Graph Convolutional Network Assisted Oceanic Variable Prediction with Multi-Characteristics Modeling of Spatio-Temporal Evolution

Min Ye, Bohan Li, Jie Nie, *Member, IEEE*, Qi Wen, Zhiqiang Wei, *Member, IEEE* and Lie-Liang Yang, *Fellow, IEEE*

Abstract—Changes in oceanic variables, such as sea surface temperature (SST) and chlorophyll-a (Chl-a), have important implications for marine ecosystems and global climate change. The deep learning methods relying on convolutional neural networks can be employed to extract the spatial correlation for the prediction of oceanic variables. However, these methods are inflexible in the cases where some regions, e.g., land and islands, are invalid for the prediction of oceanic variables. By contrast, graph convolutional networks (GCN) are capable of capturing the large-scale spatial dependency existing in irregular data. Therefore, in this paper, we propose a GCN-based method for the prediction of oceanic variables, i.e., SST and Chl-a, referred it to as OVPGCN, to achieve high-accuracy. The proposed OVPGCN consists of three modules aiming to fully extract the spatial correlation and temporal dependency via modeling the multi-characteristics of the spatio-temporal dynamic evolution. In particular, three modules are implemented to extract the stationary and non-stationary variations in the recent spatio-temporal sequences, the spatial differences between different sites, and the periodic features in historical data, respectively. The well-designed OVPGCN is applied to the monthly SST and Chl-a prediction in the Bohai Sea and the Northern South China Sea (NSCS). The performance demonstrates that the proposed OVPGCN is highly effective and enables to achieve much higher prediction accuracy than the state-of-the-art methods.

Index Terms—Oceanic variables prediction, sea surface temperature (SST), chlorophyll-a (Chl-a), deep learning, graph convolutional network (GCN).

I. INTRODUCTION

OCEAN accounts for approximately the seven-tenths of the earths surface and acts as the earths thermostat and conveyor belt. The changes of oceanic variables, such as sea surface temperature (SST) and chlorophyll-a (Chl-a), have a significant impact on the global or local climate, marine ecosystem, and even vegetation [1–3]. For instance, SST can affect the precipitation distribution and its anomaly variation may lead to droughts and floods [4]. In order to quickly

respond to climate change and avoid the impending ocean crisis, it is of paramount importance to predict the spatio-temporal evolution of the key oceanic variables, as early and accurately as possible.

The methods for forecasting oceanic variables can in general be classified into two categories, namely the physics-based methods and the data-driven methods. Straightforwardly, the physics-based approaches conduct prediction via designing the models by integrating the relevant physical and dynamical equations [5], such as the coupled ocean-atmosphere general circulation models (GCMs) [6]. Due to the boom in the application of high-performance computational device, the physics-based methods have been comprehensively developed in the past decade, with the motivation to achieve high accuracy and resolution [7]. However, the existing physics-based methods suffer from extremely high computation complexity and require the prior knowledge of a large number of oceanic variables. Furthermore, the errors may be introduced by the parameter initialization, which hinders the model from reliable prediction in practical applications [8].

By contrast, the data-driven methods predict oceanic variables relying purely on the observational data, without access to the prior knowledge and the embedded physical processes. The data-driven methods exploit the historic data to learn the underlying evolution rules via mining the spatio-temporal variations of the oceanic variables. To take these advantages, the traditional learning-based methods, such as support vector regression (SVR) and artificial neural networks (ANN), have been widely introduced for the prediction of single point of oceanic variables [9–11]. However, these traditional learning-based methods are in general only feasible for handling the relatively simple nonlinear associations, but not suitable for forecasting the complex spatio-temporal evolution in the oceanic variable field consisting of multiple observation points.

In this regard, researchers have resorted to the more advanced deep learning (DL) methods for the prediction of oceanic variables. This is also the result of development of the ocean satellite observations and ocean data assimilation technologies, which make a large amount of marine environmental data with high resolution available [12]. Specifically, the recurrent neural network (RNN)-based methods, such as the long short-term memory (LSTM) and gate recurrent unit (GRU) [13, 14], have been designed for solving the temporal sequence modeling problems, where the dynamic loop mechanism and long-term memory are obtained using control gates
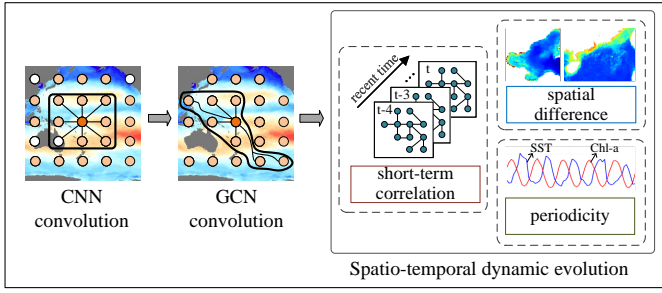
Fig. 1. Prediction of oceanic variables based on the DL methods with different convolution operations and the multi-characteristics of the spatio-temporal dynamic evolution.

[15–17]. In [18], the authors applied a LSTM model for SST prediction, showing that the LSTM outperforms SVR and also the multilayer perceptron regression. In comparison with the conventional learning-based methods, LSTM and GRU are capable of exploiting the temporal information for performance improvement, but they fail to make use of the spatial correlation within the ocean, and hence the performance can be further improved for spatio-temporal prediction problems.

To this end, by considering that convolutional operation can mine local spatial correlation based on extracting neighborhood information [19], the convolutional long short-term network (ConvLSTM) [20], which integrates the convolutional operation into LSTM, has also been proposed for precipitation prediction and widely used for the prediction of SST, sea surface height, sea ice, etc. [21–26]. For example, in [27], the daily sea ice concentration (SIC) prediction in the Northeast Passage has been improved by using a ConvLSTM-based algorithm to extract the spatio-temporal variation of SIC. Owing to the advantages of ConvLSTM over the conventional DL methods, many variant methods have been proposed for prediction of spatio-temporal processes [28–30]. Besides, the CNN-based methods, which endeavor to improve the spatial domain feature extraction, have also catched wide attentions for the prediction of oceanic variables [31–33]. As an example, in [34], the authors proposed a multi-scale CNN method for forecasting SST evolution, which allows to enlarge the receptive field and be free from the local connectivity limitations, and thereby enhancing the prediction accuracy.

A. Motivation

The above-mentioned conventional DL-based prediction methods have to be operated with regular grid structure. This makes them inflexible for some implementation and unable to filter out some noisy regions, e.g., land and islands. To tackle these shortcomings, in this paper, we formulate the prediction of oceanic variables as a graph regression problem and solve it using graph convolutional network (GCN) [35]. The underneath potential is that GCN is able to carry out the convolutional operation on the non-Euclidean and irregular data, by regarding and modeling the large-scale global connections as edges of a graph, as shown in Fig. 1.

To date, there have been some preliminary researches on the introduction of GCN for the prediction of oceanic variables.

Specifically, the authors of [36] firstly applied GCN to forecast the SST evolution associated with a typical phenomenon: El Niño-Southern Oscillation (ENSO), demonstrating that GCN outperforms CNN for the ENSO forecast up to six months ahead. In [37], the authors designed a graph based on the geographic distance information of the offshore wind nodes. Based on it, the proposed algorithm can maximize the high-dimensional spatial correlation features of these nodes to improve the prediction accuracy of wind speed. Zhang et al. [38] proposed a memory graph convolutional network (MGCN) framework, which consists of memory and graph layers, for the SST prediction. More specifically, in MGCN, the memory layer captures the temporal changes of the SST sequence, whereas the graph layers are functioned to learn the spatial relationship. Note that, in the aforementioned papers, the graph convolution is operated in the spectral-domain, by drawing the correlation relationships after the Fourier transformer of data. On the other hand, there is a class of GCN, such as GraphSAGE [39], which implement the graph convolutional operation based on the spatial relationship of nodes. Furthermore, in [40], a time-series graph network (TSGN) was proposed to jointly capture the spatial correlation and temporal dynamics, which combines GraphSAGE with LSTM.

In reality, the spatio-temporal evolution of oceanic variables are not only relied on the short-term temporal patterns, but also on the long-term changes and geographical differences. To our knowledge, the existing GCN-based approaches in the literatures for oceanic variable prediction focus mainly on the mining of the spatio-temporal features existing in the recent historical sequences, i.e., short-term correlation, as shown in Fig. 1. However, they ignore or are incapable of mining the multi-characteristics of the dynamic evolution of oceanic variable fields, such as the spatial differences, as well as the periodicity existing in the long-term temporal evolution, as shown in Fig. 1. As the result, the existing GCN-based methods fail to fully extract the joint spatio-temporal dependencies for achieving the best possible prediction performance, especially when the training samples are not sufficient. Therefore, to fill this research gap, i.e, to capture all information of spatio-temporal evolution embedded in data and achieve the maximum performance, in this paper, a novel GCN-based spatio-temporal prediction model is designed, which carries out prediction by exploiting the multi-characteristics of the spatio-temporal dynamic evolution of the oceanic variable fields, as to be detailed in our forthcoming discourses.

B. Contributions

In this paper, we propose a GCN-based spatio-temporal model to attain the high-accuracy and stable prediction of oceanic variables based on monthly data, which is termed as the OVPGCN. OVPGCN is constituted by three modules, namely the spatio-temporal correlation feature extraction (STCFE) module, bias correction (BC) module, and periodic dependency mining (PDM) module. With the aid of these modules, OVPGCN is enabled to fully extract the spatial correlations and temporal dependency existing in the data

of oceanic variables. In more detail, the gridded data of oceanic variable are mapped to the nodes of the graph data, as shown in Fig. 2. Based on the distance-based topology graph construction, the relationship graph of dynamic ocean patterns is designed to construct the dual graph as inputs to STCFE module. Furthermore, inspired by [41, 42] in the field of traffic forecasting, multiple ST-Conv blocks are introduced to extract the high-level features of both the spatio-temporal stationary and non-stationary variations in STCFE module. By contrast, the spatial differences between different sites are captured by the BC module, while the periodic characteristics of the oceanic variables are mined by the PDM module from the historical data. Finally, the outputs obtained from these three modules are integrated by a multi-module fusion (MMF) module to provide the final prediction results. In summary, the main contributions of this paper are stated as follows.

