

JGR Oceans

RESEARCH ARTICLE

10.1029/2021JC017515

Special Section:

Significant advances in ocean and climate sciences of the Pacific-Asian Marginal Seas

Key Points:

- A hybrid prediction model combining a statistical method and a deep learning model is proposed for forecasting sea surface multivariate
- Multivariate empirical orthogonal analysis is used to consider the correlations among different variables and reduce the computation
- This model performs satisfactorily under both normal and extreme conditions

Correspondence to:

W. Li and G. Han,
livei1978@tju.edu.cn;
guijun_han@tju.edu.cn

Citation:

Shao, Q., Li, W., Han, G., Hou, G., Liu, S., Gong, Y., & Qu, P. (2021). A deep learning model for forecasting sea surface height anomalies and temperatures in the South China Sea. *Journal of Geophysical Research: Oceans*, 126, e2021JC017515. <https://doi.org/10.1029/2021JC017515>

Received 26 APR 2021
Accepted 21 JUN 2021

A Deep Learning Model for Forecasting Sea Surface Height Anomalies and Temperatures in the South China Sea

Qi Shao^{1,2} , Wei Li^{1,2} , Guijun Han¹ , Guangchao Hou¹ , Siyuan Liu¹, Yantian Gong¹, and Ping Qu^{2,3}

¹School of Marine Science and Technology, Tianjin University, Tianjin, China, ²Tianjin Key Laboratory for Oceanic Meteorology, Tianjin, China, ³Tianjin Institute of Meteorological Sciences, Tianjin, China

Abstract The field of forecasting oceanic variables has traditionally relied on numerical models, which effectively consider the ocean's dynamic evolution and are of physical importance. However, to make the models more realistic, complicated processes need to be considered, which is difficult because their calculations are complex. In fact, information on the internal dynamic mechanisms and external driving forces of the ocean are already embedded in the time series of observations. Therefore, we can determine the patterns of ocean variations through data mining of these series to achieve forecasting. Furthermore, to predict variations in ocean processes more realistically, interactions between variables and spatial correlations should be effectively considered. Thus, inspired by available remote sensing data and advancements in deep learning technologies, we develop a hybrid model based on a statistical method and a deep learning model to predict multiple sea surface variables. A case study in the South China Sea shows that this model is highly promising for short-term daily forecasts of the sea surface height anomaly (SSHA) and sea surface temperature (SST). When the forecast time is 10 days, the root mean square errors of this model forecasts for SSHA and SST are approximately 0.0276 m and 0.46°C, respectively, which are much smaller than those of persistence, climatology and linear regression predictions. The anomaly correlation coefficients for SSHA and SST are approximately 0.864 and 0.633, respectively. The model performs satisfactorily under both normal and typhoon weather conditions.

Plain Language Summary Deep learning techniques of neural networks are widely used for the forecasting of ocean variables due to their good prediction performance. However, most of the models are based on a single point and a single variable, without considering the interactions between different variables, so they lack physical significance to some extent. In this study, a prediction model for multiple oceanic variables combining multivariate empirical orthogonal function (MEOF) analysis and a Conv1D-LSTM neural network is established in the South China Sea, which can effectively solve these problems. The MEOF analysis in this study has four main functions: (a) establishing the spatial correlation between different discrete points; (b) considering the correlation between different variables; (c) reducing calculations; and (d) decorrelation. To serve as the basis, the principal component series are used to train and verify the Conv1D-LSTM model. This model performs satisfactorily under both normal and extreme conditions, and it is expected to provide a reference for further research on deep learning-based methods in the field of marine prediction.

1. Introduction

The South China Sea (SCS), featuring complex terrain, is the largest marginal sea in Southeast Asia (Shaw, 1991) and connected with the East China Sea, Pacific Ocean, Sulu Sea, Java Sea, and Indian Ocean through the Taiwan Strait, Luzon Strait, Palawan and Balabac Straits, Gasper and Karimata Straits, and Strait of Malacca, respectively (Yang & Liu, 2002). As a semi-enclosed marginal ocean basin, it has complex ocean conditions because of the presence of many circulation patterns and mesoscale eddies. Additionally, it has obvious interactions with the surrounding ocean and meteorology (Hu et al., 2014). On the one hand, the SCS has abundant water exchange channels, and the water and heat exchange between the SCS and the Pacific Ocean makes it vulnerable to the Pacific Ocean's large-scale circulation. On the other hand, the Kuroshio Current flows along the eastern coast of Luzon and continues northward east of Taiwan after

making a slight excursion into the northern SCS via the Luzon Strait (D. Wang et al., 2010). Embedded in the SCS circulation, there is much variability, especially mesoscale signals (Wang & Koblinsky, 2000), which play a crucial role in the exchange of oceanic materials and energies. In view of this, prediction of the future ocean state of the SCS is still a great challenge.

Ocean dynamic models have played a dominant role in ocean forecasting for decades. They make use of a series of complex physics equations to describe different physical rules and advanced data assimilation schemes to predict ocean variables (W. Li et al., 2010; Li, Vidard, et al., 2019; Li, Wang, et al., 2019; Wu, 2016). Such methods make physical sense, and great successes have been achieved (Krishnamurti et al., 2006; W. Li et al., 2008; Oey et al., 2005; Sun et al., 2019; Vidard et al., 2006). However, to make the models more realistic, complicated processes need to be considered, which requires complex calculations. Other traditional forecasts are based on statistical techniques, learning relationships from historical observations and further using the learned relationships to infer future states. Many successes have been achieved, such as pattern searching (Agarwal et al., 2001), empirical orthogonal function (EOF) analysis (Hannachi et al., 2007), linear regression (Kug et al., 2004; Tang et al., 2000), canonical correlation analysis (Collins et al., 2004), Markov models (Xue et al., 2000; Xue & Leetmaa, 2000) and forced red noise processes (Jansen et al., 2008). However, general statistical methods are not very effective for nonlinear problems, and due to obvious physical defects, their prediction accuracy is generally not better than that of numerical models.

As one of the most popular and influential techniques, artificial neural networks (ANNs) are widely used due to their flexibility and strong ability to model complex patterns hidden in data (Yue et al., 2017; Zang et al., 2019). This technique allows deep learning models to be substantially more complex than traditional statistical models, and it is often referred to as an alternative to physics-based models (Thakur et al., 2018). Recently, considerable attention has been given to using site-specific and variable-independent ANN models to forecast ocean variables (Patil & Deo, 2017, 2018; Patil et al., 2016; Xiao et al., 2019). Site-specific models use separate ANN models to predict variables at different sites, which consider site differences but have high computational costs. In addition, they require sufficient training data at every grid point (Patil & Deo, 2018). Furthermore, the variable-independent prediction models treat each ocean variable individually, ignore the correlations between different variables and lack physical meaning (Fu et al., 2019; Patil & Deo, 2017; Y. Yang et al., 2018); however, dynamic and thermal relationships exist among different variables in the actual marine environment. Therefore, it is necessary for us to find a data-driven method with physical significance. In this way, new deep learning technologies can be used to improve forecasting skills while conforming to realistic patterns of variations in real marine environments.

In this study, we develop a hybrid model to predict multiple sea surface variables in the SCS by combining a traditional statistical method of EOF analysis and a deep learning model. It is hypothesized that short-term prediction of sea surface variables can be viewed as a sequence recognition and regression task, aiming at identifying the characteristics of the dynamic evolution of sea surface variables (Zeng et al., 2015). To conform to the real marine environment, we consider the dynamic balance of different variables in the sea surface and construct the correlation coefficient matrix among them. Then, the correlation coefficient matrix is decomposed by EOF analysis. This approach transforms the dynamic evolutionary process into a time series evolution problem. Here, EOF analysis not only establishes the correlation between different variables but also carries out computational reduction to some extent. More importantly, EOF analysis plays a role in the decorrelation of different scale processes here. Then, a deep learning technique of the Conv1D-LSTM neural network is used to train and verify the principal components (PCs). Finally, every predicted PC is combined to reconstruct the prediction fields of the sea surface multivariate.

The remainder of the paper is organized as follows: Section 2 describes the study area and data set used and introduces the theoretical framework for the prediction model in this study. Detailed methodologies and performance evaluation criteria are presented in Section 3. In Section 4, we present the proposed model for sea surface height anomaly (SSHA) and sea surface temperature (SST) predictions and compare the prediction performance with other models. Finally, we draw conclusions and discuss future research suggestions in Section 5.

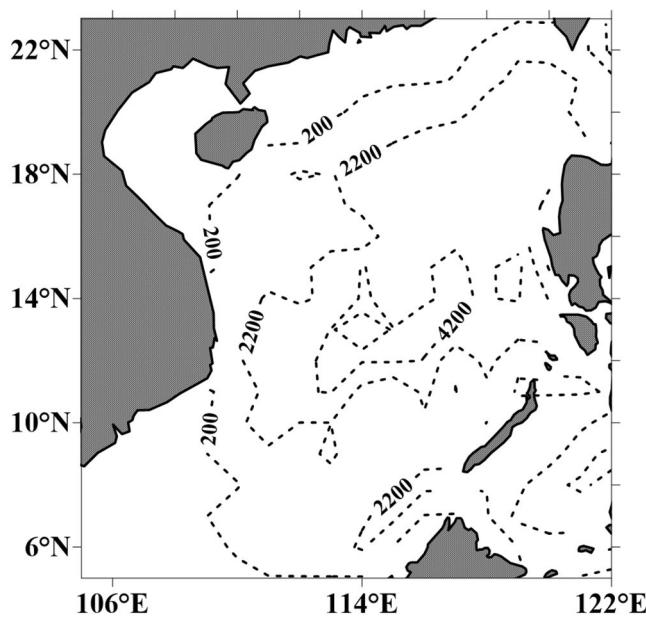


Figure 1. Study area: the South China Sea (5° – 23° N, 105° – 122° E) with the Sulu Sea excluded.

2. Study Area, Data and Proposed Model

2.1. Study Area

In this work, the SCS region, with a spatial range of 5° – 23° N and 105° – 122° E, as shown in Figure 1, is the spatial domain of concern. Since we hope that the spatial orthogonal modes obtained by EOF analysis are optimal spectral decomposition results (Chu et al., 2003), the Sulu Sea is excluded in this study because it may degrade the analysis results for the SCS region as a whole. In addition, over the shelf seas, the data still contain aliases from fast processes, such as tides and internal waves (Qian et al., 2018; Yuan et al., 2006), and may also be contaminated by land echo noise. Thus, the shelf data from depths of less than 200 m are masked out in this study.

2.2. Data

Considering the sea surface dynamics and the availability and quality of satellite remote sensing data, SSHA and SST data are used in this study.

