water

MDPI

Article

The Prediction of Oceanic Mesoscale Eddy Properties and Propagation Trajectories Based on Machine Learning

Xin Wang [1], Huizan Wang [2], Donghan Liu [1] and Wenke Wang [2,*]

[1] College of Computer, National University of Defense Technology, Changsha 410073, China; wangx14@lzu.edu.cn (X.W.); liudh@nudt.edu.cn (D.L.)

[2] College of Meteorology & Oceanography, National University of Defense Technology, Changsha 410073, China; wanghuizan17@nudt.edu.cn

* Correspondence: wangwenke@nudt.edu.cn; Tel.: +86-151-1149-0700

Received: 23 July 2020; Accepted: 8 September 2020; Published: 9 September 2020

check for updates

Abstract: Mesoscale eddies play an important role in ocean circulation, material energy exchange and variation of ocean environments. Machine learning methods can efficiently process massive amounts of data and automatically learn the implicit features, thus providing a new approach to eddy prediction research. Using the mesoscale eddy trajectory data derived from multimission satellite altimetry, we propose relevant machine learning models based on long short-term memory network (LSTM) and the extra trees (ET) algorithm for the prediction of eddy properties and propagation trajectories. Characteristic factors, including attribute features and past eddy displacements, were exploited to construct prediction models with high effectiveness and few predictors. To study their effects at different forecasting times, we separately trained the models by rebuilding the corresponding relationship between eddy samples and labels. In addition, the variation characteristics and the predictability of eddy properties and propagation trajectories were discussed from the prediction results. Cross-validation shows that at different prediction times, our method is superior to previous methods in terms of the mean absolute error (MAE) of eddy properties and the root mean square error (RMSE) of propagation. The stable variation in eddy properties makes the prediction more dependent on the historical time series than that of a propagation forecast. The short-term propagation prediction of eddies contained more noise than long-term predictions, and the long-term predictions revealed a more significant trend. Finally, we discuss the effect of eddy properties on the prediction ability of the eddy propagation trajectory.

Keywords: mesoscale eddies; eddy properties; propagation trajectory; performance evaluation; machine learning

1. Introduction

Mesoscale eddies play an important role in the mixing and transport of momentum, heat, mass and biogeochemical tracers in the global ocean [1–3]. In addition, eddies have a critical influence on rainfall, near-surface wind, clouds and marine ecosystems in their vicinity [4–6]. The accurate prediction of eddies is of great scientific and applied significance for understanding eddy propagation and evolution characteristics and improving simulations and predictions of regional weather and climate change [7,8].

In general, the main methods used for oceanic mesoscale prediction can be divided into three categories: dynamic statistical methods, numerical methods and machine learning methods. Traditionally, ocean dynamic models have been used to forecast the evolution of ocean eddies.

Robinson et al. [9] reported the first use of a mesoscale eddy observation network in the prediction of eddy evolution at two weeks; this model was an anisotropic statistical model for mixed spatiotemporal target analysis. Li et al. [7] developed a multiple linear regression model to predict the eddy propagation trajectory at 1–4 weeks. This simple empirical model combined the oceanic parameters that mainly represent β effects and mean flow advection with eddy propagation positions. With the development of computer technology, numerical prediction has been applied to study oceanic eddies. In 1994, Masina et al. [10] proposed an objective analysis method and established a regular-grid quasi-geostrophic numerical model in the initial field of the Adriatic Sea area. Based on the experimental results, they concluded that topographic changes had a considerable impact on the prediction ability of the model. Shriver et al. [11] combined the navy stratified ocean model (NLOM) with the optimal interpolation (OI) method, which improved the resolution of the prediction system from 1/16° to 1/32°. The nonlinear characteristics of mesoscale eddies and the sensitivity of numerical models to the initial conditions and background errors make forecasting difficult, which has always been a major challenge for marine numerical models [7,8].

With the continuous advancement of artificial intelligence, machine learning has led to remarkable achievements in many fields, such as pattern recognition and target detection, based on powerful feature extraction and modelling capabilities [12–14]. In recent years, with the rapid acquisition of long time series of marine remote sensing data over large areas to obtain a sufficient sample size, many breakthroughs have been made in the study of eddies. The automatic identification and extraction of eddies based on remote sensing data has become an effective means of studying eddies [15,16]. Chelton et al. identified and tracked global eddies on the basis of sea surface height (SSH) from remote sensing satellites [17,18]. Ashkezari et al. used machine learning to forecast the lifetime of eddies in stable evolution [19]. Ma et al. combined a convolutional neural network (CNN) and LSTM to reconstruct the future field of sea level anomalies (SLAs) and applied an eddy detection algorithm for 1–7 day eddy prediction [8]. However, the forecasts were strongly affected by the SLA noise in each grid, and the corresponding relationship between the time before and after the target eddies formed was ignored. Accurately tracking and predicting target eddies in a complex flow field remain major challenges.

In this paper, we constructed a numerical vector containing the most relevant features of eddy prediction targets according to the eddy properties and propagation trajectories. A LSTM network is used to learn the spatiotemporal variation characteristics, and the et algorithm is applied to learn the relevant one-dimensional characteristics at the current time. Prediction models of eddy properties and propagation trajectories are established for the comprehensive and accurate forecasting of target eddies at different prediction times. The experimental results show that the proposed method outperforms those in previous studies. By comparing the LSTM method with other machine learning regression methods, it can be concluded that the prediction of eddy properties is more dependent on the historical time series data than the propagation trajectory. Finally, the variation characteristics and predictability of eddy properties and propagation trajectories are discussed. The eddy properties have different impacts on model performance in propagation prediction.

This paper is organized as follows. In Section 2, the data sources, implementation flow of the method, and model structure used in this paper are introduced. Section 3 focuses on analyzing and comparing the results of multiple methods, and an extended discussion is given in Section 4. Finally, Section 5 presents some conclusions and future research prospects.

2. Data and Methodology

This paper used the global mesoscale oceanic eddy trajectory dataset published by Chelton et al. [18], which is still being updated and available on the Archiving, Validation, and Interpretation of Satellite Oceanographic data (AVISO) (https://www.aviso.altimetry.fr/en/data/data-access). The trajectory dataset includes records for the eddies derived from multimission altimetry datasets, with location, speed, radius and other associated metadata each day. A detailed description of the dataset was given by Chelton et al. (2011) [18] and Mason et al. (2014) [20]. The South China Sea is a region with numerous eddies with obvious spatiotemporal variation characteristics [21–23]. In this study, the South China Sea (5°–25° N, 105°–125° E) was taken as the experimental area. Based on daily eddy trajectory and attribute data over 20 years (1993–2012), eddy properties and propagation trajectories were predicted. The overall process framework of our method in this paper is shown in Figure 1.