1) An OVPGCN framework for the prediction of oceanic variables is proposed, which enables to extract the multi-characteristics of the spatio-temporal dynamic evolution of oceanic variables. In contrast to the existing works focusing mainly on extracting the short-term spatio-temporal features, OVPGCN invokes a PDM module to mine the periodic variations of oceanic variables and a BC module to capture the spatial differences between various sites of the ocean.

2) A GCN-based framework is designed in OVPGCN to overcome the problem of data irregularity, via graph representation and modeling of large-scale global connections. The developed STCFE module in OVPGCN takes the dual graphs as inputs and is constituted by several spatio-temporal convolutional blocks, which enhance the feature extraction of spatio-temporal correlations.

3) The proposed OVPGCN is applied to both SST and Chl-a prediction. Extensive experiments are carried out to validate the functions of OVPGCN, and to demonstrate its superiority to the state-of-the-art methods.

The rest of the paper is organized as follows. Section II details the proposed OVPGCN framework for prediction of oceanic variables. Section III reports the experimental results and provides our analysis for SST and Chl-a prediction. Finally, we summarize the research observations and discuss the possible future work in Section IV.

C. Formulation of the Oceanic Variable Prediction Problem

The format of the oceanic variable dataset is usually gridded, according to the geographical locations, which can be naturally mapped to the nodes of the graph data, as shown in Fig. 2. In our work, we define $V = (v_1, v_2, \ldots, v_N)$ as the node set of an undirected graph $G = (V, E)$, where each node v_i is a valid grid, defined by its latitude and longitude, and E is a set of edges, defining the connections between nodes. Furthermore, all connectivity information is stored in an adjacency matrix $\boldsymbol{A} \in R^{N \times N}$, with $\boldsymbol{A}_{i,j} > 0$ if $(v_i, v_j) \in E$ and $\boldsymbol{A}_{i,j} = 0$ if $(v_i, v_j) \notin E$.

II. METHODS

In our study, the univariate prediction of oceanic variables is investigated. Taking the SST prediction as an example,
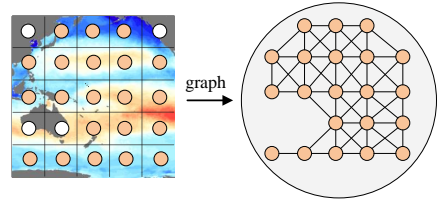

Fig. 2. An example of the gridded data mapped to a graph data. The locations of coasts and islands are considered to be invalid grids, as shown by the white grid points.

where SST is the only input variable and also the only output variable. Let $\boldsymbol{\mathcal{X}}_{1:\tau} = (\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_\tau) \in \mathbb{R}^{N \times \tau}$ be the dynamic feature sequences of N nodes over τ time steps, where $\boldsymbol{X}_t = (x_1^t, x_2^t, \ldots, x_N^t)^T \in \mathbb{R}^N$ denotes the feature value of this individual variable of the N nodes at time t. Then, the goal of oceanic variable prediction is to predict the future sequences over the following T_l time steps, which can be formulated as:

$$\widehat{\boldsymbol{\mathcal{X}}}_{\tau+1:\tau+T_l} = \arg\max_{\boldsymbol{\mathcal{X}}_{\tau+1:\tau+T_l}} P\left(\boldsymbol{\mathcal{X}}_{\tau+1:\tau+T_l} \mid \boldsymbol{\mathcal{X}}_{1:\tau}\right) \quad (1)$$

where $\widehat{\boldsymbol{\mathcal{X}}}_{\tau+1:\tau+T_l} = \left(\hat{\boldsymbol{X}}_{\tau+1}, \ldots, \hat{\boldsymbol{X}}_{\tau+T_l}\right)$ represents the predicted sequences from $\tau + 1$ to $\tau + T_l$.

A. Overview of Proposed OVPGCN

To further exploit the multi-characteristics embedded in the spatio-temporal dynamic evolution of oceanic variables, we propose an OVPGCN model based on the graph structure to achieve the prediction of oceanic variables. The overall framework of OVPGCN is shown in Fig. 3. In detail, the inputs to the OVPGCN are the dynamic graph sequences $\boldsymbol{\mathcal{X}}_{1:\tau}$, which are mapped from the historical gridded data of the oceanic variable. As shown in Fig. 3, $\boldsymbol{\mathcal{X}}_{1:\tau}$ are first fed into the graph construction (GC) module to generate three dynamic feature sequences of graph, and construct their corresponding adjacency matrices. Then, the spatio-temporal correlation feature extraction (STCFE), bias correction (BC), and the periodic dependency mining (PDM) modules are tailored to mine the multi-characteristics of the spatio-temporal evolution of the oceanic variables. The functionalities of these three modules can be highlighted as follows: (i) the most recently occurred ocean flows explicitly have an influence on the future ocean flows. Hence, the STCFE module is designed to mine the future complex spatio-temporal correlation with the past historical data. (ii) The nodes (sites) of different areas in the oceanic variable dataset have the spatial differences that may be caused by different factors, such as seabed topography and regional propagation conditions. When the number of samples is limited, it is often difficult for the STCFE module to fully explore the spatial differences of nodes. In this regard, the BC module is introduced to capture the spatial differences of the different nodes, and hence improve the prediction performance and robustness of the SCEE module. (iii) The PDM module is implemented to mine the periodic characteristic of the oceanic variable from the historical data so as to improve the accuracy of long-term prediction. After the processing of these three
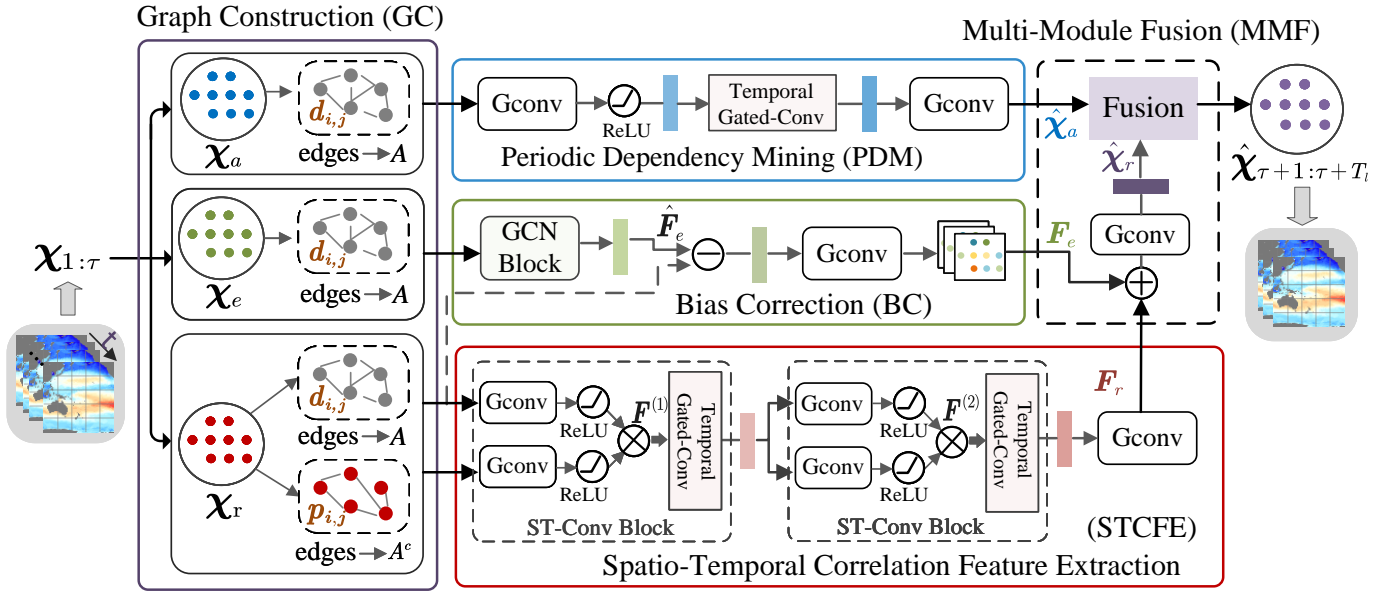
Fig. 3. Overall network structure of OVPGCN, including the GC, STCFE, BC, PDM, and MMF modules. Detailed definitions of these symbols, such as $\boldsymbol{\mathcal{X}}_a$, $\boldsymbol{\mathcal{X}}_e$, $\boldsymbol{\mathcal{X}}_r$, are given in Section II-B.
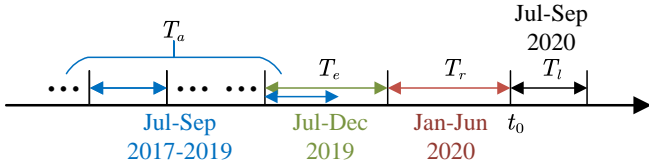

Fig. 4. An example for showing the construction of the time series segments. t_0 denotes the current or reference time and T_l denotes the size of predicting window. T_r corresponds to the time segment contained in sequence $\boldsymbol{\mathcal{X}}_r$, T_e corresponds to the time segment contained in sequence $\boldsymbol{\mathcal{X}}_e$, and T_a corresponds to the time segment contained in sequence $\boldsymbol{\mathcal{X}}_a$.

modules, their outputs are effectively merged to obtain the final prediction results in the multi-module fusion (MMF) module. As seen in Fig. 3, these colored rectangles represent the feature vectors generated after the previous operation. For instance, the first green rectangle in the BC module signifies the graph feature vector obtained by passing $\boldsymbol{\mathcal{X}}_e$ through the 'GCN Block'. Below we analyze these modules one-by-one in detail.

B. Graph Construction

In this part, we define the three feature sequences of graph nodes and their corresponding adjacency matrices, as shown in Fig. 3, to obtain the inputs to the STCFE, BC and PDM modules. It is assumed that the reference time is t_0 and the size of prediction window is T_l. Firstly, the feature sequences of graph nodes to be input to these three modules are denoted respectively as the recent segment ($\boldsymbol{\mathcal{X}}_r$), earlier segment ($\boldsymbol{\mathcal{X}}_e$), and annual-periodic segment ($\boldsymbol{\mathcal{X}}_a$), which are detailed as follows:

1) $\boldsymbol{\mathcal{X}}_r$ is comprised of the feature sequences of graph nodes over T_r time steps before t_0. As shown by the red part of Fig.