The SSHA data used in this study are a $1/4^{\circ}$ daily product. This data set is produced by the Copernicus Marine and Environment Monitoring Service (CMEMS) and is available at <http://marine.copernicus.eu>. It is mainly constructed by fusing the information obtained from multiple altimeter satellites (ESR-1/2, Topex/Poseidon, ENVISAT and Jason-1/Jason-2), which cover the global ocean and provide the daily SSHA values from January 1, 1993 to present.

The SST data used in this study are the National Oceanic and Atmospheric Administration (NOAA) $1/4^{\circ}$ daily Optimum Interpolation Sea Surface Temperature (daily OISST, version 2) data, which are available at <https://www.ncdc.noaa.gov/oisst>. It is a reanalysis data set constructed by combining observations from different platforms, including satellites, ships and buoys, on a regular global grid. Optimum interpolation is applied to fill in the gaps and make it a spatially complete SST data set. To obtain a continuous long time series of SSTs, we choose the AVHRR-Only SSTs, which use satellite SSTs from only the AVHRR (Reynolds & Chelton, 2010), cover the global ocean and provide the daily SST values from September 1, 1981 to present.

Given the availability of the above satellite data sets, the training and testing data sets are restricted to a period of 26 years, from January 1, 1993 to December 31, 2018. Part of the data set ranging from 1993 to 2013 is used to train the hybrid prediction model, and the other part ranging from 2014 to 2018 serves as the testing data set. The spatial resolution is $0.25^{\circ} \times 0.25^{\circ}$, and the temporal resolution is 1 day.

2.3. Proposed Model

In this study, a novel prediction scheme based on EOF analysis and a deep learning method is proposed for predicting the daily SSHA and SST in the SCS. This proposed model consists of four parts: data preprocessing, Conv1D-LSTM prediction, reconstruction and correction. The structure of each component in the proposed model is demonstrated in Figure 2, and the function of each component is defined as follows:

1. First, after obtaining the original SSHA and SST satellite data, the original sample data set is divided into a training set and a testing set. Then, multivariate empirical orthogonal function (MEOF) analysis, equal to data preprocessing, is employed to decompose the training set into orthogonal spatial modes (EOFs) and corresponding PCs; the PCs of the testing set is obtained by projecting the testing set to the EOFs obtained above. The details of data preprocessing are described in Section 3.1.
2. Second, a deep learning model of the Conv1D-LSTM network is used to predict the PCs with a certain variance proportion. It is trained by the training set and verified by the testing set. Detailed information about the Conv1D-LSTM network is explained in Section 3.2.

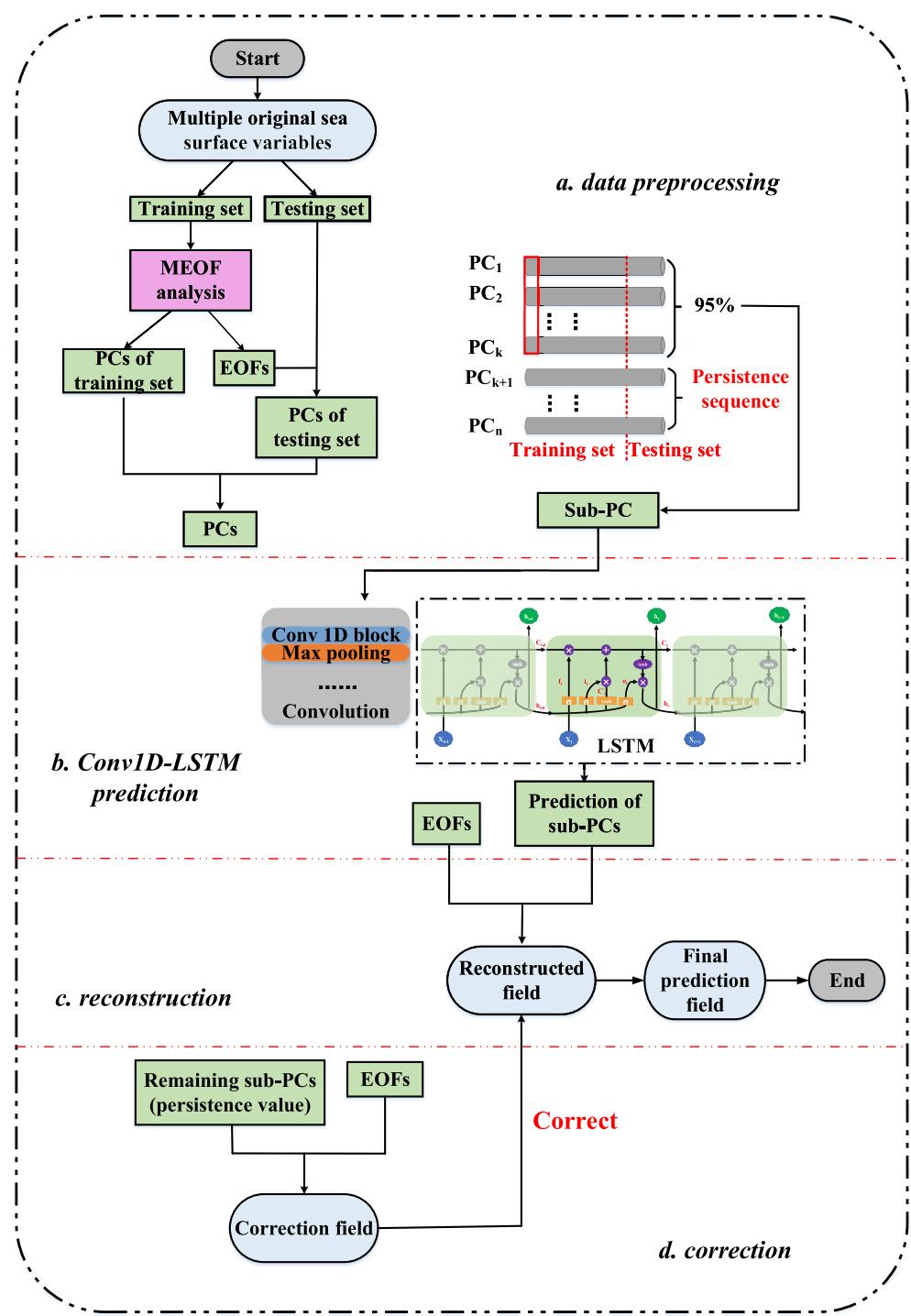


Figure 2. Framework of the proposed prediction model.

3. Third, the reconstructed field can be obtained by combining the predicted PCs obtained in step (2) with the EOFs obtained in step (1).
4. Finally, to avoid stepped forecasting, the residual between the true (or observational) field and the reconstructed field obtained in step (3) at the starting time, which is also represented by unused higher-order modes, is provided by persistence prediction to serve as a correction for the reconstructed field.

3. Related Methodology

3.1. Multivariate Empirical Orthogonal Function Analysis

As a statistical approach, EOF analysis was first introduced to meteorology and climate research in the 1950s (Lorenz, 1956). It is powerful for analyzing spatial and temporal variability by rapidly decomposing and condensing large amounts of data, as well as restoring the basic structure of data fields. In this study, MEOF analysis is used to reduce the computational effort of the data set, extract the major spatial features, establish the correlation between different variables, and consider the spatial correlation. Moreover, due to its orthogonality, MEOF analysis plays a role in the decorrelation of different scale processes to some extent so that predictions can be achieved with completely independent PCs.

The sea surface multivariate matrix of the training set can be expressed as:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{1,1}^{\text{SSHA}} & \dots & \mathbf{X}_{1,m}^{\text{SSHA}} & \dots & \mathbf{X}_{1,M}^{\text{SSHA}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{X}_{N,1}^{\text{SSHA}} & \dots & \mathbf{X}_{N,m}^{\text{SSHA}} & \dots & \mathbf{X}_{N,M}^{\text{SSHA}} \\ \mathbf{X}_{1,1}^{\text{SST}} & \dots & \mathbf{X}_{1,m}^{\text{SST}} & \dots & \mathbf{X}_{1,M}^{\text{SST}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{X}_{N,1}^{\text{SST}} & \dots & \mathbf{X}_{N,m}^{\text{SST}} & \dots & \mathbf{X}_{N,M}^{\text{SST}} \end{bmatrix} \quad (1)$$

where the dimensions of this sample matrix are $2N \times M$ (with N representing the spatial dimension of each variable and M representing the temporal dimension).

Since different variables involve different scales and units, more attention is paid to the normalized ocean surface variables, and the sample matrix is further expressed as:

$$\mathbf{X}' \equiv \boldsymbol{\sigma}^{-1}(\mathbf{X} - \bar{\mathbf{X}}) \quad (2)$$

where $\boldsymbol{\sigma}$ is a standardization deviation matrix, and $\bar{\mathbf{X}}$ is the climatology.

To decompose the normalized sample matrix (Equation 2) by orthogonal mode vectors, we need to construct a correlation coefficient matrix:

$$\text{Cor}(\mathbf{X}') \equiv \frac{1}{M} \mathbf{X}' \mathbf{X}'^T \quad (3)$$

This correlation coefficient matrix effectively considers not only the spatial correlation of the same sea surface variable at different spatial points but also the correlation of different sea surface variables. This approach is consistent with the dynamic processes of real marine environmental variations.

Then, according to the expansion of \mathbf{X}' , Equation 3 can be expressed as follows:

$$\text{Cor}(\mathbf{X}') \equiv \frac{1}{M} \mathbf{X}' \mathbf{X}'^T = \frac{1}{M} (\mathbf{V}(\mathbf{PC})) (\mathbf{V}(\mathbf{PC}))^T \quad (4)$$

where \mathbf{V} represents the EOFs, and its dimensions are $2N \times 2N$; and \mathbf{PC} represents the corresponding principal components, and its dimensions are $2N \times M$.