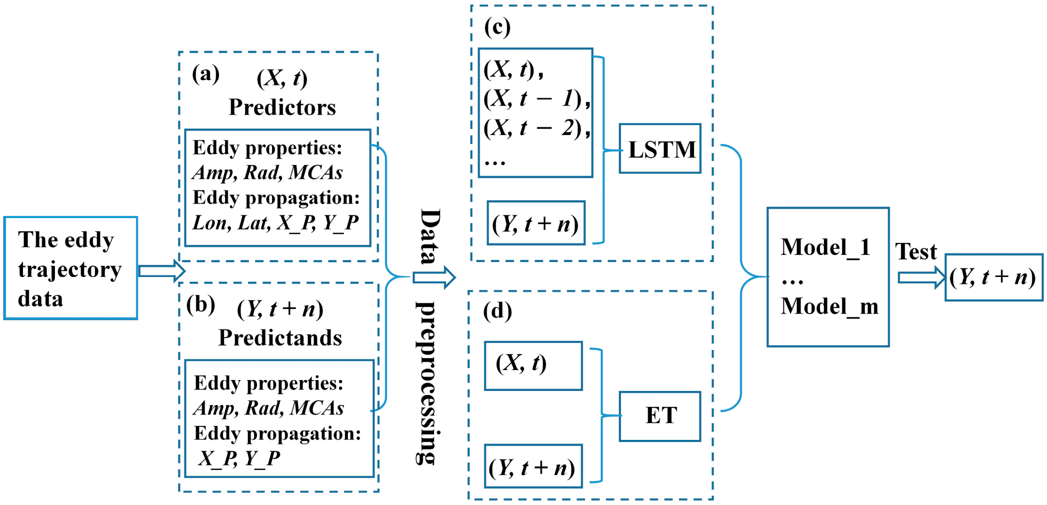


Figure 1. Framework of eddy predictions. (**a**) Characteristic factors, including the features of eddy properties and propagation trajectory characteristics at time t. (**b**) Prediction targets, including the properties and propagation trajectories of eddies at time $t + n$, where n refers to the time step of forecasting. (**c**) Eddy property prediction based on the LSTM network. (**d**) Eddy propagation trajectory prediction based on the integrated ET model.

The characteristic factors we selected can be divided into two categories. One category describes the relevant properties of an eddy itself, including the amplitude (*Amp*), radius (*Rad*) and maximum circum-average (MCA) speed (*MCAs*). Such characteristics reflect the real status of an eddy on the two-dimensional sea surface. The other category involves the eddy latitude and longitude (*Lon*, *Lat*), past zonal displacement (*X_P*) and meridional displacement (*Y_P*) associated with the propagation positions of eddies. These parameters encompass the recent patterns of eddy propagation that are affected by β effects and the mean advection. In the evolution process of eddies, the properties and propagation trajectories change dynamically and are interrelated. The two-dimensional properties of eddies are related to movement and reflect the current state and changes in eddies. Based on the corresponding relationship $((X, t), (Y, t + n))$ between a feature vector (X, t) and the predicted values $(Y, t + n)$ at different prediction time steps n, we built machine learning models to explore the implicit relations between these characteristic factors and the labels for eddy prediction. The proposed prediction algorithm includes the following steps.

(1) The characteristic factors X related to eddies at the current time t and the predicted target Y at time $t + 1$ are combined to form a one-to-one corresponding relationship $((X, t), (Y, t + n))$ between eddy samples and labels. This information is then input into the model after preprocessing.

(2) If the predicted target belongs to eddy properties, the samples $(X, t), (X, t - 1), (X, t - 2) \ldots$ from previous time series are used to learn the spatiotemporal features of the associated eddy or eddies. An LSTM network for eddy property prediction is constructed to obtain three prediction models for the eddy amplitude, radius and MCAs.

(3) If the predicted value Y is the eddy propagation trajectory, the characteristic factors (X, t) at the current time t are used to learn the eddy features at the current moment, and the ET model is built to obtain two prediction models of zonal and meridional displacement for eddies.

(4) When the forecasting time step is n, the above steps (1), (2) and (3) are repeated. Then, the $5{\cdot}n$ matrices of independent models for different prediction targets are obtained, and the training models are parallelized to form an integrated learning system for eddy prediction.

2.1. LSTM Network for Learning Temporal Features

An LSTM network is an improved version of a recurrent neural network (RNN) that successfully solves the long-term dependence issue in prediction. LSTM methods have become the most popular RNN models in voice recognition, natural language processing and many other fields [24–26]. The structure of LSTM is relatively complex, and a cell is added to the hidden layer of the original RNN to preserve the long-term status. As shown in Figure 2a, LSTM has three inputs at time t: the input value X_t of the network at the current moment and the output value h_{t-1} and the cell state C_{t-1} at time $t-1$. There are two outputs of LSTM: the LSTM output value h_t at the current moment and the current cell C_t. The LSTM network uses a "gate" structure to control information flows. Figure 2b shows the process of information transfer in LSTM unit A2. The input gate I_t of LSTM determines the input information at the current time t and adds the information to the memory information flow. The function of the forget gate F_t is to decide what information should be retained by memory and passed to the next memory layer; simultaneously, other information is discarded. The update gate U_t is used to calculate the total amount of information passed between the current input and the previous memory output; the output gate O_t controls the amount of information used for the next update. With this gate structure, the LSTM network can complete information transfer tasks with temporal relationships. The main operation process of the model is as follows:

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

$$I_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$U_t = \tanh(W_u \cdot [h_{t-1}, x_t] + b_u) \tag{3}$$

$$C_t = I_t \cdot U_t + F_t C_{t-1} \tag{4}$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = O_t \cdot \tanh(C_t) \tag{6}$$

where W_i, W_f, W_u and W_o are the weights of the input gate, forget gate, update gate and output gate, respectively. b_i, b_f, b_u and b_o are the biases of the input gate, forget gate, update gate and output gate, respectively. σ and tanh represent the activation functions.

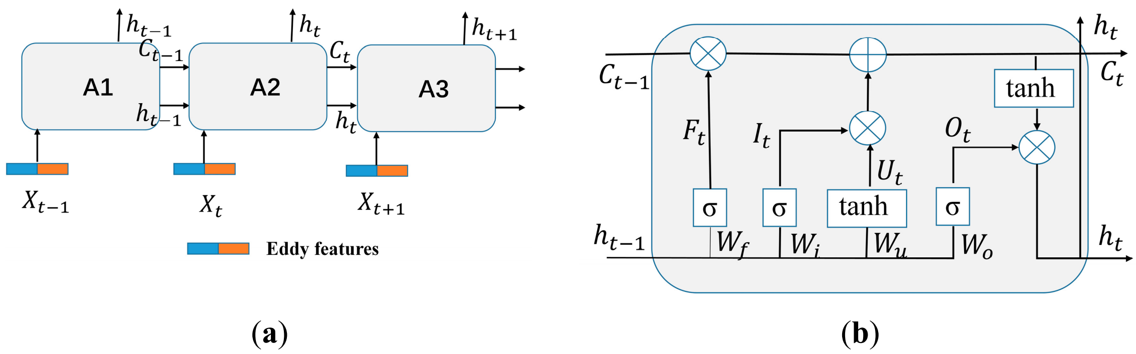


Figure 2. (**a**) Structure of the LSTM model for eddy forecasting; (**b**) schematic diagram of the information transfer inside an LSTM unit.

In this paper, an LSTM network was used to predict eddy properties. Mesoscale eddy motion (including the speed, shape, radius, displacement, etc.) is a process with spatiotemporal variations.

In the movement process, the positions and properties of eddies at different moments are characterized, and these properties are related to the previous motion state. Thus, the future state of a mesoscale eddy can be predicted by the characteristic parameters at multiple moments in the preceding sequence. In our study, the one-dimensional characteristic parameters of an eddy at time t are taken as the inputs of the LSTM model. For example, given a sequence of spatiotemporal variables $X_{t-b}(b = 1, 2, 3 \ldots)$, the future state Y_{t+n} of the mesoscale eddy is predicted. The time series prediction model can be expressed as follows:

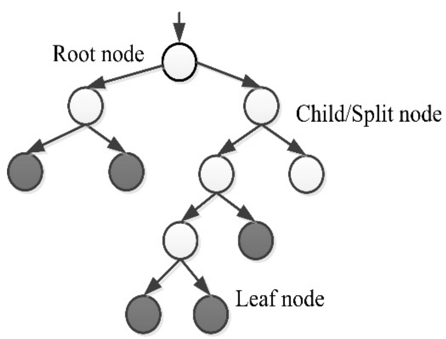
$$Y_{t+n} = f(X_t, X_{t-b})\ n, b = 1, 2, 3 \ldots \tag{7}$$

where Y_{t+n} is the eddy status after n time steps from the current prediction time t, f represents the model that needs to be learned through historical data, X_t is the vector of eddy characteristic factors at time t, and X_{t-b} is the vector of eddy characteristic factors corresponding to b time steps before time t.