4, T_r is a segment of the historical time series directly adjacent to the time for prediction. In detail, $\boldsymbol{\mathcal{X}}_r$ is constructed as

$$\boldsymbol{\mathcal{X}}_r = (\boldsymbol{X}_{t_0-T_r+1}, \boldsymbol{X}_{t_0-T_r+2}, \ldots, \boldsymbol{X}_{t_0}) \in \mathbb{R}^{N \times T_r} \quad (2)$$

2) $\boldsymbol{\mathcal{X}}_e$ is comprised of the feature sequences of graph nodes over T_e time steps close to T_r, as shown by the green part of Fig. 4. In detail, $\boldsymbol{\mathcal{X}}_e$ is constructed as

$$\boldsymbol{\mathcal{X}}_e = (\boldsymbol{X}_{t_0-T_r-T_e+1}, \boldsymbol{X}_{t_0-T_r-T_e+2}, \ldots, \boldsymbol{X}_{t_0-T_r}) \in \mathbb{R}^{N \times T_e} \quad (3)$$

3) $\boldsymbol{\mathcal{X}}_a$ is comprised of the feature sequences of graph nodes over T_a time steps close to T_e. As shown by the blue part of Fig. 4, T_a is a segment of historical time series having the same month attributes as the forecasting time window T_l. Specifically, in Fig. 4, T_l is the time interval between July and September in 2020, and T_a represents the time segment between July and September in 2017, 2018 and 2019. In detail, $\boldsymbol{\mathcal{X}}_a$ is constructed as

$$\boldsymbol{\mathcal{X}}_a = (\boldsymbol{X}_{t_0-12*(T_a/T_l)+1}, \ldots, \boldsymbol{X}_{t_0-12*(T_a/T_l)+T_l},$$
$$\boldsymbol{X}_{t_0-12*(T_a/T_l-1)+1}, \ldots, \boldsymbol{X}_{t_0-12*(T_a/T_l-1)+T_l}, \ldots, \quad (4)$$
$$\boldsymbol{X}_{t_0-12+1}, \ldots, \boldsymbol{X}_{t_0-12+T_l}) \in \mathbb{R}^{N \times T_a}.$$

The original adjacency matrix $\boldsymbol{A}$ in Fig. 3 is generated based on the spatial distance between nodes. The spatial distance between nodes i and j is given by

$$d_{i,j} = \sqrt{\left(v_i^{lat} - v_j^{lat}\right)^2 + \left(v_i^{lon} - v_j^{lon}\right)^2} \quad (5)$$

where v_i^{lat} and v_i^{lon} denote the relative position explained by the latitude and longitude of node i. Then, the element of adjacency matrix $\boldsymbol{A}$ is computed as:

$$\boldsymbol{A}_{i,j} = \begin{cases} \frac{1}{d_{i,j}}, d_{i,j} \leq d_{min} \\ 0, \text{ otherwise} \end{cases} \quad (6)$$

where d_{min} is a threshold. The nodes located with the distance larger than d_{min} are considered to be disconnected in the graph.

In the ocean, the spatiotemporal correlation between nodes is not only dependant on the distance relied, but also contributed by the variation relationship of ocean patterns. Specifically, the correlations among nodes in graphs measured using topological distances are fixed. However, in the spatio-temporal dynamic evolution of the ocean, the correlations among nodes during different time segments exhibit variability. For instance, during the summer months with typhoons compared to the regular autumn months, nodes of the same distance on the central node have different correlations with that central node. Therefore, in our model, an additional adjacency matrix $\boldsymbol{A}^c$ is introduced to construct a graph $G^c = (V, E)$ to depict the ocean-pattern relationships. The ocean-pattern relationships graph can avoid the limitations of the above-mentioned distance-dependant graph and closely connect the nodes having similar variations. Specifically, the ocean-pattern relationship of these nodes is measured using the Pearson correlation coefficient of

$$p_{i,j} = \frac{\sum_{t=t_1}^{t_0} (\mathcal{X}_{r(t,i)} - \bar{\mathcal{X}}_{r(i)})(\mathcal{X}_{r(t,j)} - \bar{\mathcal{X}}_{r(j)})}{\sqrt{\sum_{t=t_1}^{t_0} \left(\mathcal{X}_{r(t,i)} - \bar{\mathcal{X}}_{r(i)}\right)^2}\sqrt{\sum_{t=t_1}^{t_0} \left(\mathcal{X}_{r(t,j)} - \bar{\mathcal{X}}_{r(j)}\right)^2}} \tag{7}$$

where $t_1 = t_0 - T_r + 1$, $\mathcal{X}_{r(i)}$ are the historical sequences of node i in $\mathcal{X}_r$, $\mathcal{X}_{r(t,i)}$ represents the value of $\mathcal{X}_{r(i)}$ at time t. $\bar{\mathcal{X}}_{r(i)}$ represents the average value of $\mathcal{X}_{r(i)}$ in the period from t_1 to t_0. Then, $\boldsymbol{A}^c$ has the element value of

$$\boldsymbol{A}_{i,j}^c = \begin{cases} p_{i,j}, & p_{i,j} \neq 0 \\ 0, & p_{i,j} = 0 \end{cases} \tag{8}$$

C. Spatio-Temporal Correlation Feature Extraction

The objective of the STCFE module is mainly to capture the spatio-temporal correlation existing between the recent segment $\mathcal{X}_r$ and the sequence to be predicted. In our paper, by introducing the relationship graph $G^c = (V, E)$ of dynamic ocean patterns to this module, it is capable of avoiding the deficiency of the distance-based topology graph, which captures the spatial-temporal correlation based only on the distance from neighbors. Furthermore, as shown in Fig. 3, multiple ST-Conv blocks are designed to extract the high-level features of both the spatio-temporal stationary and non-stationary variations. Owing to these characteristics, the capability of the STCFE module to mine the spatio-temporal correlation features can be improved.

In detail, as shown in Fig. 3, the STCFE module consists of the multiple ST-Conv Blocks and a graph convolutional (Gconv) layer. In each of the ST-Conv Blocks, there are two parallel Gconv layers for capturing the spatial dependency with neighborhoods and a 'Temporal Gated-Conv' layer for exploiting the temporal dependency with nearby observations along the temporal dimension. Specifically, the Gconv layer performs the spatial-domain graph convolutional operation [43], expressed as:

$$\boldsymbol{h}_{v_i}' = \boldsymbol{W}_1 \boldsymbol{h}_{v_i} + \boldsymbol{W}_2 \sum_{j \in \mathcal{N}(i)} e_{i,j} \cdot \boldsymbol{h}_{v_j} \tag{9}$$

where $\boldsymbol{h}_{v_i}'$ is the updated representation of node v_i by aggregating its own features $\boldsymbol{h}_{v_i}$ with neighbors features $\boldsymbol{h}_{v_j}$, $e_{i,j}$ denotes the edge weight between node v_i and node v_j, $\mathcal{N}(i) = \{v_j \in V | e_{i,j} \in E\}$ represents the neighborhood of node v_i, and finally $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ are the learnable parameters.

In our method, the function $f_G(\cdot)$ is used to represent the Gconv layer, which performs the operations on the distance topology graph, with the value of $e_{i,j}$ given by $\boldsymbol{A}_{i,j}$. By contrast, the function $f_{G^c}(\cdot)$ is for the Gconv layer of the lower branch in ST-ConvBlock, which is operated on the relationship graph G^c of the dynamic ocean patterns, with the value of $e_{i,j}$ is denoted by $\boldsymbol{A}_{i,j}^c$. Therefore, the feature map to the input of the first 'Temporal Gated-Conv' layer can be obtained as:

$$\boldsymbol{F}^{(1)} = (ReLU(f_G(\mathcal{X}_r)) + ReLU(f_{G^c}(\mathcal{X}_r))) \odot \boldsymbol{W}_3 \tag{10}$$

where $\odot$ denotes the element-wise Hadamard product, $\boldsymbol{W}_3$ is a learnable parameter matrix and the rectified linear unit (ReLU) activation function is used.

Inspired by [44], the entire convolution structure on time axis is introduced to capture the temporal dynamic features of ocean variations. In detail, the 'Temporal Gated-Conv' layer contains a 1-D causal convolution (1-D Conv) operation followed by the gated linear units (GLU) as a non-linearity function [41], which can be written as:

$$f_\Gamma(\boldsymbol{F}^{(k)}) = (\boldsymbol{F}^{(k)} + (\Phi * \boldsymbol{F}^{(k)})) \odot \sigma(\Phi * \boldsymbol{F}^{(k)}) \tag{11}$$

where the function $f_\Gamma(\cdot)$ is used to represent the operation of the Temporal Gated-Conv layer, $\boldsymbol{F}^{(k)}$ denotes the feature map of the k_{th} ST-Conv Block input to the 'Temporal Gated-Conv' layer, '$*$' represents the convolution operation, Φ is the parameters of the convolution kernel, and $\sigma(\cdot)$ is the sigmoid function.

After stacking the two ST-Conv Blocks in Fig.3, an extra Gconv layer is implemented as the output layer in the STCFE module, giving the output expressed as:

$$\boldsymbol{F}_r = f_G(f_\Gamma(\boldsymbol{F}^{(2)})) \tag{12}$$

D. Bias Correction

In [34], the natural differences among different sites are considered so as to correct the prediction results by adding a self-defined and fixed bias correction map. However, designing this kind of maps requires the specialized domain knowledge, and different maps have to be designed for different prediction areas. Moreover, the performance of prediction varies in different time periods due to these natural differences. Therefore, to mitigate the challenge while simultaneously improve the accuracy and robustness, we design a BC module, as shown in Fig. 3, to learn the spatial differences, and correspondingly output an adaptive correction map for correcting the prediction results.

