Attention Adaptive Temporal Graph Convolutional Network for Long-Term Seasonal Sea Surface Temperature Forecasting

Ling Xiao, Peihao Yang, *Graduate Student Member, IEEE*, Yuxue Wang, Sheng Li, and Baoqin Chen

***Abstract*—Sea surface temperature (SST) plays a crucial role in the global meteorological system, particularly as long-term seasonal variations are significant for analyzing SST anomalies and supporting long-term climate decision-making. Current forecasting methods are primarily focused on short-term or fine-grained predictions and often fail to effectively capture long-term seasonal trends. To address this, we introduce a novel attention adaptive temporal graph convolutional network (AA-TGCN) specifically designed for long-term seasonal SST forecasting. Unlike traditional adaptive graph convolutional networks, the AA-TGCN incorporates an attention mechanism to capture internode correlations and utilizes a pruning strategy from SGAT to eliminate noisy connections, thereby improving the model's inductive learning capabilities. Moreover, the network employs a TCN-like architecture to expand its receptive field, enhancing its ability to grasp long-term trends, and employs differential embedding to further refine the prediction accuracy of seasonal fluctuations. Practical applications in the Bohai Sea and parts of the South China Sea demonstrate that AA-TGCN outperforms existing technologies on multiple scales, particularly achieving significant improvements in the South China Sea regions.**

***Index Terms*—Attention mechanism pruning, dynamic graph, long-term seasonal forecasting, sea surface temperature (SST).**

I. Introduction

THE ocean, as the largest ecosystem on Earth, exerts a critical influence on global climate patterns. Even minor variations in sea temperatures can trigger widespread ecological and climatic effects. For instance, anomalies in sea surface temperature (SST) are closely linked to El Niño and La Niña phenomena, which are associated with increased frequency and intensity of extreme weather events worldwide [1], [2], [3]. These phenomena also impact marine biodiversity and global natural disasters [4], [5]. Consequently, accurate prediction of ocean temperature changes is vital for scientists, policymakers, and environmentalists.

Received 3 May 2024; revised 8 August 2024; accepted 28 September 2024. Date of publication 30 September 2024; date of current version 24 October 2024. This work was supported by the National Science Fund of China under Grant 12101138. *(Corresponding author: Baoqin Chen.)*

Ling Xiao, Yuxue Wang, Sheng Li, and Baoqin Chen are with the School of Mathematics and Computer, Guangdong Ocean University, Zhanjiang 524088, China (e-mail: xiaoling@stu.gdou.edu.cn; 969887241@qq.com; lish_ls@gdou.edu.cn; cbqchen@gdou.edu.cn).

Peihao Yang is with the College of Ocean and Meteorology, Guangdong Ocean University, Zhanjiang 524088, China (e-mail: z13760576015@163.com).

Digital Object Identifier 10.1109/JSTARS.2024.3471826

Currently, predictions of SST can be categorized into two main approaches: numerical simulation and data-driven methods. Numerical simulations are based on fundamental principles of oceanography and complex physical models. They utilize dynamic models of sea-air interaction combined with physical equations and various oceanographic parameters to precisely simulate changes in SST [6], [7], [8]. This method takes into account multiple factors including ocean currents, temperature distribution, wind forces, and solar radiation. Although numerical models can profoundly reveal the complexities of SST variations, their modeling and computation processes require substantial computational resources and are highly sensitive to initial setup and parameter selection. These characteristics increase the uncertainty in prediction outcomes.

Compared to traditional physical models, data-driven approaches rely on historical data to forecast future trends in SST. The use of machine learning algorithms in SST prediction has already achieved notable progress. These methods can identify latent patterns and trends in temperature changes by analyzing vast amounts of historical climate data, thereby providing accurate predictions. The strength of this technology lies in its ability to handle complex nonlinear relationships and adapt to dynamically changing environmental conditions, making it an effective tool in both research and practical applications. Initially, traditional machine learning algorithms, such as support vector machine [9] and autoregressive models [10], were utilized to predict SST. These algorithms have been well developed and are capable of stable SST prediction. However, they do not fully exploit the hidden information in historical data.

Nowadays, deep learning algorithms are extensively applied in SST prediction [11]. For instance, recurrent neural networks (RNNs) and their variants are used to capture temporal correlations, convolutional neural networks (CNNs) and their variants for spatial correlations, and various methods incorporating attention mechanisms [12], [13], [14], [15], [16], [17]. Yet, these algorithms have not considered the structural aspects of space. For instance, when the selected area includes land or islands, the performance of convolutional long short-term memory network (ConvLSTM) [18], which simultaneously considers spatiotemporal information, may degrade [19]. Recently, graph neural networks (GNNs) have been widely used for SST prediction and have shown promising results. They effectively take into account the structural, irregular, and connective properties of

space. This includes graph convolutional network (GCN), graph attention network (GAT), graph sample and aggregate networks (GraphSAGE), and their variants [20], [21], [22], [23], [24], [25]. For example, Peng et al. [21] proposed the enhanced adaptive graph convolutional network (EA-GCN), which is a multitime guided gated recurrent unit (GRU) and GCN combination for long-term fine-grained SST prediction, Gao et al. [22] developed the global spatiotemporal graph attention network (GSTGAT), an adaptive graph convolutional algorithm using convolution and GAT for short-term SST prediction, and Sun et al. [25] introduced the time-series graph network (TSGN) based on long short-term memory network (LSTM) aggregation method for SST prediction. These algorithms have been broadly implemented in SST forecasting and have achieved state-of-the-art results.

Most existing methods for forecasting SST predominantly utilize daily, weekly, and monthly averages SST to make short- to medium-term forecasts-for example, predicting temperatures for the upcoming day, three days, week, or month, or providing granular daily SST forecasts over an extended period such as the next 60 days. However, these approaches often overlook long-term seasonal variations, are biased towards predicting short-term fluctuations, and struggle to capture the patterns of long-term seasonal fluctuations in SST. In contrast, focusing on long-term predictions, such as forecasting the monthly average SST for the next twelve months, allows for a more thorough exploration and analysis of the enduring trends and seasonal patterns in climate change. This is extremely valuable for developing long-term policies and plans, managing marine resources, and assessing the impact of climate change on ecosystems. Such long-term forecasts provide a crucial tool for policymakers and researchers, aiding in their understanding of and response to continuous environmental changes. For instance, long-term seasonal fluctuation forecasting is crucial for understanding global climate patterns, assessing and predicting future climate changes, climate sensitivity, and the anomalous phenomena of El Niño and La Niña events [26], [27]. Additionally, the seasonal fluctuations of SST are closely linked to seasonal variations in rainfall and the seasonal changes in mesoscale air-sea coupling intensity [28], [29]. A thorough exploration of the long-term seasonal fluctuation patterns of SST and their predictability is essential for forecasting future changes in marine ecosystems [30].

Long-term monthly average SST forecasting and daily average SST forecasting focus on different aspects. Daily SST variations are continuous, involve large volumes of data, and are susceptible to short-term noise influences. In contrast, monthly average SST forecasting concentrates on long-term trends and seasonal changes in specific regions, with smaller data volumes and greater susceptibility to external factors. To date, there are no effective machine learning methods specifically tailored for long-term monthly average SST forecasting. To address this gap, we propose the attention adaptive temporal graph convolutional network (AA-TGCN) designed to capture the long-term trends and seasonal variations of SST. Unlike traditional adaptive graph convolutional networks, AA-TGCN does not use a node embedding matrix to build adaptive weights. Instead, it incorporates an attention mechanism, utilizing learnable attention coefficients to dynamically change the weights of the edges. This approach not only allows the model to focus on more critical edges but also provides the model with inductive learning capabilities suitable for dynamic graphs. Additionally, we employ a special pruning method to prevent overfitting. We also introduce differencing to simulate the monthly variations of SST, enabling the model to capture these trends effectively. Finally, we use a structure similar to temporal convolutional networks (TCN) to better integrate the seasonal differences throughout the year. Experiments were conducted in select maritime regions of the Bohai Sea and the South China Sea to explore the effectiveness of the model.

In the following sections, we will provide detailed descriptions of our proposed model in Section II, elucidate our experimental setup, data, and results in Section III, analyze the experimental outcomes in Section IV. Finally, Section V concludes this article.