2.2. The Extra Trees Algorithm

Tree-based ensemble methods are popular approaches for supervised classification and regression problems [27,28]. We use the et algorithm to establish a prediction model of the eddy propagation trajectory and associate the property characteristics with changes in propagation positions. The relationship between the response and explanatory variables is simulated by an ensemble learning algorithm to evaluate the impact on eddy prediction. The ET regression algorithm is similar to the random forest algorithm, which involves the construction of a large number of decision trees. The random forest algorithm is an ensemble machine learning approach based on decision trees and was jointly proposed by Leo Breiman and Adele Cutler in 2001 [29]. This model is composed of a set of regression decision trees $\{h(x, \theta_t), t = 1, 2, \cdots, T\}$, where θ_t is a random variable subject to independent and identically distributed conditions, x is an independent variable and T is the number of decision trees. A composite model of the composition is shown in Formula 8. The strategy for each decision tree is to select an attribute that divides the sample features among nodes through the specified measurement and determine the structure of the decision tree composed of root, child and leaf nodes, as shown in Figure 3a. Finally, the mean of all decision trees is taken as the final regression result based on ensemble learning:

$$\bar{h}(x) = \frac{1}{T}\sum_{t=1}^{T}\{h(x, \theta_t)\} \tag{8}$$



1. Input: Training subset $S_i = (X_1, X_2, \ldots, X_n)$, where one of the eddy samples $X_i = (f_1, f_2, \ldots, f_7)$ is a 7-dimensional vector, the number of features to select randomly, K, and the minimum number of samples required to split a node, s_{min}.
2. If n < s_{min}
Stop splitting and define the node as a leaf node.
3. Else
Select a random subgroup of K features $(f_1, f_2, \ldots, f_K)$ among the original 7 eddy features.
4. For each feature f_i in the subgroup **do:**
Find $max(f_i)$ and $min(f_i)$ as the maximal and minimal values of the feature f_i in subset S_i. Obtain a random cut-point, C_K, uniformly in the range $[\min(f_i), max(f_i)]$. Set $[f_K < C_K]$ as a candidate split
End for
5. Select a spilt $[f_i < C_i]$ such that Score(C_i) = $minScore(C_k)$, (k = 1,2, …, K)
6. Output: Best split $[f_i < C_i]$ at the child node.

(**a**) (**b**)

Figure 3. (**a**) Illustration of a decision tree. (**b**) Extra trees algorithm for eddy forecasting.

To overcome the low precision and overfitting problems of decision trees, bagging and random subspace tasks are introduced into the random forest. Bagging refers to the random extraction of multiple training samples from the original samples, and regression decision subtrees are constructed for each group of training samples. Bagging not only uses randomization to build regression decision subtrees but also ensures a certain degree of correlation among independent trees. The random subspace process involves constructing a decision subtree, in which each splitter node randomly extracts the feature subspace from the total feature space as the candidate feature set of the node and splits the optimal feature(s) from it. This method ensures that the differences in the nodes and feature subsets among trees are different, guarantees the independence and diversity of trees, and improves the randomness of node splitting. The extra tree algorithm includes two key improvements over the random forest algorithm. The first improvement is complete random splitting based on nodes. The second improvement is that each tree grows with the entire sample set rather than employing guided sampling. In this way, the randomness of the model is stronger than that of a random forest, which means that more complete features can be learned from the data, thus increasing its robustness.

In this paper, the extra tree algorithm has been applied to predict eddy propagation trajectories. The training set $S_i = (X_1, X_2, \ldots, X_n)$, where $X_i = (f_1, f_2, \ldots, f_7)$ is a vector containing 7 eddy features, is considered. In each decision tree, S_i represents the dataset trained at child node i. At each node i, the model selects the best split based on S_i, and random feature grouping is performed by the algorithm described in Figure 3b to reduce the variance compared to that in other randomization schemes. Variance is generated if a model is too sensitive to small fluctuations in the training set, and high variance will lead to overfitting; however, explicit randomization in feature subset selection and cutting point selection will reduce the variance [30].

3. Experiment and Results

3.1. Model Training and Evaluation

Our model was trained on two Titan Xp GPUs (NVIDIA, Santa Clara, CA, USA) with 12 GB memory with the Ubuntu 16.04 system and Keras and Sklearn as the machine learning framework. The LSTM network consisted of several processing layers that were used to continuously extract abstract features from the input data and match these features with the targets learned through regression tasks. Each layer included numerous neurons that calculated the weighted combination of inputs and trained the model with the relevant dataset to optimize the model parameters through a nonlinear activation function. In our model, the training started with the initialization of the weights of the nodes between the layers and the updating of parameters through gradient descent. Based on iterative processing with the training set and minimizing the loss function, the optimal solution was obtained [24,25]. In the training of the ET model, the two parameters that had the biggest impact on performance were the number of generated decision trees and the number of feature subsets. We trained the model with default parameters combined with local fine-tuning.

For multi-time-step prediction, all models used historical observation data to establish a one-to-one correspondence with the prediction target. The intermediate prediction results were not used for recursive prediction or training at longer time steps to avoid the accumulation of prediction errors. The prediction at each time step was direct and independent. The models adopted similar parameter settings and network frameworks for different prediction times and targets. Of all the samples, 75% were used for model training, and 25% were used for cross-validation to evaluate model performance. To perform homogeneous comparisons with other studies, we used the MAE and RMSE in the performance evaluation of eddy properties and propagation trajectory prediction, respectively; the corresponding formulas are as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{y}_i - y_i\right| \tag{9}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{10}$$

where N is the number of samples, $\hat{y}_i$ is the predicted value, and y_i is the real value.

3.2. The Prediction of Eddy Properties

Eddy properties are an important embodiment of the characteristic signal and evolution. Due to the temporal availability, resolution and coverage improvements associated with satellite altimetry data, an increasing number of studies have been performed to detect and extract the features of eddy properties based on the SSH, which is an essential part of follow-up eddy tracking research. However, few prediction studies have been based on eddy trajectory data. Machine learning models require large numbers of samples to train and learn the implicit relationship between the data and the target and convert the complex evolution process into a simple expression. In this study, we obtained numerous effective samples based on the eddy positions and property information recorded in the eddy trajectory dataset and used machine learning to study the evolution trend of eddies. Histograms of the attribute information, including amplitude, radius and MCA information, from the South China Sea between 1993 and 2012 are shown in Figure 4; the means of the three attributes are 8.2 cm, 112.1 km and 30.8 cm/s, respectively.

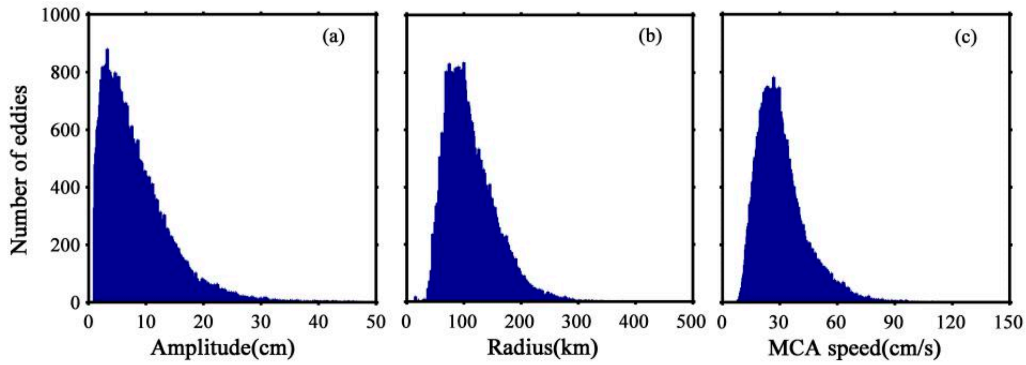


Figure 4. Histograms of eddy properties from 1993–2012. (**a**) Amplitude (cm); (**b**) Radius (km); (**c**) MCA speed (cm/s).

