A Spatio-Temporal Attention Graph Convolutional Networks for Sea Surface Temperature Prediction

Desheng Chen, Jiabao Wen*, Caiyun Lv

School of Electrical and Information Engineering, Tianjin University, Tianjin (P.R. China)

Received 16 June 2022 | Accepted 10 February 2023 | Early Access 23 February 2023

ABSTRACT

Sea surface temperature (SST) is an important index to detect ocean changes, predict SST anomalies, and prevent natural disasters caused by abnormal changes, dynamic variation of which have a profound impact on the whole marine ecosystem and the dynamic changes of climate. In order to better capture the dynamic changes of ocean temperature, it’s vitally essential to predict the SST in the future. A new spatio-temporal attention graph convolutional network (STAGCN) for SST prediction was proposed in this paper which can capture spatial dependence and temporal correlation in the way of integrating gated recurrent unit (GRU) model with graph convolutional network (GCN) and introduced attention mechanism. The STAGCN model adopts the GCN model to learn the topological structure between ocean location points for extracting the spatial characteristics from the ocean position nodes network. Besides, capturing temporal correlation by learning dynamic variation of SST time series data, a GRU model is introduced into the STAGCN model to deal with the prediction problem about long time series, the input of which is the SST data with spatial characteristics. To capture the significance of SST information at different times and increase the accuracy of SST forecast, the attention mechanism was used to obtain the spatial and temporal characteristics globally. In this study, the proposed STAGCN model was trained and tested on the East China Sea. Experiments with different prediction lengths show that the model can capture the spatio-temporal correlation of regional-scale sea surface temperature series and almost uniformly outperforms other classical models under different sea areas and different prediction levels, in which the root mean square error is reduced by about 0.2 compared with the LSTM model.

I. Introduction

The dynamics of the oceans, which make up about two-thirds of the planet, have an extremely important impact on climate, marine ecology, and the lives of the people around them. Sea surface temperature (SST) is an important index to detect ocean changes, predict SST anomalies, and prevent natural disasters caused by abnormal changes. Therefore, it’s significant to predict the dynamic change of SST in the future. Moreover, SST has played an indispensable role in the ocean-atmosphere interaction, that is, the exchange of matter, energy, and momentum between the ocean and the atmosphere [1] [2] [3]. As a result, changes in SST have incalculable impacts on global climate and marine ecosystem [4] [5] [6] [7]. Besides, SST predictions also has implications for applications related to the ocean, such as weather forecasting, fisheries, and marine environmental protection. Therefore, it’s critically necessary to predict dynamic changes of SST in the future to help people identify and prevent severe weather events such as drought in advance [8] [9], and it’s also of great significance for scientific research and applications [10]. However, due to the influence of many complex factors, such as sea surface heat flow, radiation, and solar wind, the prediction of SST is quite indefinite and challenging.

In recent years, SST prediction methods have been widely applied and attracted much attention further become an attractive field of marine research. Three kinds of methods are generally used to predict SST including the numerical method based on the mathematical model, the data-driven method using the historical model to predict SST in the future, and the method combining the two methods [11]. The numerical methods generally use kinetic and thermodynamic equations to exposit the dynamic changes of SST and then solve a series of differential equations which are difficult to solve due to they are usually sophisticated and require a large amount of computation. The data-driven method is mainly used to predict the future SST value from the perspective of data. This method builds the model by learning the relationships and patterns from the historical SST observation data and further uses the learned relationships model to approximate the future SST data. The data-driven method is less complicated than the numerical method and is suitable for the prediction of SST in high-resolution areas. The data-driven method mainly predicts the future SST from the perspective of statistical data analysis, machine learning, and artificial intelligence algorithms. Among them, statistical data analysis techniques primarily contain the Markov model [12] [13], Empirical canonical correlation analysis [14] and regression model.

Keywords

Data Forecast, Gated Recurrent Unit, Graph Convolutional Network, Sea Surface Temperature, Spatiotemporal Attention Graph Convolutional Network.
The contribution of this paper can be summarized in the following two aspects: (1) A STAGCN model is designed to capture the global spatial and temporal dependence simultaneously for accurate SST prediction, which combines the GCN deep learning model with the GRU learning model and introduces an attention mechanism. (2) The concept of a graph is applied to the field of SST prediction, and the topology structure network of ocean location points graph is constructed to obtain the spatial characteristics from SST time series by GCN model. Using 38-year time-series satellite data from some areas of the East China Sea, the experiments show that the STAGCN model can achieve preferable prediction results than the GRU model and the GCN model demonstrating that the STAGCN model has the ability to capture both time and space correlation from SST series data simultaneously, and can achieve desirable prediction effect for the short-term prediction of future SST.