In detail, as shown in Fig. 3, $\mathcal{X}_e$ is the input to the 'GCN Block' in the BC module, which produces the feature vector $\hat{\boldsymbol{F}}_e$ having the same size as the recent-time segment $\mathcal{X}_r$. The 'GCN Block' consists of three cascaded Gconv layers, each of which is followed by a ReLU function. Next, the network generated $\hat{\boldsymbol{F}}_e$ is subtracted from the real data $\mathcal{X}_r$ to obtain a spatial difference feature matrix. Then, this matrix is sent to

the Gconv layer, which further mines the spatial relationship to finally produce the bias correction map as

$$\boldsymbol{F}_e = f_G(\boldsymbol{\mathcal{X}}_r - \hat{\boldsymbol{F}}_e) \tag{13}$$

E. Periodic Dependency Mining

The function of this module is to capture the annual periodic features of the oceanic variables in historical data, so as to generate the future sequence $\widehat{\boldsymbol{\mathcal{X}}}_a$ for improving the long-term prediction results. The motivation here can be highlighted as follows. The evolution of oceanic variables presents the periodic nature and can be exploited to enhance prediction accuracy. However, the existing prediction methods only use the recent data for forecasting, and hence it is difficult to mine the periodicity for achieving better performance. For example, when the 6-lead prediction is considered, using the data from December to February to predict an oceanic variable in August may lead to undesirable performance. This is because the evolution characteristics of an oceanic variable in winter is different from that in the predicted month in summer. In contrast, the August to be predicted has similar evolutionary characteristics to the historical Augusts. Note that here a possible solution is that we can directly feed all the historical data to the model, and let the model automatically learn the periodicity. For example, we can input the historical data of 3 years instead of 3 months (December to February) to make prediction. However, for this kind of schemes, it may be challenging to train the network in the case of insufficient data samples and furthermore the performance is limited by the model complexity. Therefore, in this paper, we only use the same historically segmental time series as the segment to be predicted so that the future sequences can be predicted using a simple network. To be more specific, as shown in Fig. 3, the PDM module consists of a Gconv layer followed by a ReLU function, a 'Temporal Gated-Conv' layer and an output Gconv layer to jointly process the graph-structured time series. The predicted sequence $\widehat{\boldsymbol{\mathcal{X}}}_a$ is computed as:

$$\hat{\boldsymbol{\mathcal{X}}}_a = f_G(f_\Gamma(ReLU(f_G(\boldsymbol{\mathcal{X}}_a)))) \tag{14}$$

F. Multi-Module Fusion

Finally, as shown in Fig. 3, the MMF module is operated to integrate the outputs of the STCFE, BC, and PD, to provide the final prediction for the whole lead time series. First, the output $\boldsymbol{F}_e$ of BC module and the output $\boldsymbol{F}_r$ of STCFE module are combined, and the output of which is further processed using a Gconv layer, yielding the predicted sequence

$$\hat{\boldsymbol{\mathcal{X}}}_r = f_G(\boldsymbol{F}_e + \boldsymbol{F}_r) \tag{15}$$

After the bias correction, the impact of periodic characteristics is then taken into account. In practice, the variation of some forecasting sub-regions presents the periodic patterns and seasonal differences, and in this case, the output of the PD module is more important. On the other side, the periodic characteristics is not obvious in some sub-regions. In this case, the output of the STCFE module mainly mining the short-term spatio-temporal evolution is more crucial. Consequently, when the outputs of $\hat{\boldsymbol{\mathcal{X}}}_r$ and $\hat{\boldsymbol{\mathcal{X}}}_a$ are fused for forecasting different

sub-regions and time series, they should be weighted differently via the learning from the historical data. Hence, when taking account of the above-mentioned, the final prediction results provided by the MMF module can be written as

$$\widehat{\boldsymbol{\mathcal{X}}}_{\tau+1:\tau+T_l} = \boldsymbol{W}_r \odot \widehat{\boldsymbol{\mathcal{X}}}_r + \boldsymbol{W}_a \odot \widehat{\boldsymbol{\mathcal{X}}}_a \tag{16}$$

where $\boldsymbol{W}_r$ and $\boldsymbol{W}_a$ are learning parameters, reflecting the degrees of influence of $\hat{\boldsymbol{\mathcal{X}}}_r$ and $\hat{\boldsymbol{\mathcal{X}}}_a$ on the final prediction results.

To train the OVPGCN network, the loss function introduced is expressed as

$$\mathcal{L} = \sum_{t=1}^{T_l} \left\| \widehat{\boldsymbol{\mathcal{X}}}_{\tau+t-1:\tau+t} - \boldsymbol{\mathcal{X}}_{\tau+t-1:\tau+t} \right\|^2 \tag{17}$$

where $\boldsymbol{\mathcal{X}}_{\tau+t-1:\tau+t}$ is the data of ground truth. For clarity, the implementation procedures of our OVPGCN algorithm are summarized as Algorithm 1.

Algorithm 1: Training of OVPGCN Algorithm

Input: Historical sequence $\boldsymbol{\mathcal{X}}_{1:\tau}$, prediction step size T_l, label $\boldsymbol{\mathcal{X}}_{\tau+1:\tau+T_l}$, maximum epoch γ, and learning rate λ.

Output: Predicted future sequence $\widehat{\boldsymbol{\mathcal{X}}}_{\tau+1:\tau+T_l}$.

1 **Step 1:** *Graph Construction (GC)*
2 Construct three feature sequences: $\boldsymbol{\mathcal{X}}_r$, $\boldsymbol{\mathcal{X}}_e$ and $\boldsymbol{\mathcal{X}}_a$;
3 Construct their corresponding adjacency matrices: $\boldsymbol{A}$ or $\boldsymbol{A}^c$, using Eq.(5)-(8).
4 **Step 2:** *Training OVPGCN*
5 Initialize STCFE, BC, PDM and MMF modules;
6 **for** *each traing epoch* **do**
7   a) Generate spatio-temporal correlation features $\boldsymbol{F}_r$, based on Eq. (9)-(12), using STCFE module;
8   b) Generate bias correction map $\boldsymbol{F}_e$, as equation (13), using the BC module;
9   c) Generate the predicted sequence $\widehat{\boldsymbol{\mathcal{X}}}_a$, based on Eq. (14), using the PDM module;
10   d) Fuse $\boldsymbol{F}_r$, $\boldsymbol{F}_e$, and $\widehat{\boldsymbol{\mathcal{X}}}_a$, according to Eq. (15) and (16), using the MMF module;
11   e) Predict $\widehat{\boldsymbol{\mathcal{X}}}_{\tau+1:\tau+T_l}$ and update modules' parameters using gradient descent.
12 **end**

III. EXPERIMENTS

A. Data and Study Area

In this study, we consider the spatio-temporal prediction of two oceanic variables, namely SST and Chl-a, that are commonly considered in the open literature. The studying areas are the Bohai Sea and the Northern South China Sea (NSCS). Note that, the Bohai Sea is located in the north of China, bounded by 37° - 41°N and 117.5° - 121.5°E, and the NSCS is bordered by the mainland of China in the west area, bounded by 17° 25° N and 108° 121°E.

The monthly mean SST data used in the experiment are obtained from the ERAInterim of European Centre for Medium-range Weather Forecasts (ECMWF). The temporal coverage of
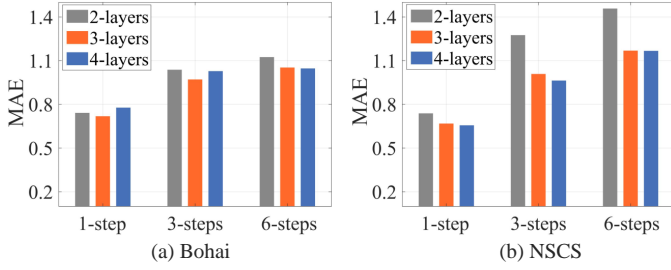
Fig. 5. MAE of the prediction results when different numbers of Gconv layers are respectively employed.

TABLE I
ABLATION EXPERIMENT OF ST-CONV BLOCK IN BOHAI SEA

Module	MAE(°C)		
	1-step	3-steps	6-steps
(i) w/o TGL, w/o G^c	0.725	0.971	1.098
(ii) w TGL, w/o G^c	0.712	0.967	1.015
(iii) w TGL, w G^c	0.690	0.873	0.989

TABLE II
ABLATION EXPERIMENT OF ST-CONV BLOCK IN NSCS

Module	MAE(°C)		
	1-step	3-steps	6-steps
(i) w/o TGL, w/o G^c	0.676	0.978	1.227
(ii) w TGL, w/o G^c	0.628	0.965	1.083
(iii) w TGL, w G^c	0.578	0.959	1.037

the SST datasets is from January 1979 to August 2019 with a spatial resolution of $0.125°$. The monthly mean surface Chl-a 5.0 data are downloaded from the Ocean Colour Climate Change Initiative (OC-CCI) with a spatial resolution of 4 km. The temporal coverage of the Chl-a datasets spans the period from January 1998 to December 2020. As there are missing data in the Chl-a 5.0 dataset, the Data Interpolating Empirical Orthogonal Functions (DINEOF) method [45] is introduced in this study to generate the estimate for the missing Chl-a data because of various improper conditions, such as the cloud obscuring measurements and the contamination of high sun glint. The effectiveness of DINEOF has been studied in detail in [46].

B. Experiment Settings for Prediction

The experimental settings for the SST and Chl-a prediction are as follows. All our experiments are operated on a single NVIDIA 1070 GPU, and the PyTorch framework is used to build the neural network. The learning rate is set to 0.001. The size of the mini-batch is set to 4, and the model is trained by 100 epochs. We select 70% of the data as the training samples, while the remaining 30% of the data are used as the testing samples. During training, an adaptive moment estimation (Adam) optimizer is applied. The implementation of our model and that of the baseline methods related to GCN are based on PyTorch Geometric (PYG) library [47]. Note that in our work, d_{min} is set to $\sqrt{2}$. As shown in Fig. 4, T_r and T_e are set to 6, and T_a is three times the size of T_l. We use 1-step, 3-steps, and 6-steps prediction to represent the lead prediction of 1 month, 3 months, and 6 months, i.e., we have T_l equaling to 1, 3, and 6. Additionally, $p_{i,j}$ is set to 0 if the absolute value of $p_{i,j}$ less than 0.3, to reduce computation.

1) Baseline Models: The following neural network methods are selected as the baselines for comparison with our model, which are:

(a) CNN [31]: The CNN model is a three-layer network structure for forecasting ENSO [31], which is highly correlated to SST prediction. Note that the CNN model has been slightly modified by removing the max-pooling layer and fully connected layer to achieve the prediction for all grid points in two-dimensional area.

(b) CNN-LSTM [48]: The CNN-LSTM, consisting of three convolutional layers, a flattening layer, and an LSTM layer, was proposed to predict Chl-a.

(c) ConvLSTM [22]: ConvLSTM was proposed to overcome the drawbacks of LSTM in handling the spatio-temporal data. The ConvLSTM-assisted baseline model is the same as that in [22].

(d) GCN [35]: GCN treats data as an undirected graph and uses a layer-wise propagation rule that is based on a first-order approximation of the spectral convolutions on graphs. This basline model consists of three layers of GCN and ReLU functions.

(e) GraphSAGE [39]: GraphSAGE carries out the graph convolution operations by aggregating the information of neighboring nodes. This basline model consists of three layers of GraphSAGE and ReLU functions.

