DOI: 10.3217/978-3-99161-057-1-126

Spatial-Temporal Graph Model for Environmental Temperature and Traffic Flow Prediction of City Regions

Chenglong Lin[1,2], Yang Xu[1,2, 0000-0002-8394-9224*]
[1]Key Lab of Smart Prevention and Mitigation of Civil Engineering Disasters of the Ministry of Industry and Information Technology, Harbin Institute of Technology, Harbin 150090, China
[2]School of Civil Engineering, Harbin Institute of Technology, Harbin 150090, China
*Corresponding Author: Dr. Yang XU, xyce@hit.edu.cn

ABSTRACT: Recent research for correlation prediction from spatial-temporal monitoring data of bridge groups has explored graph neural networks and state space models, offering new angles and advanced algorithms. However, current research still faces significant challenges: (1) constructing suitable graph structures to accurately reflect complex spatial-temporal correlations, (2) designing an effective spatial-temporal neural network to capture spatial-temporal dependencies during the service state evolution of bridge groups, and (3) fully making use of spatial-temporal monitoring data to boost prediction accuracy and efficiency. To tackle these challenges, this study introduces a graph selective state space model for spatial-temporal prediction of environmental temperature and traffic flow for bridge groups. Firstly, a spatial-temporal graph structure is set up to account for data characteristics in both spatial and temporal aspects and forecast the dynamic evolution of bridge group system. Then, a state space model is built to produce a structured state space sequence and introduce a selective mechanism to dynamically adjust model behaviors and optimize computational resources. Lastly, through decomposing and reintegrating spatial-temporal features of monitoring data for bridge groups under different complexities, validation experiments are performed to show the efficacy, universality, and efficiency using multi-type, multi-scale, and multi-granularity spatial-temporal monitoring data of environmental temperature and traffic flow.

KEY WORDS: Spatial-Temporal Correlation; Time Series Prediction; Graph Model; Environmental Temperature; Traffic Flow.

1 INTRODUCTION

The rapid progression of urbanization in China has made the consistent serviceability of bridge groups a pivotal factor in determining the efficiency and safety of urban traffic. As urban transport system essentials, bridge groups consist of multiple interconnected bridges with interdependent serviceability. These bridges inside a city region are exposed to various related factors including environmental conditions and traffic flow, which can notably affect their structural integrity and operational efficiency. Accurately predicting the serviceability of bridge groups is crucial for traffic safety and maintenance strategy optimization. Traditional methods employed for predicting the service state of bridge groups often focus on individual bridges, neglecting the spatial-temporal correlations amongst different bridges in the group. This can cause inaccurate predictions and thus insufficient maintenance strategies, potentially causing significant safety risks and leading to considerable economic losses.

Recent deep learning advances have created new ways to predict the service state of bridge groups. Graph neural networks (GNNs) have shown great ability in dealing with data that has complex relational structures, like the interactions between different bridges in a group. Additionally, state space models (SSMs[1-4]) have drawn increasing interests due to their capability to model dynamic systems and capture the temporal evolution of structural service states. But even with these advances, current research faces big challenges. A main challenge is building a suitable graph structure that can truly reflect the complex spatial-temporal correlations embedded in the monitoring data of bridge groups. Another challenge is designing an effective neural network architecture that can capture both spatial and temporal dependencies during the dynamic evolution of service states for regional bridges. Furthermore, it is crucial to leverage the abundant spatial-temporal data available for bridge groups to enhance prediction accuracy and efficiency.