The rest of the paper is organized as follows. Section II proposes the novel STAGCN model for SST prediction in detail. Section III describes the experiments and analyzes the results. Finally, Section IV gives the conclusion of the paper.

II. Method

A. Problem Clarification

In this study, the prediction of SST is to predict the sea surface temperature within a certain time in the future according to the historical SST time-series information.

The undirected graph $G = (V,E)$ with no weights was used to describe a topological network composed of ocean observation points, and each location represents a node in the graph, where $V = V_1, V_2, \ldots, V_N$, means all oceanic positions that correspond to the $N$ vertices of the graph. $E$ represents the correlation between ocean points corresponding to the edges between nodes in the graph, reflecting the connection between nodes at different positions. The connection relationship between nodes in the whole graph is represented by the adjacencies matrix $A$, the number of the rows and columns of which are determined by the number of nodes, and each value of which represents the connection relationship between nodes. The value of each item in the adjacency matrix $A$ is either 0 or 1. 0 indicates that there is no direct connection between nodes and 1 expresses that there has a linkage between nodes.

The feature $X^{pxN}$ of the node in the topology graph corresponds to the SST value of each location point on the ocean positions network, where $N$ represents the length of the time dimension in oceanographic satellite data, and $X \in R^{p \times N}$ represents the SST value at time $t$.

The SST prediction problem can be regarded as looking for a mapping function $f$, which map the SST value in the historical time to the SST value in the next $T$ periods under the conditions of the topology diagram $G$ of the ocean location points network and feature matrix $X$. Equation (1) refers to the SST prediction process:

$$f(\cdot) = f(\{X_{t-\theta}, \ldots, X_{t-1}, X_t\})$$

The left side is the historical SST value with the length $n$, from which the model learns the variation trend of SST, and the right side of the equation is the predicted future SST value with length $T$ under the mapping condition.

B. The STAGCN Model

To capture global dynamic changes of the SST information, a novel STAGCN model combines graph convolutional network capturing spatial correlation and gated recurrent unit obtaining temporal dependence with attention model at the same time is proposed in this paper. In the STAGCN model, a layer of GCN network can be used to
obtain a better prediction effect, and Equation (2) refers to the specific convolution process:

\[
f(X, A) = \sigma(\bar{A}ReLU(\bar{A}XW_{0}))W_{1}
\]

where \(f(X, A)\) express the final output of the GCN model, \(X\) is the eigenmatrix, \(A\) means the adjacency matrix of the graph convolution. \(ReLU(\cdot)\) represents an activation function used to add linear factors to improve model expression, \(\bar{A}\) is the form of the adjacent matrix after further renormalization to avoid causing gradient explosion. \(W_{0}\) represents the weight that needs to be trained in the progress of the graph convolutional network, whose first dimension \(F\) means the time series length of ocean data and second dimension \(G\) means the number of neural units in the output layer.

Specifically, the area we study is ocean location points graphs composed of position points determined by longitude and latitude, and the topological connection relationship between location points can be captured by the GCN network. The shaded part of the dotted line in Fig. 1 is a topological graph of a point in the simulated ocean location. It is assumed that the red dot 1 represents the central node in the topology diagram, and the green dots (the green dot in the shaded part of the dotted line) around it are the adjacent nodes. The interaction degree between the central node and the adjacent nodes can be obtained by GCN model. The GRU model was adopted in this study to obtain the temporal dependence of SST data. The attention model is used to screen which moments of SST data are relevant, that is, to distinguish the importance of data at different moments, which improves the accuracy of the prediction and realize the SST prediction task based on the structure of the ocean location points network. The specific spatial-temporal prediction process of the STAGCN model is shown in Fig. 2, where the distance function (Equation) was applied to calculate the adjacency matrix \(A_{i}\) that represents the connection relation between the position (a) (b)

Fig. 1. Topology graph of an ocean location point (in the dotted shaded part). (a) The green nodes indicate the location points have connection relationship with the central point. (b) The topological structure between the central node 1 and its adjacent nodes is established through the GCN model to obtain the spatial features.

