art baselines.

Extra Trees, Random Forest, Gradient Boosting and Multiple Linear Regression. Traditional machine learning models. One model was built for each prediction day, for a total of seven models.

MesoLSTM and MesoGRU. The MesoGRU model is a variant of the GRU model, consisting of 4 layers of GRU. It utilized historical data from the previous three days to predict future data for the next day. To achieve multi-step prediction, this model employed a sliding window approach for training and testing. For this experiment, the hidden state was set to 128. Additionally, the experimental setup for the MesoLSTM model was the same as the MesoGRU model, except for the substitution of GRU units with LSTM units.

Seq2Seq. The Seq2Seq model employs an encoder-decoder architecture composed of 4 recurrent neural networks (RNNs). The encoder encoded historical data of the past 10 days of eddies, while the decoder performed daily predictions using a sliding window approach, making a total of 7 predictions. The network was configured with 128 hidden states.

EddyTPNet-CP. EddyTPNet-CP, as a variant of EddyTPNet, removed the encoding part of the Eddy attribute module and the spatio-temporal module. However, as the decoder of

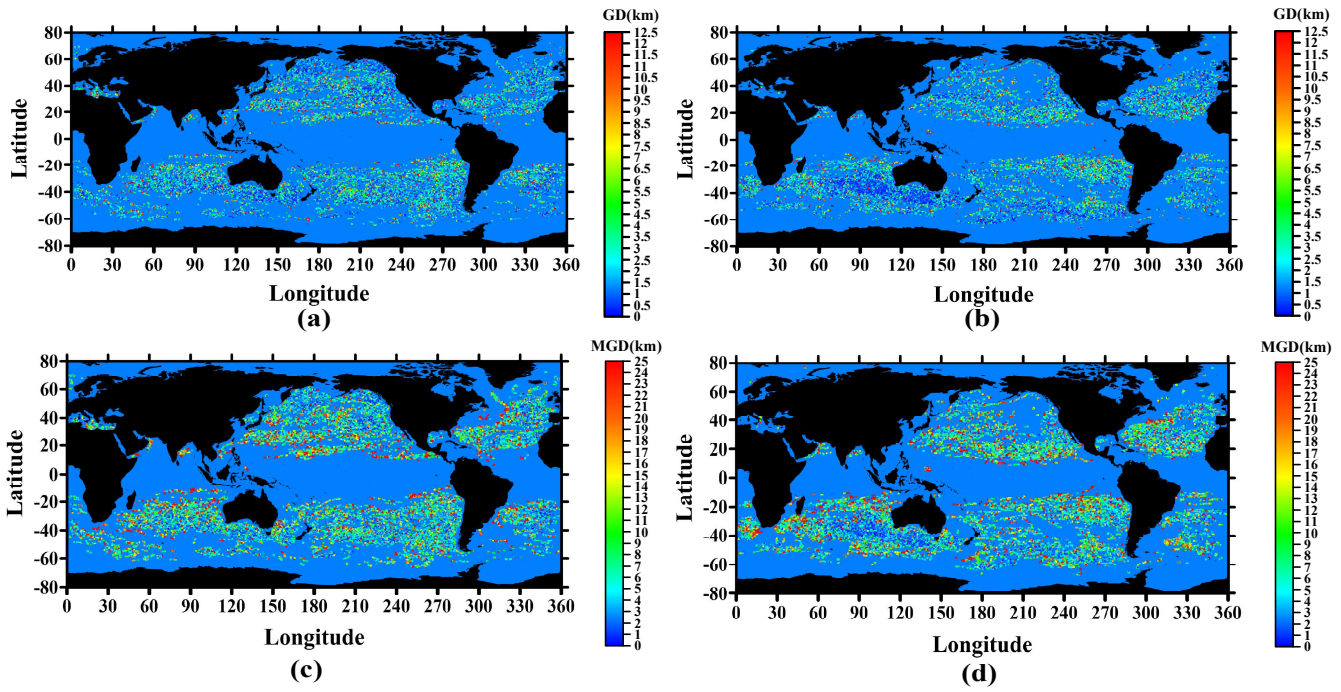

Fig. 7. Global geodetic distance distribution of test sets. (a) and (b) represent the Geodetic Distance (GD) distribution of the next day prediction results for Anticyclonic eddies and Cyclonic eddies respectively. (c) and (d) denote the Mean Geodetic Distance (MGD) distribution of the 7-day forecast results for Anticyclonic eddies and Cyclonic eddies respectively.

TABLE II
PREDICTION ERRORS OF MODELS UNDER DIRECTIONAL DIVERGENCE PHYSICAL INFORMATION.

Method	MGD(km)		SGD(km)		Precision(%)	
	Angular	Without angular	Angular	Without angular	Angular	Without angular
MesoGRU	8.02	8.87	56.12	62.05	82.20	79.29
MesoLSTM	8.02	8.86	56.13	62.00	82.20	79.37
Seq2Seq	7.44	7.67	52.04	53.70	84.19	83.80
EddyTPNet-CP	7.30	7.77	51.08	54.37	84.58	83.47
EddyTPNet-ET	7.28	7.75	50.96	54.22	84.66	83.54
EddyTPNet-ST	7.20	7.67	50.37	53.60	84.90	83.84
EddyTPNet	**7.18**	7.63	**50.28**	53.38	**84.99**	83.97

EddyTPNet, it still possesses the ability of multi-step prediction. The training and prediction processes remain the same as the CP module while maintaining the same hyperparameters as EddyTPNet.

EddyTPNet-ST. The model conformed to the basic architecture of the Seq2Seq and was based on EddyTPNet with the EP module removed and other parameters set in the same way as EddyTPNet.