(f) TSGN [40]: TSGN uses a LSTM network to aggregate the features of time series data and establishes a graph neural network model to jointly capture the graph-relied spatial correlation and temporal dynamics. TSGN was used to predict SST in the Pacific Northwest [40].

During the experiments, the six baseline models are trained using the same hyperparameters, as above-mentioned during the OVPGCN training.

2) Performance Metrics: Three performance metrics are selected to measure the performance of the SST and Chl-a prediction, which are the root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE), given as

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(p_i - o_i)^2}{n}} \tag{18}$$

$$\text{MAE} = \frac{\sum_{i=1}^{n}|p_i - o_i|}{n} \tag{19}$$

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{p_i - o_i}{o_i}\right| \tag{20}$$

respectively, where p_i is the predicted value, o_i is the true value, and n denotes the total number of testing samples.

C. Ablation Study

1) Analysis of Gconv Layers: The SST datasets of the Bohai Sea and the NSCS are selected for the ablation experiment analysis. The influence of the different number of

TABLE III
RMSE(°C) PERFORMANCE OF STCFE, BC, AND PD MODULE ON THE SST DATA

STCFE	BC	PD	Bohai Sea			NSCS		
			1-step	3-steps	6-steps	1-step	3-steps	6-steps
			0.942	1.272	1.413	0.870	1.298	1.601
✓			0.876	1.116	1.255	0.728	1.197	1.296
✓	✓		0.835	1.069	1.171	0.816	0.829	0.987
✓	✓	✓	0.829	0.948	1.024	0.648	0.687	0.761

TABLE IV
MAE(°C) PERFORMANCE OF STCFE, BC, AND PD MODULE ON THE SST DATA

STCFE	BC	PD	Bohai Sea			NSCS		
			1-step	3-steps	6-steps	1-step	3-steps	6-steps
			0.725	0.971	1.098	0.676	0.978	1.227
✓			0.690	0.873	0.989	0.578	0.959	1.037
✓	✓		0.666	0.847	0.915	0.665	0.626	0.761
✓	✓	✓	0.643	0.739	0.796	0.471	0.537	0.606

graph convolutional layers on the prediction performance is first investigated. Then, the ablation experiments of the ST-Conv Block are carried out. The above-mentioned experiments are conducted based only on the STCFE module to easily select the appropriate parameters. Furthermore, we carry out extensive ablation study in the context of OVPGCN to demonstrate the effectiveness of the STCFE, BC, and PD modules.

In this part, only the input $\mathcal{X}_r$ is utilised to analyze the influence of the number of Gconv layers on the prediction result. The comparison of MAE with different numbers of Gconv layers is shown in Fig. 5. As shown by the results, for both the Bohai Sea and NSCS data, the network with 2-layers achieves the relatively poorer performance, especially for the Bohai Sea data. By contrast, the prediction performance of the networks with 3-layers and 4-layers are similar. The results reveal that adding more layers may not always improve the prediction performance, which is due to the degradation problem of the deep learning networks [21]. Therefore, when also considering the implementation complexity, the network with the 3 Gconv layers is chosen as a suitable network for further experiments.
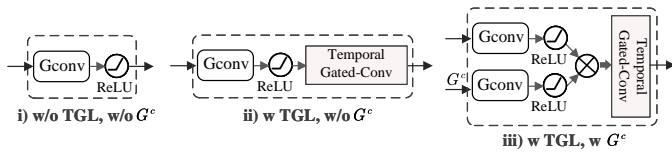

Fig. 6. The structure of three modified ST-Conv Blocks.

2) Analysis of ST-Conv Block: To analyze the necessity of the 'Temporal Gated-Conv' layer and two parallel Gconv layers with the added relationship graph G^c, three modified ST-Conv Blocks in the STCFE module are considered, which are shown in Fig. 6 and explained as follows:

i) **w/o TGL, w/o** G^c: It denotes the modified ST-Conv Block including only one Gconv layer with a ReLU function. Thus, the network in the STCFE module consisting of two cascade Block i) and one Gconv layer is referred as a basic network;

ii) **w TGL, w/o** G^c: It adds the 'Temporal Gated-Conv' layer on the basis of Block i);

iii) **w TGL, w** G^c: It adds the relationship graph G^c and two parallel Gconv layers on the basis of Block ii). This structure corresponds to the architecture of the ST-Conv Block utilized in the OVPGCN model as depicted in Fig. 3.

The results are shown in Table I and Table II. From the comparison of i) and ii), we can see that the MAE of prediction for both study area is reduced by adding the 'Temporal Gated-Conv' layer. Specifically, the MAE is decreased by 0.013, 0.004, and 0.083, respectively, when 1-step, 3-steps and 6-steps prediction are carried out for the Bohai Sea data. Correspondingly, for the NSCS data, the MAE is decreased by 0.048, 0.013, and 0.144, respectively. Hence, the 'Temporal Gated-Conv' layer is capable of mining the temporal dependencies as well as the correlations existing in both the temporal domain and spatial domain, which can be exploited to improve the prediction accuracy. As shown in Table I and Table II, using the Block iii), i.e., the ST-Conv Block in our OVPGCN, achieves the best performance among the three cases. Therefore, by further introducing the dynamic graph G^c, the prediction performance can be improved. This is because when multiple ST-Conv blocks are implemented, the network can enhance the connectivity of nodes and take full advantages of the stationary and non-stationary correlation characteristics of the spatio-temporal data.

3) Performance of STCFE, BC, and PDM Modules: To show the performance of these modules, a range of experiments are conducted. The experimental results of RMSE and MAE are shown in Table III and Table IV, respectively, for the Bohai Sea and NSCS. Note that the first rows in the tables correspond to a initial standard 3-layer Gconv network without adding any other modules, which is defined as a basic network. By contrast, the other rows in Table III and Table IV correspond to the basic network with added STCFE, STCFE and BC, and STCFE, BC and PDM, respectively, as indicated in the tables. From the results shown in Table III and Table IV, several groups of observations can be obtained as follows.

TABLE V
PREDICTION RESULTS OF SST

Models	Metrics	Bohai Sea			NSCS		
		1-step	3-steps	6-steps	1-step	3-steps	6-steps
CNN	RMSE	1.163	1.623	1.703	0.966	1.710	1.976
	MAE	0.916	1.260	1.325	0.769	1.405	1.653
	MAPE(%)	0.317	0.438	0.461	0.257	0.469	0.551
CNN-LSTM	RMSE	1.085	1.240	1.408	0.819	1.057	1.163
	MAE	0.858	0.977	1.097	0.654	0.855	0.944
	MAPE(%)	0.300	0.342	0.384	0.219	0.285	0.315
ConvLSTM	RMSE	0.998	1.229	1.280	0.764	0.954	1.059
	MAE	0.776	0.943	1.014	0.587	0.738	0.831
	MAPE(%)	0.270	0.330	0.356	0.197	0.247	0.278
GCN	RMSE	1.251	1.560	1.590	1.409	1.734	1.782
	MAE	0.908	1.165	1.209	1.014	1.311	1.346
	MAPE(%)	0.315	0.406	0.422	0.339	0.439	0.451
GraphSAGE	RMSE	0.933	1.267	1.314	0.858	1.335	1.506
	MAE	0.719	0.971	1.053	0.668	1.009	1.169
	MAPE(%)	0.250	0.339	0.369	0.223	0.374	0.392
TSGN	RMSE	0.879	1.162	1.215	0.793	1.295	1.366
	MAE	0.676	0.893	0.937	0.608	0.943	1.008
	MAPE(%)	0.236	0.313	0.328	0.204	0.316	0.338
OVPGCN	RMSE	**0.829**	**0.948**	**1.024**	**0.600**	**0.687**	**0.761**
	MAE	**0.643**	**0.739**	**0.796**	**0.471**	**0.537**	**0.606**
	MAPE(%)	**0.226**	**0.260**	**0.280**	**0.158**	**0.180**	**0.203**

First, the basic network performs the worst, and adding the STCFE module to the basic network is able to increase the prediction performance of the basic network. Specifically, when the STCFE module is employed for the 1-step prediction, the RMSE can be improved by 0.066 and 0.142, and the MAE can be improved by 0.035 and 0.098, respectively, for the Bohai Sea and NSCS data. Hence, the results verify that the STCFE module is capable of mining the spatio-temporal dynamic correlations existing in the data sequences. Furthermore, when the BC module is employed plus the STCFE module, the prediction performance can be enhanced in most cases. Specifically, when the 3-steps prediction is considered, the RMSE can be improved by 0.047 and 0.368, and the MAE can be improved by 0.026 and 0.333, respectively, for the Bohai Sea and NSCS data. Therefore, it is effective to correct the spatial differences. However, as seen in Table III and Table IV, the prediction performance of using 1-step prediction for NSCS data gets worse in terms of both RMSE and MAE. The reason behind this may be that the 1-step prediction is relatively simple and the temperature change in the NSCS is smaller than that in the Bohai Sea, leading to the over-fitting problem of the model. Finally, when further adding the PDM module, the prediction results demonstrate that the periodic patterns existing in the data can be mined for improving the prediction accuracy. Therefore, with the aid of the STCFE, BC and PD modules, the multi-characteristics of the ocean spatio-temporal evolution can be fully exploited for improving the prediction performance.

D. Experiment Results and Analysis for SST Prediction

Next, we compare our method with six baseline algotirhms, as mentioned in Section III-B, for the SST prediction in the Bohai Sea and NSCS.

Table V shows the RMSE, MAE, and MAPE values obtained by the different prediction algorithms based on the
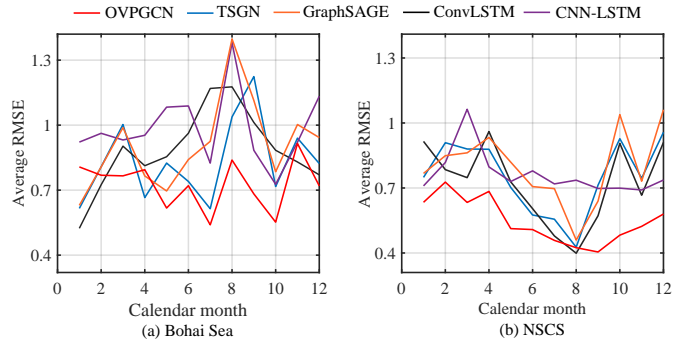

Fig. 7. RMSE comparison of SST prediction in different calendar months between OVPGCN, TSGN, GraphSAGE, ConvLSTM and CNN-LSTM: (a) Bohai Sea and (b) NSCS.

datasets of the Bohai Sea and NSCS. Note that the Kelvin (K) is used to calculate the MAPE of SST to avoid some cases that the true values are close to $0°C$, resulting in a very large MAPE value. From Table V, we can see that for given prediction steps, CNN yields the poorest performance, as it has a simple structure with finite feature extraction capability, and in particular, it is unable to exploit the chronological dependency of the input data. Both CNN-LSTM and ConvLSTM outperform CNN, owing to their introduction of the LSTM network to mine the temporal dependency in the SST data. Moreover, as shown in Table V, ConvLSTM is slightly better than CNN-LSTM.