In the experiment, we used the LSTM network for the prediction of eddy properties. The daily variations in eddy properties are generally stable and small. Similarity algorithms based on the relevant threshold of parameters, such as those corresponding to eddy properties, have been widely used in studies of eddy tracking problems [18,20] to reflect the variation characteristics of eddy properties and to achieve a simple learning mode. However, with increasing prediction time, the stability of this change gradually decreases. The model can learn complex trends by using the characteristic information and changes from previous time steps. The "gate" structure of the LSTM network can be used for the transmission and updating of important features in the time series of historical information and for selectively forgetting invalid information. A historical time series that is too long will waste resources and require a long run time; thus, such series are not ideal for sample testing or the processing of missing data. In the prediction of eddy properties, we use historical data sets with different time series lengths to train and test the performance of the LSTM model. In eddy radius forecasting, the model performance is best when the prediction time is short and the time series length is 5; when the forecasting time was comparatively long, the time series length was 7, which led to the lowest prediction error. For eddy amplitude and MCA forecasting, the prediction results were similar for different time series lengths, which suggested that the time series length had little impact on the performance of the model. Notably, the prediction error of the model was very small, and the slight change in the time series length was not reflected in the overall index.

We calculated the prediction results for 14,000 eddy samples with the best parameters of the LSTM model. Figures 5–7 show the spatiotemporal distribution of the errors in eddy properties within the prediction time of 1–7 days. In the first two days of prediction, the error of most eddy samples is very small. As the prediction time increases, the number of eddies with large absolute errors increases, and the number of eddies with small errors decreases, resulting in an increase in the average absolute error for all tested eddies. The distribution errors of the eddy properties are also related to their absolute values, and the errors generated by the eddies with small attribute characteristics are relatively small. Table 1 lists the MAEs of eddy properties at different forecasting times, and a comparison of eddy prediction studies in recent years is presented. Ma et al. [8] found that the polarity of eddies had little impact on the prediction of eddy properties, so we did not distinguish the eddy polarity and took the mean value of the prediction errors of cyclonic and anticyclonic eddies obtained by Ma et al. as the measurement standard of the models in this paper.

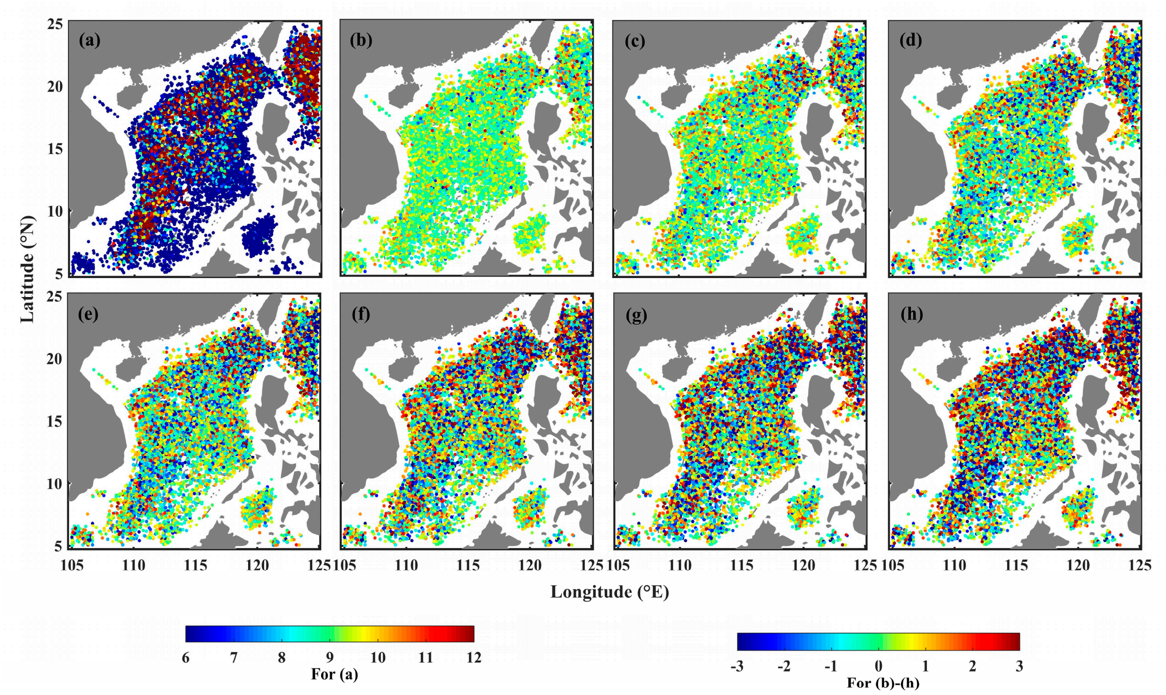


Figure 5. (**a**) Scatterplot of the spatial distribution of the eddy amplitude (cm) in the test set. (**b**–**h**) The forecasting errors (cm) of eddy amplitude from the first day to the seventh day.

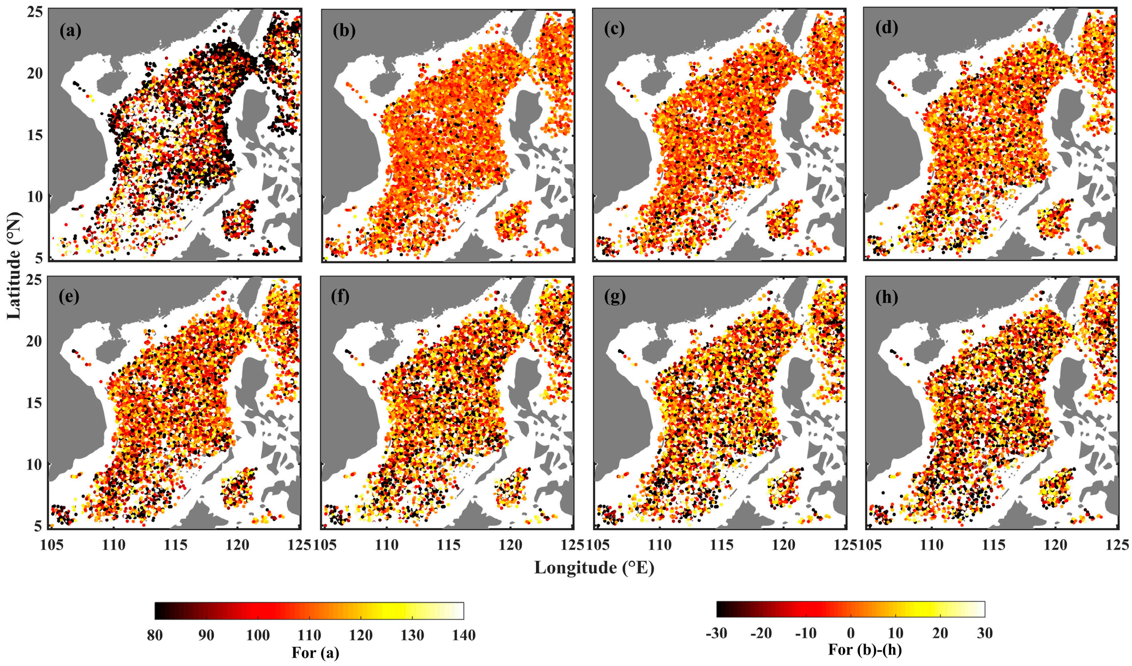


Figure 6. (**a**) Scatterplot of the spatial distribution of the eddy radius (km) in the test set. (**b**–**h**) The forecasting errors (km) of eddy radius from the first day to the seventh day.