EddyTPNet-EP. This model was a model where only the EP module and the CP module are retained, the encoder-decoder structure is split, and the parameter settings remain unchanged. The eddy properties were used as input to the EP module, with self-attention as the connection between the EP module and the CP module.

4) Evaluation metrics: Applying the above model, the evaluation indicators used for the assessment of the results were accuracy, geographical MGD, and summed geodetic distance (SGD) respectively.

MGD and SGD. MGD and SGD are the mean geodetic distance and the sum of 7-day geodetic distances, respectively. They can reflect the error between the predicted and true values by the true Earth distance.

Precision. As grid data with 12.5km or 25km resolution can be applied to real scenarios, we define trajectories with MGD errors below 12.5km as prediction accuracy.

B. Comparison and analysis of prediction results

To facilitate marine exploration tasks, it is customary to deploy equipment in advance, hence the establishment of a 7-day prediction window. The decoding module utilizes a recurrent encoder to generate the next predicted value, ensuring that long-term predictions do not compromise short-term accuracy. The trajectory is only predicted for the following day, and the distribution of geographic distances (GD) is depicted in Fig. 7(a-b), showcasing highly accurate prediction results. Furthermore, Fig. 7(c-d) presents the distribution of the average geographic distance (MGD) within the 7-day

prediction window, highlighting the strong predictive efficacy of the proposed EddyTPNet.

Here, we label the error as MGD $>$ 12.5km, and the poor forecast area mainly exists near the shore, the equator, and poles, as shown in Fig.8. Nevertheless, it can be seen that the error distributions are relatively uniform and there are no problems with regional failures. To further demonstrate the effectiveness and complexity of the method, several comparative experiments were conducted.

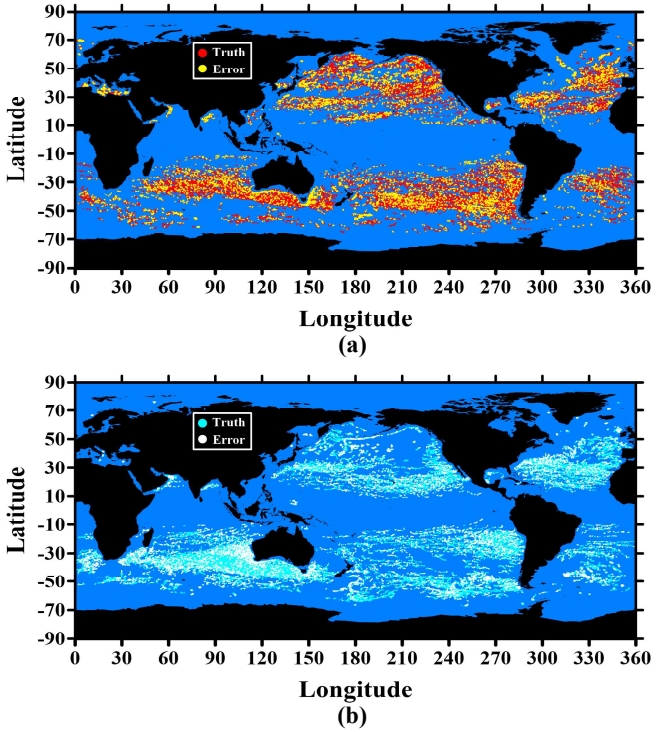

Fig. 8. Global distribution of test sets and errors. (a) denotes Anticyclonic eddies and (b) denotes Cyclonic eddies. Among them, the trajectory of the Mean Geodetic Distance (MGD) $>$ 12.5km is marked as an error.

1) Global Comparative Experiment: The EddyTPNet variants are first compared. On the one hand, EddyTPNet-10-7, EddyTPNet-20-7, and EddyTPNet-30-7 are used to represent the prediction ability of EddyTPNet for different lengths of historical sequences, which provide the encoder with 10, 20, and 30 days of historical spatio-temporal and eddy state information, respectively. After introducing the directional divergence physics information, Table I shows that 30 days has the best prediction effect on 7 days, but in general, different lengths of historical sequences do not have much effect on EddyTPNet. The reason is that the orientation information embedded by the neural network largely replaces the encoding process. Therefore, the subsequent experiments are still based on 10 days predicting 7 days. On the other hand, EddyTPNet-CP, EddyTPNet-ST, and EddyTPNet-EP are used to demonstrate the importance of each module. Table I also indicates that each module is effective in improving the network accuracy when directional dispersion is included, and the ST module is more important than the ET module, as illustrated in Table II.

The experimentation now shifts towards comparing the MesoGRU, MesoLSTM, and Seq2Seq models. Combined with Fig.9 and Table II, it can be found that MesoLSTM and MesoGRU are less expressive and not sensitive enough to the motion distance of eddy currents. In contrast, the prediction accuracy of the Seq2Seq model, the basic architecture of EddyTPNet, is much higher than that of LSTM and GRU, indicating that the encoder part can remember and extract the historical propagation distance of eddy motion.

The most expressive EddyTPNet has good prediction performance in predicting both motion distance and motion trajectory direction. The physical information of the direction dispersion is crucial for the accuracy of the eddy trajectory prediction. Table II presents the statistical results for different models with and without the introduction of directionally dispersive physical information. The results imply that the introduction of directionally dispersive physical information can improve the effectiveness of the deep neural network, and the angular dispersion information shown in Fig.9 can greatly fit the dispersion ratios of the true values, which solves the problem of the imbalance between the westward and eastward eddy data.

Further, some examples of single trajectories are plotted in Fig.10. A closer look reveals that the prediction results without the introduction of the directional dispersion physical information are more inclined to smooth and less varied trajectories. whereas, with the introduction of the directional dispersion physical information, the network can accurately capture the instantaneous variations of the eddies. The results indicate that EddyTPNet has the highest prediction accuracy and the smallest prediction error of 84.99%, while MGD and SGD are 7.18 km and 50.28 km, respectively.