All the above-mentioned three methods extract the spatial correlation features through the convolution operation. By contrast, the other four methods in Table V carry out the graph convolution operation, via modeling the large-scale global connections as the edges of the graph. As the results shown in Table V, the performance obtained by GraphSAGE is significantly better than that obtained by GCN. This implies that the spatial domain convolution in GraphSAGE is better
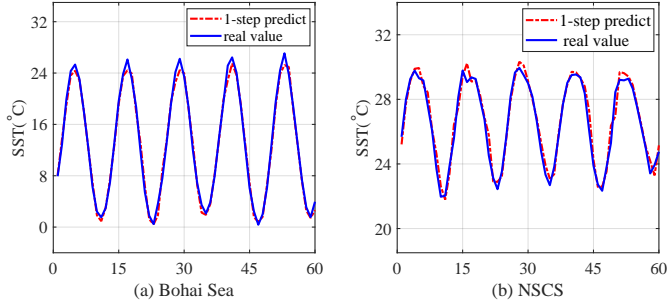
Fig. 8. Temporal trend of the real and OVPGCN predicted SST values for 1-step prediction: (a) Bohai Sea and (b) NSCS.
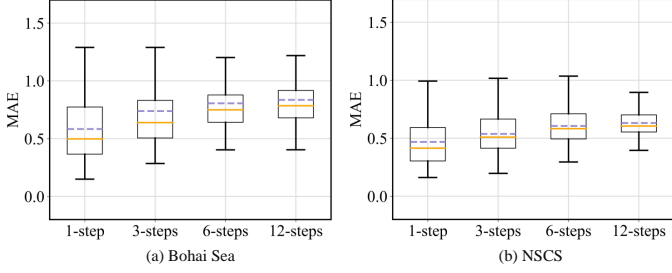

Fig. 9. The box plot of the MAE comparison for 1-step, 3-steps, 6-steps, and 12-steps SST prediction by OVPGCN.
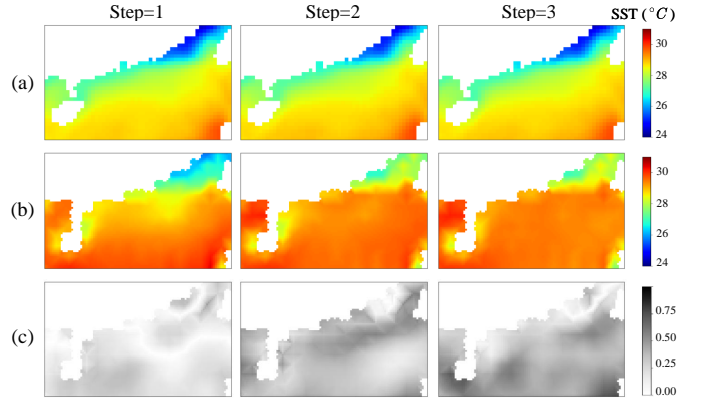

Fig. 10. Example of qualitative results obtained by the OVPGCN for the SST predictions of next 3-steps (months) in NSCS. (a) Predicted SST fields. (b) True SST fields. (c) Deviation between observed and predicted SST fields.

than the spectral domain convolution in GCN, when they are employed to perform SST prediction.

From Table V, we can explicitly see that our proposed OVPGCN outperforms all the legacy baseline algorithms and attains the best performance. Specifically, for the Bohai Sea, it outperforms the TSGN, which provides the best prediction results among the six baseline algorithms, about 0.050, 0.214 and 0.191 of the RMSE, when the 1-step, 3-steps, and 6-steps predictions are implemented. In terms of MAPE, our OVPGCN method outperforms TSGN by 0.010, 0.053, and 0.048, also when these predictions are considered. Similarly, for the NSCS area, when the 1-step, 3-steps, and 6-steps predictions are implemented, it outperforms the TSGN by 0.193, 0.608, and 0.605 of RMSE, and by 0.046, 0.136, and 0.135 of MAPE. Based on the experimental results shown in Table V, we can conclude that our approach is in general capable of outperforming the six legacy schemes for the SST prediction, owing to its capability to effectively mine the information about the multi-characteristics of oceanic spatio-temporal evolution.

To further illustrate the performance of the proposed OVPGCN, Fig. 7 compares the average RMSE attained by OVPGCN and other four baseline algorithms in different calendar months for 1-step prediction. For both the Bohai sea and NSCS, compared to the other four baseline algorithms, it can be observed that the proposed OVPGCN has the lowest RMSE in most cases. Furthermore, these four baseline algorithms exhibit notable variations in performance across different calendar months. For instance, for the Bohai Sea data, TSGN's predictive performance is closest to that of OVPGCN, with its best performance attained in April but the highest RMSE exhibited in September. By contrast, for the NSCS data,

the ConvLSTM generally outperforms the other four baseline algorithms. Furthermore, the OVPGCN performs better on the NSCS data than it does on the Bohai Sea data. The reason behind is that, the seasonal SST variation in the Bohai Sea is larger than that in the NSCS, resulting in that it is more difficulty to predict the SST in the Bohai Sea area. From these results, it can be implied that the proposed OVPGCN model enables a relatively stable SST prediction performance throughout a year. Correspondingly, the temporal trends of the real and predicted SST values of OVPGCN from the last five years of the test sample are present in Fig. 8. It can be observed that the predicted results by the OVPGCN fit well with the real SST data. As shown by the results, the real SST exhibits clearly a monthly periodic characteristics. Our proposed method can efficiently extract this periodicity, resulting in a good periodic fitting with the real SST data.

Next, Fig. 9 shows the box plot for the MAE comparison of OVPGCN for 1-step, 3-steps, 6-steps, and 12-steps SST prediction. It can be observed that, as the step size of prediction increases, the MAE error of the OVPGCN model does not exhibit a rapid increase. Moreover, for short, medium, and long-term prediction, the distribution of MAE remains relatively concentrated, without significant deviations. Hence, the OVPGCN model can achieve stable long-term SST prediction.

Fig. 10 shows an example of the spatial distributions of SST prediction in NSCS for the next 3 steps, namely for predicting the SST of July, August and September using the data before July. From top to bottom, the predicted SST fields by the proposed OVPGCN, the true SST fields, and the deviation between them are shown sequentially. The results demonstrate that the difference between the observed true values and the predicted results are in general small and acceptable, even when the prediction deviation becomes relatively larger, as the step size of prediction becomes bigger.

E. Experiment Results and Analysis for Chl-a Prediction

Chl-a is a paramount variable in the ocean, which is used as a key indicator of ecosystem changes. Therefore, in this section, we depict the performance of OVPGCN for Chl-a

TABLE VI
PREDICTION RESULTS OF CHL-A

Models	Metrics	Bohai Sea			NSCS		
		1-step	3-steps	6-steps	1-step	3-steps	6-steps
CNN	RMSE	2.570	2.957	3.057	1.002	1.087	1.103
	MAE	1.525	1.764	1.829	0.271	0.307	0.313
	MAPE(%)	1.394	1.612	1.673	0.255	0.289	0.295
CNN-LSTM	RMSE	2.147	2.248	2.110	0.778	0.792	0.811
	MAE	1.350	1.440	1.319	0.295	0.313	0.299
	MAPE(%)	1.258	1.230	0.384	0.284	0.303	0.288
ConvLSTM	RMSE	2.052	2.398	2.683	0.835	0.916	0.989
	MAE	1.179	1.578	1.774	0.252	0.316	0.329
	MAPE(%)	1.092	1.469	1.650	0.240	0.302	0.315
GCN	RMSE	2.549	2.859	2.868	1.127	1.157	1.150
	MAE	1.302	1.668	1.693	0.294	0.328	0.327
	MAPE(%)	1.190	1.538	1.562	0.277	0.311	0.310
GraphSAGE	RMSE	2.139	2.583	2.582	0.870	0.972	0.971
	MAE	1.239	1.554	1.577	0.254	0.292	0.294
	MAPE(%)	1.146	1.440	1.460	0.242	0.278	0.280
TSGN	RMSE	2.117	2.532	2.523	0.875	0.966	0.970
	MAE	1.200	1.507	1.507	0.256	0.288	0.291
	MAPE(%)	1.108	1.395	1.394	0.242	0.274	0.277
OVPGCN	RMSE	**1.727**	**1.852**	**1.911**	**0.738**	**0.718**	**0.711**
	MAE	**0.991**	**1.081**	**1.106**	**0.216**	**0.222**	**0.225**
	MAPE(%)	**0.918**	**1.003**	**1.023**	**0.206**	**0.212**	**0.215**

prediction and compare it with other benchmarks. Table VI shows the RMSE, MAE and MAPE results obtained by our proposed OVPGCN and the baseline prediction algorithms on the Chl-a prediction. From the results of Table VI, we can explicitly see that our proposed OVPGCN algorithm outperforms all the other benchmarks in terms of these three performance metrics. Specifically, for the Bohai Sea, it outperforms the TSGN, respectively, by 0.390, 0.680 and 0.612 in terms of RMSE, when the 1-step, 3-steps, and 6-steps predictions are implemented. For the NSCS area, it outperforms the TSGN, respectively, by 0.137, 0.248, and 0.259 in terms of RMSE, again when the 1-step, 3-steps, and 6-steps predictions are considered. Besides, from the results shown in Table VI, we can observe that for the multi-steps Chl-a prediction, the MAE, RMSE, and MAPE values obtained in the prediction of the NSCS area are lower than the corresponding values in the prediction of the Bohai Sea area. This is because about 84% of the Chl-a values in the NSCS data fall in a small range of 0-1 $mg \cdot m^{-3}$, making the model obtain very accurate predictions reflected by the low these three metric values.