Fig. 2. The specific spatial-temporal prediction process of the STAGCN model.
points, the GCN-GRU represents the process of combining the output of the GCN model and the GRU model, and the FC express the fully connected layer.

First of all, the obtained adjacency matrix and the SST feature data \( X_i (i = t-n, \ldots, t-1, 1, t) \) of \( n \) historical time series were taken as input, from which the GCN model capture the spatial information. Then, The input of the GRU model is replaced by the output of the GCN model to obtain the temporal characteristics of SST data. Equation (3) to (6) refer to the update gate, reset gate, cell state and output state at time \( t \) in the STAGCN model respectively.

\[
\begin{align*}
    u_t &= \sigma(W_u \cdot f_{GCN}(A[h_{t-1}, X_t]) + b_u) \\
    r_t &= \sigma(W_r \cdot f_{GCN}(A[h_{t-1}, X_t]) + b_r) \\
    c_t &= \tanh(W_c \cdot f_{GCN}(A[(r_t \cdot h_{t-1}), X_t]) + b_c) \\
    h_t &= (u_t \cdot h_{t-1} + (1 - u_t) \cdot c_t)
\end{align*}
\]

where \( u_t \) and \( r_t \) express the update gate and reset gate, \( h_{t-1} \) means the output at previous time, \( h_t \) means the state output at present \( t \) moment, \( c_t \) represents the information be reserved from previous moment and present time. Function \( f_{GCN}(\cdot) \) express the graph convolution, \( W_f, W_r, W_c \) are the connection weights between the output of graph convolution and the previous output \( h_{t-1}, b_u, b_r \) and \( b_c \) are the corresponding thresholds. Moreover, the final hidden state information of the GRU model is used as the input of the attention model, which is used to obtain the importance of the changes information of SST series data. Finally, we get the prediction from the full connection layer. In the attention model, multi-layer perception is used as the scoring function, in which \( w_i (i = t-n, \ldots, t-1, t) \) is the weight matrix of multi-layer perception. The scores \( e_i (i = t-n, \ldots, t-1, t) \) of the multi-layer perception output is brought into the Softmax function to get the attention distribution probability. The last hidden state and its weight are weighted to obtain the final context vector \( C \).

To sum up, we propose that the STAGCN model has the ability to obtain the global spatial dependence and temporal dynamic changes which can obtain a preferable SST prediction effect. The GCN model is applied to capture spatial information by building the structure of the interrelation between the position nodes. The GRU model is used to obtain the temporal dependence from the SST series data with spatial characteristics. Moreover, the attention model captures the global variation trends of the SST information which is significant to achieve accurate SST prediction tasks.

III. Experiments

A. Research Area and Data

Rich in natural resources, the East China Sea is the confluence of many rivers covering a wide area including China’s Bohai Sea in the north and the Taiwan Strait in the south and is the strategic maritime area for China, Japan, South Korea, and other countries. Therefore, studying the dynamic changes of SST data in the East China Sea plays an extremely important role in national marine transportation and people’s production and life in the surrounding countries. The research area selected in this study is the part area of the East China Sea with sea areas of 26.875°N - 32.125°N and 123.125°E - 127.125°E covering most areas of the East China Sea and no land area, as is shown in Fig. 3. The selected area includes most of the East China Sea to facilitate the acquisition of grid-based SST data and the establishment of the topological structure of the ocean position points graph for SST prediction.

In this study, the data used in the experiments were derived from the daily Optimum interpolated SST (OISST Daily Edition 2.1) with a spatial resolution of 1/4° in the N ational Oceanic and Atmospheric Administration (NOAA) platform. In this study, AVHRR-only daily SST data contained a total of 13,879 days SST data from 1982/01/01-2019/12/31 covered the spatial span of 26.875°N-32.125°N, 123.125°E-127.125°E (the most of the east sea) are used as the experimental dataset, with a total of 22°17’ position points as the study nodes.