The daily prediction errors and accuracies of each model were calculated separately to fully demonstrate the advantages of each model in the time series task, as shown in Table III.

The prediction accuracy of each model decreases over time due to the accumulation of errors. Nevertheless, the results show that EddyTPNet still outperforms the other models in different periods, and still achieves a trajectory prediction rate of 71.26% on the seventh day. Furthermore, in contrast to other models, EddyTPNet exhibits a consistent increase in daily prediction errors, with a difference of approximately 1.4 km. This observation suggests that the model demonstrates greater stability in forecasting long sequences.

2) Local Regional Prediction Results: To examine the generalizability of the proposed method, we opted to focus our research on the South China Sea and the North Atlantic region [55]. The South China Sea stands out due to its distinctive geographical location and intricate topography, which has led to the development of prediction methods specifically tailored to this area's eddy trajectories. Moreover, the North Atlantic also serves as a significant source region for eddies of varying lifespan scales. For our study, we gathered data from the South China Sea region ($5^\circ - 25^\circ N, 105^\circ - 125^\circ E$) and the North Pacific region ($20^\circ - 40^\circ N, 20^\circ - 40^\circ W$) spanning a period of 30 years (1993-2022). This data was divided into 31,050 and 80,751 trajectory data, respectively. Subsequently, these datasets were partitioned into training, validation, and testing

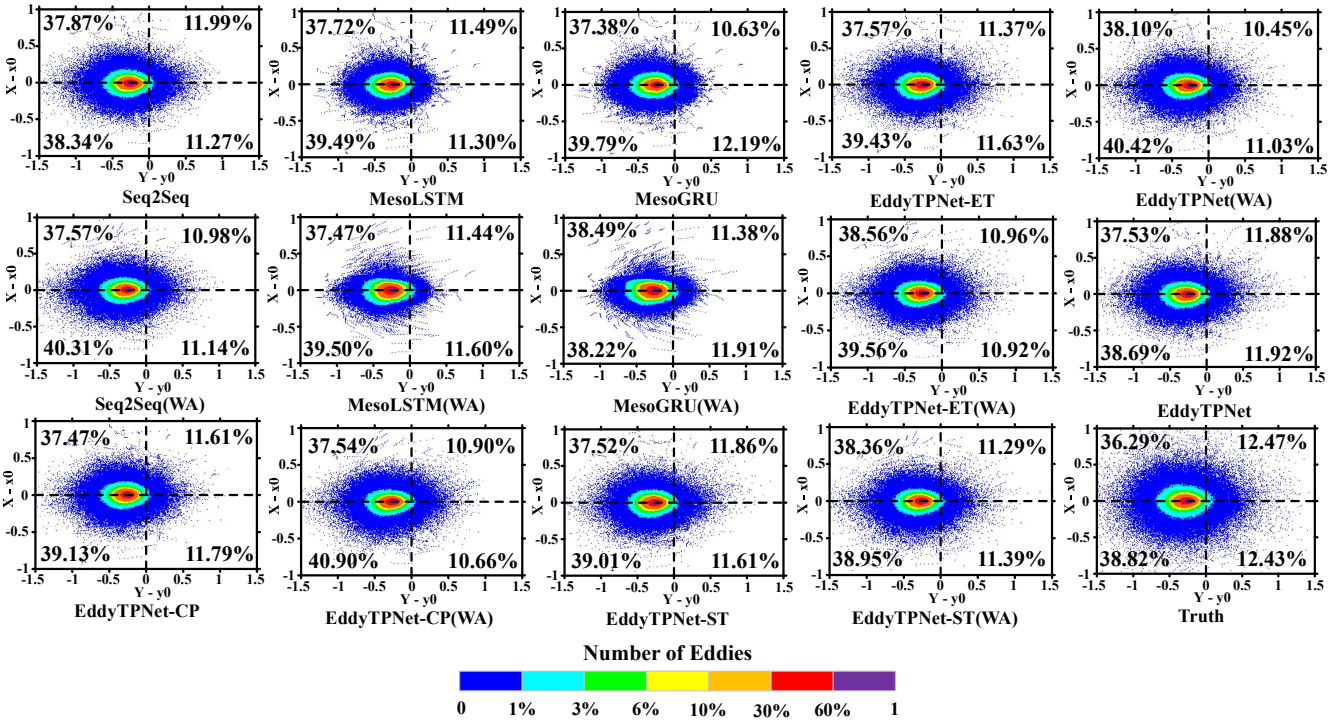

Fig. 9. Distribution of eddy trajectories predicted by different methods. Each grid is populated by the number of coordinates consisting of the meridional and zonal displacements of the eddy. Where the color indicates the number of eddies per unit grid, decreasing from purple to blue, respectively. The percentages in the figure indicate the proportion of the number of eddies accounted for by the four quadrants in the Cartesian coordinate system. (WA) denotes no angle dispersion information.

TABLE III
DAILY PREDICTION ERRORS UNDER DIFFERENT MODELS.