From the results given in Table VI, we can conclude that in general, the performance of GraphSAGE is better than that of CNN, explaining that the graphic convolutional operation is capable of mining the spatial correlation for Chl-a prediction. Moreover, TSGN is in general superior to GraphSAGE, which reflects that implementing an LSTM network to aggregate the neighboring node features enables to capture the temporal information. The results in Table VI show that the performance achieved by the six baseline methods, especially in the cases for the three CNN-based methods, is not very stable when the predictions of multi-step are carried out. Specifically, this is reflected by the inconsistent results of the multi-step predictions. As shown in Table VI, TSGN performs better than ConvLSTM in both the 3-steps and 6-steps predictions, but worse than ConvLSTM in the 1-step prediction. The
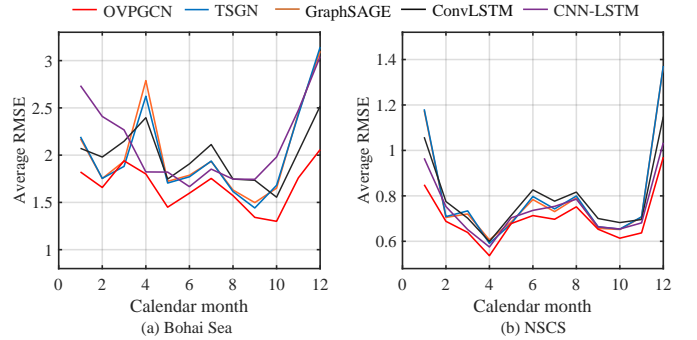

Fig. 11. RMSE comparison of Chl-a prediction in different calendar months between OVPGCN, TSGN, GraphSAGE, ConvLSTM and CNN-LSTM: (a) Bohai Sea and (b) NSCS.

unstable performance is also reflected in the differences for the prediction of different areas. While CNN achieves the poorest performance in the prediction of Bohai Sea, it outperforms GCN in the prediction of NSCS. By contrast, TSGN outperforms CNN-LSTM, in terms of MAE and MAPE, but generates a higher RMSE than CNN-LSTM when the NSCS prediction is considered. The reason behind may be that the relatively small number of Chl-a training samples makes it difficult for these baseline algorithms to be fully optimized. Additionally, the range of Chl-a data in the NSCS is relatively large, with the values distributed in 0.01 - 126.85 $mg \cdot m^{-3}$, but most of them below 1 $mg \cdot m^{-3}$, which results in the inconsistency of the three performance indicators. Finally, the proposed OVPGCN is capable of exploiting the spatial differences and mining the spatial-temporal correlations from multiple perspectives, making it achieve stable prediction, even when there are only a small number of samples are available. Consequently, no matter which area is predicted and what the prediction lead time is, the OVPGCN is highly effective to mine the embedded spatio-temporal information to attain the
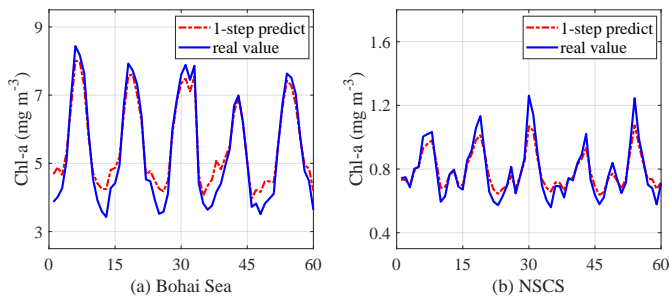
Fig. 12. Temporal trend of the real and predicted SST values for 1-step prediction: (a) Bohai Sea and (b) NSCS.

Fig. 13. The box plot of the MAE comparison for 1-step, 3-steps, 6-steps, and 12-steps Chl-a prediction by OVPGCN.

best possible prediction.

Fig. 11 compares the average RMSE yielded by the OVPGCN and other four baseline algorithms in different calendar months for the 1-step Chl-a prediction. It can be seen that the proposed OVPGCN achieves the lowest prediction RMSE than the other four baseline algorithms in both the Bohai Sea and NSCS, especially in cold seasons. Compared to the OVPGCN, the performance of these baseline models exhibits significant fluctuations in RMSE during the 12 calendar months in the Bohai Sea, showing instability. Additionally, OVPGCN performs better in the Chl-a prediction for the NSCS than that for the Bohai Sea, which is similar to the SST predictions in Section III-D. In conclusion, the proposed OVPGCN algorithm is high-reliability for Chl-a prediction and enables a stable Chl-a prediction performance in all seasons. It can also be seen from the Fig.12, which presents the temporal trend of real and predicted Chl-a values for 1-step prediction.

Fig. 13 shows the box plot for the MAE comparison of OVPGCN for 1-step, 3-steps, 6-steps, and 12-steps Chl-a prediction. From Fig. 13, we can see that for the long-term 12-steps prediction, the predictive performance of OVPGCN does not show a significant decrease, indicating that proposed OVPGCN is capable of making use of the temporal correlation and the annual periodic feature for improving the accuracy of Chl-a prediction.

Finally, the spatial distributions of OVPGCN used for Chl-a prediction in the NSCS is demonstrated in Fig. 14. As illustrated in Fig. 14, the Chl-a values of the nodes located in the deep offshore waters, which are far away from land, are small and have low fluctuations. As a result, the deviation of the prediction results does not change much, as the step-size of prediction becomes bigger. However, in some coastal areas near land, such as the areas in the blue and red boxes, the prediction error is relatively big and becomes larger as the step-size of prediction increases. This is because mining the characteristics of the spatio-temporal evolution in coastal waters is challenging due to the large Chl-a spatial gradient and the complicated time-varying property under the influence of human activities. Nevertheless, the overall performance of OVPGCN is still acceptable. However, the prediction performance in the coastal regions can be possibly further enhanced with a specifically-designed algorithm, if we have comprehensive knowledge and information about the Chl-a in
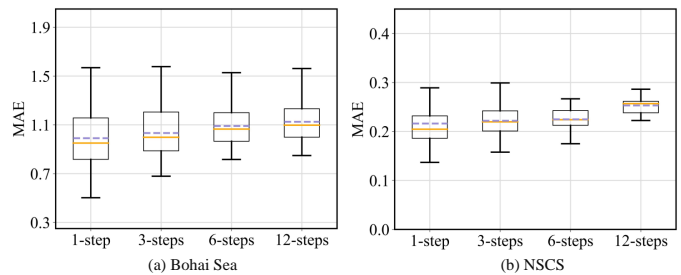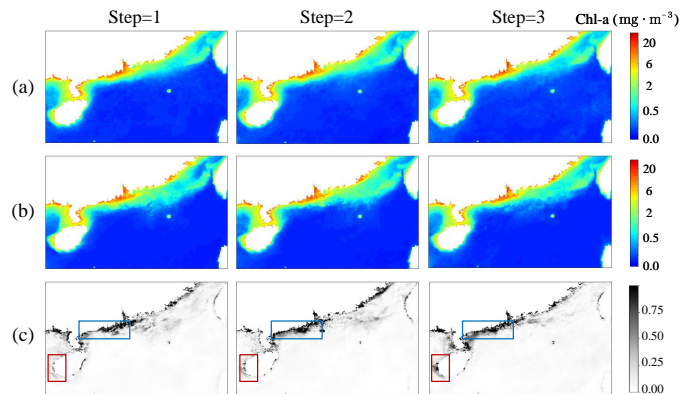

Fig. 14. Example of qualitative results of the OVPGCN model for the Chl-a predictions of next 3-steps (months) in NSCS. (a) Predicted Chl-a fields. (b) Observed Chl-a fields. (c) Deviation between observed and predicted Chl-a fields.

these regions, which is left to our future research study.

IV. CONCLUSION

Accurate prediction of oceanic variables is paramount for maintaining a healthy ocean ecosystem and providing early warning of extreme climate variations. In this paper, a novel OVPGCN algorithm has been developed for monthly oceanic variables prediction, with the objective to fully extract the spatial correlation and temporal dependency by modeling the multi-characteristics of spatio-temporal dynamic evolution of oceanic variables. In our method, the time-series of oceanic variable fields are represented by graph to overcome the data irregularity and to model the large-scale global connections. The PDM module is designed to extract the periodic characteristics of oceanic variables. The BC module is designed to capture the spatial differences between various regions in the ocean. Whereas, the STCFE module is proposed to enhance the feature extraction of the spatio-temporal correlations, which has the dual graph inputs and several spatio-temporal convolutional blocks. With the aid of these dedicated modules, our studies demonstrate that OVPGCN is capable of providing high-reliability prediction results. Specifically, to show its effectiveness, the monthly SST and Chl-a data from the Bohai Sea and NSCS were used to train and test the algorithm. For the 6-steps SST prediction, OVPGCN achieves a remarkable reduction of 15.72% and 28.14% in RMSE compared to the state-of-the-art methods in the Bohai Sea and NSCS, respectively. Likewise, for the 6-steps Chl-a prediction,

OVPGCN reduces the RMSE by 9.43% and 12.33% in the Bohai Sea and NSCS compared to the current state-of-the-art approaches. In summary, the experimental results demonstrate that OVPGCN outperforms all the six baseline algorithms in both SST and Chl-a prediction. Moreover, even when there are only a small number of data samples available, when the data are unevenly distributed, or when the data values are distributed in a wide range, OVPGCN is still capable of performing efficiently and providing a stable Chl-a prediction.

In this paper, the OVPGCN relies solely on training data to learn the underlying characteristics, but it is difficult to interpret physically. In our future work, we will endeavor to develop a physics-guided prediction framework based on OVPGCN, to take the advantage of the prior and available knowledge about the ocean physical laws.

REFERENCES

[1] M. M. Omand, J. J. Leichter, P. J. Franks, R. Guza, A. J. Lucas, and F. Feddersen, "Physical and biological processes underlying the sudden surface appearance of a red tide in the nearshore," *Limnology and Oceanography*, vol. 56, no. 3, pp. 787–801, 2011.

[2] P. W. Boyd, S. T. Lennartz, D. M. Glover, and S. C. Doney, "Biological ramifications of climate-change-mediated oceanic multi-stressors," *Nature Climate Change*, vol. 5, no. 1, pp. 71–79, 2015.

[3] M. A. Litzow, M. E. Hunsicker, N. A. Bond, B. J. Burke, C. J. Cunningham, J. L. Gosselin, E. L. Norton, E. J. Ward, and S. G. Zador, "The changing physical and ecological meanings of north pacific ocean climate indices," *Proceedings of the National Academy of Sciences*, vol. 117, no. 14, pp. 7665–7671, 2020.