To tackle these challenges, this study proposes a graph selective state space model for spatial-temporal prediction of environmental temperature and traffic flow for bridge groups. The model uses the core capacities of GNNs and SSMs to achieve comprehensive and accurate predictions of spatial-temporal variables. The main contributions of this study are as follows:

- An adaptive multi-granularity data fusion is designed to integrate multi-granularity data of recent, cyclic, and trend information from bridge groups. This fusion method not only captures various temporal patterns but also assesses their impacts on operational states of bridges. This approach strengthens the model's ability to handle complex spatial-temporal dependencies, thereby improving prediction accuracy.
- A novel spatial-temporal graph convolution module (STGCM) is introduced to consider the spatial-temporal correlation of environmental temperature and traffic flow.
- A graph selective state space module (GSSSM) is developed to model how spatial-temporal dependencies in regional monitoring data evolve for bridge group, which dynamically adjusts the model's learning behavior and optimizes computational resources. The GSSSM prioritizes the most relevant parts of regional monitoring data for environmental temperature and traffic flow,

DOI: 10.3217/978-3-99161-057-1-126

thereby enhancing the model capacity to capture spatial-temporal correlations.

2 RELATED WORKS

2.1 Spatial-Temporal Graph Neural Networks

Most spatial-temporal graph neural networks (STGNNs) are categorized into three types: those based on recurrent neural networks (RNNs), those based on convolutional neural networks (CNNs), and those based on attention mechanisms.

Seo et al. (2018)[5] presented a graph convolutional recurrent network, capturing spatial-temporal dependencies by filtering inputs and hidden states in a recurrent unit via graph convolution. Later works adopted different strategies, for example, Li et al. (2017)[6] developed a diffusion convolutional recurrent neural network, integrating diffusion convolution with gated recurrent units (GRUs) to grasp spatial and temporal dependencies at the same time, and Zhang et al. (2018)[7] proposed gated attention networks, which combine GNNs with a convolutional subnetwork to assess the importance of each attention head. Another parallel work was that Jain et al. (2016)[8] used node-level and edge-level RNNs in their work to handle different aspects of temporal information. The main drawbacks of RNN-based approaches are that they become inefficient for long sequences and the gradients are more likely to explode when combined with graph convolution networks (GCNs). CNN-based approaches combine graph convolutions with standard 1D convolutions, for example, Yu et al. (2018)[9] proposed a spatial-temporal graph convolution network (STGCN), combining GCNs with temporal convolutional networks (TCN[10]) to capture spatial and temporal dependencies. A recent work about attention-based approaches is that Li et al. (2023)[11] proposed GCN-Informer, which combines GCNs with Informer to capture long-term dependencies in time series data. Although these three types of approaches are computationally efficient, they generally need to stack multiple layers or use global pooling to expand the neural network model's receptive field, which is a limitation that the proposed Spatial-Temporal Graph Selective State Space Model (STGSSSM) addresses by employing stacked dilated casual convolutions to capture temporal dependencies more effectively with fewer layers.

2.2 State Space Models

State Space Models (SSMs) are powerful tools for modeling dynamic systems, offering flexible frameworks to capture temporal evolution through state transitions influenced by external inputs. The recent fusion of SSMs with GNNs has achieved promising results in managing complex spatial-temporal graph-structured data.

For instance, Zhao et al. (2024)[12] proposed the graph state space network, the first model to incorporate SSMs into the spectral filter design of GNNs, thus overcoming the limitations of conventional methods in dealing with complex graph spectra. Behrouz et al. (2024)[13] introduced graph mamba (selective state space model) networks, an SSM-based framework for graph learning, which uses neighborhood tokenization, token ordering, bidirectional SSM encoder, and local encoding to efficiently handle long-range dependencies and heterophilic graphs. Wang et al. (2024)[14] presented graph-mamba, a novel graph model that uses a mamba module for efficient context selection, achieving linear complexity and superior performance in long-range graph tasks. Yuan et al. (2024)[15] proposed dynamic graph-mamba, a dynamic graph structure learning framework that combines mamba and a kernelized dynamic message-passing operator to efficiently learn dynamic graph structures and capture long-range dependencies.

Although these approaches can dynamically adjust the learning behavior of the model and has high computational efficiency, they are relatively singular in feature extraction and lack in-depth extraction of spatial-temporal features, which is a limitation that the proposed STGSSSM addresses by utilizing STGCM to extract diverse spatial-temporal features.