Days	MesoLSTM		MesoGRU		Seq2Seq		EddyTPNet-CP		EddyTPNet-ET		EddyTPNet-ST		EddyTPNet	
	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)
1	3.24	96.30	3.25	96.28	3.07	96.44	3.02	96.51	3.02	96.50	3.00	96.50	**3.00**	**96.51**
2	4.89	92.44	4.90	92.41	4.53	93.26	4.44	93.38	4.45	93.22	4.41	93.33	**4.41**	**93.34**
3	6.51	87.74	6.52	87.77	5.98	89.39	5.87	89.44	5.88	89.45	5.80	89.63	**5.80**	**89.69**
4	8.04	82.74	8.05	82.79	7.39	84.97	7.26	85.24	7.25	85.22	7.17	85.49	**7.16**	**85.50**
5	9.60	77.16	9.60	77.38	8.86	80.12	8.71	80.44	8.69	80.54	8.58	80.87	**8.56**	**80.89**
6	11.15	71.73	11.15	71.92	10.36	74.87	10.17	75.54	10.13	75.70	10.00	76.15	**9.97**	**76.19**
7	12.68	66.43	12.67	66.72	11.86	69.72	11.62	70.58	11.56	71.00	11.40	71.21	**11.37**	**71.37**

sets with a ratio of 7:1.5:1.5. In both regions, we employed various methods, including Extra Tree [27], Multiple Linear Regression [40], Random Forest [27], Gradient Boosting [27], and MesoGRU [42], to predict eddies trajectories. However, there are complex background fields in different oceanic regions, which result in significant differences in local areas. To accommodate these regional differences, EddyTPNet migrates the model trained on a global scale to local regions for fine-tuning experiments, with an initial learning rate set to 0.00001 [56]. Ultimately, results with prediction errors of less than 25 kilometers are considered accurate predictions. A summary of the 7-day forecast outcomes can be found in Table IV.

Meanwhile, the daily evaluations for prediction results of the South China Sea and the North Atlantic region are summarized in Table V and Table VI for more efficient comparisons. It can be seen that EddyTPNet's prediction accuracy exceeds

TABLE IV
PREDICTION RESULTS OF DIFFERENT MODELS IN THE SOUTH CHINA SEA AND THE NORTH ATLANTIC.

The South China Sea	MGD(km)	SGD(km)	Preicision(%)
Extra Trees	34.75	243.25	28.98
MesoGRU	32.09	224.63	42.98
Multiple Linear Regression	24.05	168.35	63.03
Random Forest	23.17	162.20	65.67
Gradient Boosting	23.07	161.49	65.52
EddyTPNet	**21.35**	**149.45**	**69.32**
The North Atlantic	MGD(km)	SGD(km)	Preicision(%)
Extra Trees	15.47	108.29	88.74
MesoGRU	14.92	104.44	87.41
Multiple Linear Regression	10.42	72.94	93.49
Random Forest	10.21	71.49	93.76
Gradient Boosting	11.44	80.11	93.73
EddyTPNet	**8.77**	**61.39**	**94.72**

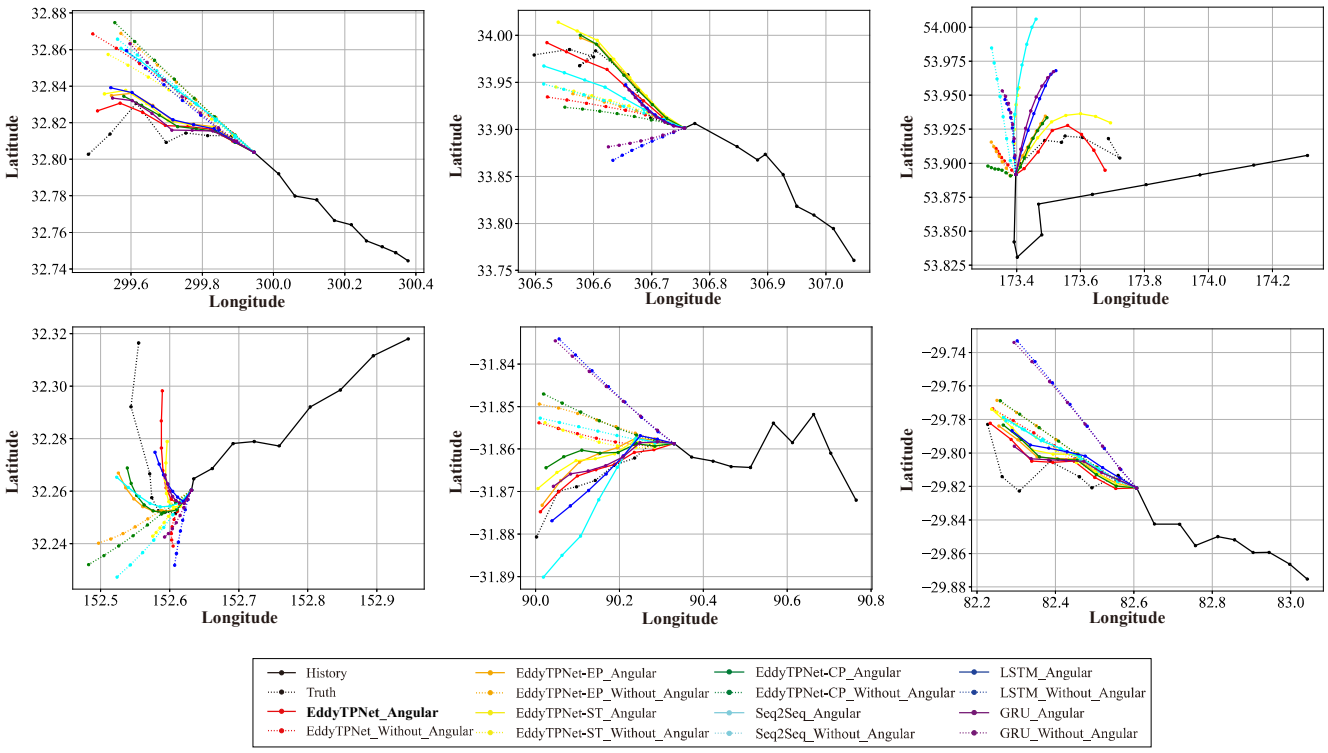

Fig. 10. Examples of trajectory prediction by different methods under directional divergence and without directional divergence.

that of the competing models. Nevertheless, there is still much room for improvement in detailed positioning accuracy. On the one hand, there are abundant types of eddy tracks in the South China Sea and the North Atlantic region, and the mechanism of each eddy is different. The insufficient number of eddy track data of each type makes it difficult to fit the eddy tracks, and it is unscientific to embed all the data into a lightweight model. On the other hand, there is a large discrepancy between the historical data and the predicted data, which reflects the complexity of the evolution of eddy motion and its dynamical mechanisms.