Finally, Jacobi decomposition is used to obtain the eigenvectors and eigenvalues of $\text{Cor}(\mathbf{X}')$:

$$\text{Cor}(\mathbf{X}') = \mathbf{V} \bar{\Lambda} \mathbf{V}^T \quad (5)$$

Equivalently, the eigenvectors and eigenvalues of $\mathbf{X}' \mathbf{X}'^T$ can be obtained as follows:

$$\mathbf{X}' \mathbf{X}'^T = \mathbf{V} \Lambda \mathbf{V}^T \quad (6)$$

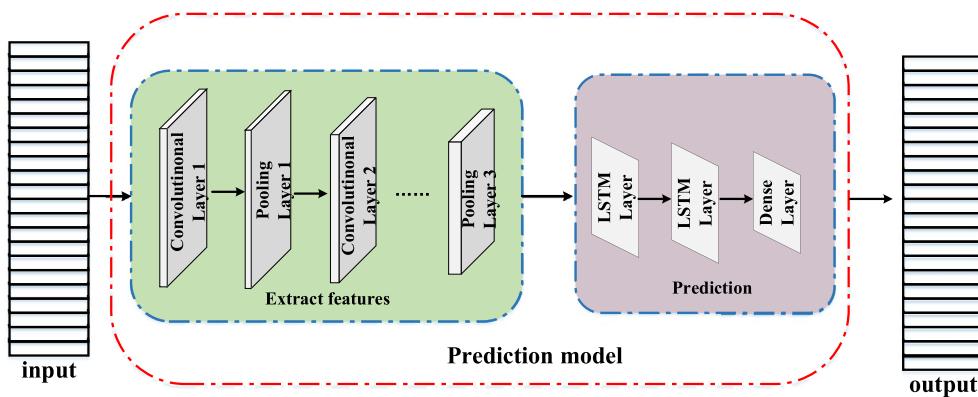


Figure 3. Structure of the Conv1D-LSTM network.

where $\Lambda = M\bar{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_N)$. Therefore, the original sample of X can be expressed as:

$$X = \bar{X} + \sigma X' = \bar{X} + \sigma V(PC) \quad (7)$$

In this study, the training data set from 1993 to 2013 is selected for MEOF analysis, and the corresponding EOFs and PCs are obtained. Then, the PCs of the testing set can be obtained by projecting the testing data onto the EOFs obtained before; therefore, the PCs of the testing set are independent of those of the training set and can serve as independent verification. Obviously, this treatment is much closer to the operational ocean forecast scenario.

Generally, to reduce the computational complexity and dimensionality, we need to use only the first few EOFs so that the total variance contribution is large enough to reconstruct the original data with the main characteristics of the spatial structure. Although it may cause errors during the data reconstruction process, such a choice can significantly reduce the computation time and greatly improve the computational efficiency. Here, we use 118 EOFs and corresponding PCs, accounting for 95% of the total variance, all of which pass the significance test (North, 1983; North et al., 1982). More importantly, each PC represents the temporal evolution of its corresponding spatial mode, and every EOF (PC) is independent of every other EOF (PC) due to the decorrelation of the MEOF analysis, so it is reasonable to predict each PC individually. Then, these PCs are used to train and validate the deep learning model, and the predicted PCs are combined with the EOFs to obtain the reconstructed field. Furthermore, to avoid a stepped forecast, the difference between the true field and the reconstructed field at the starting time, which is the remaining 5% residual, is used to make corrections in the way of persistence at every forecast time.

3.2. Conv1D-LSTM Model

One problem that deep learning systems seek to solve is the preservation of previous information for future prediction tasks (J. L. Wang et al., 2019). In this study, the Conv1D-LSTM model, which combines a convolutional neural network (CNN) and a long short-term memory (LSTM) network, serves as a time series prediction model (Shi et al., 2015). A 1-D convolutional neural network (Conv1D block) is introduced to further extract the different characteristics of PCs in the time dimension, and then, the LSTM layers are used to fuse the characteristics and make predictions. The specific structure of the Conv1D-LSTM model, obtained by adding convolution operations in the LSTM, is shown in Figure 3.

3.2.1. Convolutional Neural Network

CNN is a feedforward neural network that successfully reduces the risk of overfitting and has achieved great achievements in feature extraction and generalization (L. Yang et al., 1998). Compared to traditional neural networks, it introduces convolutional layers and pooling layers. The convolutional layer is used to extract global features, and the pooling layer is used to calculate sufficient local statistics of features to reduce the

number of overall features, prevent overfitting, and reduce the amount of calculation. The convolutional layer is the core of the CNN, and the output of the convolutional layer can be formalized as follows:

$$h^k = f((W^k * x) + b^k) \quad (8)$$

where $*$ represents the convolutional operation, $f(\cdot)$ represents the activation function, and W^k and b^k are the weights and biases of the k th feature map, respectively.

3.2.2. Long Short-Term Memory Neural Network

LSTM is a variation of the recursive neural network (RNN) and is a solution to the vanishing and exploding gradient problem of RNN. It is capable of processing one-dimensional time series and learning long-term dependencies (Hochreiter & Schmidhuber, 1997). As shown in Figure 2, LSTM has a chain-like structure. The basic building block of LSTM is a cell whose state is the key to LSTM. There are four important variables of LSTM: the internal memory state C_t , forget gate f_t , input gate i_t and output gate o_t . The calculation process of each variable at time t is displayed as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (9)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (10)$$

$$C_t = f_t * C_{t-1} + i_t * \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (11)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (12)$$

$$h_t = o_t * \tanh(C_t) \quad (13)$$

where σ is the sigmoid function; W_f , W_i , W_C , and W_o are the weights applied to the concentration of the new input x_t and the output h_{t-1} from the previous cell; and b_f , b_i , b_C , and b_o are the corresponding biases.

The Conv1D-LSTM model in this study consists of two LSTM layers, two dense layers, three convolution layers and three max pooling layers. After many experiments, the activation function of each convolution layer is determined to be leaky rectified linear units (leaky ReLUs), while the activation of the LSTM layer is determined to be a tanh function. In addition, dropout regularization is added to prevent overfitting, and adaptive moment estimation is chosen as the optimization function.

3.3. Performance Evaluation Criteria

To evaluate the performance of the proposed prediction model and to prove its own superiority by comparison with other models, three statistical metrics are used. The mean absolute error (MAE), root mean square error (RMSE) and anomaly correlation coefficient (ACC) are utilized to verify the prediction accuracy of the proposed model in this study. The meaning and specific calculation methods are as follows:

$$MAE_i = \sum_{j=1}^{MM} (|T_{ij} - P_{ij}|) / MM \quad (14)$$

$$RMSE_i = \sqrt{\frac{\sum_{j=1}^{MM} (T_{ij} - P_{ij})^2}{MM}} \quad (15)$$

$$ACC = \left\{ \sum_{j=1}^{MM} \left[\frac{\sum_{i=1}^N (T_{ij} - \bar{T}_j)(P_{ij} - \bar{P}_j)}{\sqrt{\sum_{i=1}^N (T_{ij} - \bar{T}_j)^2 \sum_{i=1}^N (P_{ij} - \bar{P}_j)^2}} \right] \right\} / MM, (\bar{T}_j = (\sum_{i=1}^N T_{ij}) / N, \bar{P}_j = (\sum_{i=1}^N P_{ij}) / N,) \quad (16)$$

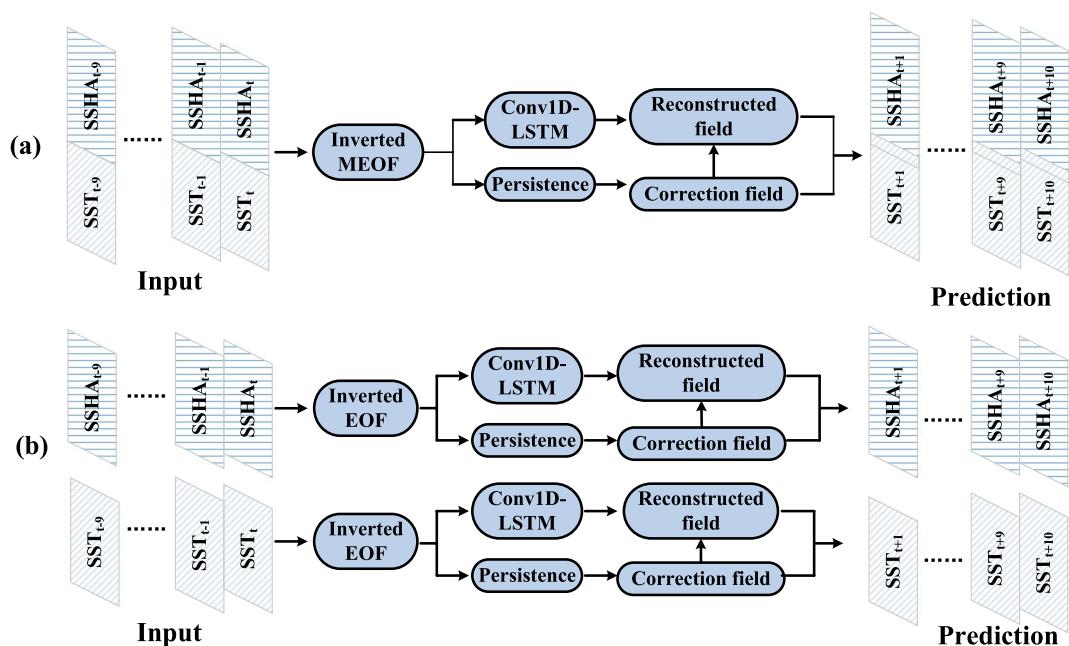


Figure 4. Flowchart of the prediction model: (a) the model with multivariate empirical orthogonal function (MEOF) analysis and (b) the model with empirical orthogonal function (EOF) analysis.

where MAE_i and $RMSE_i$ represent the MAE and RMSE values of the i th grid point, respectively; ACC represents the spatial anomaly correlation coefficient; T_{ij} is the true value of the i th grid point in the j th sample; P_{ij} is the predicted value of the i th grid point in the j th sample; N is the number of grid points; and MM is the number of testing samples.

4. Experiments and Results

4.1. Comparisons Between EOF Analysis and MEOF Analysis Results

Figure 4a shows the prediction model with MEOF analysis; hereafter, we call it MEOF-Conv1D-LSTM. In this model, the dynamic balance of SSHA and SST is considered, and the correlation coefficient matrix between them is used for EOF decomposition. As a comparison, Figure 4b shows the prediction model with EOF analysis, referred to as EOF-Conv1D-LSTM, which does not consider the correlation between different variables and treats each variable as independent. Both the input and output are 10 days.

To compare the differences between MEOF analysis and EOF analysis on the sea surface variables, the spatial distributions of EOFs for the training set obtained by these two decomposition methods are compared first. Figure 5 shows the spatial distributions and corresponding proportions of the first three modes of SSHA and SST. In the decomposition of MEOF analysis, the spatial modes of SSHA and SST obtained by this method have a certain correlation because the correlation between different variables is taken into account, which exactly reflects the baroclinic structure of the SCS.

Then, Figure 6 displays the histograms of SSHA differences and SST differences between forecasts and true values (observations). Compared with the prediction model with EOF analysis, the statistics show that the model with MEOF analysis has a lower bias and higher proportions of SSHA differences within ± 0.05 m and SST differences within $\pm 0.5^{\circ}\text{C}$. Especially for the statistical results of SSHA, the errors of the prediction model with MEOF analysis are always lower than those of the model with EOF analysis during the whole prediction window. When the lead time is 10 days, the significant difference between the MEOF analysis and EOF analysis is the largest. The bias of the model with MEOF analysis is 0.0077 m, whereas that of the model with EOF analysis is 0.0092 m.