II. Methods

A. Problem Formulation

For forecasting SST using GNNs, we typically divide the study area into a grid based on latitude, longitude, and resolution, creating N data points. For instance, Fig. 1 illustrates the division of a section of China's Bohai Sea. The target geographic range extends from 117°E to 122°E and from 37°N to 41°N, subdivided into 16×20 grid points with a resolution of 0.25° by 0.25°. Due to the Bohai Sea being an inland sea, some areas within this range are land, resulting in 162 effective grid points. In addition, we selected a region in the South China Sea, between longitudes 112–117°E and latitudes 12–16°N. This region was divided into a grid of 16×20 points. Due to the absence of land or islands, all 320 grid points within the region are valid points. These grid points are structured into a graph model $G = (V, E)$, where V represents the nodes of the graph, and E represents the edges connecting these nodes. The connections between nodes can be represented by an adjacency matrix $A \in R^{N\times N}$. If there is an edge e_{ij} between node i and node j, then $A_{ij} = 1$, otherwise $A_{ij} = 0$. We enhance this adjacency matrix by adding a self-loop to each node, resulting in the final adjacency matrix $\hat{A} = A + I_n$, where I_n is the identity matrix.

In the graph model, SST data are typically represented as a series of time-sequenced observations, $X = X_1, X_2, \ldots, X_t$, where X_t indicates the SST data for all nodes at time t. Our objective is to utilize the SST data from the past p time steps to train a model that can predict the SST data for the next q time steps. Formally, the model function can be expressed as

$$X_{t+1}, X_{t+2}, \ldots, X_{t+q} = \mathcal{F}(X_{t-p+1}, X_{t-p+2}, \ldots, X_t) \quad (1)$$

where $\mathcal{F}$ represents the forecasting function developed to estimate future SST values based on past observations.

B. AA-TGCN Structure

Fig. 2. illustrates the structure of the AA-TGCN model, which is primarily composed of temporal graph convolutional network (TGCN) blocks. Initially, the SST data X are processed to calculate its difference ΔX. Subsequently, both the original

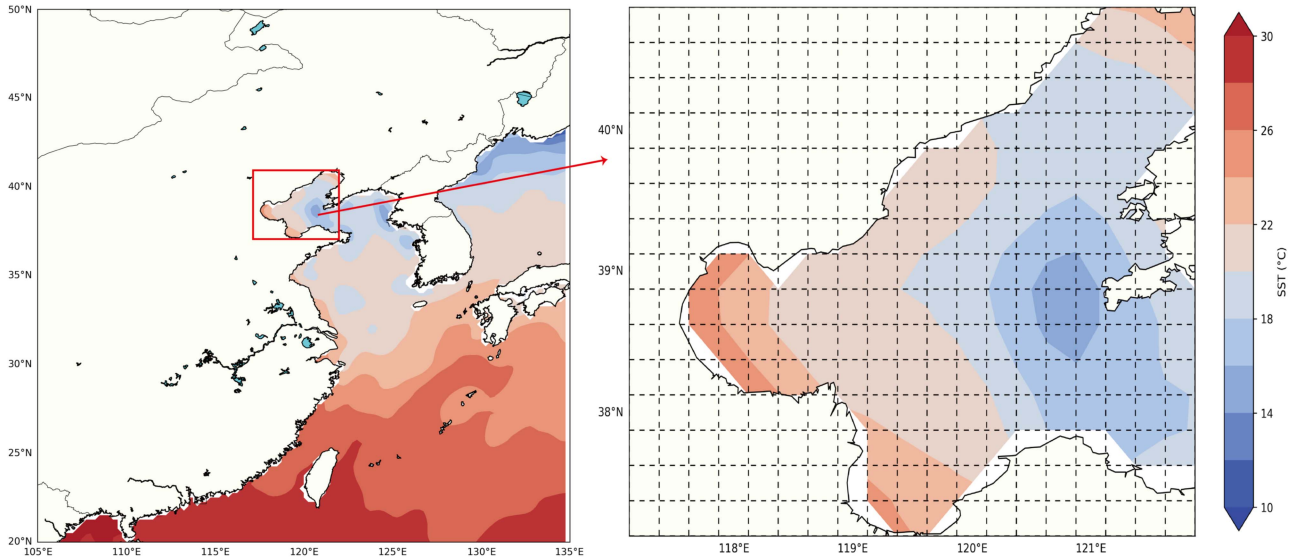

Fig. 1. Grid partitioning in part of the Bohai Sea, divided into 16×20 grid points, with 162 valid points.

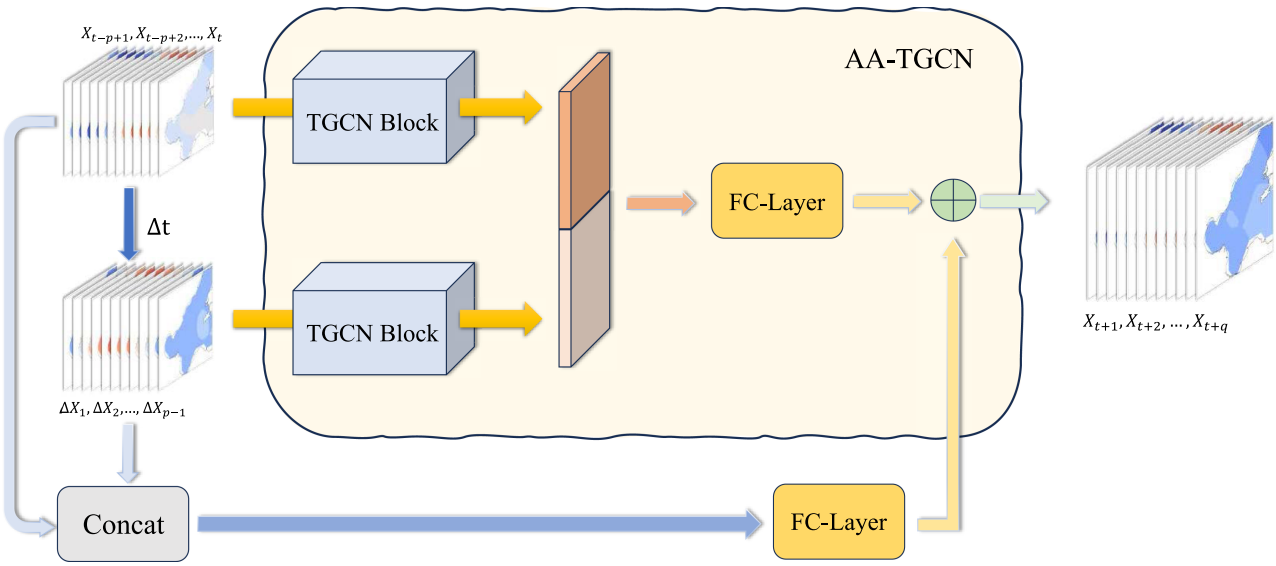

Fig. 2. AA-TGCN model, Δt indicates differencing the data in time.

data X and its difference ΔX are fed into the TGCN blocks for encoding, and the encoded information from both is concatenated and input into a fully connected layer for decoding. Before obtaining the final output, X and ΔX are concatenated and input into another fully connected layer, where they are added to the encoded information to implement a residual connection. This process results in the predicted output. Each TGCN module utilizes an attention mechanism and TCN [31] to jointly capture temporal and spatial information. The attention mechanism, with its unique pruning strategy, constructs a dynamic graph structure to capture spatial correlations in the data. TCN forms the main structure of TGCN. By expanding the receptive field of the model, it enables the model to capture temporal information. We effectively integrated the attention mechanism into the TCN structure, allowing each layer of TGCN to capture both temporal context information and spatial structure information.

C. TGCN Block

In studying the long-term monthly seasonal variations of SST, we note distinct differences from daily variations. Specifically, daily changes tend to be continuous, whereas monthly average SST variations are generally larger. To effectively capture these fluctuations, particularly seasonal ones, it is crucial that the encoded information comprehensively includes past variation data. The key lies in ensuring that each output maximizes the receptive field. Therefore, we have designed a structure similar to TCN, utilizing dilated convolution techniques. The purpose of

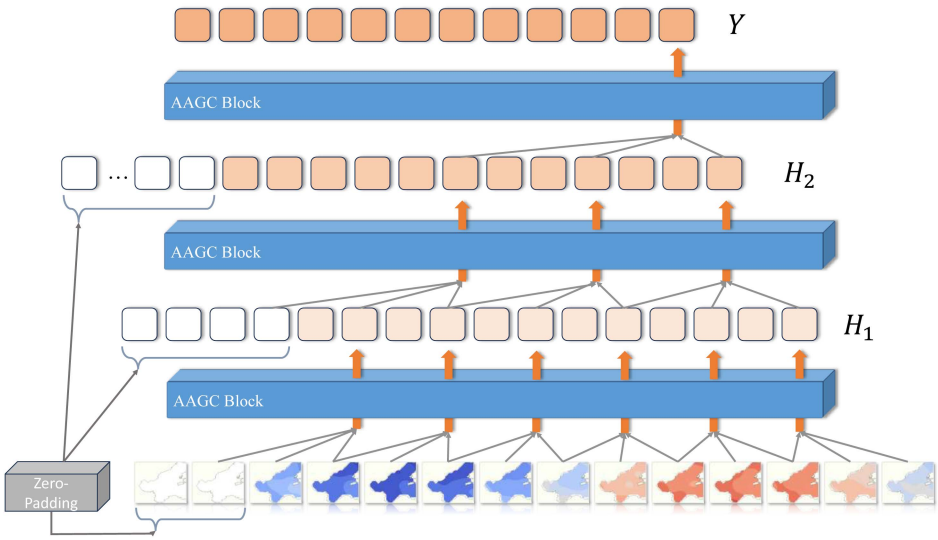

Fig. 3. TGCN block structure, where H_1 and H_2 encode information, and Y represents the output.