Consequently, we try to determine the inconsistent eddy trajectories and conduct elementary statistics on eddy trajectories, and preliminary classification of the phenomenon of complex eddy trajectories. The inconsistent trajectory cases that can be used to prove our hypothesis are shown in Fig.11, here we temporarily define them as anomalous eddy trajectories [57].

It is necessary to trace the complex ocean dynamics mechanism [58] for the cause of the anomalous eddy trajectory. The conclusion can be drawn that the introduction of prior knowledge has indeed solved the influence caused by uncertain factors. It is undeniable that the exploration of anomalous eddy trajectories still requires further research. In general, the proposed EddyTPNet with the introduction of directional divergence physical information can achieve promising prediction results.

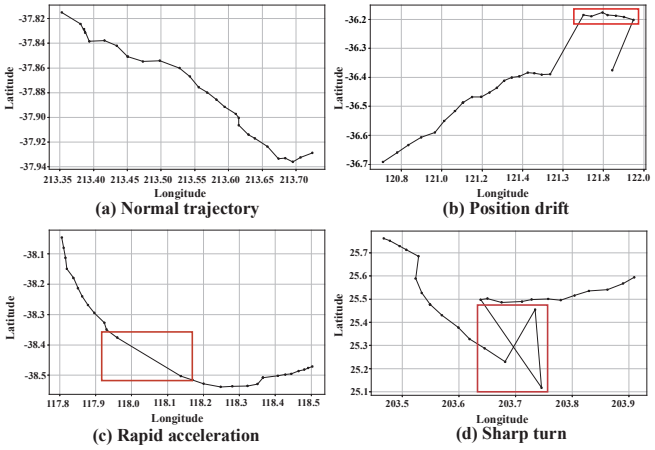

Fig. 11. Schematic diagram of the classification of eddy trajectories.

V. Conclusions and Future Work

The trajectory prediction of the oceanic eddy is a scientific issue that has not been well tackled in oceanography. With the idea of artificial intelligence for science in mind, this paper proposes a deep learning model for the estimation of eddy trajectories in oceans. It combines different modules using recurrent neural networks (GRU, LSTM) and self-attention modules to forecast the trajectory location of eddies up to 7 days in the future, based on 10 days of data (the data is mainly the amplitude, the radius, the speed average, the current displacement, the current location, the time). Moreover, a map created by averaging the past eddy trajectories (angular divergences) at every grid location was used to help the network. The results are promising and the comparison with state-of-the-art are performed.

TABLE V
DAILY PREDICTION RESULTS OF DIFFERENT MODELS IN THE SOUTH CHINA SEA.

Days	Extra Trees		MesoGRU		Mutiple Linear Regression		Random Forest		Gradient Boosting		EddyTPNet	
	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)
1	14.81	84.89	21.4	70.74	9.61	93.80	9.21	93.82	10.59	93.71	**8.28**	**94.74**
2	21.83	69.29	21.23	70.50	14.43	85.10	14.50	85.32	14.71	85.77	**12.74**	**87.27**
3	29.44	51.93	21.53	69.90	19.29	75.10	19.62	74.67	19.08	96.00	**17.25**	**78.34**
4	35.09	40.34	36.64	38.45	23.39	64.80	24.21	63.31	23.10	65.89	**21.32**	**69.15**
5	41.33	33.21	36.79	38.04	27.76	56.68	29.14	53.11	27.34	56.85	**25.87**	**60.82**
6	47.59	25.89	36.82	37.81	31.76	48.30	33.61	45.28	31.41	48.60	**29.90**	**52.92**
7	53.13	20.46	50.25	23.08	35.96	40.79	38.09	37.81	35.34	41.76	**34.12**	**46.46**

TABLE VI
DAILY PREDICTION RESULTS OF DIFFERENT MODELS IN THE NORTH ATLANTIC.

Days	Extra Trees		MesoGRU		Mutiple Linear Regression		Random Forest		Gradient Boosting		EddyTPNet	
	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)	MGD (km)	Presion (%)
1	7.29	95.72	9.92	93.91	4.57	98.06	4.86	97.76	7.23	97.96	**4.06**	**98.07**
2	10.36	92.19	9.90	93.84	6.65	96.28	6.80	96.24	8.59	96.29	**5.72**	**96.73**
3	13.08	87.72	9.95	93.87	8.61	94.45	8.58	94.37	10.05	94.39	**7.29**	**95.08**
4	15.59	82.90	16.97	81.85	10.47	91.97	10.25	92.19	11.34	92.17	**8.74**	**93.23**
5	18.32	77.30	16.93	81.44	12.35	88.97	11.93	89.88	12.83	89.52	**10.26**	**90.91**
6	20.57	71.20	17.10	81.30	14.23	85.64	13.67	86.86	14.32	86.61	**11.86**	**88.50**
7	23.09	64.96	23.71	62.79	16.07	81.38	15.40	83.49	15.75	83.25	**13.43**	**85.26**

The limitation of this study is that despite directional divergence information being used to fit the eddy movement trend, there are still eastward moving trajectories with low accuracy. As shown in Fig.9, the number of predicted eastward trajectories is significantly reduced. Possible reasons and future research may exist in:

(1) Since most trajectories are westward, the neural network neutralizes the predictions of east-west trajectories to ensure a high fit. Due to the difference in dynamic mechanism, the prediction distance of the westward trajectory is shortened and the ability to represent the eastward trajectory is decreased;

(2) The DDGEP data can guarantee the motion trend of most eddies while can not accurately represent the few eastward trajectories as the data are averaged. The more detailed dynamic mechanism and motion rules should be embedded to enhance the prediction ability;

(3) The discovery of anomalous eddy trajectories will lead to a deeper understanding of the movement of eddies. The causes of the abrupt changes and the environmental impact are also worthy of further study.