2.3 Spatial-Temporal Traffic Forecasting

Traffic forecasting is crucial for optimizing urban transportation systems. Recent progress in spatial-temporal graph neural networks have greatly boosted the prediction accuracy by capturing complex spatial-temporal dependencies. Spatial-temporal traffic forecasting has been also investigated by RNN-based, CNN-based, and attention-based approaches.

For example, Zhao et al. (2019)[16] combined GCN with GRU to model temporal dynamics, effectively capturing long-term dependencies and achieving validation on the Los-loop dataset. Wu et al. (2019)[17] developed Graph WaveNet, using dilated causal convolutions to grasp long-term dependencies. Wu et al. (2020)[18] proposed the multivariate time series graph neural network (MTGNN), which integrated GCN and TCN for adaptive graph learning, delivering high performance on the large-scale METR-LA and PEMS-BAY datasets. In the attention-based approach, Zheng et al. (2020)[19] proposed graph multi-attention network, which utilized multi-attention mechanisms to weigh the significance of different nodes and time steps, achieving high accuracy on Xiamen and PeMS datasets. Guo et al. (2021)[20] developed an attention-based spatial-temporal graph neural network, combining dynamic GCN with transformers to adapt to dynamic traffic patterns, and showing robustness on the PEMS dataset.

2.4 Spatial-Temporal Temperature Forecasting

Environmental temperature greatly affects the service state of bridge groups, and spatial-temporal graph models can conduct highly efficient and accurate temperature predictions, which are summarized from RNN-based, CNN-based, and attention-based aspects.

Zhao et al. (2024)[21] presented an adaptive spatial-temporal graph recurrent network model. It used dynamic graph structures with a spatial-temporal recurrent network for sea surface temperature forecasting. Yang et al. (2023)[22] introduced a hierarchical graph recurrent network which utilized adaptive node embedding and hierarchical graph convolution to predict global sea surface temperatures accurately. Yu et al. (2021)[23] proposed a spatial-temporal graph neural network model that integrated graph attention networks and GRU for air temperature forecasting. Guo et al. (2025)[24] introduced spatial-temporal fusion graph neural networks with mixed adjacency, a model that leveraged spatial-temporal fusion GNNs with mixed adjacency and integrated GNNs with self-attention mechanisms to capture both long-term temporal periodicity and short-term spatial-temporal dependencies for temperature forecasting. Xu et al. (2024)[25] proposed the dynamic graph former model, a physics-guided

DOI: 10.3217/978-3-99161-057-1-126

dynamic graph neural network for weather forecasting which combined GNNs and the Reformer architecture for temperature forecasting.

Despite the progress in traffic and temperature forecasting based on novel deep learning models, some challenges remain in multi-granularity data fusion, in-depth extraction of spatial-temporal features, dynamic adjustment of learning behavior and computational efficiency. This study addresses these limitations by integrating the multi-granularity data fusion, STGCM and GSSSM to enhance the prediction accuracy and efficiency.

3 METHODOLOGY

In this section, the mathematical definition is first formulated for the investigated issue of environmental temperature and traffic flow prediction of bridge groups in this study. Then, the overall schematic of the proposed STGSSSM is presented, along with detailed descriptions of its two core modules, i.e., spatial-temporal graph convolution module (STGCM) and graph selective state space module (GSSSM), which are combined together to capture the spatial-temporal dependencies.

3.1 Problem Definition

A graph of bridge groups can be represented as $\mathbf{G}_t = (\mathbf{V}, \mathbf{E}_t, \mathbf{A})$, where $\mathbf{V}$ denotes the set of nodes, representing bridges in a bridge group; $\mathbf{E}_t$ denotes the set of edges (connections), describing the relationships between different bridges at time step *t*; and $\mathbf{A}$ denotes the adjacency matrix with elements A_{ij} representing the connection weight between nodes v_i and v_j. The input feature matrix of the bridge group at time step *t* is denoted as $\mathrm{X}_t \in \mathbb{R}^{N\times D}$, where *N* is the number of bridges in a bridge group, and *D* represents the number of feature dimensions, i.e., *D* = 1 represents the investigated variable of environmental temperature or traffic flow.