this structure is to expand the model's receptive field, allowing it to detect cross-quarterly fluctuations while ensuring predictions are based solely on past and present information to prevent "looking ahead." As shown in Fig. 3. Specifically, our model uses the past 12 months of monthly average SST data to predict the next 12 months' monthly SST. To cover the entire 12-month sequence with the receptive field and capture long-term dependencies, specific dilation rates, kernel sizes, and the number of convolution layers are selected to meet the required receptive field size. The relationship among these parameters can be determined using the following formula:

$$r = 1 + (K-1) \cdot \sum_{i=0}^{n-1} d^i \tag{2}$$

where r is the receptive field size, K is the convolutional kernel size, n is the number of convolutional layers, and d^i represents the dilation rate for the ith layer. In this experiment, we set the dilation base to 2, the kernel size to 3, and the total number of layers in the model to 3. The specific operational procedure is as follows. First, the input SST sequence is padded in the first layer to ensure that its length remains unchanged postconvolution. Concurrently, a size-3 AAGC block processes the sequence in a sliding manner. With a dilation rate of 1 in this layer, no additional expansion is required. The same procedure is applied in the second layer, where the dilation rate increases to 2. This increment means introducing a gap of one data point between every three consecutive data points, thereby enlarging the receptive field while maintaining consistency in sequence processing. The third layer follows a similar approach to further expand the receptive field, accommodating the long-term dependencies of the sequence. These operations can be described by the following formula:

$$H(t) = (x * \omega)(t) = \sum_{\tau=0}^{k-1} x\left(t - d \cdot \tau\right) \cdot \omega(\tau) \tag{3}$$

where H represents the encoded information, x is the input sequence, w denotes the weights of the AAGC block, t indicates the time step, k is the kernel size, and d is the dilation factor. Through this design, we ensure that each output from the model has a maximized receptive field, effectively capturing the long-term dependencies present in the SST data. This enables more accurate predictions of the monthly average SST fluctuations for the upcoming 12 months.

D. AAGC Block

To better capture the long-term dynamic correlations in SST, we have developed a method called attention adaptive graph convolution (AAGC). Unlike traditional adaptive graph convolution (AGC), which requires the initialization of a node embedding matrix to represent inter-node relationships before training, our AAGC method does not rely on a predefined node embedding matrix. This traditional approach necessitates access to the entire graph for transductive learning and needs retraining if there are changes in the graph's nodes, limiting its applicability in future or cross-disciplinary studies. Our AAGC method utilizes an attention mechanism to directly capture the interactions between nodes, thereby endowing the model with a stronger capability for inductive learning. Given that long-term monthly average SST predictions are susceptible to external influences and exhibit significant fluctuations, identifying and eliminating disadvantageous connections (or edges) becomes crucial. Traditional GAT often perform poorly in this respect, as they are prone to overfitting due to noise and fluctuations. Therefore, we have implemented a special sparse attention mechanism using the LO-norm as a strategy to identify and remove noisy edges in the graph. This approach effectively eliminates redundant information within the graph, enhancing model performance and suitability for long-term SST forecasting tasks. This is described in detail below.

A conventional multilayer GAT can be represented by the following:

$$h_i^{l+1} = \sigma\left(\sum_{j\in N_i} a_{ij}^{(l)} h_j^{(l)} W^{(l)}\right) \tag{4}$$

$$a_{ij}^{(l)} = \frac{\exp(\text{LeakyReLU}(b^{(l)}[W^{(l)}x_i \| W^{(l)}x_j]))}{\sum_{k\in N_i}\exp(\text{LeakyReLU}(b^{(l)}[W^{(l)}x_i \| W^{(l)}x_k]))} \tag{5}$$

where $h_i^{(l+1)}$ is the new feature of vertex i at layer $l+1$ obtained by aggregating neighbor information, $b^{(l)}$ represents the learnable attention coefficient at layer l, and W is the learnable weight matrix. The essence of the model is to identify significant edges. Inspired by the SGAT model [32], we introduced a binary mask to control the edges $e_{ij} \in E$, expressed as

$$\bar{A} = A \odot Z, Z \in \{0,1\}^M \tag{6}$$

where M is the number of edges, with $Z = 1$ if edge conditions are satisfied, otherwise $Z = 0$. Since Z is a set of binary variables and nondifferentiable, we employ a hard concrete estimator [33] to approximate the distribution of discrete variables. This approximation facilitates the optimization of the original problem and the generation of the binary mask with added randomness, represented as

$$\hat{\mathcal{R}}(W, \log\alpha) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_{u\sim U(0,1)}\mathcal{L}(f_i(X, A\odot g, W), y_i) + \lambda\sum_{(i,j)\in E}\sigma\left(\log\alpha_{ij} - \beta\log\frac{-\gamma}{\zeta}\right) \tag{7}$$

$$f(\log\alpha, u) = \sigma((\log u - \log(1-u) + \log\alpha)/\beta)(\zeta-\gamma) + \gamma \tag{8}$$

$$g = \min(1, \max(0, g(f(\log\alpha, u)))) \tag{9}$$

where $\mathbb{E}$ denotes the expectation, u is uniformly distributed between [0,1], f is the encoding function, and λ is a regularization parameter. σ is the sigmoid function, $\beta = \frac{2}{3}, \gamma = -0.1$, and $\zeta = 1.1$ are parameters typical for the hard concrete distribution. The log attention coefficients $log\alpha_{ij}$ are computed as follows:

$$\log\alpha_{ij} = b^T(x_i W_1^{(1)} \| x_j W_1^{(1)}) \tag{10}$$

where b^T is a learnable parameter, and $W_1^{(1)}$ is the weight matrix for the first layer and first head. From this, we derive the mask Z_{ij} from the hard concrete distribution $q(z|\log\alpha)$. During the testing phase, randomness is eliminated, and a deterministic mask $\hat{Z}$ is generated using

$$\hat{Z} = \min(1, \max(0, \sigma((\log\alpha)/\beta)(\zeta-\gamma)+\gamma)). \tag{11}$$

This process generates a continuous mask within the range of 0–1, effectively pruning redundant connections. Subsequently, normalization is applied

$$a_{ij} = \frac{A_{ij}z_{ij}}{\sum_{k\in N_i} A_{ik}z_{ik}}. \tag{12}$$

To preserve the information of the nodes themselves, we set $z_{ii} = 1$. In multilayer graph attention networks, using different attention coefficients for each layer is redundant for our predictive tasks; instead, assigning the same attention coefficient across layers can significantly reduce computational time. This uniform attention coefficient simplifies understanding the importance of edges and facilitates targeted learning by the model. This approach has been validated in SGAT implementations. Therefore, we derive the following formula:

$$h_i^{l+1} = \sigma\left(\sum_{j\in N_i} a_{ij} h_j^{(l)} \mathrm{W_k}^{(l)}\right). \tag{13}$$

We then obtain a multihead, multilayer attention model expression as follows:

$$h_i^{l+1} = \sigma\left(\sum_{j\in N_i} a_{ij} h_j^{(l)} \mathrm{W_k}^{(l)}\right). \tag{14}$$

Through these steps, we have established an attention adaptive graph convolution (AAGC) block based on adaptive attention weights. The structure of this block can be straightforwardly visualized in Fig. 4. After the sea temperature data are processed and converted into node data, the attention mechanism is used to obtain the relevant edge weights between nodes, resulting in the initial graph structure. At this stage, the graph structure contains many noisy edges. Through a special pruning operation using the hard concrete distribution, sparse sampling is performed to obtain the pruned edges, thus achieving the final graph structure. Finally, the graph structure is combined with the sea temperature node data, and neighbor aggregation is performed on the nodes to obtain new feature representations.