Notwithstanding these limitations, this exploratory study offers some insight into the feasibility of solving the prediction problem of ocean phenomena based on knowledge-fused time series deep neural networks. Within this knowledge-fused deep learning framework, there are various potentials to obtain a better outlook for oceanic eddy trajectory prediction and explore the model's generalization capability in other oceanic phenomena prediction.

VI. ACKNOWLEDGMENTS

The altimetric Mesoscale Eddy Trajectories Atlas (META3.2 DT) was produced by SSALTO/DUACS and distributed by AVISO+ with support from CNES, in collaboration with Mediterranean Institute for Advanced Studies (IMEDEA). The dataset includes allsat version and twosat version and can be downloaded with the linage META3.2 DT. In the spirit of reproducibility, the Python code of the proposed prediction methodology can be accessed with the linkage EddyTPNet.

REFERENCES

[1] X. Wang, X. Wang, C. Li, Y. Zhao, and P. Ren, "Data-attention-yolo (day): A comprehensive framework for mesoscale eddy identification," *Pattern Recognition*, vol. 131, p. 108870, 2022.

[2] Q. Liu, Y. Liu, and X. Li, "A deep learning model for eddy tracking based on multi-source remote sensing imagery," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, 2021, pp. 7576–7579.

[3] D. B. Chelton, M. G. Schlax, and R. M. Samelson, "Global observations of nonlinear mesoscale eddies," *Progress in Oceanography*, vol. 91, no. 2, pp. 167–216, 2011.

[4] Q. Wang, C. Pang, and C. Dong, "Role of submesoscale processes in the isopycnal mixing associated with subthermocline eddies in the philippine sea," *Deep Sea Research Part II: Topical Studies in Oceanography*, vol. 202, p. 105148, 2022.

[5] D. Dong, P. Brandt, P. Chang, F. Schütte, X. Yang, J. Yan, and J. Zeng, "Mesoscale eddies in the northwestern pacific ocean: Three-dimensional eddy structures and heat/salt transports," *Journal of Geophysical Research: Oceans*, vol. 122, no. 12, pp. 9795–9813, 2017.

[6] Y. Du, X. Yang, J. Yang, S. Tan, W. Ma, Z. Li, and X. Li, "Effects of temperature on sea surface radar backscattering under neutral and nonneutral atmospheric conditions for wind retrieval applications: a numerical study," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 4, pp. 2727–2743, 2020.

[7] Q. Xia, G. Li, and C. Dong, "Global oceanic mass transport by coherent eddies," *Journal of Physical Oceanography*, vol. 52, no. 6, pp. 1111–1132, 2022.

[8] J. Martínez-Moreno, A. M. Hogg, M. H. England, N. C. Constantinou, A. E. Kiss, and A. K. Morrison, "Global changes in oceanic mesoscale currents over the satellite altimetry record," *Nature Climate Change*, vol. 11, no. 5, pp. 397–403, 2021.