The graph-based spatial-temporal prediction task is to use the graph of bridge groups $\mathbf{G}_t$ and feature matrix X_t to learn a mapping function *f* (the proposed STGSSSM) which can accurately predict the prospective environmental temperature and traffic flow of bridge groups.

Assuming a future time step *T*, the prediction process can be expressed as:

$$[\mathrm{X}_{t+1}, \cdots, \mathrm{X}_{t+T}] = f[(\mathrm{X}_{t-T+1}, \cdots, \mathrm{X}_t); (\mathrm{G}_{t-T+1}, \cdots, \mathrm{G}_t)] \quad (1)$$

where $[\mathrm{X}_{t+1}, \cdots, \mathrm{X}_{t+T}] \in \mathbb{R}^{N\times D\times T}$ is the predicted environmental temperature or traffic flow in the future *T* time steps, $(\mathrm{X}_{t-T+1}, \cdots, \mathrm{X}_t) \in \mathbb{R}^{N\times D\times T}$ is the observed environmental temperature or traffic flow in the current *T* time steps.

3.2 Overall Schematic of STGSSSM

The proposed STGSSSM comprises two primary modules of STGCM and GSSSM for environmental temperature and traffic flow prediction of bridge groups, as depicted in Figure 1.

Three temporal granularities of input data, recent data X^r, cyclic data X^c and trend data X^q are individually fed into STGCM, which consists of *N* ST-Blocks and an output layer that consists of the *ReLU* activation function and a linear layer, each ST-Block employs a series of layers including 1×1 convolution, gated temporal convolution (Gated TC), self-attention diffusion graph convolution (SADGC), residual connection (Add), and batch normalization (BN); and adaptively fused after being passed through *N* layers of spatial-temporal graph convolution blocks (ST-Block) with skip connections and the output layer; In each ST-Block, the result of Gated TC is processed via a 1×1 convolution, and the resulting output is skip-connected to the result of the last ST-Block.

Furthermore, X^r is fed into GSSSM, which consists of *M* GSSS-Blocks. Each GSSS-Block incorporates a series of layers including layer normalization (Layer Norm), *m* parallel dynamic filter graph convolution (Dynamic Filter GC), concatenation and linear layer (Concat & Linear), graph state space selection mechanism (GSSS-Mechanism) which consists of two main algorithms of Parameter Calculation and Graph Selective Scan, linear layer and residual connection (Linear & Add). *N*, *M* and *m* are hyperparameters.

The outputs of three STGCMs are first fused using an adaptive fusion module and then adaptively fused with the output of GSSSM. The final fusion result of STGCM and GSSSM is then passed through a fully connected layer to obtain the final output X_{pred}.

The regularized mean squared error (MSE) is used for training the proposed STGSSSM and defined as

$$L = \frac{1}{B}\left(\sum_{b=1}^{B}\left(\frac{1}{TND}\sum_{i=t+1}^{t+T}\sum_{j=1}^{N}\sum_{k=1}^{D}\left(\hat{X}_{i,j,k}^{b} - X_{i,j,k}^{b}\right)^2\right) + \lambda R(\Theta)\right) \quad (2)$$

where $X_{i,j,k}^{b}$, $\hat{X}_{i,j,k}^{b}$ represents the model-predicted and ground-truth data for the *b*-th segment; *i*, *j*, *k* denote the indexes of time step, bridge node, and feature dimension; *B* and *D* denote the number of batch size and considered feature dimension (i.e., *D* = 1 represents environmental temperature or traffic flow); *λ* is the regularization coefficient, and *R*(Θ) is the regularization term.