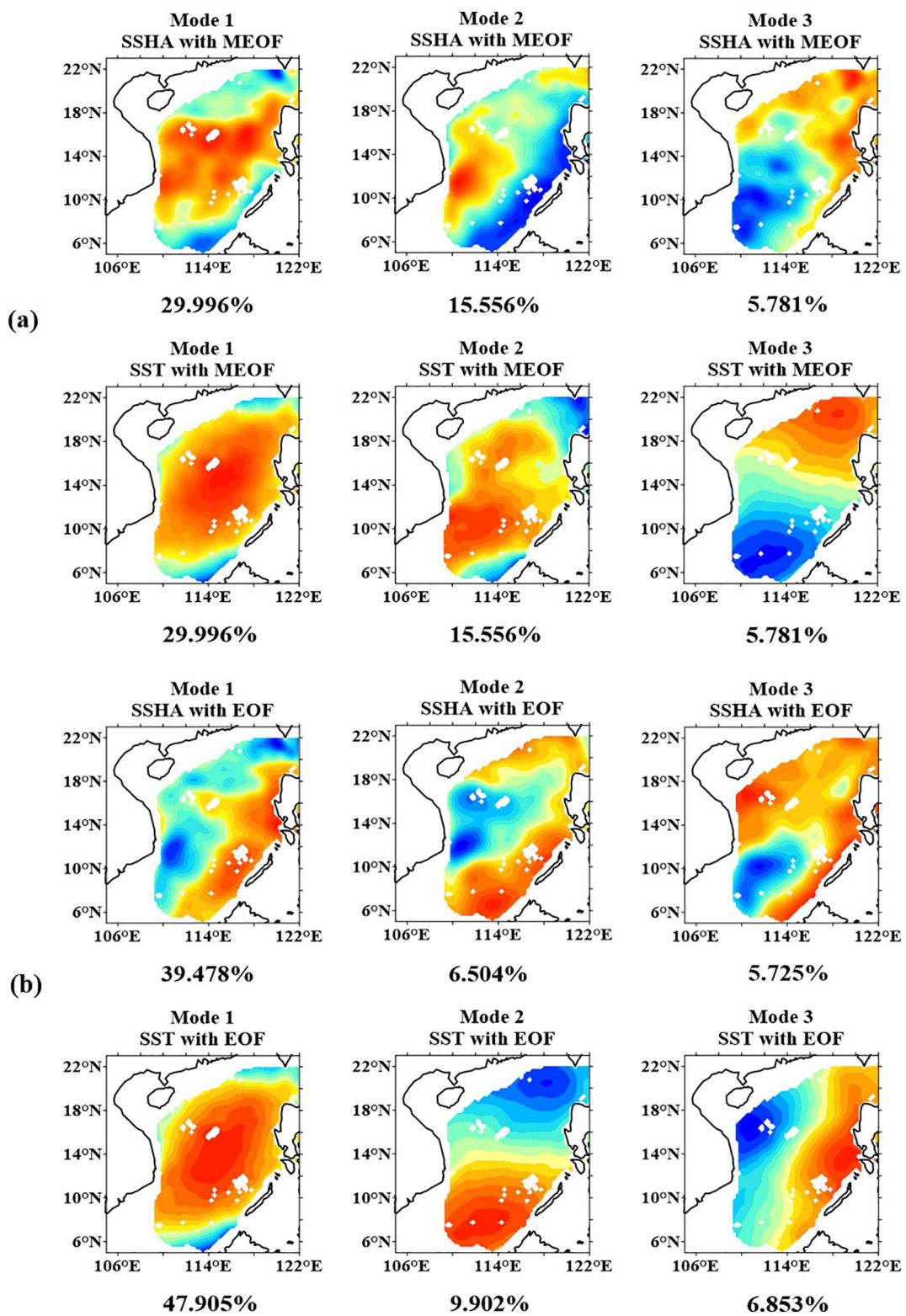


Figure 5. First three empirical orthogonal functions (EOFs) and the corresponding proportion of variance. (a) Spatial modes of sea surface height anomaly (SSHA, m) and sea surface temperature (SST, °C) with multivariate empirical orthogonal function (MEOF) analysis and (b) those with EOF analysis. For each pair of panels, the top panels represent SSHA and the bottom panels represent SST.

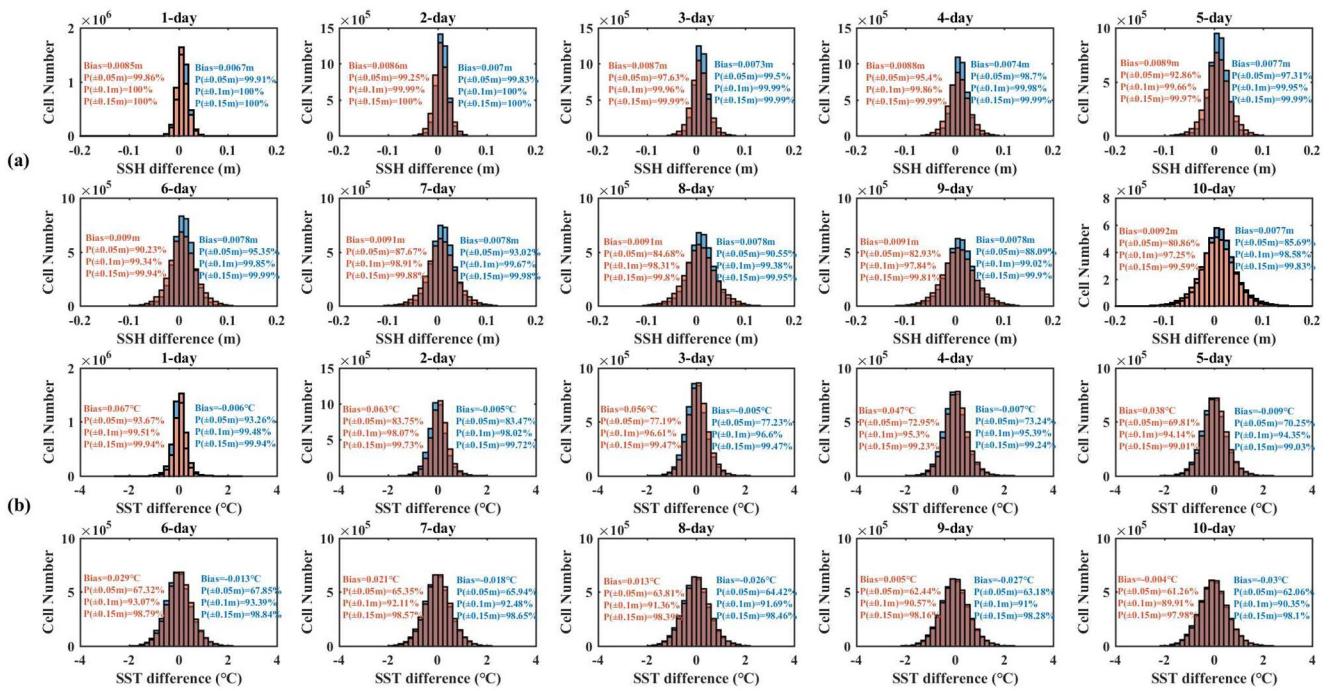


Figure 6. Statistics of the sea surface height anomaly (SSHA) differences (m) and SST differences (°C) from the prediction model with multivariate empirical orthogonal function (MEOF) analysis (blue bars and text) and the model with empirical orthogonal function (EOF) analysis (red bars and text) between the forecasts and the observations from 2014 to 2018. (a) Shows the SSHA differences (m) and (b) shows the SST differences (°C) with leading times from 1 to 10 days.

Furthermore, the spatial distributions of the anomaly RMSEs of these two models are shown in Figure 7 for SSHA and SST. The RMSEs of the forecasts with the verification set of 2014–2018 are calculated at each grid point from daily predictions over a 10-day range. As shown in this figure, these two models both show one area with large RMSEs, and this area is from the Luzon Strait to the northeast of the study domain. Although the patterns of RMSEs are similar for these two models, the error of the prediction model based on MEOF analysis is smaller.

Figures 6 and 7 show that the performance of SSHA in these two prediction models is quite different, while the performance of SST is similar. The reason for this result may be as follows: in principle, the variation in SSHA is determined by SST and the underwater thermohaline structure, but SST is difficult to be determined by SSHA. Therefore, EOF analysis of a single variable is unable to consider the contribution of SST to SSHA.

From the discussion above, the model considering the correlation of different variables has better prediction performance. More importantly, the results suggest that we should not treat each variable individually when using a data-driven forecasting method, and considering variables individually leads to a decrease in prediction accuracy due to a lack of physical meaning.

4.2. Prediction of SSHA and SST for Different Lead Times

To evaluate the performance of the MEOF-Conv1D-LSTM model, persistence prediction, climatology and linear regression results are used to predict SSHA and SST from 2014 to 2018. Persistence prediction is a benchmark comparison and forecast reference widely accepted in atmospheric and oceanic sciences, and it is based on the assumption that the initial oceanic state will persist for the entire forecast time (Li, Vidard, et al., 2019; Li, Wang, et al., 2019). The climatology result is taken as another comparison, which uses the k-year averaged historical data as the forecast (Huang & Barnston, 1996), and we use the average data from 1993 to 2013 in this study. Then, to better compare the differences between linear and nonlinear methods,