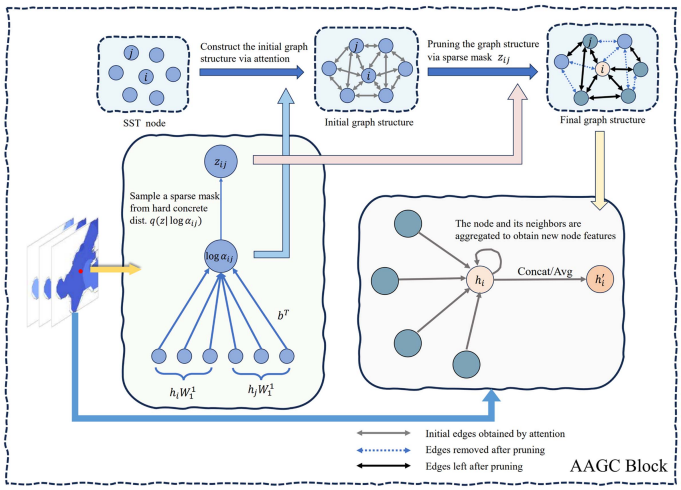

Fig. 4. AAGC block structure.

E. Complexity

We will analyze the time and space complexity of AA-TGCN model in detail.

1) Time Complexity: The base AAGC module of AA-TGCN is a variant similar to the GAT model, so its time complexity is similar to that of GAT. During the training phase, GAT mainly

involves two multiplication operations. First is the feature mapping of the vertices, i.e., mapping each node's features from F dimensions to H dimensions, where H is the dimension of the hidden layer. The time complexity for this is $O(N \times F \times H)$. The next step is the computation of attention weights, where attention weights are computed between each node and its neighbors, resulting in a time complexity of $O(E \times H)$, where E is the number of edges of the graph. Additionally, AAGC requires the computation of a mask Z for each attention weight, with a time complexity of $O(E)$. Thus, the time complexity for each layer of the AAGC module is $O(K(N \times F \times H + E \times H) + E)$, where K is the number of attention heads.

The main structure of AA-TGCN is the TGCN module. This module, similar to the TCN structure, in our design, the AAGC module's kernel size is 3, and there are three layers, with each layer's output being 12 new feature representations for each node. Hence, the time complexity for the entire TGCN module is $O(3 \times 12 \times (K(N \times F \times H + E \times H) + E))$, where F is the kernel size of 3. For the entire AA-TGCN, which includes two TGCN modules for the original data and differential data, the time complexity analysis is similar. Therefore, simplifying the model's time complexity, it becomes $O(K(N \times H + E \times H) + E)$.

2) Space Complexity: For the space complexity of AA-TGCN, we start by considering the model's base components. The space complexity of the AAGC module is primarily determined by the following factors. First is the feature storage of sea temperature node data, with a space complexity of $O(N \times F)$. Second is the feature space complexity after transformation, which is $O(N \times H \times K)$. Next is the space complexity of the computed attention weights, $O(E \times K)$, followed by the storage of model weight parameters, $O(F \times H \times K)$. Lastly, there is the storage of the mask Z, with a space complexity of $O(E)$. Thus, the space complexity of the AAGC module is $O(N \times H \times K + E \times K + F \times H \times K + E)$. Similarly, under the TGCN structure, each layer's space storage is constant, hence for the AA-TGCN model, the space complexity can be simplified to $O(N \times H \times K + E \times K + H \times K + E)$.

III. Experiments

A. Datasets

Our study utilizes the optimum interpolation sea surface temperature (OISST) V2.1 dataset provided by the National Oceanic and Atmospheric Administration (NOAA), based in Boulder, Colorado, USA. This dataset is available online[1] [34]. It offers daily, weekly, and monthly average SST data globally from September 1981, covering from 89.875° South to 89.875° North latitude and from 0.125° East to 359.875° West longitude, with a spatial resolution of 0.25° × 0.25°.

For this research, we selected monthly average SST data for specific regions in the Bohai Sea and the South China Sea, spanning from January 1982 to December 2021. Specifically, the data range for the Bohai Sea is from 117° to 122° East longitude and from 37° to 41° North latitude, while for the South China Sea, it is from 112° to 117° East longitude and from 12° to 16° North latitude.

To train and validate our model, we used data from 1982 to 2001 as the training set, from 2002 to 2011 as the validation set, and from 2012 to 2021 as the test set for analysis.

B. Settings

Before conducting comparative experiments, we need to select different hyperparameters to test the sensitivity of the model to parameter selection. We have chosen the following hyperparameters for the experiment: hidden layer dimensions: [16, 32, 64]; learning rates: [0.001, 0.0001]; number of attention heads: [4, 8, 16]. We use grid search to obtain experimental results for different combinations. Finally, we determine the optimal hyperparameters to be a learning rate of 0.001, a hidden layer dimension of 32, and four attention heads. Additionally, other parameter choices exhibit robust performance.

In this study, we compared the performance of attention adaptive AA-TGCN with EA-GCN, sparse graph attention network (SGAT), GAT, TSGN, ConvLSTM, TCN, LSTM, and support vector regression (SVR). All models were trained using the same sliding window length, predicting the monthly average SST for the next 12 months based on the past 12 months' data. A 12-month period was chosen for the sliding window to cover all seasonal variations, providing comprehensive background information for the models.

To ensure the fairness of the experiments and the validity of the model structure comparison, all models were set to have the same dimensionality in their hidden layers, fixed at 16. The AA-TGCN's AAGC block, along with SGAT and GAT, implemented a dual-layer structure equipped with eight attention heads. EA-GCN, due to differences in the research field, did not have daily average SST data input into the model for fairness. TSGN aggregated using two layers of LSTM units. ConvLSTM and LSTM used a two-layer structure. TCN and AA-TGCN used the same kernel size, dilation rate, and number of layers. SVR optimized its parameters through grid search.

We used the Python programming language and frameworks such as Torch and DGL to build and test all neural network models. The experiments were conducted on a high-performance computing platform equipped with a 5.4 GHz Intel Core processor, an 8 GB NVIDIA GeForce RTX 4060 graphics card, 32 GB of 5600 MHz DDR5 RAM, and a 1 TB solid-state drive.

The models were trained using the Adam optimizer with an early stopping mechanism (patience set at 20 iterations) to prevent overfitting and ensure model stability. The learning rate was set at 0.001, with a batch size of 8.

To assess the performance of the different models on the SST prediction task, we used mean squared error (MSE) and mean absolute error (MAE) as performance metrics. Lower values of these metrics indicate better predictive performance of the model.

C. Result

Tables I and II, respectively, present the experimental results of five different methods (AA-TGCN, EA-GCN SGAT, TSGN,

[1] [Online]. Available: www.ncei.noaa.gov/

TABLE I
PREDICTION PERFORMANCE OF VARIOUS MODELS AT DIFFERENT FORECASTING SCALES IN THE BOHAI SEA

Sea Area	Metrics	Model	Monthly											
			1	2	3	4	5	6	7	8	9	10	11	12
Bohai Sea	MSE	SVR	0.6636	0.7675	0.8425	0.9051	0.9518	0.9934	1.0218	1.0488	1.0670	1.0800	1.0870	1.0878
		LSTM	0.8587	0.8066	1.0109	1.0859	1.0797	1.0333	1.1032	1.0421	1.1227	1.1248	1.0699	1.0784
		TCN	0.6187	0.7934	0.8346	0.8672	0.9007	0.9398	1.0951	0.9439	0.9688	0.9542	1.0250	0.9875
		ConvLSTM	0.8091	1.1040	0.9767	1.0144	1.0239	0.9846	1.1049	1.0891	1.1094	1.0791	1.1219	1.1096
		GAT	0.6839	0.8300	0.8934	1.0613	1.0121	1.0774	1.1593	1.0809	1.0902	1.0230	1.0372	1.0613
		TSGN	0.6373	0.8429	0.8431	0.9546	1.0506	0.9742	1.0226	0.9419	1.0530	1.0359	0.9932	1.0191
		SGAT	0.6587	0.7311	0.7833	0.8398	0.8696	0.9226	0.9334	0.9457	0.9576	0.9590	0.9656	0.9584
		EA-GCN	0.8640	0.9095	0.9373	0.8835	1.0092	0.9596	1.1181	1.0054	0.9613	1.0581	1.0401	0.9368
		AA-TGCN	**0.5957**	**0.6913**	**0.7548**	**0.7751**	**0.8653**	**0.8731**	**0.8798**	**0.8949**	**0.9210**	**0.9026**	**0.8979**	**0.8972**
	MAE	SVR	0.6394	0.6875	0.7219	0.7490	0.7700	0.7878	0.7995	0.8114	0.8195	0.8251	0.8278	0.8280
		LSTM	0.7384	0.7037	0.7978	0.8318	0.8235	0.8027	0.8259	0.8073	0.8414	0.8380	0.8128	0.8207
		TCN	0.6225	0.6973	0.7132	0.7294	0.7471	0.7580	0.8228	0.7645	0.7730	0.7698	0.7837	0.7771
		ConvLSTM	0.6977	0.8318	0.7825	0.7889	0.8000	0.7771	0.8273	0.8254	0.8283	0.8171	0.8337	0.8374
		GAT	0.6481	0.7155	0.7422	0.8156	0.8073	0.8246	0.8729	0.8325	0.8338	0.8037	0.8111	0.8144
		TSGN	0.6318	0.7277	0.7134	0.7710	0.8069	0.7776	0.7937	0.7601	0.8100	0.7995	0.7867	0.7912
		SGAT	0.6471	0.6709	0.6960	0.7162	0.7322	0.7551	0.7595	0.7666	0.7728	0.7716	0.7764	0.7746
		EA-GCN	0.7368	0.7653	0.7710	0.7286	0.7965	0.7653	0.8347	0.7873	0.7602	0.8127	0.8059	0.7565
		AA-TGCN	**0.6117**	**0.6473**	**0.6727**	**0.6761**	**0.7263**	**0.7295**	**0.7293**	**0.7348**	**0.7490**	**0.7416**	**0.7428**	**0.7375**