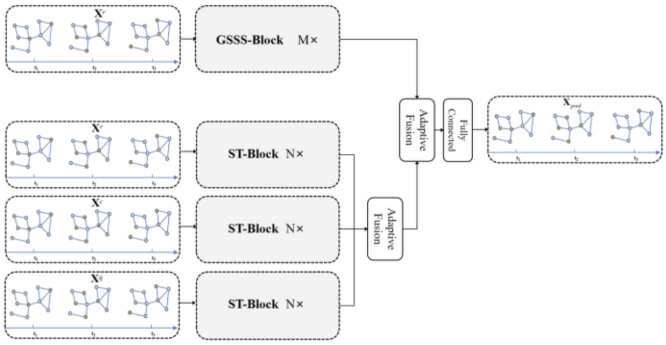

Figure 1. Overall schematic of the proposed STGSSSM.

4 EXPERIMENT STUDIES

4.1 Datasets

Two real-world datasets are utilized here to assess the proposed approach for spatial-temporal prediction of environmental temperature and traffic flow for bridge groups. Training samples for time series are typically obtained by sliding a window of length *P* + *T* across the original time series. Here, the first *P* time steps are used as historical data, and the subsequent *T* time steps are used as future data; and the datasets are detailed as follows:

- **KnowAir[26]**: This temperature dataset contains spatial-temporal temperature data from weather stations across 184 main cities in China from September 1, 2016, to

DOI: 10.3217/978-3-99161-057-1-126

January 31, 2017. It contains 184 nodes (weather stations) and 11688 time steps of temperature data with a 1-hour time interval; and the adjacency matrix of the graph is constructed based on the geographical locations of the weather stations. In the released KnowAir dataset, P = 7×24, and T = 12.

- **PEMS04[9]**: This traffic dataset consists of spatial-temporal traffic flow data on California road network from January 1, 2018, to February 28, 2018. It contains 307 nodes (traffic monitoring stations) and 16992 time steps of data with a 5-minute time interval; and the adjacency matrix of the graph is constructed based on the geographical locations of the traffic monitoring stations. In the released PEMS04 dataset, P = 7×24×12, and T = 24 or 36.

4.2 Experimental Setup

All the datasets are split with a ratio of 7:1:2 for training/validation/testing sets, respectively. Before starting the model training process, all the data samples are normalized into range [0,1] with min-max normalization. The batch size is set as 48, and AdamW is employed as the optimizer. Model training epochs is set as 200. The number of ST-Blocks is set as N = 8 with a sequence of dilation ratios R = 1 or 2, and the number of GSSS-Blocks is set as M = 4. The diffusion step K = 2 in Eq. (5), and the number of parallel Dynamic Filter GC, m = 3 in Eq. (7). Hyperparameters in Eq. (3) of MSE loss function are set as the regularization coefficient λ of $1e^{-2}$ and the learning rate of $1e^{-4}$.

All experiments are conducted under the software environment of PyTorch 1.12.1 and Python 3.8.10 with a NVIDIA GeForce RTX 3090 24GB GPU to accelerate neural computing. The training process requires nearly 4h for obtaining a well-trained model with acceptable accuracy.

4.3 Comparative Studies with Baseline Models

Two recent Graph-based spatial-temporal predictive model of STGCN[9] and MTGNN[18] are utilized as baselines for comparative studies. STGCN integrates graph convolutional layers to model spatial dependencies and gated temporal convolutional layers to capture temporal dynamics in traffic data. MTGNN considers the features in the time-series to be multivariate and captures spatial dependencies through a learned graph structure, which uses a mix-hop propagation layer to handle spatial dependencies and a dilated inception layer for temporal dependencies.

In comparative studies with STGCN and MTGNN models, to ensure fairness, several trials are conducted to select an acceptable model with good accuracy though not necessarily optimal. For STGCN, the graph and temporal convolution kernel sizes are set to 3. Chebyshev polynomial approximation and first-order approximation are used in STGCN. For MTGNN, the propagation depth and dilation factor are set to 2 and 1, respectively. During training, the Adam optimizer is employed with a $1e^{-3}$ learning rate, a regularization coefficient of $1e^{-4}$, and a dropout rate of 0.3. These adjustments made sure that the baseline models performed well on our datasets, which enable a fair comparison of model performances and a better evaluation of the proposed STGSSSM's strengths.