In this study, the experimental data is composed of two parts: the adjacency matrix and the eigenmatrix respectively. The former is an adjacency matrix of size 374*374, which depicts the spatial dependence between position points. Equation (7) refers to each value in the adjacency matrix \( W \) is derived by some scaling of the distance between each position point in the ocean anchor point network.

\[
W_{ij} = \begin{cases} 
\exp\left(-\frac{d_{ij}^2}{\sigma^2}\right), & i \neq j \text{ and } \exp\left(-\frac{d_{ij}^2}{\sigma^2} \geq \epsilon\right) \\
0, & \text{otherwise}
\end{cases}
\]

where \( W_{ij} \) is the weight of the edges in the position graph determined by the distance between position \( i \) and \( j \) \( (d_{ij}) \), \( \sigma^2 \) and \( \epsilon \) are thresholds to control spacial arrangement and sparse arrangement of the adjacency matrix \( W \), which are set to 0.1 and 0.4 respectively after testing in the experiments. The latter part of the experimental data is the eigenmatrix, which describes the dynamic changes of the SST value over time at the position nodes.

B. Experiment Setup

In order to better explain the superiority of the STAGCN model proposed in predicting SST, we chose five comparative models including the autoregressive moving average model (ARIMA), linear support vector machine model (SVR), graph convolutional network model (GCN), and gated recurrent unit model (GRU). In the experiments, we divided the dataset into the training set and test set, and the ratio of the two is 8:2. The Adam optimizer is used to train the model. The STAGCN model, GCN model, and GRU model use TensorFlow 1.5.0 (GPU version) as runtime environment during the training and testing progress. The ARIMA and SVR are respectively implemented using Statsmodels 0.12.2 and Scikit-learn 0.24.1 [31].

The real SST and the predicted SST of different nodes at time \( t \) are respectively expressed by \( Y_t \) and \( \hat{Y}_t \). In the network training, the loss function value should be minimized as far as possible, which is beneficial for the predicted SST value of each ocean position node to be a better fit to the actual SST value. The loss function defined in the STAGCN model is shown in Equation. The first term \( ||\cdot|| \) is the 2-norm of the real value and the predicted value, and the second term is the regularization term \( \lambda \hat{L}_{reg} \) with hyperparameter \( \lambda \), which improved the prediction performance of the model and prevented overfitting occurrence during training.
To prove the desirable prediction performance of the STAGCN model, five measurement criteria are used to compare the SST prediction performance of the proposed model with other models including root mean square error (RMSE), mean absolute error (MAE), accuracy (Accuracy), coefficient of determination ($R^2$) and explained variance score ($\text{Var}$).

$$RMSE = \sqrt{\frac{1}{TN} \sum_{t=1}^{T} \sum_{n=1}^{N} (\hat{y}_t^n - y_t^n)^2}$$  \hspace{1cm} (8)

$$MAE = \frac{1}{TN} \sum_{t=1}^{T} \sum_{n=1}^{N} |\hat{y}_t^n - y_t^n|$$  \hspace{1cm} (9)

$$\text{Accuracy} = 1 - \frac{||Y - \hat{Y}||_F}{||Y||_F}$$  \hspace{1cm} (10)

$$R^2 = 1 - \frac{\sum_{t=1}^{T} \sum_{n=1}^{N} (y_t^n - \hat{y}_t^n)^2}{\sum_{t=1}^{T} \sum_{n=1}^{N} (y_t^n - \bar{y})^2}$$  \hspace{1cm} (11)

$$\text{Var} = 1 - \frac{\text{var}(Y - \hat{Y})}{\text{var}(Y)}$$  \hspace{1cm} (12)

where Equation (8) to (12) refer to the calculation process of RMSE, MAE, Accuracy, ($R^2$) and ($\text{Var}$) respectively, $T$ means the length of the SST time series, $N$ represents the total number of position points. $TN$ is the number of recorded changes in temperature values at all ocean locations with a time length of $T$. $y_t^n$ represents the real SST value and the $\hat{y}_t^n$ means predicted SST value in the position point $n$ at the moment $t$. The entire sets of $y_t^n$ and $\hat{y}_t^n$ can be defined as $Y$ and $\hat{Y}$ respectively. $\bar{y}$ represents the average value of collection $Y$. Five criteria are used to measure the merits of the model from the perspective of error, accuracy, and fitting degree. The $F$ in the accuracy indicator refers to the $F$-norm.