[9] D. Zhao, Y. Xu, X. Zhang, and C. Huang, "Global chlorophyll distribution induced by mesoscale eddies," *Remote Sensing of Environment*, vol. 254, p. 112245, 2021.
[10] M. Naeije, K. Wakker, R. Scharroo, and B. Ambrosius, "Observation of mesoscale ocean currents from geosat altimeter data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 47, no. 5, pp. 347–368, 1992.
[11] J. Ji, C. Dong, X. Liu, T. Liu, Y. Yu, K. T. L. K. Sian, B. J. Bethel, and H. Zhao, "Influence of oceanic mesoscale eddy on the atmospheric boundary layer based on an idealized model," *Deep Sea Research Part II: Topical Studies in Oceanography*, vol. 202, p. 105146, 2022.
[12] C. Dong, L. Liu, F. Nencioli, B. J. Bethel, Y. Liu, G. Xu, J. Ma, J. Ji, W. Sun, H. Shan *et al.*, "The near-global ocean mesoscale eddy atmospheric-oceanic-biological interaction observational dataset," *Scientific Data*, vol. 9, no. 1, p. 436, 2022.
[13] A. S. J. Wyatt, J. J. Leichter, L. Washburn, L. Kui, P. J. Edmunds, and S. C. Burgess, "Hidden heatwaves and severe coral bleaching linked to mesoscale eddies and thermocline dynamics," *Nature Communications*, vol. 14, no. 1, p. 25, 2023.
[14] Q. Wang, C. Dong, J. Dong, H. Zhang, and J. Yang, "Submesoscale processes-induced vertical heat transport modulated by oceanic mesoscale eddies," *Deep Sea Research Part II: Topical Studies in Oceanography*, vol. 202, p. 105138, 2022.
[15] Y. Liu, Q. Zheng, and X. Li, "Characteristics of global ocean abnormal mesoscale eddies derived from the fusion of sea surface height and temperature data by deep learning," *Geophysical Research Letters*, vol. 48, no. 17, p. e2021GL094772, SEP 16 2021.
[16] B. Lim and S. Zohren, "Time-series forecasting with deep learning: a survey," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2194, p. 20200209, 2021.
[17] Y. Zhou, C. Lu, K. Chen, and X. Li, "Multilayer fusion recurrent neural network for sea surface height anomaly field prediction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–11, 2021.
[18] G. Chen and G. Han, "Contrasting short-lived with long-lived mesoscale eddies in the global ocean," *Journal of Geophysical Research: Oceans*, vol. 124, no. 5, pp. 3149–3167, 2019.
[19] F. Teng, C. Dong, L. K. Sian, K. T. Choy, J. Ji, and W. Zhu, "Wind work on oceanic mesoscale eddies in the northeast tropical pacific ocean," *Frontiers in Marine Science*, 2023.
[20] J. Shi, D. Lü, Y. Wang, Y. Du, Y. Pang, D. Yang, X. Wang, X. Dong, and X. Yang, "Recent progress of earth science satellite missions in china," *Chin. J. Space Sci*, vol. 42, pp. 712–723, 2022.
[21] J. Souza, C. de Boyer Montégut, C. Cabanes, and P. Klein, "Estimation of the agulhas ring impacts on meridional heat fluxes and transport using argo floats and satellite data," *Geophysical Research Letters*, vol. 38, no. 21, 2011.
[22] F. Nencioli, C. Dong, T. Dickey, L. Washburn, and J. C. McWilliams, "A vector geometry–based eddy detection algorithm and its application to a high-resolution numerical model product and high-frequency radar surface velocities in the southern california bight," *Journal of atmospheric and oceanic technology*, vol. 27, no. 3, pp. 564–579, 2010.
[23] Y. Du, X. Yang, J. Yang, S. Tan, W. Ma, Z. Li, and X. Li, "Effects of temperature on sea surface radar backscattering under neutral and nonneutral atmospheric conditions for wind retrieval applications: A numerical study," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 4, pp. 2727–2743, 2020.
[24] Y. Du, X. Yang, J. Yang, and X. Li, "A numerical study of sst effects on ocean radar backscattering," in *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2020, pp. 5725–5728.
[25] A. R. Robinson and W. G. Leslie, "Estimation and prediction of oceanic eddy fields," *Progress in Oceanography*, vol. 14, pp. 485–510, 1985.
[26] X. Jiang, C. Dong, Y. Ji, C. Wang, Y. Shu, L. Liu, and J. Ji, "Influences of deep-water seamounts on the hydrodynamic environment in the northwestern pacific ocean," *Journal of Geophysical Research: Oceans*, vol. 126, no. 12, p. e2021JC017396, 2021.
[27] X. Wang, H. Wang, D. Liu, and W. Wang, "The prediction of oceanic mesoscale eddy properties and propagation trajectories based on machine learning," *Water*, vol. 12, no. 9, 2020.
[28] J. Wang, "A nowcast/forecast system for coastal ocean circulation using simple nudging data assimilation," *Journal of Atmospheric and Oceanic Technology*, vol. 18, no. 6, pp. 1037–1047, 2001.
[29] C. Dong, G. Xu, G. Han, B. J. Bethel, W. Xie, and S. Zhou, "Recent developments in artificial intelligence in oceanography," *Ocean-Land-Atmosphere Research*, vol. 2022, 2022.
[30] H. Wang and X. Li, "Deepblue: Advanced convolutional neural network applications for ocean remote sensing," *IEEE Geoscience and Remote Sensing Magazine*, 2023.
[31] Y. Ren, X. Li, and H. Xu, "A deep learning model to extract ship size from sentinel-1 sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2021.
[32] Y. Ren, X. Li, X. Yang, and H. Xu, "Development of a dual-attention u-net model for sea ice and open water classification on sar images," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2021.
[33] Z. Liao, Q. Gu, S. Li, and Y. Sun, "A knowledge transfer-based adaptive differential evolution for solving nonlinear equation systems," *Knowledge-Based Systems*, vol. 261, p. 110214, 2023.
[34] X. Li and R. Fablet, "Foreword to the special issue on the remote sensing of the world oceans," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 11, pp. 4895–4897, 2016.
[35] Y. Liu, X. Li, and Y. Ren, "A deep learning model for oceanic mesoscale eddy detection based on multi-source remote sensing imagery," in *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2020, pp. 6762–6765.
[36] Y. Liu, L. Gao, Q. Liu, and X. Li, "Dual-branch neural network for mesoscale eddy identification based on multi-variables remote sensing data," in *2023 Photonics & Electromagnetics Research Symposium (PIERS)*. IEEE, 2023, pp. 275–279.
[37] Z.-H. Zhou, *Machine learning*. Springer Nature, 2021.
[38] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
[39] Z. Zhang, X. Yang, L. Shi, B. Wang, Z. Du, F. Zhang, and R. Liu, "A neural network framework for fine-grained tropical cyclone intensity prediction," *Knowledge-Based Systems*, vol. 241, p. 108195, 2022.
[40] J. Li, G. Wang, H. Xue, and H. Wang, "A simple predictive model for the eddy propagation trajectory in the northern south china sea," *Ocean Science*, vol. 15, no. 2, pp. 401–412, 2019.
[41] J. Wen, J. Yang, Y. Li, and L. Gao, "Harmful algal bloom warning based on machine learning in maritime site monitoring," *Knowledge-Based Systems*, vol. 245, p. 108569, 2022.
[42] X. Wang, X. Wang, M. Yu, C. Li, D. Song, P. Ren, and J. Wu, "Mesogru: Deep learning framework for mesoscale eddy trajectory prediction," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2021.
[43] Y.-R. Wang and X.-M. Li, "Arctic sea ice cover data from spaceborne synthetic aperture radar by deep learning," *Earth System Science Data*, vol. 13, no. 6, pp. 2723–2742, 2021.
[44] J. Dong, B. Fox-Kemper, H. Zhang, and C. Dong, "The scale of submesoscale baroclinic instability globally," *Journal of Physical Oceanography*, vol. 50, no. 9, pp. 2649–2667, 2020.
[45] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," *Nature Reviews Physics*, vol. 3, no. 6, pp. 422–440, JUN 2021.
[46] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[47] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, 1994.
[48] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 30, 2017.
[49] J. Ji, C. Dong, B. Zhang, Y. Liu, B. Zou, G. P. King, G. Xu, and D. Chen, "Oceanic eddy characteristics and generation mechanisms in the kuroshio extension region," *Journal of Geophysical Research: Oceans*, vol. 123, no. 11, pp. 8548–8567, 2018.
[50] T. Rohr, C. Harrison, M. C. Long, P. Gaube, and S. C. Doney, "The simulated biological response to southern ocean eddies via biological rate modification and physical transport," *Global Biogeochemical Cycles*, vol. 34, no. 6, 2020.
[51] G. Chen, X. Chen, and C. Cao, "Divergence and dispersion of global eddy propagation from satellite altimetry," *Journal of Physical Oceanography*, vol. 52, no. 4, pp. 705–722, 2022.
[52] N. Mohajerin and S. L. Waslander, "Multistep prediction of dynamic systems with recurrent neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3370–3383, 2019.
[53] L. Ge, B. Huang, X. Chen, and G. Chen, "Medium-range trajectory prediction network compliant to physical constraint for oceanic eddy," *IEEE Transactions on Geoscience and Remote Sensing*, 2023.
[54] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," *Advances in neural information processing systems*, vol. 27, 2014.