Evaluation metrics of mean absolute error (MAE), root mean squared Error (RMSE), and mean absolute percentage error (MAPE) are utilized to measure and assess the accuracy of various methods.

The MAE is defined as

$$\mathrm{MAE}_{test} = \frac{1}{TND}\sum_{i=t+1}^{t+T}\sum_{j=1}^{N}\sum_{k=1}^{D}\left|\hat{X}_{i,j,k} - X_{i,j,k}\right| \tag{3}$$

The RMSE is defined as

$$\mathrm{RMSE}_{test} = \sqrt{\frac{1}{TND}\sum_{i=t+1}^{t+T}\sum_{j=1}^{N}\sum_{k=1}^{D}\left(\hat{X}_{i,j,k} - X_{i,j,k}\right)^2} \tag{4}$$

The MAPE is defined as

$$\mathrm{MAPE}_{test} = \frac{100\%}{TND}\sum_{i=t+1}^{t+T}\sum_{j=1}^{N}\sum_{k=1}^{D}\left|\frac{\hat{X}_{i,j,k} - X_{i,j,k}}{\hat{X}_{i,j,k}}\right| \tag{5}$$

Table 1 shows the performance comparison of the proposed STGSSSM with baseline models of STGCN and MTGNN. In the experimental results, 12 hours with a time interval of 1 hour corresponds to 12 time steps for temperature prediction; 6 hours with a time interval of 5 minutes corresponds to 72 time steps for traffic flow prediction; 9 hours corresponds to 108 time steps for traffic flow prediction.

Table 1. Performance comparison of the proposed STGSSSM with baseline models of STGCN and MTGNN.

KnowAir	12h		
	MAE	RMSE	MAPE
STGCN	0.35	0.51	3.77%
MTGNN	0.33	0.48	3.34%
STGSSSM	**0.28**	**0.44**	**3.02%**
PEMS04	6h		
	MAE	RMSE	MAPE
STGCN	14.21	20.85	12.95%
MTGNN	**13.28**	19.67	12.34%
STGSSSM	13.39	**19.47**	**12.12%**
PEMS04	9h		
	MAE	RMSE	MAPE
STGCN	17.60	26.39	14.96%
MTGNN	16.76	25.13	14.39%
STGSSSM	**16.37**	**24.01**	**14.23%**

4.4 Ablation Study

To assess the effectiveness of each model component within STGSSSM, three kinds of model variants are first designed, and their forecasting performance is evaluated on the KnowAir and PEMS04 datasets: (1) the full model of STGSSSM, (2) GSSSM without STGCM, (3) STGCM without GSSSM.

To further evaluate the effects of three temporal granularities of recent, cyclic, and trend data, another three kinds of model variants are designed: (4) STGSSSM without cyclic and trend data, (5) STGSSSM without cyclic data, (6) STGSSSM without trend data.

Table 2. Model performances of ablation study for different modules of STGSSSM.

KnowAir	12h		
	MAE	RMSE	MAPE
STGSSSM	**0.28**	**0.44**	**3.02%**
GSSSM w/o STGCM	0.39	0.54	4.15%
STGCM w/o GSSSM	0.31	0.47	3.27%
STGSSSM w/o cyclic and trend	0.40	0.55	4.27%
STGSSSM w/o cyclic	0.30	0.45	3.16%
STGSSSM w/o trend	0.29	0.44	3.13%
PEMS04	6h		
	MAE	RMSE	MAPE