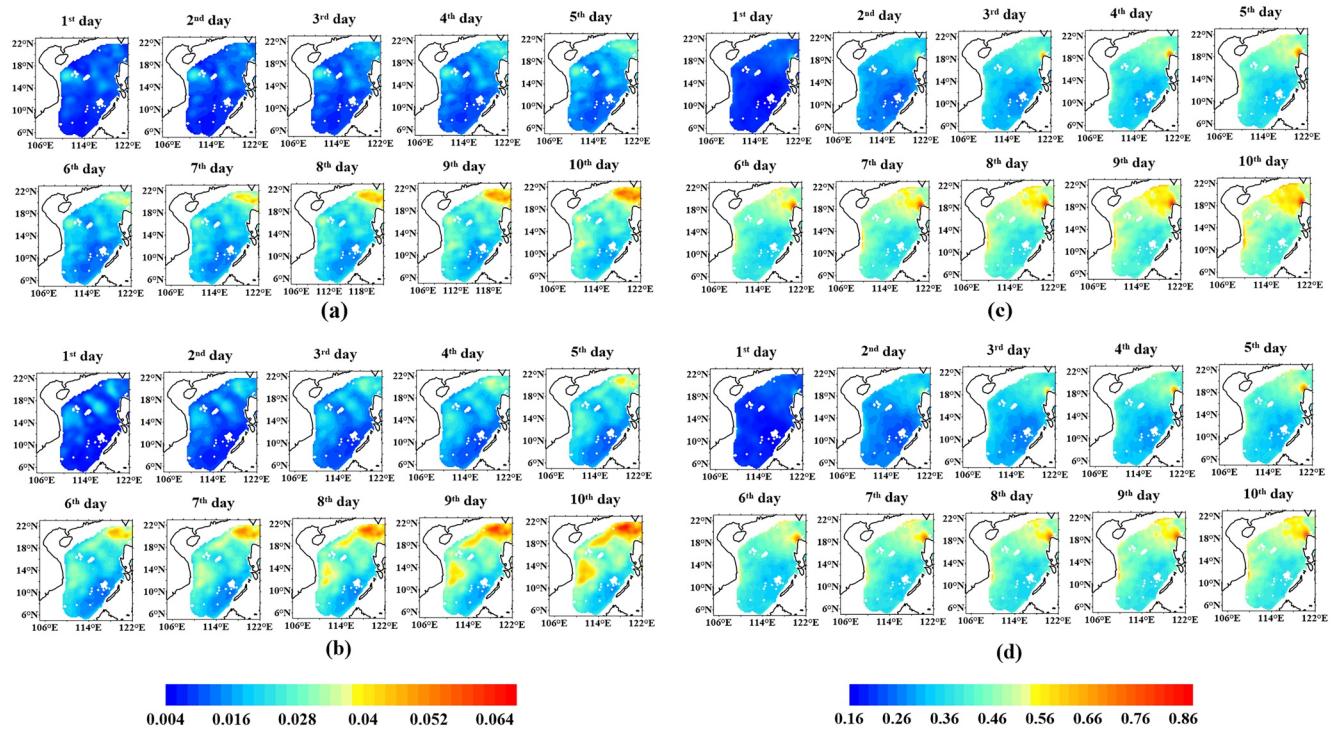


Figure 7. Spatial maps of sea surface height anomaly root mean square error (SSHA RMSE; left panels; unit: m) and sea surface temperature (SST) RMSE (right panels; unit: °C). (a and c) are the model with MEOF analysis, and (b and d) are the model with EOF analysis. Values are calculated from daily forecasts over a 10-day range from 2014 to 2018.

we choose to perform linear regression on PCs. For each PC, the linear regression model is established as follows:

$$y_i^q = k_{1i}^q * x_1^q + k_{2i}^q * x_2^q + k_{3i}^q * x_3^q + k_{4i}^q * x_4^q + k_{5i}^q * x_5^q \\ + k_{6i}^q * x_6^q + k_{7i}^q * x_7^q + k_{8i}^q * x_8^q + k_{9i}^q * x_9^q + k_{10i}^q * x_{10}^q + c_i^q \quad (17)$$

where y_i^q represents the predict and, in which q represents the q th PC, i represents the lead time, k is the regression coefficient, x is the predictor, and c is the constant term.

The qualitative comparison of the MEOF-Conv1D-LSTM prediction model with the persistence prediction model, climatology and linear regression model is further quantified. The details of basin average MAE and RMSE, and spatial ACC are shown in Figure 8, which are computed with the true remote sensing observational data as a basis. The left panels show the SSHA results, while the right panels show the SST results. It is clear from Figure 8 that the MEOF-Conv1D-LSTM model shows not only better agreement with the true remote sensing values in terms of higher ACC but also lower MAE and RMSE than the prediction by persistence, climatology and linear regression results. More importantly, the above conclusions are correct for both SSHA and SST.

For the prediction of SSHA, when the lead time is 1–2 days, the performance of the MEOF-Conv1D-LSTM model outperforms the climatology results and linear regression model but not persistence prediction. This is because SSHA is a slowly changing variable, and persistence prediction is the best description of its variation over a short time. However, the error of persistence prediction increases with forecasting times. The MAE, RMSE and ACC of the MEOF-Conv1D-LSTM model for SSHA forecasts at the end of the forecast window are approximately 0.0217 m, 0.0276 m and 0.864, respectively, which are better than those of other models. For the prediction of SST, the MEOF-Conv1D-LSTM model outperforms the others throughout the forecast period. The MAE, RMSE and ACC of the MEOF-Conv1D-LSTM model for SST forecasts at the end

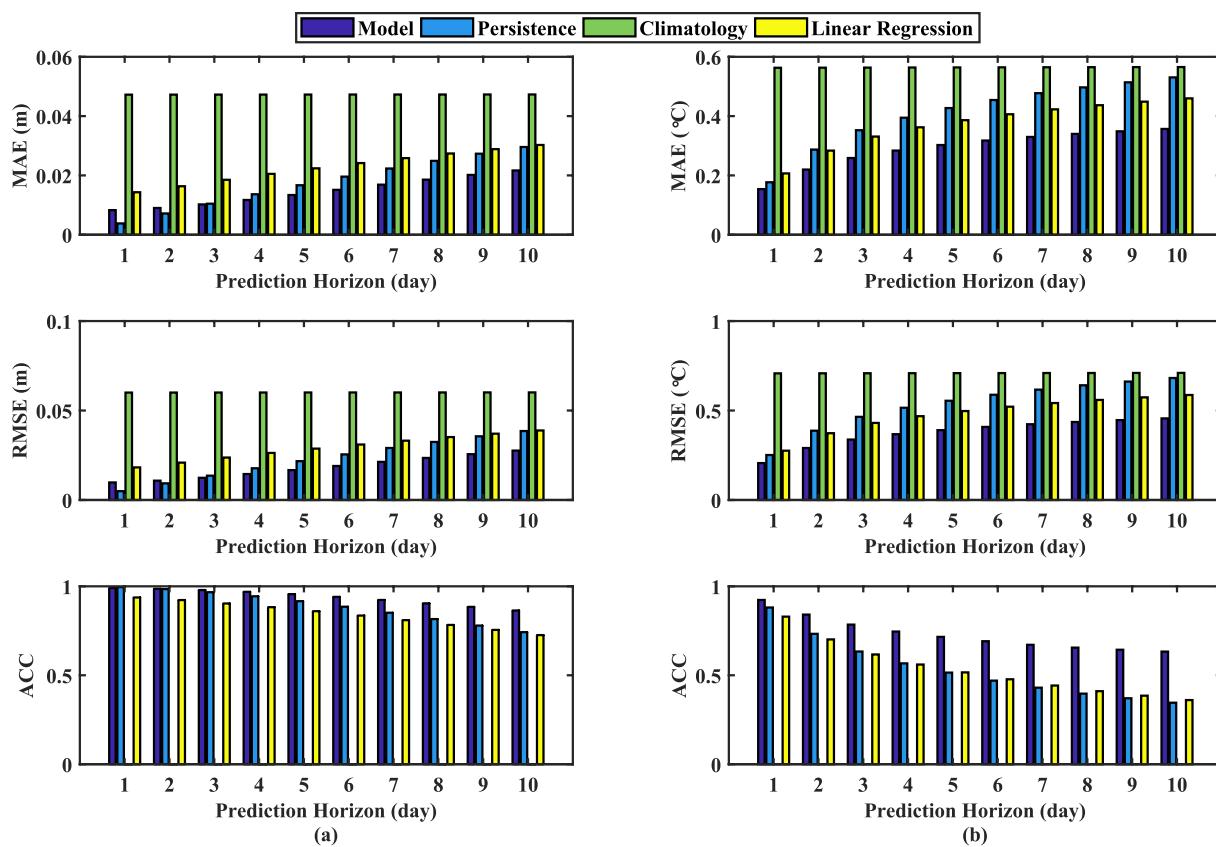


Figure 8. The basin average mean absolute error (MAE) and root mean square error (RMSE), and spatial anomaly correlation coefficient (ACC) of the MEOF-Conv1D-LSTM model (“Model” in the legend), persistence model, climatology results and linear regression results for lead times varying from 1 to 10 days; (a) sea surface height anomaly (SSHA) and (b) sea surface temperature (SST).

of the forecast window are approximately 0.36°C, 0.46°C and 0.633, respectively. Notably, these statistics are the average of all grid points in the study area. Moreover, the statistical significance of the correlation coefficients derived above is determined by the t statistic test, and they are statistically significant.

The above discussion indicates that, in general, the MEOF-Conv1D-LSTM prediction model can make daily SSHA and SST predictions more efficiently than the traditional persistence, climatology results and linear regression model. Perhaps this is because MEOF analysis considers the correlation between different sea surface variables, and the Conv1D-LSTM model also effectively captures the nonlinear process of PC evolution. Both persistence prediction and linear regression models are linear prediction methods, so the comparison with these models emphasizes the need to pursue more nonlinear approaches. Climatology results generally represent a multiyear average, which is insufficient to describe the short and medium-term variations in oceanic multiscale processes.

4.3. Case Study

In this section, to demonstrate the prediction performance of the MEOF-Conv1D-LSTM model under different conditions, prediction examples under normal weather conditions and typhoon weather conditions are presented.

Figure 9 shows a set of 10-day anomaly predictions for SSHA and SST under normal weather conditions, corresponding to 10 February to 19 February 2018. (a and b) show the anomalies of SSHA and SST, respectively. A strong similarity between true values and the MEOF-Conv1D-LSTM predictions can be seen in these figures, which indicates that this model has good prediction performance under normal weather