TABLE II
PREDICTION PERFORMANCE OF VARIOUS MODELS AT DIFFERENT FORECASTING SCALES IN THE SOUTH CHINA SEA

Sea Area	Metrics	Model	Monthly											
			1	2	3	4	5	6	7	8	9	10	11	12
South China Sea	MSE	SVR	0.5154	0.6493	0.7293	0.7702	0.7981	0.8058	0.8104	0.8056	0.7913	0.7749	0.7623	0.7590
		LSTM	0.4499	0.5702	0.6888	0.6381	0.6850	0.7877	0.7778	0.6327	0.7465	0.6420	0.6565	0.7411
		TCN	0.3508	0.4641	0.4682	0.4648	0.5360	0.5276	0.5686	0.5372	0.5874	0.5991	0.6002	0.6306
		ConvLSTM	0.4041	0.5005	0.4869	0.5184	0.5500	0.5508	0.5621	0.5657	0.5733	0.6030	0.5612	0.5664
		GAT	0.6136	0.5538	0.7042	0.8053	0.7533	0.7938	0.6761	0.6006	0.5717	0.5582	0.7482	0.5405
		TSGN	0.3959	0.5521	0.4843	0.5394	0.5461	0.5635	0.6544	0.6232	0.6064	0.5928	0.6170	0.6307
		SGAT	0.5001	0.6361	0.7442	0.6158	0.7262	0.5586	0.6484	0.6897	0.5616	0.5847	0.6417	0.6081
		EA-GCN	0.3799	0.4486	0.5681	0.6211	0.5678	0.6009	0.6488	0.5984	0.6242	0.6305	0.6119	0.6064
		AA-TGCN	**0.3249**	**0.3944**	**0.4065**	**0.4429**	**0.4463**	**0.4728**	**0.4713**	**0.4797**	**0.4609**	**0.4875**	**0.5013**	**0.4758**
	MAE	SVR	0.5627	0.6278	0.6633	0.6857	0.7031	0.7107	0.7124	0.7099	0.7045	0.6983	0.6932	0.6927
		LSTM	0.5434	0.6049	0.6643	0.6469	0.6639	0.7059	0.6967	0.6404	0.6951	0.6454	0.6504	0.6824
		TCN	0.4697	0.5523	0.5443	0.5418	0.5912	0.5768	0.5971	0.5848	0.6132	0.6173	0.6232	0.6376
		ConvLSTM	0.5006	0.5605	0.5569	0.5720	0.6022	0.5932	0.5985	0.6027	0.6101	0.6266	0.6002	0.6026
		GAT	0.6402	0.5920	0.6656	0.7141	0.6978	0.7206	0.6545	0.6173	0.6015	0.5962	0.6944	0.5822
		TSGN	0.5152	0.5865	0.5588	0.5789	0.5920	0.5964	0.6450	0.6238	0.6201	0.6162	0.6256	0.6376
		SGAT	0.5637	0.6597	0.6977	0.6317	0.6766	0.5948	0.6513	0.6650	0.5970	0.6125	0.6441	0.6274
		EA-GCN	0.4931	0.5327	0.6185	0.6355	0.6066	0.6244	0.6467	0.6205	0.6403	0.6411	0.6310	0.6240
		AA-TGCN	**0.4545**	**0.5018**	**0.5023**	**0.5287**	**0.5332**	**0.5438**	**0.5376**	**0.5497**	**0.5356**	**0.5576**	**0.5626**	**0.5474**

GAT, ConvLSTM, TCN, LSTM, SVR) on datasets from the Bohai Sea and the South China Sea.

We visualized the MSE and MAE results for the different models across these regions, as shown in the Fig. 5.

The results are presented in Tables I and II, where the bold numbers indicate the best outcomes. It is evident from the data that our AA-TGCN model outperforms other models across various temporal scales of long-term prediction. Specifically, compared to the next best model, AA-TGCN shows an average improvement of 5% in MSE and 4% in MAE for the Bohai Sea dataset; and an even more significant improvement in the South China Sea dataset, with increases of 15% in MSE and 8% in

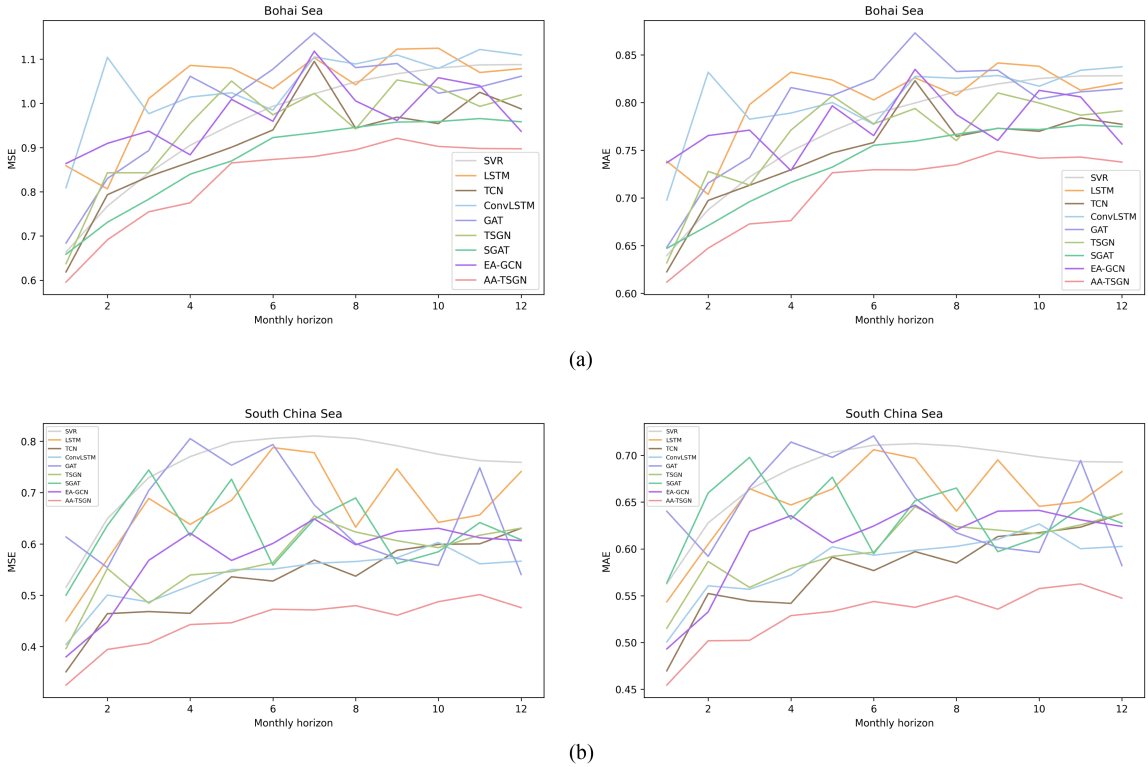

Fig. 5. Prediction results of different models at various forecasting scales in different regions, left subfigure shows MSE evaluation, right subfigure shows MAE evaluation. (a) Bohai Sea. (b) South China Sea.

MAE. Notably, we observed that all models generally perform better in the South China Sea than in the Bohai Sea, which can be attributed to different geographic and meteorological conditions of the two maritime areas. The Bohai Sea, located in northern China, is a semienclosed sea surrounded by land. This geographic characteristic means its hydrological and meteorological conditions are heavily influenced by the surrounding land, which itself varies greatly. In contrast, the South China Sea is an open sea connected to the Pacific Ocean, with a larger area and more frequent water exchanges, providing a relatively stable environment.

Further analysis of different models in various maritime areas showed that the LSTM model performs worse compared to the TCN model, which demonstrates that, despite both being time series models, TCN is more suitable for long-term unstable sequence prediction. This is because TCN can aggregate overall information by expanding the receptive field.