In the neural network model, the setting of model parameters is crucial to the model training, such as the number of neurons in the hidden layer, whose size determines the computational complexity and predictive performance of the whole model. Therefore, in order to improve the training efficiency and the accuracy of prediction, we used different hidden units to conduct experiments on the test set and select the corresponding number of hidden units with the best predicted effect from the predicted results. We set the number of neurons in the hidden layer as the value in [8, 16, 32, 64, 128], respectively. The variation trend of error indicators under different hidden units condition as shown in 4, from which we can see that the RMSE and MAE reach the minimum value when the number of hidden layer units is 32. At the same time, we can see from 4 that the prediction error including RMSE and MAE has a similar decreasing trend while hidden units increases from 8 to 32, and then it shows an increasing trend when hidden units exceed 32. On the contrary, the prediction accuracy has an opposite trend, rising first and then falling. The trend of the model measurement indexes shows that there is a critical value of the number of hidden layer cells in the model. When the critical value is exceeded, the complexity of the model will increase with the increase of the number of hidden layer cells, and the performance of the model will decline simultaneously. Since the prediction performance of the model has the best performance when the number of hidden units is 32, the following experiments will be carried out under the condition that the number of hidden units is set to 32.

C. Experiment Results and Discussion

We performed prediction experiments on 1 day, 7 days, 14 days, and 30 days SST values in the future and measure the performance of the proposed STAGCN model, the ARIMA model, SVR model, GCN model, and GRU model with five performance metrics. The SST prediction result of the STAGCN model is analyzed from the perspectives of prediction accuracy, temporal and spatial prediction ability, and long-term prediction ability.

Table I shows that the result of the five metrics using to measure the prediction result of the STAGCN model compared with other models on the East China Sea dataset for the 1-day, 7-days, 14-days, and 30-days prediction tasks, with boldface sections representing the optimal values of the various metrics of the model. show that the value is negative, indicating that the model cannot predict well and can be ignored. Compared with other models, the proposed STAGCN model achieves excellent prediction performance under almost all conditions, indicating that the STAGCN model can capture the global spatio-temporal correlation, thus achieved an accurate prediction of SST. For different prediction lengths, the prediction effects of the STAGCN model are preferable to other models, demonstrating that the STAGCN model has the ability to predict SST accurately in both the short and long term.

As can be seen from Table I, the prediction performance of the GRU model is better than that of the GCN model. The RMSE of the GRU model is approximately 0.07 lower than that of the GCN model and the accuracy of the GRU model is increased by 0.4% compared with the GCN model for future one-day SST prediction. The forecasted effect of the GCN model is worse than that of the GRU model probably because the data itself has obvious time series features, but the GCN model merely captures the spatial dependency without considering the temporal characteristics from SST data.
In order to better illustrate that the STAGCN model proposed in this paper can capture the temporal and spatial dependence from SST data and obtain satisfactory prediction effect simultaneously, we visualized the error index RMSE of STAGCN model, GRU model, and GCN model in predicting SST for next 1 day, 7 days, 14 days and 30 days and analyzed their prediction performance. The visualization results are shown in Fig. 5. As can be seen from the bar chart, the RMSE of the models increases with the increase of the predicted length, but the error of the STAGCN model is smaller than that of the other two models demonstrating that the STAGCN model has the ability to obtain spatio-temporal correlation from SST series data. For example, the RMSE values of STAGCN model are about 0.16, 0.21, 0.23 and 0.35 lower than that of the GCN model for 1-day, 7-day, 14-day and 30-day SST forecasting, indicating that the STAGCN model is capable to obtain spatial features from sequence data. The RMSE values of the GRU model considered single temporal characteristics are raised by about 0.07, 0.14, 0.14 and 0.17 for future 1-day, 7-day, 14-day and 30-day SST prediction respectively compared with the STAGCN model, indicated that the STAGCN model can obtain the time correlation.

Fig. 5. Comparison of SST prediction errors between STAGCN model, GCN model and GRU model. (a) Comparison of RMSE values between STAGCN model and GCN model. (b) Comparison of RMSE values between STAGCN model and GCN model.