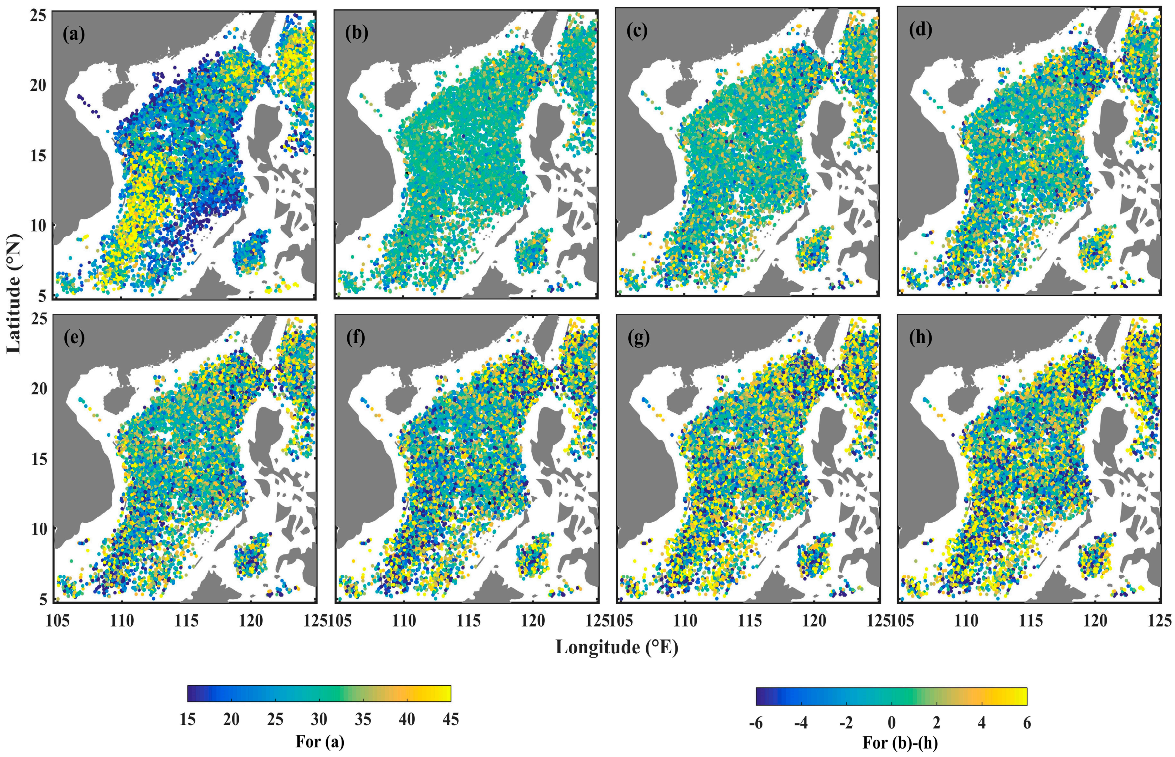


Figure 7. (**a**) Scatterplot of the spatial distribution of the eddy MCAs (cm/s) in the test set. (**b**–**h**) The forecasting errors (cm/s) of eddy MCA speed from the first day to the seventh day**.**

Table 1. Prediction errors of eddy properties in our work and the results in () are from Ma et al. [8].

Forecasting Days	1st	2nd	3rd	4th	5th	6th	7th	14th	21th	28th
Amplitude (cm)	0.6 (0.8)	0.9 (1.2)	1.1 (1.6)	1.4 (1.9)	1.6 (2.2)	1.8 (2.4)	2.0 (2.7)	2.8	3.2	3.5
Radius (km)	10.6 (11.4)	14.3 (15.3)	17.1 (18.3)	19.2 (21.2)	21.1 (23.2)	22.8 (24.9)	23.3 (26.5)	27.2	28.9	29.8
MCA speed (cm/s)	1.3	1.9	2.6	3.1	3.6	4.0	4.5	6.1	7.1	7.7

The results showed that the MAE gradually increased over time. The amplitude of the MAE increased gradually from 0.6 cm on the first day to 2.0 cm on the seventh day. The prediction performance of eddy amplitude was significantly improved compared with that of Ma et al. [8], with the MAE of the radius increasing from 10.6 km to 23.3 km and the MAE of the MCA speed increasing from 1.3 cm/s to 4.5 cm/s. The results reflect the excellent performance of the proposed model, which not only improved the overall MAE values but also resulted in longer prediction windows and different properties. These results highlight the efficiency of the proposed LSTM network in predicting eddy properties and mitigating traditional prediction problems. We reconstructed the corresponding relationship between the samples and labels at different prediction times and trained the model separately. Thus, the results of the model at the current prediction time had no direct correlation with the results at the previous prediction time, indicating that the models were independent and parallel. The approach avoids the accumulation of prediction errors and yields excellent prediction performance.

3.3. Eddy Propagation Trajectory Prediction

ET is a common regression algorithm in machine learning that performs well for various datasets. Compared with traditional regression algorithms, the training speed and prediction accuracy of the proposed method have notable advantages. Figure 8 shows the test results for each model. From the scatter diagram, the models show consistent prediction accuracy at almost all displacements, and the eddies are evenly distributed on both sides of the fitting line. In the prediction in the first week, the RMSEs obtained for zonal displacement and meridional displacement were 31.1 km and 29.4 km, respectively. In the same forecasting time, the meridional displacement and zonal displacement showed a consistent correlation coefficient, which indicated that the model prediction of zonal displacement is equivalent to that of meridional displacement. With an increasing forecasting time, the correlation coefficient of eddies increased from 0.74 to 0.90. The results indicated that the fitting ability and predictions of the model gradually improved. These findings were verified by the motion characteristics of the eddies in the South China Sea. Influenced by the planetary β effect and the seawater flow in the background field, the eddies generally move southwest in the evolution process, and the trend of the movement becomes more obvious as time increases [31,32].

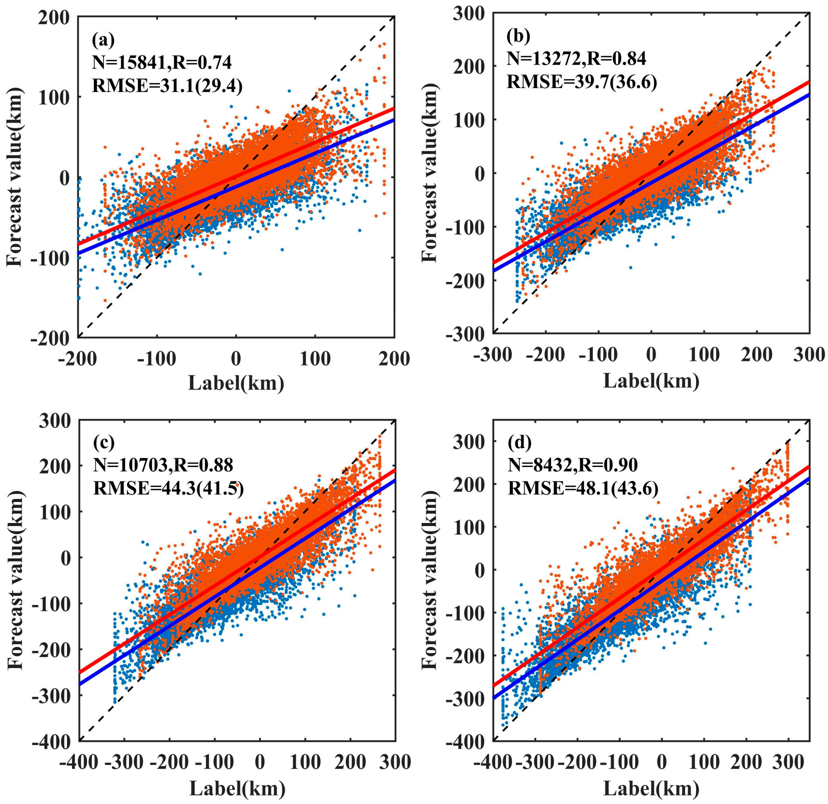


Figure 8. Scatterplots of eddy propagation trajectory prediction. (**a**–**d**) show the results at the forecasting time of 1–4 weeks. The least squares regression lines of zonal displacement (blue line) and meridional displacement (red line) with the correlation coefficient R, the number of test eddies N, and RMSE are shown.