Additionally, we can observe that ConvLSTM performs poorly in the Bohai area but better in the South China Sea area. This is because the presence of some land in the Bohai area causes issues with convolutional effects. The EA-GCN model uses GRU, a variant of LSTM, to aggregate temporal information, which might be the primary reason for its poor performance. The SGAT model has consistent predictive performance in the Bohai Sea but exhibits significant fluctuations in the South China Sea. This variation may be due to SGAT's pruning operations that control information aggregation, which are particularly effective in the unstable and externally influenced Bohai Sea environment, removing noise and enhancing model stability. However, in the South China Sea, where the environment is more stable and exchanges with other water bodies are frequent, SGAT's pruning may inadvertently remove nodes containing useful information. It was observed that the AA-TGCN model remains stable across different maritime regions, demonstrating its robustness.

We also found that SVR performed better than other models in the Bohai Sea in some respects, possibly due to the lower temporal and spatial resolution and smaller data size in the Bohai Sea, which may prevent deep learning models from fully exploiting their advantages and lead to overfitting. In contrast, SVR is comparatively more robust. The significant performance enhancement of the AA-TGCN model in the South China Sea validates its capability in capturing stable long-term seasonal trends, underscoring its superior robustness. These findings not only highlight the potential of the AA-TGCN model but also provide valuable insights for more accurate temperature prediction in different maritime regions in the future.

IV. Discussion

A. Seasonal Error Analysis

We conducted tests on various models in different seasons for the Bohai Sea in Northern China and the South China Sea in the south to analyze the predictive effectiveness of each model and their relationship with seasonal fluctuations. Using the year 2020 as a case study, we predicted the entire year's monthly

TABLE III
PREDICTION PERFORMANCE OF DIFFERENT AA-TGCN VARIANTS IN THE BOHAI SEA AT VARIOUS FORECASTING SCALES DURING ABLATION STUDIES

Model	AA-TGCN		R-LSTM		R-Adapt		D-Differences	
Metrics	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
1	0.5957	0.6117	0.7237	0.6796	0.8444	0.7354	0.7528	0.6837
2	0.6913	0.6473	0.8400	0.7213	0.8044	0.6902	0.6928	0.6519
3	0.7548	0.6727	0.8383	0.7226	0.8081	0.6935	0.8089	0.6977
4	0.7751	0.6761	0.9650	0.7740	0.9022	0.7461	0.9109	0.7436
5	0.8653	0.7263	1.0581	0.8126	0.9208	0.7557	0.8797	0.7289
6	0.8731	0.7295	0.9620	0.7709	0.9687	0.7731	0.9468	0.7595
7	0.8798	0.7293	1.0104	0.7918	0.9693	0.7718	0.9039	0.7447
8	0.8949	0.7348	1.1202	0.8487	0.9555	0.7681	0.9550	0.7628
9	0.9210	0.7490	1.0464	0.8055	1.0200	0.7919	0.9378	0.7596
10	0.9026	0.7416	1.0887	0.8255	0.8718	0.7291	0.9653	0.7672
11	0.8979	0.7428	1.0554	0.8100	1.0162	0.7959	0.9191	0.7494
12	0.8972	0.7375	1.0359	0.8042	0.8687	0.7297	0.9140	0.7416

average SST using data from 2019 and grouped the results by season to calculate the MAE for each season. Additionally, we compared the seasonal fluctuation absolute errors between 2019 and 2020, as illustrated in Figs. 6 and 7.

From these figures, it is evident that the error distribution areas of different models often overlap significantly with the areas of seasonal fluctuations, indicating that the accuracy of seasonal SST predictions is linked to these annual seasonal variations. Specifically, while models generally capture the long-term SST trends over different periods, they struggle to predict seasonal SST trends accurately.

Comparing different models, we observed that GAT shows instability in its predictions under complex fluctuation conditions. TSGN, which uses a unique node aggregation method (LSTM), exhibits uneven prediction results. Interestingly, EA-GCN uses a variant of LSTM, GRU, as its main structure, while ConvLSTM uses LSTM as its main structure. The error distributions of the three are very similar, especially in the South China Sea region. And both SGAT and AA-TGCN employ specialized pruning techniques for node aggregation, resulting in similar areas of error distribution, with our AA-TGCN fitting the long-term seasonal fluctuations of SST more accurately.

These comparisons validate that our model outperforms the others in terms of prediction stability, error distribution, and overall performance, confirming its superiority in forecasting seasonal SST variations accurately.

B. Ablation Experiment

We conducted a series of ablation experiments to explore and validate the effectiveness and necessity of our model. These experiments primarily involved removing or replacing certain modules within the model and performing predictions for the Bohai Sea area. The specific experimental results are summarized in Table III, where R-LSTM denotes replacing the TGCN structure with an LSTM structure that includes the AAGC block, R-Adapt indicates replacing the attention adaptive dynamic graph with an embedding-based adaptive dynamic graph. D-Differences refers to the deletion of the differencing auxiliary. We have visualized these results, as shown in the Fig. 8.

The results indicate that R-Adapt outperformed AA-TGCN in the October and December forecasting periods. This could be attributed to the complexity of influences on sea temperature in long-term, large-scale predictions. Unlike AA-TGCN, which controls edges with a pruning objective, R-Adapt adjusts edge weights by initializing a learnable N*D dimensional node embedding matrix, allowing it to capture more information after selecting an appropriate D. However, this method is also more susceptible to noise, potentially leading to overfitting, as indicated by the fluctuating predictive performance shown in the figures. On the other hand, AA-TGCN might remove informative edges through its pruning process. Nonetheless, such modifications are necessary; we observed that AA-TGCN generally shows better performance and stability on other temporal scales, whereas R-Adapt's results are more volatile. Most importantly, AA-TGCN changes the nature of the model from a transductive to an inductive learning approach, allowing it to predict unseen data without retraining. In contrast, transductive learning would require retraining upon the introduction of new nodes. This broadens the applicability of AA-TGCN, affirming the necessity of this modification. Furthermore, examining changes in other modules, we found that R-LSTM performed poorly. Unlike the continuous changes in fine-grained long-term predictions, coarse-grained long-term predictions are more volatile.

The former can be easily remembered by an LSTM structure, while the latter may disrupt the memory capability of LSTM's memory cells due to significant fluctuations. Thus, AA-TSGN more effectively integrates overall information and captures long-term seasonal trends. The removal of the D-Differences differencing module also led to fluctuations in predictive