In order to better illustrate that the STAGCN model proposed in this paper can capture the temporal and spatial dependence from SST data and obtain satisfactory prediction effect simultaneously, we visualized the error index RMSE of STAGCN model, GRU model, and GCN model in predicting SST for next 1 day, 7 days, 14 days and 30 days and analyzed their prediction performance. The visualization results are shown in Fig. 5. Fig. 5 (a) and Fig. 5 (b) are visual comparison effects of the STAGCN model compared with the GCN model and the GRU model on RMSE metric respectively. As can be seen from the bar chart, the RMSE of the models increases with the increase of the predicted length, but the error of the STAGCN model is smaller than that of the other two models demonstrating that the STAGCN model has the ability to obtain spatio-temporal correlation from SST series data. For example, the RMSE values of STAGCN model are about 0.16, 0.21, 0.23 and 0.35 lower than that of the GCN model for 1-day, 7-day, 14-day and 30-day SST forecasting, indicating that the STAGCN model is capable to obtain spatial features from sequence data. The RMSE values of the GRU model considered single temporal characteristics are raised by about 0.07, 0.14, 0.14 and 0.17 for future 1-day, 7-day, 14-day and 30-day SST prediction respectively compared with the STAGCN model, indicated that the STAGCN model can obtain the time correlation.

Fig. 6. The change of RMSE and Accuracy of STAGCN model for future SST prediction under different prediction interval conditions.

### TABLE I. THE PREDICTION RESULT OF THE STAGCN MODEL COMPARED WITH OTHER MODELS

<table>
<thead>
<tr>
<th>Time</th>
<th>ARIMA</th>
<th>SVR</th>
<th>GCN</th>
<th>GRU</th>
<th>LSTM</th>
<th>STAGCN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 day</td>
<td>RMSE</td>
<td>7.2835</td>
<td>0.9280</td>
<td>0.5034</td>
<td>0.4217</td>
<td>0.5528</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>5.9413</td>
<td>0.7567</td>
<td>0.3574</td>
<td>0.2825</td>
<td>0.4316</td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>0.6625</td>
<td>0.9588</td>
<td>0.9759</td>
<td>0.9799</td>
<td>0.9757</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>-</td>
<td>0.9670</td>
<td>0.9933</td>
<td>0.9953</td>
<td>0.9884</td>
</tr>
<tr>
<td></td>
<td>Var</td>
<td>-</td>
<td>0.9671</td>
<td>0.9933</td>
<td>0.9953</td>
<td>0.9910</td>
</tr>
<tr>
<td>7 day</td>
<td>RMSE</td>
<td>7.9466</td>
<td>1.2042</td>
<td>0.8746</td>
<td>0.8034</td>
<td>0.7909</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>6.5981</td>
<td>0.9860</td>
<td>0.6450</td>
<td>0.5836</td>
<td>0.6077</td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>0.6304</td>
<td>0.9464</td>
<td>0.9581</td>
<td>0.9615</td>
<td>0.9651</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>-</td>
<td>0.9447</td>
<td>0.9798</td>
<td>0.9830</td>
<td>0.9761</td>
</tr>
<tr>
<td></td>
<td>Var</td>
<td>-</td>
<td>0.9477</td>
<td>0.9799</td>
<td>0.9831</td>
<td>0.9786</td>
</tr>
<tr>
<td>14 day</td>
<td>RMSE</td>
<td>7.7821</td>
<td>1.3285</td>
<td>1.0467</td>
<td>0.9493</td>
<td>0.8941</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>6.4279</td>
<td>1.0855</td>
<td>0.7860</td>
<td>0.7086</td>
<td>0.6853</td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>0.6386</td>
<td>0.9408</td>
<td>0.9497</td>
<td>0.9544</td>
<td>0.9604</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>-</td>
<td>0.9325</td>
<td>0.9711</td>
<td>0.9762</td>
<td>0.9691</td>
</tr>
<tr>
<td></td>
<td>Var</td>
<td>-</td>
<td>0.9351</td>
<td>0.9711</td>
<td>0.9762</td>
<td>0.9691</td>
</tr>
<tr>
<td>30 day</td>
<td>RMSE</td>
<td>8.3632</td>
<td>1.5406</td>
<td>1.3253</td>
<td>1.1428</td>
<td>1.1605</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>7.0629</td>
<td>1.2533</td>
<td>1.0052</td>
<td>0.8662</td>
<td>0.8860</td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>0.6128</td>
<td>0.9314</td>
<td>0.9364</td>
<td>0.9452</td>
<td>0.9492</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>-</td>
<td>0.9095</td>
<td>0.9538</td>
<td>0.9657</td>
<td>0.9469</td>
</tr>
<tr>
<td></td>
<td>Var</td>
<td>0.0040</td>
<td>0.9117</td>
<td>0.9539</td>
<td>0.9666</td>
<td>0.9471</td>
</tr>
</tbody>
</table>
To reflect the prediction effect of the STAGCN model, we visualized the variation trend of the RMSE and the accuracy in SST forecasting for the next 1 day, 7 days, 14 days and 30 days. The results are shown in Fig.6. As can be seen from the figure, with the increase of prediction length, the error of the STAGCN model increases relatively large and the accuracy decreases slightly. Although the RMSE of the STAGCN model does not have a certain stability, its prediction accuracy is relatively stable, indicating that the STAGCN model can also achieve long-term SST prediction while it has better prediction ability for short-term SST prediction than long-term SST prediction.