[55] A. Ioannou, S. Speich, and R. Laxenaire, "Characterizing mesoscale eddies of eastern upwelling origins in the atlantic ocean and their role in offshore transport," *Frontiers in Marine Science*, vol. 9, p. 835260, 2022.
[56] X. Zhang, H. Wang, S. Wang, Y. Liu, W. Yu, J. Wang, Q. Xu, and X. Li, "Oceanic internal wave amplitude retrieval from satellite images based on a data-driven transfer learning model," *Remote Sensing of Environment*, vol. 272, p. 112940, 2022.
[57] Y. Liu, Q. Zheng, and X. Li, "Characteristics of global ocean abnormal mesoscale eddies derived from the fusion of sea surface height and temperature data by deep learning," *Geophysical Research Letters*, vol. 48, no. 17, p. e2021GL094772, 2021.
[58] E. J. Drenkard and K. B. Karnauskas, "Strengthening of the pacific equatorial undercurrent in the soda reanalysis: Mechanisms, ocean dynamics, and implications," *Journal of Climate*, vol. 27, no. 6, pp. 2405–2416, 2014.

Xinmin Zhang is pursuing the M.S. degree in computer application technology from the Qingdao University, Qingdao.His research interests include satellite remote sensing of the ocean, big data oceanography, and deep learning.

Baoxiang Huang (Member, IEEE) received the B.S. degree in traffic engineering from the Shandong University of Technology, China, in 2002, the M.S. degree in mechatronic engineering from Shandong University, China, in 2005, and the Ph.D. degree in computer engineering from Ocean University of China, in 2011, China. She was an academic visitor of Nottingham University. Currently, she is a Professor at the College of Computer Science and Technology, Qingdao University, China. Her research interests include remote sensing image processing and analysis, big data oceanography, and artificial intelligence.

Ge Chen received the B.S. degree in marine physics, the M.S. degree in satellite oceanography, and the Ph.D. degree in physical oceanography from the Ocean University of China (OUC), Qingdao, China, in 1988 and 1990, respectively. After graduation, he worked as a Post-Doctoral Fellow at The French Research Institute for the Exploitation of the Sea (IFREMER), Brest, France, from 1994 to 1996. Since 1997, he has been a Professor of satellite oceanography and meteorology with the OUC. In 2001, he received the National Science Fund for Outstanding Young Scientists awarded by the Natural Science Foundation of China and became the Chair Professor of the Cheung Kong Scholars Program nominated by the Chinese Ministry of Education. He is also the Chief Scientist for Ocean Science Satellite Missions at the National Laboratory for Ocean Science and Technology, Qingdao. He is the author of more than 180 peer-reviewed scientific articles published in internationally recognized journals. His research interests include satellite remote sensing of the ocean and big data oceanography.

Linyao Ge received the B.S. degree in computer application technology and the M.S. degree in computer technology from Qingdao University,Qingdao, China, in 2018 and 2021, respectively.He is pursuing the Ph.D. degree in computer application technology from the Ocean University of China, Qingdao.His research interests include satellite remote sensing of the ocean, big data oceanography, and deep learning.

Milena Radenkovic received the Dipl.−Ing. degree in electric and electronic engineering from the University of Nis, Nis, Serbia, in 1998, and the Ph.D. degree in computer science from the University of Nottingham, Nottingham, U.K., in 2002. She has authored more than 80 papers in premium conferences and journal venues. Her research interests include intelligent mobile and disconnection-tolerant networking, complex temporal graphs, self-organized security, distributed predictive analytics with applications to autonomous vehicles, mobile social networks, smart manufacturing, and predictive telemetry. Dr. Radenkovic was a recipient of multiple EPSRC and EU grants for her research. She has organized and chaired multiple ACM and IEEE conferences and served on many program committees. She is an Editor for premium journals such as the Ad Hoc Networks (Elsevier), the IEEE Transactions on Parallel and Distributed Computing, and ACM Multimedia.

Guojia Hou received his B.S. degree in computer science and technology from Linyi University in 2010, and Ph.D. degree in computer application technology from the Ocean University of China in 2015, respectively. He is currently an Associate Professor at the School of Computer Science and Technology, Qingdao University. His current research interests include underwater vision, image/video processing, and image quality assessment.