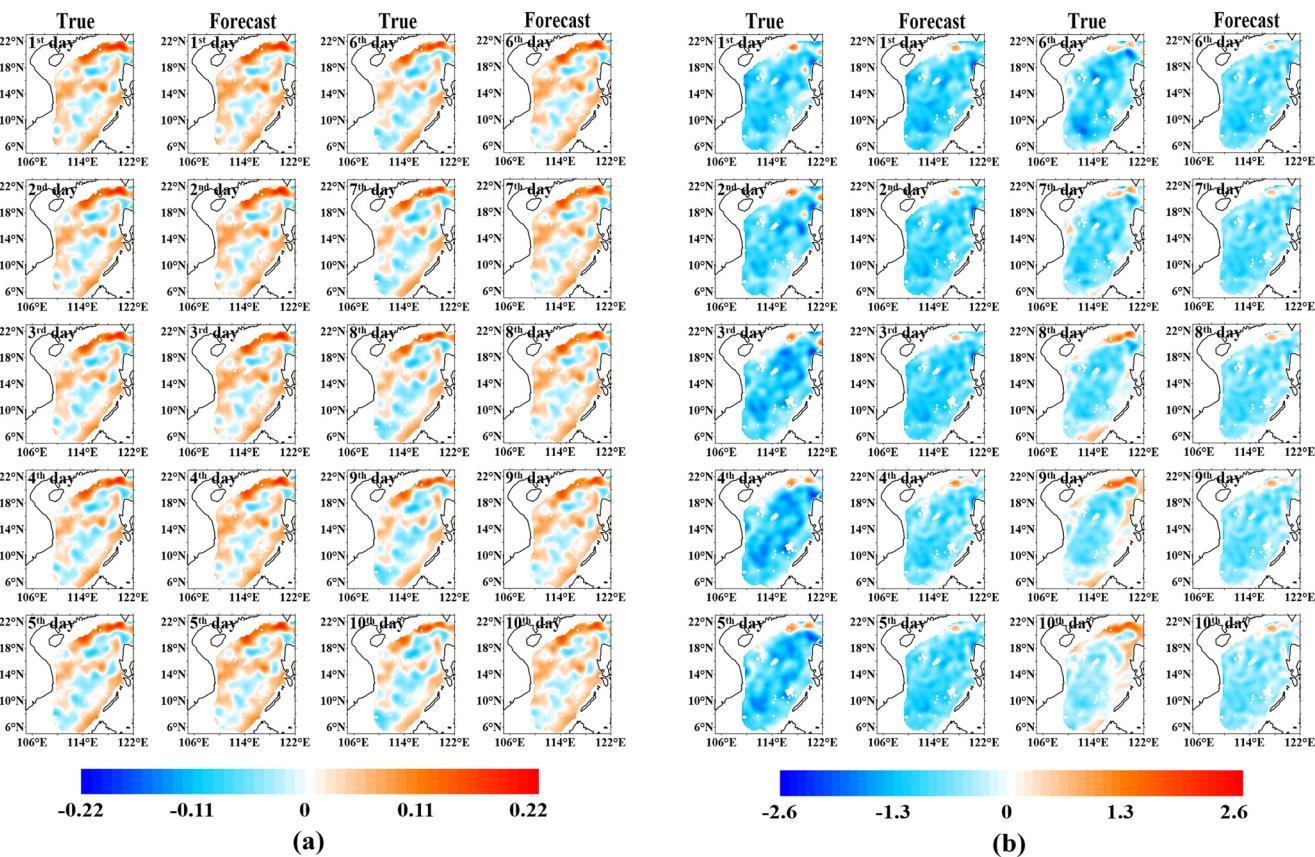


Figure 9. Snapshots of sea surface height anomaly (SSHA, m) and sea surface temperature (SST, °C) forecasts for 1–10 days under normal weather conditions, corresponding to February 10 to February 19, 2018. The anomaly fields of (a) SSHA and (b) SST are shown.

conditions. However, it is not very good at predicting small-scale information, and we think this is caused by the inaccurate prediction of the higher-order modes. In this study, we predict PCs with the first 95% to obtain the reconstructed field. To avoid the phenomenon of stepped prediction, we use the remaining 5% (i.e., the higher-order modes) to correct the reconstructed field in persistence way at the starting time. This does not provide a good representation of the higher-order mode information.

In addition, Figure 10 shows an example under Typhoon Damrey, corresponding to October 28 to November 6, 2017. The forecasts are very similar to the true fields in terms of the overall pattern and the characterization of eddies, although the forecast accuracy is not as good as that under normal conditions.

The forecast increment fields relative to the forecast initialization in the SCS in the above two cases are shown in Figure 11. (a and c) show SSHAs under normal conditions and Typhoon Damrey, respectively, and (b and d) show SSTs. The black lines in (d) represent the typhoon path. Typhoon Damrey entered the SCS on 2 November, and its center was located in the southeastern SCS and moved westward. Subsequently, the typhoon intensified and reached its peak on 4 November. We know that during the passage of a typhoon, the SSTs on the right side of the typhoon's path decreased significantly. Fortunately, the MEOF-Conv1D-LSTM model predicts this cooling trend, although with a slightly lower magnitude. This is because both the internal evolutionary process and external drivers of SSHA and SST are embedded in the data itself. That is, the satellite remote sensing data we used not only represent the values of ocean variables at the current time but also carry past meteorological information and ocean dynamics. In addition, to predict extreme weather conditions more accurately, modeling with external forces as a predictor is a better approach.

Finally, statistical analyses are made to the forecast results of the independent test samples under normal and extreme weather conditions separately. We counted 35 typhoons that passed through the study area

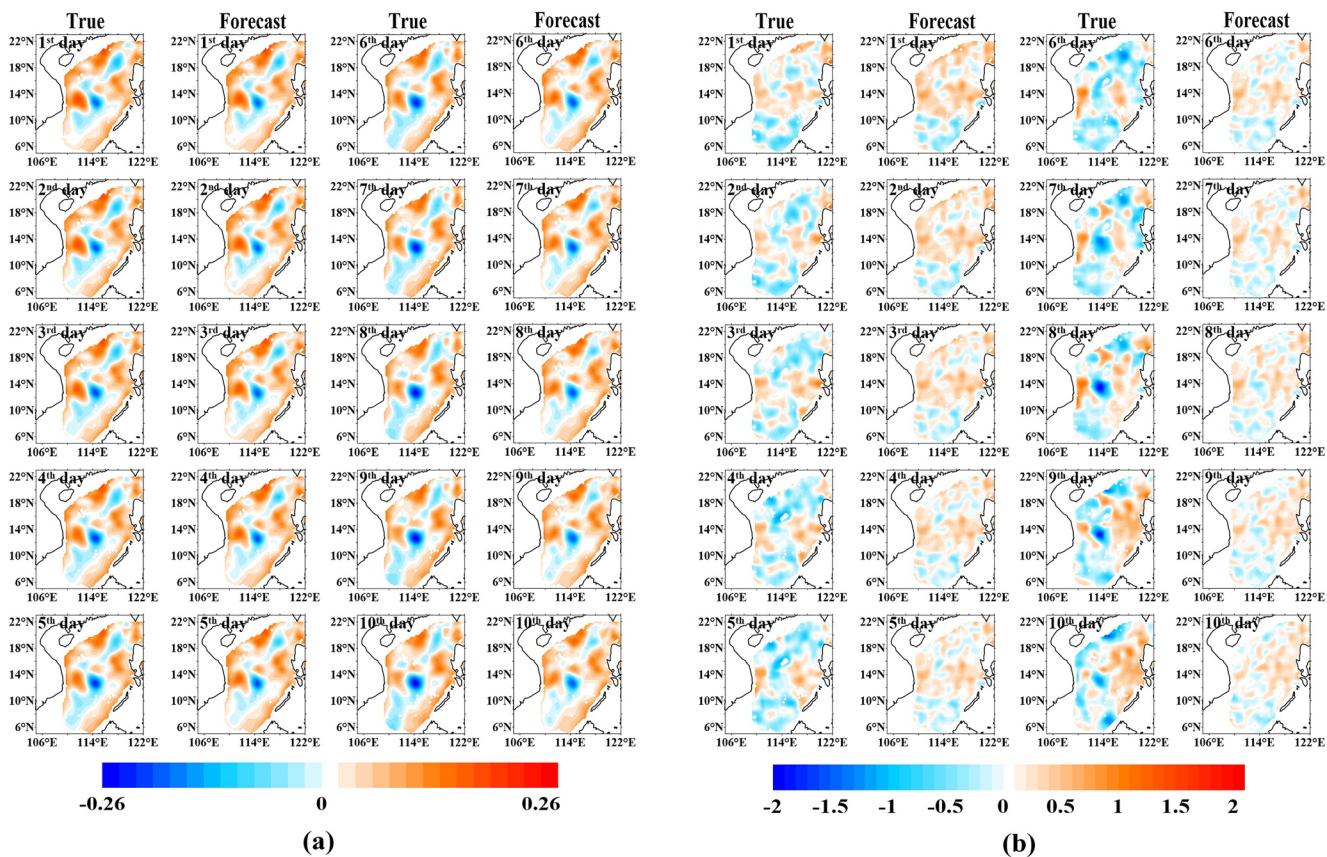


Figure 10. Same as Figure 9 but under Typhoon Damrey corresponding to October 28 to November 6, 2017.

during the period 2014–2018 (i.e., the testing set). Here, we use the time range from 10 days before the typhoon (including the typhoon-forming period) enters the study area to 10 days after it leaves as the time interval of impact. Table 1 shows the statistical results of SSHA and SST under normal and typhoon-affected weather conditions. From the table, we can see that the forecast performance of SSHA under these two conditions is comparable. When the prediction horizon is 10 days, the forecast MAE, RMSE and ACC are 0.0215 m, 0.0274 m and 0.862, respectively, under normal weather conditions and 0.0220 m, 0.0279 m and 0.869 under typhoon conditions. Similarly, the forecast performance of SST under these two conditions is also comparable. At the end of the forecast window, the forecast MAE, RMSE and ACC under normal weather conditions are 0.360°C, 0.461°C and 0.637, respectively, while the forecast MAE, RMSE and ACC under typhoon conditions are 0.348°C, 0.443°C and 0.625, respectively.

5. Conclusions

Our study describes how a hybrid model based on traditional statistical and deep learning methods carries out the simultaneous prediction of multiple sea surface variables on weather and subseasonal time scales. This hybrid model is developed by combining MEOF analysis and the Conv1D-LSTM model, called the MEOF-Conv1D-LSTM model. The MEOF analysis here has four main contributions: (a) establishing the spatial correlation between different discrete points in the study area; (b) considering the correlation between different sea surface variables, thereby making the prediction model more consistent with the variation process of the real marine environment; (c) extracting the main spatial orthogonal mode information to reduce the computational effort; and (d) de-correlating the PCs due to the orthogonal property of MEOF analysis, indicating that it is feasible to predict each PC separately. Then, the Conv1D-LSTM neural network is trained and verified by the PC series.