Fig.7 (a) and Fig.7 (b) are RMSE and Accuracy results of SST prediction by different methods at different prediction horizons. It’s observed that the STAGCN model achieved the lowest RMSE and the highest Accuracy for different predicted lengths.

To better explain the prediction performance of the STAGCN model, an ocean location point was randomly selected from the East China Sea dataset, in which we predicted the future SST for the next 1 day, 7 days, 14 days and 30 days, and visualize the prediction effect of all the selected days and the next 90 days. Fig. 8, Fig. 9, Fig. 10, and Fig. 11 show the visualization results of the SST for the 1-day, 7-days, 14-days and 30-days forecast intervals.

The prediction results of STAGCN model with the prediction length of 1 day, 7 days, 14 days and 30 days show that STAGCN model has poor prediction at peak and peak valley. The main reason may be that the GCN model in STAGCN model captures spatial features by constantly moving its defined smoothing filter, which will lead to excessive peak smoothing of the overall prediction results. At the same time, the model has a corresponding delay in the overall prediction. Although the prediction performance of STAGCN model decreases with the increase of prediction length, the fitting degree of predicted value and real value of STAGCN model is still high, and good prediction results can be obtained.
STAGCN model can capture the temporal correlation and regional spatial characteristics from the TIME series of SST, and obtain the global spatiotemporal dynamic change trend by using the attention mechanism, which reduces the prediction error and improves the accuracy of prediction, and realizes the long-term and short-term prediction task of regional scale SST.

IV. Conclusion

The sea surface temperature is an important index to detect ocean changes, predict SST anomalies, and prevent natural disasters caused by abnormal changes, the dynamic variation of which have a profound impact on the whole marine ecosystem and the changes of climate. Therefore, it’s essential to predict the future sea surface temperature. In order to achieve accurate SST prediction, a prediction model combining the GCN model with the GRU model and introduces the attention mechanism named the STAGCN model is proposed in this paper. We use the graph network to model the network of ocean location points. Nodes on the graph represent each ocean location point, edges on the graph represent that there have connections between location points. The GCN model is used to obtain the spatial correlation from the SST time series by constructing the spatial topology structure on the ocean points graph, which is obtained by the distance function between the nodes. The STAGCN model takes the GRU model to capture time dependence in the way of filtering and retaining historical and current SST information. Meanwhile, the attention model is applied to capture the importance of SST information from the output state and combines the global spatio-temporal characteristics from SST information. In this study, the experimental results of predicting the future short-term and long-term SST with STAGCN model on the data set indicating that the STAGCN model can achieve desirable prediction performance compared with the ARIMA model, SVR model, GCN model, and GRU model. In conclusion, the STAGCN model can acquire preferable forecasting results for future short-term and long-term SST prediction in the way of capturing global spatial characteristics and temporal dependence from SST series data.

References


Desheng Chen
He is currently pursuing the Ph.D. degree at the School of Electrical and Information Engineering, Tianjin University, China. His research interests include marine exploration, marine intelligent equipment control, artificial intelligence and Big data computing.

Jiabao Wen
He received the Ph.D. degree in information and communication engineering from Tianjin University, China, in 2021. He is currently a postdoctor at School of Electrical and Information Engineering, Tianjin University. His research interests include ocean information processing, pattern recognition, cloud computing.

Caiyun Lv
She received the Master degree in Electronics and Communication Engineering from Tianjin University, China, in 2022. Her research interests include ocean prediction and ocean information processing.