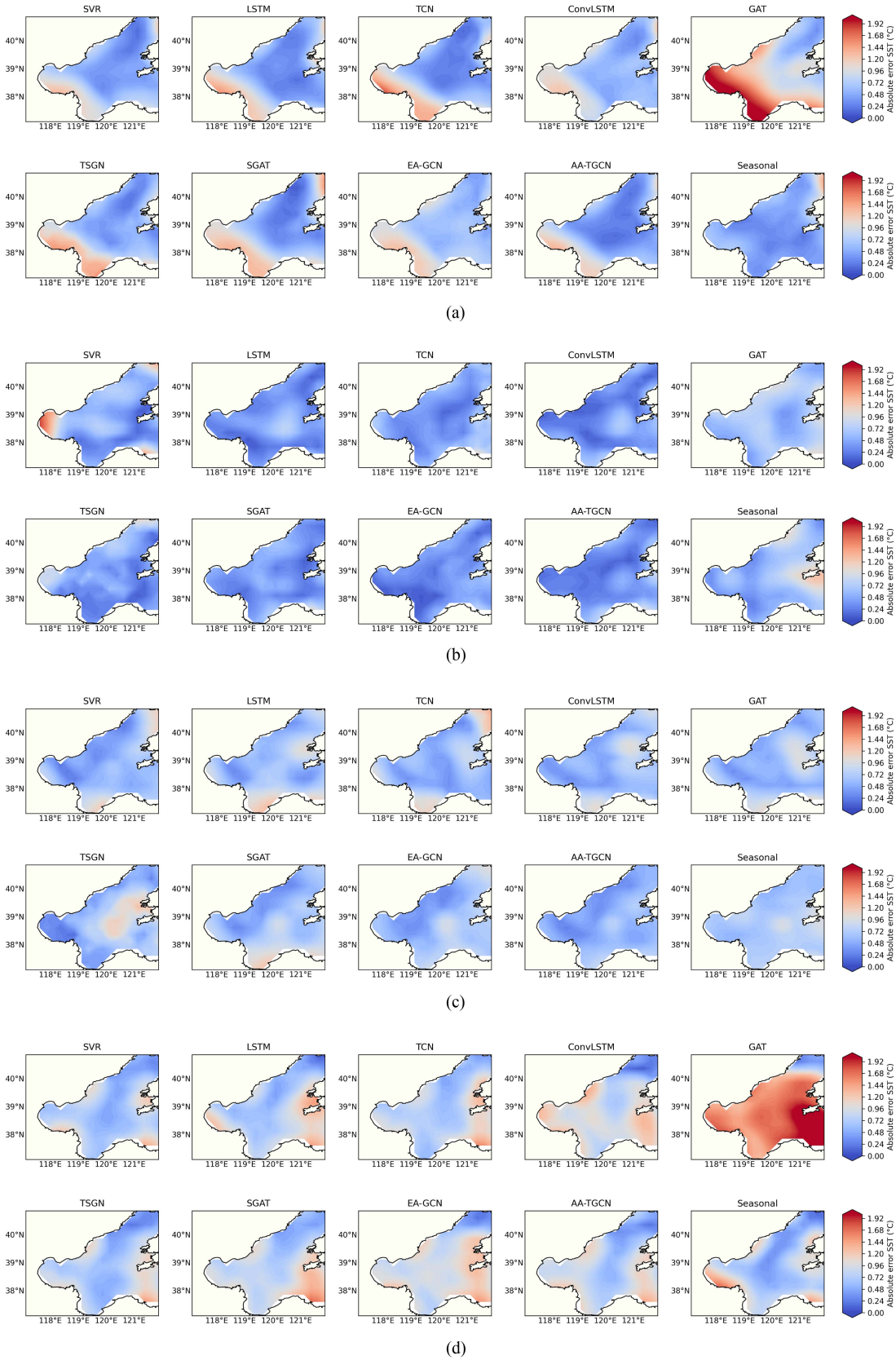

Fig. 6. Average absolute prediction error by season in the Bohai Sea using different models, "Seasonal" refers to the distribution of absolute differences for the same season over two years. (a) Spring. (b) Summer. (c) Autumn. (d) Winter.

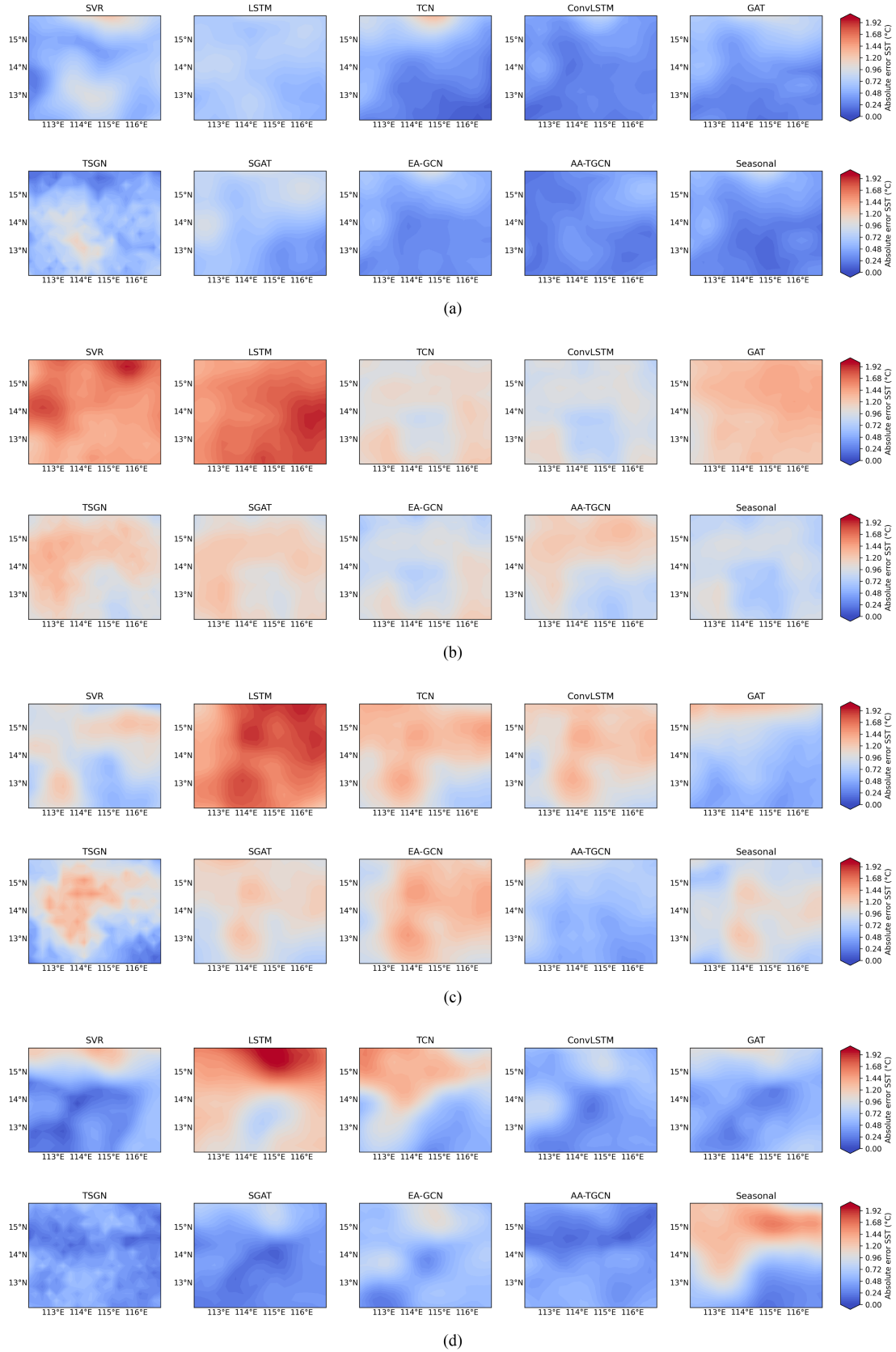

Fig. 7. Average absolute prediction error by season in the South China Sea using different models, "Seasonal" refers to the distribution of absolute differences for the same season over two years. (a) Spring. (b) Summer. (c) Autumn. (d) Winter.

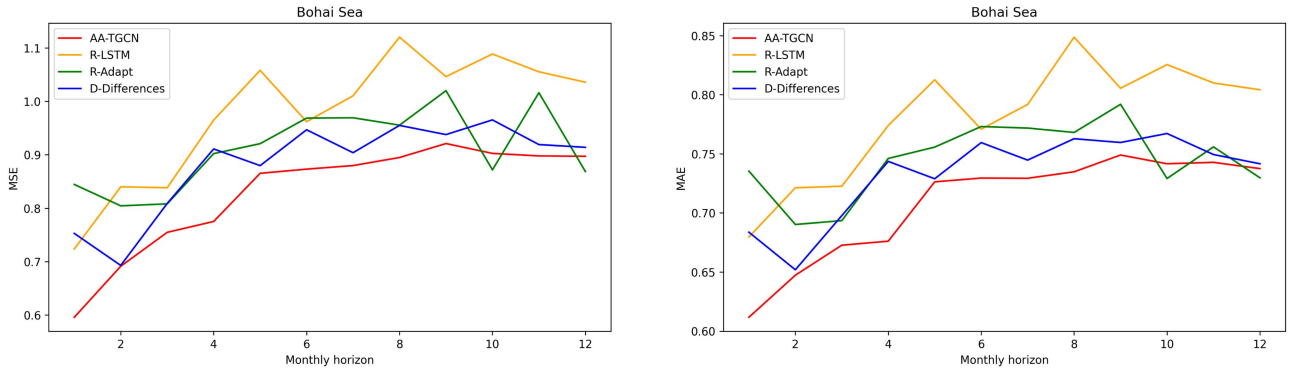

Fig. 8. Prediction performance of different AA-TGCN variants in the Bohai Sea at various forecasting scales during ablation studies, left subfigure is MSE evaluation, right subfigure is MAE evaluation.

performance by losing the ability to enhance memory through differencing. Overall, AA-TGCN demonstrated superior stability and effectiveness, and each module's design was thoroughly justified and necessary.

V. Conclusion

In this article, we introduce a novel AA-TGCN model aimed at addressing the challenge of long-term seasonal prediction of SST. We initially incorporate an attention mechanism to learn internode correlations and utilize a unique binary mask for pruning edges, which enhances the model's inductive learning capabilities beyond those of traditional adaptive GNNs. Subsequently, we design a structure similar to TCN to maximize the model's receptive field, enabling it to observe annual SST seasonal fluctuations and long-term trends. Through the application of first-order differencing, we further improve our model's ability to capture SST seasonal variations. Experiments conducted in the Bohai Sea and South China Sea validate its excellent performance in long-term seasonal SST forecasting.

References

[1] Y. Yao and C. Wang, "Variations in summer marine heatwaves in the South China Sea," *J. Geophysical Res.: Oceans*, vol. 126, no. 10, Oct. 2021, Art. no. e2021JC017792.