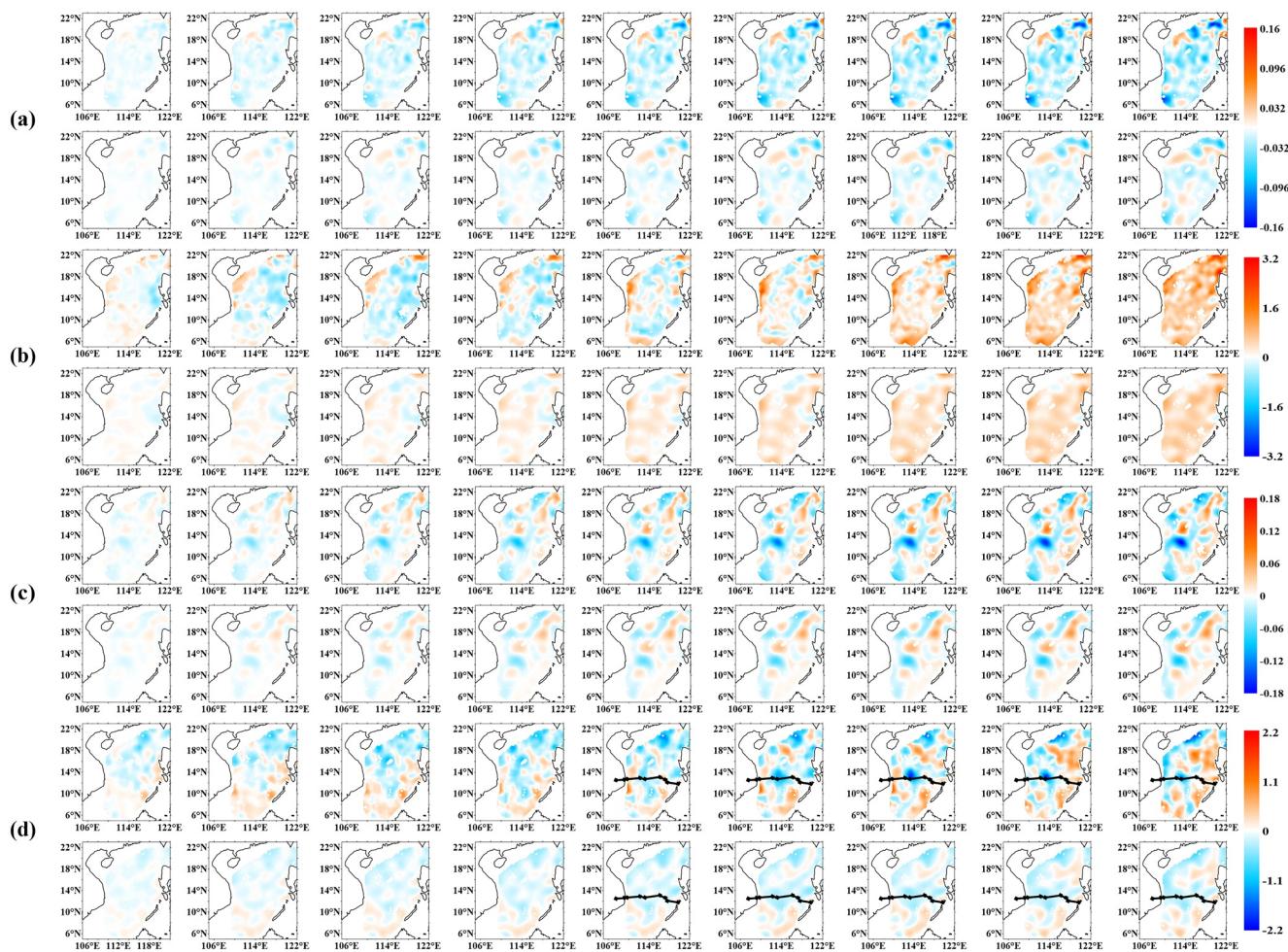


Figure 11. Difference fields relative to the initialization. (a and c) represent sea surface height anomaly (SSHA) under normal conditions and Typhoon Damrey, respectively, and (b and d) represent sea surface temperature (SST). The black line in (d) represents the typhoon path. For each pair of panels, the top panels represent true values and the bottom panels represent forecasts.

To evaluate the performance of the model, SSHA and SST in the SCS are selected for experiments in this study, considering the dynamic balance between different sea surface variables and the availability and quality of satellite remote sensing data. The training set of the prediction model is restricted to a period of 21 years, from January 1, 1993 to December 31, 2013. Then, a 5-year data set from January 1, 2014 to December 31, 2018 is used as an independent testing sample to validate the prediction results. First, a comparison of the prediction models based on MEOF analysis and EOF analysis shows significant improvement by modeling SSHA and SST together, indicating that we should not treat each variable individually when using a data-driven forecasting approach. Then, MAE, RMSE and ACC are used as metrics to evaluate the skill of the MEOF-Conv1D-LSTM model compared with persistence, climatology and linear regression prediction as references; the results show that the MEOF-Conv1D-LSTM model has good performance in the 10-day forecast range. The RMSEs of this model's forecasts for SSHA and SST at the end of the forecast window are approximately 0.0276 m and 0.46°C, respectively, which are lower than those of the persistence, climatology and linear regression. The ACCs for SSHA and SST are approximately 0.864 and 0.633, respectively. Moreover, when climatology is added back, the CCs are approximately 0.919 and 0.850 for SSHA and SST, respectively, which are much higher than those in the persistence and linear regression models. Finally, forecast examples under normal weather conditions and typhoon conditions are presented. These two examples show that this model can predict not only the sea surface variables under normal sea conditions but also the variations and trends of sea surface variables during the passage of typhoons, although the accuracy

Table 1*Statistical Results for SSHA and SST in Normal and Typhoon-Affected Weather Conditions*

Variable	Condition	Metric	Prediction horizon (days)									
			1	2	3	4	5	6	7	8	9	10
SSHA	normal	MAE (m)	0.0084	0.0091	0.0103	0.0117	0.0134	0.0151	0.0168	0.0185	0.0200	0.0215
		RMSE (m)	0.0098	0.0109	0.0125	0.0144	0.0166	0.0189	0.0211	0.0233	0.0254	0.0274
		ACC	0.990	0.986	0.979	0.969	0.955	0.940	0.922	0.900	0.883	0.862
	typhoon	MAE (m)	0.0081	0.0088	0.0101	0.0117	0.0135	0.0153	0.0171	0.0189	0.0205	0.0220
		RMSE (m)	0.0096	0.0106	0.0123	0.0144	0.0167	0.0191	0.0215	0.0138	0.0259	0.0279
		ACC	0.990	0.986	0.980	0.970	0.958	0.943	0.926	0.908	0.888	0.869
SST	normal	MAE (°C)	0.155	0.221	0.260	0.285	0.305	0.320	0.333	0.343	0.352	0.360
		RMSE (°C)	0.206	0.291	0.338	0.369	0.392	0.411	0.427	0.439	0.451	0.461
		ACC	0.924	0.843	0.787	0.749	0.720	0.696	0.676	0.660	0.647	0.637
	typhoon	MAE (°C)	0.152	0.217	0.255	0.280	0.297	0.311	0.323	0.332	0.340	0.348
		RMSE (°C)	0.205	0.288	0.334	0.363	0.383	0.400	0.414	0.425	0.434	0.443
		ACC	0.921	0.837	0.780	0.740	0.710	0.680	0.660	0.644	0.634	0.625

may be slightly worse. Furthermore, the forecast results of independent test samples under normal weather conditions and extreme weather conditions are statistically analyzed, and the statistical results show that this model has considerable predictive ability under normal weather conditions and typhoon conditions. In subsequent work, meteorological factors, lateral boundary conditions and other external drivers in the study area will also be considered in the model to improve the capability to predict extreme events.

Notably, the anomaly field evolution of these sea surface variables between the forecast initialization and forecast window is determined by a combination of thermodynamic coordination within the ocean system and boundary and surface meteorological driving forces. In fact, the internal evolutionary processes and external drivers of the sea surface are already embedded in the data itself. In other words, the satellite remote sensing data we used not only represent the values of ocean variables at the current time but also carry past meteorological information and ocean dynamics. Taking these processes into account can help us identify prominent contributors to the superiority of the MEOF-Conv1D-LSTM model over persistence, climatology and linear statistical models, as shown above.

This new forecasting technique could potentially be applied to other sea surface variables, such as sea surface salinity, or even entire three-dimensional ocean variables, as well as atmospheric system or ocean-atmosphere coupled models. To better describe and simulate the processes in the marine environment, we should learn from numerical prediction models. The relationship between different variables and different systems should be fully considered, and then, a data-driven method could be used to realize more accurate and simpler predictions of marine environment variables based on satisfying the conservation of mass and energy. This will help us to better understand the relations between the prediction models and the physical processes.

Finally, the prediction model in this study also has some advantages in real-time operational applications. Notably, the training set and the testing set in this study are completely independent. In the prediction process, we need to store only the EOFs obtained from the decomposition of the training set, and when the newly arrived real-time data are available, we can obtain the corresponding PCs by inverted MEOF analysis; then, the Conv1D-LSTM model can be used for prediction. Finally, the real-time forecast results can be reconstructed by combining the predicted values of PCs and the EOFs stored previously.

Data Availability Statement

The SSHA data are produced by CMEMS and are available at <http://marine.copernicus.eu>, and the SST data are produced by NOAA and available at <https://www.ncdc.noaa.gov/oisst>.

Acknowledgments

The authors are grateful to the editors and reviewers for providing valuable suggestions, which improved this study. The authors would like to thank the following data and tool providers: Copernicus Marine Environment Monitoring Service (CMEMS) for providing SSHA data (<http://marine.copernicus.eu>), NOAA for providing OISST data (<https://www.ncdc.noaa.gov/oisst>), Google for providing the open source machine learning framework TensorFlow (<https://www.tensorflow.org/>), Keras (<https://keras.io/>) for making it easier to use TensorFlow to conduct Conv1D-LSTM deep learning experiment, and Scikit-learn (<http://scikit-learn.org/stable/>) for providing the machine learning library based on which the experiments are conducted. The authors are also grateful for the data information of Typhoon from <http://typhoon.zjwater.gov.cn/>. This research is cosponsored by grants from the National Natural Science Foundation (41876014) of China, the National Key Research and Development Program (2016YFC1401800), and the Open Project of Tianjin Key Laboratory of Oceanic Meteorology (2020TKLOMVB04).