Figure 9a shows the statistical distribution of the zonal and meridional displacements of the eddies during weeks 1–4. The quantity ratio of zonal displacement to the east and west in the first week was 0.38:1, and this value decreased to 0.27:1 in the 4th week. During the period of weeks 1–4, more than 72.5% of eddies shifted to the west in the zonal displacement. The quantity ratio of meridional displacement displayed no obvious trend from 1–4 weeks. As the forecasting time increased, the RMSE of the reverse-normalized evaluation index increased. The zonal displacement RMSE increased from 31.1 km to 48.1 km, and the meridional displacement increased from 29.4 km to 43.6 km. The results suggest that the short-term propagation forecasts of eddies contain more noise than the long-term forecasts and that the long-term motion trend is more significant; this result may be due to the joint action of local oceanic environmental factors and the planetary β effect [23,32–34].

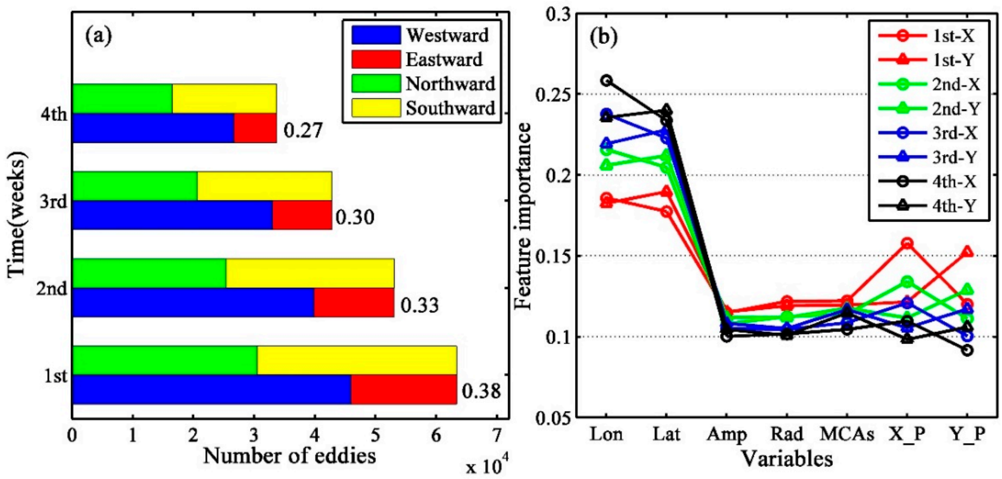


Figure 9. (**a**) Frequency statistics for the displacement direction of eddies from 1–4 weeks; (**b**) comparison of the feature importance of the predictive variables.

The model measures the importance of features according to the number of times the feature is selected in the splitting of each tree; the mean value of all trees is used as the indicator of feature importance. After training, our model gives the importance of each variable. Figure 9b provides a comparison of feature importance for the variables in the 1–4 week forecasting windows, and the relative contribution of each predictor in each prediction period is shown. The larger the feature importance value is, the greater the contribution of this feature to the individual forecasting model. Among the two types of variables, the feature importance of the eddy property variables in the prediction steadily changes, accounting for more than 30% of the total contribution. Overall, the importance of such variables decreases with increasing forecasting time. Similarly, the eddy displacement information (*X_P*, *Y_P*) from the previous week decreases in feature importance as the forecasting time increases. The longitude contributes the most to the prediction of latitudinal displacement, and the latitude contributes the most to the meridional displacement. This result indicates that the latitude and longitude are important characteristic factors in our models, and the longer the prediction time is, the closer the relationship between the positions and eddy displacement. Feature importance analysis reflects the eddy movement in a relatively short time for the propagation of uncertainty to a certain extent. The uncertainty increases the difficulty of producing accurate predictions, and the results verify that the eddy properties have an impact on the prediction of the eddy propagation trajectory. If eddy movement is treated as particle motion considering the two-dimensional property characteristics, the actual motion status of the vortex can be accurately determined, resulting in a high prediction accuracy.

The South China Sea is a semi-enclosed sea affected by the East Asian monsoon and Kuroshio intrusion. In this region, mesoscale eddies have an obvious geographical distribution and multiple characteristics due to variable external forces and complex topography [21–23]. To evaluate the predictive performance of the model, we chose the same region (12–23° N, 108–121° E) as considered in previous studies for further comparison [7]. Table 2 shows the comparison of our ET model with other regression methods, including LSTM, random forest and gradient boosting methods, which are commonly applied to solve different regression problems. The results suggest that our method is superior to the other machine learning methods, as well as the method of Li et al. (2019) [7]. Li et al. constructed a multiple linear regression model based on eddy dynamics. However, the nonlinear motion of eddies and the external non-uniform forces make it difficult to predict the eddy propagation trajectory. Our ET model can effectively capture the nonlinear motion characteristics of eddies through its unique integrated learning mode and effectively learn the patterns of eddy propagation, thereby reducing the prediction error associated with the nonlinear characteristics of eddies.

Table 2. Comparison of prediction performance for various methods in eddy zonal displacement (meridional displacement).

Forecasting Weeks	LSTM	Random Forest	Extra Trees	Gradient Boosting	Multiple Linear Regression [7]
1st	34.8 (28.7)	31.4 (26.7)	28.8 (23.8)	38.5 (33.2)	32.7 (29.5)
2nd	58.1 (46.8)	41.9 (36.8)	36.9 (30.6)	58.8 (49.9)	55.1 (47.3)
3rd	74.5 (53.6)	48.9 (41.6)	41.9 (34.1)	72.9 (60.8)	72.5 (61.4)
4th	86.4 (60.9)	56.3 (47.4)	47.2 (37.2)	85.4 (67.8)	89.2 (73.5)

4. Discussion

In this study, we established the corresponding relationship between historical data and the predicted targets and used machine learning methods to learn the implicit variation patterns of the data. LSTM network prediction and an ET model were used to predict eddy properties and propagation trajectories, respectively. In addition, we performed a comparison experiment to identify the best prediction models for various eddy prediction problems. The results showed that the LSTM network is most suitable for the prediction of eddy properties and that the ET model is ideal for the prediction of the eddy propagation trajectory. To some extent, the dependence of eddy property prediction on multiple historical time series is greater than that on eddy propagation prediction, which may be related to the stability variations of eddy properties and the nonlinear motion characteristics of the eddy propagation trajectory.