[4] Y. Xue, I. Diallo, W. Li, J. David Neelin, P. C. Chu, R. Vasic, W. Guo, Q. Li, D. A. Robinson, Y. Zhu *et al.*, "Spring land surface and subsurface temperature anomalies and subsequent downstream late spring-summer droughts/floods in north america and east asia," *Journal of Geophysical Research: Atmospheres*, vol. 123, no. 10, pp. 5001–5019, 2018.

[5] E. P. Chassignet, H. E. Hurlburt, E. J. Metzger, O. M. Smedstad, J. A. Cummings, G. R. Halliwell, R. Bleck, R. Baraille, A. J. Wallcraft, C. Lozano *et al.*, "US GODAE: global ocean prediction with the hybrid coordinate ocean model (HYCOM)," *Oceanography*, vol. 22, no. 2, pp. 64–75, 2009.

[6] T. N. Stockdale, M. A. Balmaseda, and A. Vidard, "Tropical Atlantic SST prediction with coupled ocean–atmosphere GCMs," *Journal of climate*, vol. 19, no. 23, pp. 6047–6061, 2006.

[7] P. Bauer, A. Thorpe, and G. Brunet, "The quiet revolution of numerical weather prediction," *Nature*, vol. 525, no. 7567, pp. 47–55, 2015.

[8] M. R. Allen, J. Kettleborough, and D. Stainforth, "Model error in weather and climate forecasting," in *ECMWF Predictability of Weather and Climate Seminar*. European Centre for Medium Range Weather Forecasts, Reading, UK, http://www , 2002, pp. 279–304.

[9] K. Patil and M. Deo, "Basin-scale prediction of sea surface temperature with artificial neural networks," *Journal of Atmospheric and Oceanic Technology*, vol. 35, no. 7, pp. 1441–1455, 2018.

[10] I. D. Lins, M. Araujo, M. das Chagas Moura, M. A. Silva, and E. L. Droguett, "Prediction of sea surface temperature in the tropical Atlantic by support vector machines," *Computational Statistics & Data Analysis*, vol. 61, pp. 187–198, 2013.

[11] X. Li, J. Sha, and Z.-L. Wang, "Application of feature selection and regression models for chlorophyll-a prediction in a shallow lake," *Environmental Science and Pollution Research*, vol. 25, no. 20, pp. 19 488–19 498, 2018.

[12] Q. Shao, Y. Zhao, W. Li, G. Han, G. Hou, C. Li, S. Liu, Y. Gong, H. Liu, and P. Qu, "A simple statistical intra-seasonal prediction model for sea surface variables utilizing satellite remote sensing," *Remote Sensing*, vol. 14, no. 5, p. 1162, 2022.

[13] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural computation*, vol. 31, no. 7, pp. 1235–1270, 2019.

[14] C. Li, G. Tang, X. Xue, A. Saeed, and X. Hu, "Short-term wind speed interval prediction based on ensemble GRU model," *IEEE transactions on sustainable energy*, vol. 11, no. 3, pp. 1370–1380, 2019.

[15] H. Cho, U.-J. Choi, and H. Park, "Deep learning application to time-series prediction of daily chlorophyll-a concentration," *WIT Trans. Ecol. Environ*, vol. 215, pp. 157–163, 2018.

[16] Z. Zhang, X. Pan, T. Jiang, B. Sui, C. Liu, and W. Sun, "Monthly and quarterly sea surface temperature prediction based on gated recurrent unit neural network," *Journal of Marine Science and Engineering*, vol. 8, no. 4, p. 249, 2020.

[17] A.-L. Balogun and N. Adebisi, "Sea level prediction using ARIMA, SVR and LSTM neural network: assessing the impact of ensemble ocean-atmospheric processes on models accuracy," *Geomatics, Natural Hazards and Risk*, vol. 12, no. 1, pp. 653–674, 2021.

[18] Q. Zhang, H. Wang, J. Dong, G. Zhong, and X. Sun, "Prediction of sea surface temperature using long short-term memory," *IEEE geoscience and remote sensing letters*, vol. 14, no. 10, pp. 1745–1749, 2017.

[19] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 international conference on engineering and technology (ICET)*. Ieee, 2017, pp. 1–6.

[20] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," *Advances in neural information processing systems*, vol. 28, 2015.

[21] B. Mu, C. Peng, S. Yuan, and L. Chen, "ENSO forecasting over multiple time horizons using ConvLSTM network and rolling mechanism," in *2019 international joint conference on neural networks (ijcnn)*. IEEE, 2019, pp. 1–8.

[22] C. Xiao, N. Chen, C. Hu, K. Wang, Z. Xu, Y. Cai, L. Xu, Z. Chen, and J. Gong, "A spatiotemporal deep learning model for sea surface temperature field prediction using time-series satellite data," *Environmental Modelling & Software*, vol. 120, p. 104502, 2019.

[23] T. Song, N. Han, Y. Zhu, Z. Li, Y. Li, S. Li, and S. Peng, "Application of deep learning technique to the sea surface height prediction in the South China Sea," *Acta Oceanologica Sinica*, vol. 40, pp. 68–76, 2021.

[24] S. Zhou, W. Xie, Y. Lu, Y. Wang, Y. Zhou, N. Hui, and C. Dong, "ConvLSTM-based wave forecasts in the South and East China seas," *Frontiers in Marine Science*, vol. 8, p. 680079, 2021.

[25] L. Han, Q. Ji, X. Jia, Y. Liu, G. Han, and X. Lin, "Significant wave height prediction in the South China Sea based on the ConvLSTM algorithm," *Journal of Marine Science and Engineering*, vol. 10, no. 11, p. 1683, 2022.

[26] C. Li, Y. Feng, T. Sun, and X. Zhang, "Long term Indian ocean dipole (IOD) index prediction used deep learning by ConvLSTM," *Remote Sensing*, vol. 14, no. 3, p. 523, 2022.

[27] Q. Liu, R. Zhang, Y. Wang, H. Yan, and M. Hong, "Daily prediction of the Arctic sea ice concentration using reanalysis data based on a convolutional LSTM network," *Journal of Marine Science and Engineering*, vol. 9, no. 3, p. 330, 2021.

[28] Y. Wang, J. Zhang, H. Zhu, M. Long, J. Wang, and P. S. Yu, "Memory in memory: A predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 9154–9162.

[29] Y. Zhou, C. Lu, K. Chen, and X. Li, "Multilayer fusion recurrent neural network for sea surface height anomaly field prediction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–11, 2021.

[30] H. Wu, Z. Yao, J. Wang, and M. Long, "Motionrnn: A flexible model for video prediction with spacetime-varying motions," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 15 435–15 444.

[31] Y.-G. Ham, J.-H. Kim, and J.-J. Luo, "Deep learning for multi-year ENSO forecasts," *Nature*, vol. 573, no. 7775, pp. 568–572, 2019.

[32] M. Ye, J. Nie, A. Liu, Z. Wang, L. Huang, H. Tian, D. Song, and Z. Wei, "Multi-year ENSO forecasts using parallel convolutional neural networks with heterogeneous architecture," *Frontiers in Marine Science*, p. 1092, 2021.

[33] J. G. Fernández, I. A. Abdellaoui, and S. Mehrkanoon, "Deep coastal sea elements forecasting using UNet-based models," *Knowledge-Based Systems*, vol. 252, p. 109445, 2022.

[34] G. Zheng, X. Li, R.-H. Zhang, and B. Liu, "Purely satellite data–driven deep learning forecast of complicated tropical instability waves," *Science advances*, vol. 6, no. 29, p. eaba1482, 2020.

[35] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[36] S. R. Cachay, E. Erickson, A. F. C. Bucker, E. Pokropek, W. Potosnak, S. Bire, S. Osei, and B. Lütjens, "The world as a graph: Improving EI Niño forecasts with graph neural networks," *arXiv preprint arXiv:2104.05089*, 2021.

[37] X. Geng, L. Xu, X. He, and J. Yu, "Graph optimization neural network with spatio-temporal correlation learning for multi-node offshore wind speed forecasting," *Renewable Energy*, vol. 180, pp. 1014–1025, 2021.

[38] X. Zhang, Y. Li, A. C. Frery, and P. Ren, "Sea surface temperature prediction with memory graph convolutional networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2021.

[39] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," *Advances in neural information processing systems*, vol. 30, 2017.

[40] Y. Sun, X. Yao, X. Bi, X. Huang, X. Zhao, and B. Qiao, "Time-series graph network for sea surface temperature prediction," *Big Data Research*, vol. 25, p. 100237, 2021.

[41] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," *arXiv preprint arXiv:1709.04875*, 2017.

[42] S. Guo, Y. Lin, N. Feng, C. Song, and H. Wan, "Attention based spatial-temporal graph convolutional networks for traffic flow forecasting," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 33, no. 01, 2019, pp. 922–929.

[43] C. Morris, M. Ritzert, M. Fey, W. L. Hamilton, J. E. Lenssen, G. Rattan, and M. Grohe, "Weisfeiler and leman go neural: higher-order graph neural networks," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 33, no. 01, 2019, pp. 4602–4609.

[44] J. Gehring, M. Auli, D. Grangier, D. Yarats, and Y. N. Dauphin, "Convolutional sequence to sequence learning," in *International conference on machine learning*. PMLR, 2017, pp. 1243–1252.

[45] A. Alvera-Azcárate, A. Barth, M. Rixen, and J.-M. Beckers, "Reconstruction of incomplete oceanographic data sets using empirical orthogonal functions: application to the Adriatic sea surface temperature," *Ocean Modelling*, vol. 9, no. 4, pp. 325–346, 2005.

[46] C. Ma, J. Zhao, B. Ai, and S. Sun, "Two-decade variability of sea surface temperature and chlorophyll-a in the Northern South China Sea as revealed by reconstructed cloud-free satellite data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 11, pp. 9033–9046, 2021.

[47] M. Fey and J. E. Lenssen, "Fast graph representation learning with Pytorch Geometric," *arXiv preprint arXiv:1903.02428*, 2019.

[48] R. Barzegar, M. T. Aalami, and J. Adamowski, "Short-term water quality variable prediction using a hybrid CNN–LSTM deep learning model," *Stochastic Environmental Research and Risk Assessment*, vol. 34, no. 2, pp. 415–433, 2020.