[2] W. Cai, B. Ng, G. Wang, A. Santoso, L. Wu, and K. Yang, "Increased ENSO sea surface temperature variability under four IPCC emission scenarios," *Nature Climate Change*, vol. 12, no. 3, pp. 228–231, Mar. 2022.

[3] E. C. J. Oliver et al., "Marine heatwaves," 2020.

[4] K. W. Vollset et al., "Ecological regime shift in the Northeast Atlantic Ocean revealed from the unprecedented reduction in marine growth of Atlantic salmon," *Sci. Adv.*, vol. 8, no. 9, Mar. 2022, Art. no. eabk2542.

[5] J. Garrabou et al., "Marine heatwaves drive recurrent mass mortalities in the Mediterranean sea," *Glob. Change Biol.*, vol. 28, no. 19, pp. 5708–5725, Oct. 2022.

[6] M. B. Menary, J. Mignot, and J. Robson, "Skilful decadal predictions of subpolar North Atlantic SSTs using CMIP model-analogues," *Environ. Res. Lett.*, vol. 16, no. 6, Jun. 2021, Art. no. 0 64090.

[7] X. Zhu, T. Dong, S. Zhao, and W. He, "A comparison of global surface air temperature over the oceans between CMIP5 models and NCEP reanalysis," *Front. Environ. Sci.*, vol. 9, May 2021, Art. no. 656779.

[8] Y. Wang, Z. Zhang, and P. Huang, "An improved model-based analogue forecasting for the prediction of the tropical Indo-Pacific sea surface temperature in a coupled climate model," *Int. J. Climatol.*, vol. 40, no. 15, pp. 6346–6360, Dec. 2020.

[9] M. Imani et al., "Spatiotemporal prediction of satellite altimetry sea level anomalies in the tropical Pacific Ocean," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1126–1130, Jul. 2017.

[10] D. E. Lee, D. Chapman, N. Henderson, C. Chen, and M. A. Cane, "Multilevel vector autoregressive prediction of sea surface temperature in the north tropical Atlantic ocean and the Caribbean sea," *Climate Dyn.*, vol. 47, no. 1/2, pp. 95–106, Jul. 2016.

[11] T. Zhao et al., "Artificial intelligence for geoscience: Progress, challenges and perspectives[J]," *Innovation*, vol. 5, no. 5, Art. no. 100691.

[12] Y. Han, K. Sun, J. Yan, and C. Dong, "The CNN-GRU model with frequency analysis module for sea surface temperature prediction," *Soft Computing*, vol. 27, no. 13, pp. 8711–8720, 2023, doi: 10.1007/S00500-023-08172-2.

[13] C. Xiao, N. Chen, C. Hu, K. Wang, J. Gong, and Z. Chen, "Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach," *Remote Sens. Environ.*, vol. 233, Nov. 2019, Art. no. 111358.

[14] L. Wei and L. Guan, "Seven-day sea surface temperature prediction using a 3DConv-LSTM model," *Front. Mar. Sci.*, vol. 9, Dec. 2022, Art. no. 905848.

[15] M. Jahanbakht, W. Xiang, and M. R. Azghadi, "Sea surface temperature forecasting with ensemble of stacked deep neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 1502605.

[16] Q. Zhang, H. Wang, J. Dong, G. Zhong, and X. Sun, "Prediction of sea surface temperature using long short-term memory," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1745–1749, Oct. 2017.

[17] J. Xie, J. Zhang, J. Yu, and L. Xu, "An adaptive scale sea surface temperature predicting method based on deep learning with attention mechanism," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 5, pp. 740–744, May 2020.

[18] C. Xiao et al., "A spatiotemporal deep learning model for sea surface temperature field prediction using time-series satellite data," *Environ. Model. Softw.*, vol. 120, Oct. 2019, Art. no. 104502.

[19] L. Xiao, S. Li, and B. Chen, "Gssa: A network for short to medium-term regional sea surface temperature prediction," *IEEE Geosci. Remote Sens. Lett.*, vol. 21, 2024, Art. no. 1504105.

[20] J. Pan, Z. Li, S. Shi, L. Xu, J. Yu, and X. Wu, "Adaptive graph neural network based south China sea seawater temperature prediction and multivariate uncertainty correlation analysis," *Stochastic Environ. Res. Risk Assessment*, vol. 37, no. 5, pp. 1877–1896, May 2023.

[21] H. Peng, C. Jin, W. Li, and J. Guan, "Enhanced adaptive graph convolutional network for long-term fine-grained SST prediction," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 7968–7978, 2023.

[22] Z. Gao, Z. Li, J. Yu, and L. Xu, "Global spatiotemporal graph attention network for sea surface temperature prediction," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, pp. 1–5, 2023.

[23] X. Zhang, Y. Li, A. C. Frery, and P. Ren, "Sea surface temperature prediction with memory graph convolutional networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[24] J. Kim, T. Kim, J.-G. Ryu, and J. Kim, "Spatiotemporal graph neural network for multivariate multi-step ahead time-series forecasting of sea temperature," *Eng. Appl. Artif. Intell.*, vol. 126, Nov. 2023, Art. no. 106854.

[25] Y. Sun, X. Yao, X. Bi, X. Huang, X. Zhao, and B. Qiao, "Time-series graph network for sea surface temperature prediction," *Big Data Res.*, vol. 25, Jul. 2021, Art. no. 100237.

[26] J. D. Beverley, M. Newman, and A. Hoell, "Rapid development of systematic ENSO-Related seasonal forecast errors," *Geophysical Res. Lett.*, vol. 50, no. 10, May 2023, Art. no.e2022GL102249.

[27] K. E. Lee et al., "Roles of insolation forcing and CO_2 forcing on late pleistocene seasonal sea surface temperatures," *Nature Commun.*, vol. 12, no. 1, Sep. 2021, Art. no. 5742.
[28] P. V. V. Le et al., "Climate-driven changes in the predictability of seasonal precipitation," *Nature Commun.*, vol. 14, no. 1, Jun. 2023, Art. no. 3822.
[29] B. Sun, B. Li, J. Yan, Y. Zhou, and S. Zhou, "Seasonal variation of atmospheric coupling with oceanic mesoscale eddies in the north Pacific subtropical countercurrent," *Acta Oceanologica Sinica*, vol. 41, no. 10, pp. 109–118, Oct. 2022.
[30] A. R. Jo, J.-Y. Lee, A. Timmermann, F.-F. Jin, R. Yamaguchi, and A. Gallego, "Future amplification of sea surface temperature seasonality due to enhanced ocean stratification," *Geophysical Res. Lett.*, vol. 49, no. 9, May 2022, Art. no. e2022GL098607.
[31] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," 2018, *arXiv:1803.01271*.
[32] Y. Ye and S. Ji, "Sparse graph attention networks," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 1, pp. 905–916, Jan. 1, 2023.
[33] C. Louizos, M. Welling, and D. P. Kingma, "Learning sparse neural networks through L_0 regularization," in *Proc. Int. Conf. Learn. Representations*, Vancouver, BC, Canada, 2018.
[34] B. Huang et al., "Improvements of the daily optimum interpolation sea surface temperature (DOISST) version 2.1," *J. Climate*, vol. 34, no. 8, pp. 2923–2939, Apr. 2021.

Ling Xiao was born in 2001. He received the bachelor's degree in information and computational science in 2023 from Guangdong Ocean University, Zhanjiang, China, where he is currently working toward the master's degree in computer science and technology.

His research interests include oceanic meteorological forecasting and time series prediction.

Peihao Yang (Graduate Student Member) was born in 1994. He received the master's degree in computer science and technology in 2023 from Guangdong Ocean University, Zhanjiang, China, where he is currently working toward the Ph.D. degree in marine science.

His primary research interest includes oceanic meteorological forecasting.

Yuxue Wang was born in 1998. He received the bachelor's degree in mechanical design, manufacturing, and automation in 2021 from Guangdong Ocean University, Zhanjiang, China, where he is currently working toward the master's degree in computer science and technology.

His primary research interest includes oceanic meteorological forecasting.

Sheng Li was born in 1982. He received the Ph.D. degree in fundamental mathematics from Beihang University, Beijing, China, in 2011.

He is currently a Professor with Guangdong Ocean University, Zhanjiang, China. His primary research interests include complex analysis and its applications, as well as Big Data technology and its applications.

Baoqin Chen was born in 1982. She received the Ph.D. degree in fundamental mathematics from South China Normal University, Guangzhou, China, in 2012.

She is currently a Professor with Guangdong Ocean University, Zhanjiang, China. Her primary research interests include complex analysis and its applications, as well as Big Data technology and its applications.