We incorporated the two-dimensional property characteristics of eddies into a prediction system for the eddy propagation trajectory and verified that eddy properties have a certain influence on displacement forecasting. Figure 10 shows the distributions of eddy current properties and propagation trajectory prediction performance for different prediction times. In Figure 10a, as the amplitude increases, the prediction error of eddy propagation decreases. With increasing prediction time, large-amplitude eddies tend to be easier to predict than other eddies, but the prediction error of propagation fluctuates greatly for large-amplitude eddies. The impact of the eddy radius on the propagation prediction is reflected in the fact that the prediction error decreases as the radius increases from 40 km to 120 km (Figure 10b). The RMSE values for eddies with radii of 120–180 km are basically unchanged. When the radius is larger than 220 km, the prediction error fluctuates significantly, indicating that the prediction instability of the models increases with an increasing radius. The distribution of prediction errors with eddy MCA speed is similar to that for the eddy radius (Figure 10c): the higher eddy MCA speed are, the better the prediction performance of the model. When the eddy MCA speed increase above 50 cm/s, the fluctuation causes prediction instability in the model. The above results show that the local differences and fluctuations result in different sensitives to eddy properties in propagation trajectory prediction; however, eddy properties on a smaller scale caused greater prediction errors in the model. Notably, the weak signal of eddies was affected by the background field and local environmental disturbances, resulting in an increase in the predicted instability [21,35]. When eddy attributes reach maximum values, the eddies have high kinetic energy, which makes accurate predictions more difficult to obtain. Moreover, machine learning requires a large number of samples, and underfitting can easily occur when the training sample set is insufficient. These eddies appear infrequently in the South China Sea (Figure 4), which makes the model sensitive and causes the prediction quality to decline due to local noise.

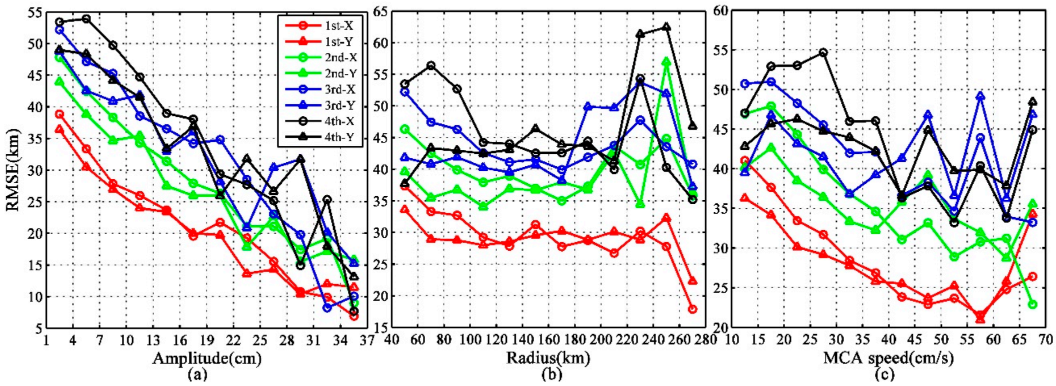


Figure 10. The relationship between eddy properties and the prediction error of the eddy propagation trajectory. The vertical axis represents the RMSE (km), and the horizontal axis represents the eddy amplitude (**a**), radius (**b**), and MCA speed (**c**).

In addition to the above results, we also explore the inter-annual variability and seasonal differences of eddy variables. For the inter-annual variation of eddy properties and trajectories, the average amplitude, radius and MCA speed tend to decrease with the increase in time. On the contrary, the errors of zonal and meridional displacement prediction of eddies tend to increase on the whole, which indicates that the mode of eddy trajectories will become more complex to some extent. This may be related to the variation of various potential factors affecting the propagation trajectories of eddies. In the context of global climate change, whether the inter-annual variation of eddy properties leads to the increase in the prediction error of propagation trajectories needs to be further studied. On the other hand, we made statistics of test eddies according to different seasons, and the predicted results are shown in Table 3. There is no significant difference between the prediction performance of summer and winter eddies in the meridional displacement, but the prediction performance of summer is greater than that of winter in the zonal displacement prediction. To a large extent, the performance of the prediction model depends on the inherent difficulty of trajectory prediction. Different trajectory patterns in summer and winter lead to differences in prediction, and other potential factors, such as the eddy generation mechanism and interaction between eddies and topography in different seasons, may also be affected [7,10]. The prediction performance of the model depends largely on the complexity and nonlinear characteristics of the motion of the eddy itself, and the motion of the eddy depends on the intensity of the motion of the ocean current field. Some studies have pointed out that the eddy movement is related to topography, which also causes the eddies in different basins to have different motion characteristics. In this paper, an end-to-end approach is adopted to directly establish the correlation between the initial characteristics of the eddy and its propagation trajectory. Although more important and complex changes such as large-scale velocity fields may be changed, the displacement characteristics of eddies are the direct result of velocity changes and other influencing factors, which encompass the recent patterns of eddy propagation that are affected by β effects and the mean advection. Since the complexity of eddy motions is the result of multiple factors, the training based on machine learning algorithm can indirectly reflect the correlation of other factors affecting the eddy propagation trajectories, so as to effectively realize the prediction of eddy trajectories.

Table 3. Prediction errors of zonal displacement (meridional displacement) in different seasons.

Forecasting Weeks	Summer	Winter
1st	29.2 (23.8)	30.1 (24.2)
2nd	36.8 (29.8)	40.0 (30.7)
3rd	40.2 (31.9)	45.9 (32.6)
4th	45.1 (33.5)	50.1 (34.6)

5. Conclusions and Prospects

In this study, only the characteristic parameters of the eddy trajectory data were used as the model inputs to comprehensively predict eddy properties and propagation trajectories. Compared with the existing methods, our method achieves a better prediction performance. At different prediction times, we re-established the corresponding relationship between samples and labels and trained the model separately. In this way, the accumulation of prediction errors was avoided to maximize prediction performance. The distinct characteristics of eddy prediction targets were considered to evaluate different machine learning methods. Based on an LSTM network, the prediction of eddy properties achieved good results for multiple historical time series. The amplitude MAE increased gradually from 0.6 cm on the first day to 2.0 cm on the seventh day, with the MAE of the radius increasing from 10.6 km to 23.3 km and the MAE of the MCA speed increasing from 1.3 cm/s to 4.5 cm/s. The prediction of eddy displacement based on the ET models reduced the prediction errors associated with the nonlinear characteristics of eddies. The RMSE between the predicted and actual longitudes (latitudes) throughout the 1–4 week horizon was 28.8–47.2 km (23.8–37.2 km).

The disturbances caused by local marine environmental factors and the planetary β effect resulted in an increase in the predictability of eddy propagation with the extension of the forecasting time. The short-term propagation predictions of eddies contained more noise than long-term predictions, and the long-term predictability reflected a more significant movement trend. The predictability of zonal displacement in the same forecasting time is equivalent to that of meridional displacement, which is the result of the combination of the planetary β effect and self-advection of eddies. Additionally, we verify that the two-dimensional properties of eddies affect the prediction performance of the propagation trajectory to different degrees. With an increasing eddy amplitude, radius and MCA speed, the prediction performance of the models improved, and the eddies characterized by maximum properties make the prediction unstable. In the future, we can combine sea surface variables and incorporate more features to improve eddy predictions.

Author Contributions: X.W. and W.W. conceived and designed the experiments; X.W. and H.W. performed the experiments and wrote the paper; W.W. and D.L. helped with field work, discussion, and revisions. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Foundation of China (61972411).

Acknowledgments: All the data used in this study are available as follows: The global mesoscale oceanic eddy trajectory dataset is from the Archiving, Validation and Interpretation of Satellite Oceanographic data (AVISO) (https://www.aviso.altimetry.fr/en/data/data-access).