References

- Agarwal, N., Kishtawal, C. M., & Pal, P. K. (2001). An analogue prediction method for global sea surface temperature. *Current Science*, 80(1), 49–55.
- Chu, P. C., Ivanov, L. M., Korzhova, T. P., Margolina, T. M., & Melnichenko, O. V. (2003). Analysis of sparse and noisy ocean current data using flow decomposition. Part I: Theory. *Journal of Atmospheric and Oceanic Technology*, 20, 478–491. [https://doi.org/10.1175/1520-0426\(2003\)20<478:aozano>2.0.co;2](https://doi.org/10.1175/1520-0426(2003)20<478:aozano>2.0.co;2)
- Collins, D. C., Reason, C. J. C., & Tangang, F. (2004). Predictability of Indian Ocean sea surface temperature using canonical correlation analysis. *Climate Dynamics*, 22, 481–497. <https://doi.org/10.1007/s00382-004-0390-4>
- Fu, Y., Zhou, X., Sun, W., & Tang, Q. (2019). Hybrid model combining empirical mode decomposition, singular spectrum analysis, and least squares for satellite-derived sea-level anomaly prediction. *International Journal of Remote Sensing*, 40(20), 7817–7829. <https://doi.org/10.1080/01431161.2019.1606959>
- Hannachi, A., Jolliffe, I. T., & Stephenson, D. B. (2007). Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology*, 27(9), 1119–1152. <https://doi.org/10.1002/joc.1499>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hu, W., Wu, R., & Liu, Y. (2014). Relation of the South China Sea precipitation variability to tropical Indo-Pacific SST anomalies during spring-to-summer transition. *Journal of Climate*, 27(14), 5451–5467. <https://doi.org/10.1175/jcli-d-14-00089.1>
- Huang, J., & Barnston, G. A. (1996). Long-lead seasonal temperature prediction using optimal climate normals. *Journal of Climate*, 9, 809–817. [https://doi.org/10.1175/1520-0442\(1996\)009<0809:llstpu>2.0.co;2](https://doi.org/10.1175/1520-0442(1996)009<0809:llstpu>2.0.co;2)
- Jansen, M. F., Dommenget, D., & Keenlyside, N. (2008). Tropical atmosphere–ocean interactions in a conceptual framework. *Journal of Climate*, 22, 550–567. <https://doi.org/10.1175/2008JCLI2243.1>
- Krishnamurti, T. N., Chakraborty, A., Krishnamurti, R., Dewar, W. K., & Clayson, C. A. (2006). Seasonal prediction of sea surface temperature anomalies using a suite of 13 coupled atmosphere–ocean models. *Journal of Climate*, 19(23), 6069–6088. <https://doi.org/10.1175/jcli3938.1>
- Kug, J., Kang, I., Lee, J., & Jhun, J. (2004). A statistical approach to Indian Ocean sea surface temperature prediction using a dynamical ENSO prediction. *Geophysical Research Letters*, 31(9), 112–124. <https://doi.org/10.1029/2003gl019209>
- Li, J., Wang, G., Xue, H., & Wang, H. (2019). A simple predictive model for the eddy propagation trajectory in the northern South China Sea. *Ocean Science*, 15(2), 401–412. <https://doi.org/10.5194/os-15-401-2019>
- Li, L., Vidard, A., Le Dimet, F.-X., & Ma, J. (2019). Topological data assimilation using Wasserstein distance. *Inverse Problems*, 35(1), 1–23. <https://doi.org/10.1088/1361-6420/aae993>
- Li, W., Xie, Y., Deng, S.-M., & Wang, Q. (2010). Application of the multigrid method to the two-dimensional Doppler radar radial velocity data assimilation. *Journal of Atmospheric and Oceanic Technology*, 27(2), 319–332. <https://doi.org/10.1175/2009jtecha1271.1>
- Li, W., Xie, Y., He, Z., Han, G., Liu, K., Ma, J., & Li, D. (2008). Application of the multigrid data assimilation scheme to the China Seas’ temperature forecast. *Journal of Atmospheric and Oceanic Technology*, 25(11), 2106–2116. <https://doi.org/10.1175/2008jtecho510.1>
- Lorenz, E. N. (1956). *Empirical orthogonal functions and statistical weather prediction*. (Vol. 1, pp. 1–49). Department of Meteorology, Massachusetts Institute of Technology, Statistical Forecasting Project Report. <https://doi.org/10.1134/S1028334X06060377>
- North, G. R. (1983). Empirical orthogonal functions and normal modes. *Journal of Atmospheric Sciences*, 41(5), 879–887. [https://doi.org/10.1175/1520-0469\(1984\)0412.0.CO](https://doi.org/10.1175/1520-0469(1984)0412.0.CO)
- North, G. R., Bell, T. L., Cahalan, R. F., & Moeng, F. J. (1982). Sampling errors in the estimation of empirical orthogonal functions. *Monthly Weather Review*, 110(7), 699–706. [https://doi.org/10.1175/1520-0493\(1982\)110<0699:seiteo>2.0.co;2](https://doi.org/10.1175/1520-0493(1982)110<0699:seiteo>2.0.co;2)
- Oey, L. Y., Ezer, T., Forristall, G., Cooper, C., Dimarco, S., & Fan, S. (2005). An exercise in forecasting loop current and eddy frontal positions in the Gulf of Mexico. *Geophysical Research Letters*, 32(12), 1–12. <https://doi.org/10.1029/2005gl023253>
- Patil, K., & Deo, M. C. (2017). Prediction of daily sea surface temperature using efficient neural networks. *Ocean Dynamics*, 67(3–4), 357–368. <https://doi.org/10.1007/s10236-017-1032-9>
- Patil, K., & Deo, M. C. (2018). Basin-scale prediction of sea surface temperature with artificial neural networks. *Journal of Atmospheric and Oceanic Technology*, 35(7), 1441–1455. <https://doi.org/10.1175/jtech-d-17-0217.1>
- Patil, K., Deo, M. C., & Ravichandran, M. (2016). Prediction of sea surface temperature by combining numerical and neural techniques. *Journal of Atmospheric and Oceanic Technology*, 33(8), 1715–1726. <https://doi.org/10.1175/jtech-d-15-0213.1>
- Qian, S., Wei, H., Xiao, J.-G., & Nie, H. (2018). Impacts of the Kuroshio intrusion on the two eddies in the northern South China Sea in late spring 2016. *Ocean Dynamics*, 68(12), 1695–1709. <https://doi.org/10.1007/s10236-018-1224-y>
- Reynolds, R. W., & Chelton, D. B. (2010). Comparisons of daily sea surface temperature analyses for 2007–08. *Journal of Climate*, 23(13), 3545–3562. <https://doi.org/10.1175/2010JCLI3294.1>
- Shaw, P.-T. (1991). The seasonal variation of the intrusion of the Philippine Sea water into the South China Sea. *Journal of Geophysical Research*, 96(C1), 821–827. <https://doi.org/10.1029/90jc02367>
- Shi, X., Chen, Z., Wang, H., & Yeung, D.-Y. (2015). Convolutional LSTM Network: A machine learning approach for precipitation nowcasting. *Proceedings of the Advances in Neural Information Processing Systems*, 7, 802–810. https://doi.org/10.1007/978-3-319-21233-3_6
- Sun, J., Wei, Z., Xu, T., Sun, M., Liu, K., Yang, Y., et al. (2019). Development of a fine-resolution atmosphere-wave-ocean coupled forecasting model for the South China Sea and its adjacent seas. *Acta Oceanologica Sinica*, 38(4), 154–166. <https://doi.org/10.1007/s13131-019-1419-1>
- Tang, B., Hsieh, W. W., Monahan, A. H., & Tangang, F. T. (2000). Skill comparisons between neural networks and canonical correlation analysis in predicting the Equatorial Pacific Sea surface temperatures. *Journal of Climate*, 13, 287–293. [https://doi.org/10.1175/1520-0420\(2000\)013<0287:scbnma>2.0.co;2](https://doi.org/10.1175/1520-0420(2000)013<0287:scbnma>2.0.co;2)

- Thakur, K. K., Vanderstichel, R., Barrell, J., Stryhn, H., Patanasatienkul, T., & Revie, C. W. (2018). Comparison of remotely-sensed sea surface temperature and salinity products with in situ measurements from British Columbia, Canada. *Frontiers in Marine Science*, 5. <https://doi.org/10.3389/fmars.2018.00121>
- Vidard, A., Balmaseda, M. A., & Stockdale, T. N. (2006). Tropical Atlantic SST prediction with coupled ocean-atmosphere GCMs. *Journal of Climate*, 19(23), 6047–6061. <https://doi.org/10.1175/jcli3947.1>
- Wang, D., Hong, B., Gan, J., & Xu, H. (2010). Numerical investigation on propulsion of the counter-wind current in the northern South China Sea in winter. *Deep Sea Research Part I: Oceanographic Research Papers*, 57(10), 1206–1221. <https://doi.org/10.1016/j.dsr.2010.06.007>
- Wang, J. L., Zhuang, H., Cherubin, L. M., Ibrahim, A. K., & Muhamed Ali, A. (2019). Medium-term forecasting of loop current Eddy Cameron and Eddy Darwin formation in the Gulf of Mexico with a divide-and-conquer machine learning approach. *Journal of Geophysical Research: Oceans*, 124(8), 5586–5606. <https://doi.org/10.1029/2019jc015172>
- Wang, L., & Koblinsky, C. J. (2000). Mesoscale variability in the South China Sea from the TOPEX/Poseidon altimetry data. *Deep-Sea Research Part I: Oceanographic Research Papers*, 47(4), 681–708. [https://doi.org/10.1016/S0967-0637\(99\)00068-0](https://doi.org/10.1016/S0967-0637(99)00068-0)
- Wu, X. (2016). Improving EnKF-based initialization for ENSO prediction using a hybrid adaptive method. *Journal of Climate*, 29(20), 7365–7381. <https://doi.org/10.1175/jcli-d-16-0062.1>
- Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J., & Chen, Z. (2019). Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach. *Remote Sensing of Environment*, 233, 1–18. <https://doi.org/10.1016/j.rse.2019.111358>
- Xue, Y., & Leetmaa, A. (2000). Forecasts of tropical Pacific SST and sea level using a Markov model. *Geophysical Research Letters*, 27(17), 2701–2704. <https://doi.org/10.1029/1999gl011107>
- Xue, Y., Leetmaa, A., & Ji, M. (2000). ENSO prediction with Markov models: The impact of sea level. *Journal of Climate*, 13, 849–871. [https://doi.org/10.1175/1520-0442\(2000\)013<0849:epwmmt>2.0.co;2](https://doi.org/10.1175/1520-0442(2000)013<0849:epwmmt>2.0.co;2)
- Yang, H., & Liu, Q. (2002). A general circulation model study of the dynamics of the upper ocean circulation of the South China Sea. *Journal of Geophysical Research*, 107(22), 14. <https://doi.org/10.1029/2001JC001084>
- Yang, L., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324. <https://doi.org/10.1109/5.726791>
- Yang, Y., Dong, J., Sun, X., Lima, E., Mu, Q., & Wang, X. (2018). A CFCC-LSTM model for sea surface temperature prediction. *IEEE Geoscience and Remote Sensing Letters*, 15(2), 207–211. <https://doi.org/10.1109/lgrs.2017.2780843>
- Yuan, D., Han, W., & Hu, D. (2006). Surface Kuroshio path in the Luzon Strait area derived from satellite remote sensing data. *Journal of Geophysical Research*, 111(11), 1–16. <https://doi.org/10.1029/2005jc003412>
- Yue, L., Shen, H., Zhang, L., Zheng, X., Zhang, F., & Yuan, Q. (2017). High-quality seamless DEM generation blending SRTM-1, ASTER GDEM v2 and ICESat/GLAS observations. *ISPRS Journal of Photogrammetry and Remote Sensing*, 123, 20–34. <https://doi.org/10.1016/j.isprsjprs.2016.11.002>
- Zang, L., Mao, F., Guo, J., Wang, W., Pan, Z., Shen, H., et al. (2019). Estimation of spatiotemporal PM_{1.0} distributions in China by combining PM_{2.5} observations with satellite aerosol optical depth. *Science of the Total Environment*, 658, 1256–1264. <https://doi.org/10.1016/j.scitotenv.2018.12.297>
- Zeng, X., Li, Y., & He, R. (2015). Predictability of the loop current variation and Eddy shedding process in the Gulf of Mexico using an artificial neural network approach. *Journal of Atmospheric and Oceanic Technology*, 32(5), 1098–1111. <https://doi.org/10.1175/jtech-d-14-00176.1>