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

References

1. Dong, C.; McWilliams, J.C.; Liu, Y.; Chen, D. Global heat and salt transports by eddy movement. *Nat. Commun.* **2014**, *5*, 3294. [CrossRef] [PubMed]
2. Zhang, Z.; Wang, W.; Qiu, B. Oceanic mass transport by mesoscale eddies. *Science* **2014**, *345*, 322–324. [CrossRef] [PubMed]
3. Li, J.; Wang, G.; Zhai, X. Observed cold filaments associated with mesoscale eddies in the South China Sea. *J. Geophys. Res. Oceans* **2017**, *122*, 762–770. [CrossRef]
4. Wang, Q.; Zhang, S.; Xie, S.-P.; Norris, J.R.; Sun, J.-X.; Jiang, Y.-X. Observed Variations of the Atmospheric Boundary Layer and Stratocumulus over a Warm Eddy in the Kuroshio Extension. *Mon. Weather Rev.* **2019**, *147*, 1581–1591. [CrossRef]
5. Gaube, P.; Barceló, C.; McGillicuddy, D., Jr.; Domingo, A.; Miller, P.; Giffoni, B.; Marcovaldi, N.; Swimmer, Y. The use of mesoscale eddies by juvenile loggerhead sea turtles (Caretta caretta) in the southwestern Atlantic. *PLoS ONE* **2017**, *12*, e0172839. [CrossRef] [PubMed]
6. Gaube, P.; Chelton, D.B.; Samelson, R.M.; Schlax, M.G.; O'Neill, L.W. Satellite Observations of Mesoscale Eddy-Induced Ekman Pumping. *J. Phys. Oceanogr.* **2015**, *45*, 104–132. [CrossRef]

7. Li, J.; Wang, G.; Xue, H.; Wang, H. A simple predictive model for the eddy propagation trajectory in the northern South China Sea. *Ocean Sci.* **2019**, *15*, 401–412. [CrossRef]
8. Ma, C.; Li, S.; Wang, A.; Yang, J.; Chen, G. Altimeter Observation-Based Eddy Nowcasting Using an Improved Conv-LSTM Network. *Remote Sens.* **2019**, *11*, 783. [CrossRef]
9. Robinson, A.R.; Carton, J.A.; Mooers, C.N.K.; Walstad, L.J.; Carter, E.F.; Rienecker, M.M.; Smith, J.A.; Leslie, W.G. A real-time dynamical forecast of ocean synoptic/mesoscale eddies. *Nature* **1984**, *309*, 781–783. [CrossRef]
10. Masina, S.; Pinardi, N. Mesoscale data assimilation studies in the middle adriatic sea. *Cont. Shelf Res.* **1994**, *14*, 1293–1310. [CrossRef]
11. Shriver, J.F.; Hurlburt, H.; Smedstad, O.; Wallcraft, A.; Rhodes, R. 1/32° real-time global ocean prediction and value-added over 1/16° resolution. *J. Mar. Syst.* **2007**, *65*, 3–26. [CrossRef]
12. Zhang, X.; Wang, T.; Chen, G.; Tan, X.; Zhu, K. Convective Clouds Extraction from Himawari–8 Satellite Images Based on Double-Stream Fully Convolutional Networks. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 553–557. [CrossRef]
13. Lukoševičius, M.; Jaeger, H. Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.* **2009**, *3*, 127–149. [CrossRef]
14. Karpathy, A.; Li, F.-F. Deep Visual-Semantic Alignments for Generating Image Descriptions. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 664–676. [CrossRef] [PubMed]
15. Duo, Z.; Wang, W.; Wang, H. Oceanic Mesoscale Eddy Detection Method Based on Deep Learning. *Remote Sens.* **2019**, *11*, 1921. [CrossRef]
16. Sun, M.; Tian, F.; Liu, Y.; Chen, G. An Improved Automatic Algorithm for Global Eddy Tracking Using Satellite Altimeter Data. *Remote Sens.* **2017**, *9*, 206. [CrossRef]
17. Chelton, D.B.; Schlax, M.G.; Samelson, R.M.; De Szoeke, R.A. Global observations of large oceanic eddies. *Geophys. Res. Lett.* **2007**, *34*. [CrossRef]
18. Chelton, D.B.; Schlax, M.G.; Samelson, R.M. Global observations of nonlinear mesoscale eddies. *Prog. Oceanogr.* **2011**, *91*, 167–216. [CrossRef]
19. Ashkezari, M.D.; Hill, C.N.; Follett, C.N.; Forget, G.; Follows, M.J. Oceanic eddy detection and lifetime forecast using machine learning methods. *Geophys. Res. Lett.* **2016**, *43*, 12234–12241. [CrossRef]
20. Mason, E.; Pascual, A.; McWilliams, J.C. A New Sea Surface Height–Based Code for Oceanic Mesoscale Eddy Tracking. *J. Atmos. Ocean. Technol.* **2014**, *31*, 1181–1188. [CrossRef]
21. Chen, G.; Hou, Y.; Chu, X. Mesoscale eddies in the South China Sea: Mean properties, spatiotemporal variability, and impact on thermohaline structure. *J. Geophys. Res. Space Phys.* **2011**, *116*, 116. [CrossRef]
22. Nan, F.; Xue, H.; Yu, F. Kuroshio intrusion into the South China Sea: A review. *Prog. Oceanogr.* **2015**, *137*, 314–333. [CrossRef]
23. Du, Y.; Wu, D.; Liang, F.; Yi, J.; Mo, Y.; He, Z.; Pei, T. Major migration corridors of mesoscale ocean eddies in the South China Sea from 1992 to 2012. *J. Mar. Syst.* **2016**, *158*, 173–181. [CrossRef]
24. Zhang, K.; Chao, W.-L.; Sha, F.; Grauman, K. Video Summarization with Long Short-Term Memory. In Proceedings of the 14th European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 11–14 October 2016; pp. 766–782.
25. Wang, D.; Su, J.; Yu, H. Feature Extraction and Analysis of Natural Language Processing for Deep Learning English Language. *IEEE Access* **2020**, *8*, 46335–46345. [CrossRef]
26. Donahue, J.; Hendricks, L.A.; Rohrbach, M.; Venugopalan, S.; Guadarrama, S.; Saenko, K.; Darrell, T. Long-Term Recurrent Convolutional Networks for Visual Recognition and Description. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 677–691. [CrossRef]
27. Galar, M.; Fernandez, A.; Barrenechea, E.; Bustince, H.; Herrera, F. A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches. *IEEE Trans. Syst. Man, Cybern. Part C* **2012**, *42*, 463–484. [CrossRef]
28. Sagi, O.; Rokach, L. Ensemble learning: A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1249. [CrossRef]
29. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
30. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **2006**, *63*, 3–42. [CrossRef]
31. Cai, S.; Long, X.; Wu, R.; Wang, S. Geographical and monthly variability of the first baroclinic Rossby radius of deformation in the South China Sea. *J. Mar. Syst.* **2008**, *74*, 711–720. [CrossRef]

32. Pratt, L.J. Introduction to geophysical fluid dynamics (Book Review). *Am. Sci.* **1996**, *84*, 80–81.
33. Morrow, R.; Birol, F.; Griffin, D.; Sudre, J. Divergent pathways of cyclonic and anti-cyclonic ocean eddies. *Geophys. Res. Lett.* **2004**, *31*. [CrossRef]
34. Li, J.; Zhang, R.; Jin, B. Eddy characteristics in the northern South China Sea as inferred from Lagrangian drifter data. *Ocean Sci.* **2011**, *7*, 661–669. [CrossRef]
35. Liu, C.; Du, Y.; Zhuang, W.; Xia, H.; Xie, Q. Evolution and propagation of a mesoscale eddy in the northern South China Sea during winter. *Acta Oceanol. Sin.* **2013**, *32*, 1–7. [CrossRef]