DOI: 10.3217/978-3-99161-057-1-126

STGSSSM	**13.39**	**19.47**	12.12%
GSSSM w/o STGCM	18.86	28.94	15.72%
STGCM w/o GSSSM	14.48	21.46	12.29%
STGSSSM w/o cyclic and trend	13.94	20.25	12.22%
STGSSSM w/o cyclic	13.51	19.57	12.01%
STGSSSM w/o trend	13.45	19.63	**11.93%**
PEMS04	9h		
	MAE	RMSE	MAPE
STGSSSM	**16.37**	**24.01**	14.23%
GSSSM w/o STGCM	18.99	29.21	15.66%
STGCM w/o GSSSM	18.38	27.86	15.11%
STGSSSM w/o cyclic and trend	16.83	24.92	14.81%
STGSSSM w/o cyclic	16.39	24.36	13.95%
STGSSSM w/o trend	16.38	24.27	**13.93%**

The results in Table 2 show the contributions of each module to the overall performance.

(1) Influence of STGCM: The STGCM, which processes recent, cyclic, and trend data through a series of spatial-temporal graph convolution operations, is essential for capturing complex spatial-temporal dependencies. The results show that removing the STGCM leads to a substantial decline in performance across both datasets. For instance, on the KnowAir dataset, the MAE, RMSE, and MAPE increase from 0.28, 0.44, and 3.02% (full model) to 0.39, 0.54, and 4.15% (GSSSM w/o STGCM), respectively. Similarly, on the PEMS04 dataset, the MAE, RMSE, and MAPE rise from 13.39, 19.47, and 12.12% (6h prediction) to 18.86, 28.94, and 15.72% (GSSSM w/o STGCM). This shows that the STGCM greatly boosts the model's ability to handle multi-granularity data and extract meaningful spatial-temporal features.

(2) Influence of GSSSM: The GSSSM, which models the evolution of spatial-temporal dependencies and optimizes computational resources, also makes a significant contribution to the model's performance. Removing the GSSSM results in a significant decline in performance. On the KnowAir dataset, the MAE, RMSE, and MAPE increase to 0.31, 0.47, and 3.27%, respectively. On the PeMS04 dataset, the MAE, RMSE, and MAPE rise to 14.48, 21.46, and 12.29% (6h prediction). This demonstrates the importance of the GSSSM in dynamically adjusting the model's behavior and focusing on the most relevant parts of the data.

(3) Influence of Multi-Granularity Data Fusion: The fusion of recent, cyclic, and trend data is another critical aspect of the model. The results illustrate that removing both cyclic and trend data (STGSSSM w/o cyclic and trend) results in a significant decline in performance. On the KnowAir dataset, the MAE, RMSE, and MAPE increase to 0.40, 0.55, and 4.27%, respectively. On the PEMS04 dataset, the MAE, RMSE, and MAPE rise to 13.94, 20.25, and 12.22% (6h prediction). This shows that the fusion of multi-granularity data is crucial for seizing diverse temporal patterns and enhancing prediction accuracy.

(4) Influence of Individual Data Components: The results also illustrate the contributions of individual data components (cyclic and trend data). Removing only the cyclic data (STGSSSM w/o cyclic) or only the trend data (STGSSSM w/o trend) leads to moderate performance declines. On the Temperature dataset, removing cyclic data results in MAE, RMSE, and MAPE of 0.30, 0.45, and 3.16%, while removing trend data results in 0.29, 0.44, and 3.13%. On the PEMS04 dataset, removing cyclic data leads to MAE, RMSE, and MAPE of 13.51, 19.57, and 12.01% (6h prediction), while removing trend data results in 13.45, 19.63, and 11.93%. These results indicate that both cyclic and trend data boost the model's performance, with cyclic data having a slightly more significant impact.

The proposed STGSSSM model gains from the integration of the multi-granularity data fusion, STGCM and GSSSM, which jointly deliver accurate and efficient spatial-temporal predictions for bridge groups.

5 CONCLUSION

This study introduced a Graph Selective State Space Model (STGSSSM) for predicting environmental temperature and traffic flow in bridge groups, integrating Graph Neural Networks (GNNs) and State Space Models (SSMs). The main contributions of the model include adaptive fusion of multi-granularity data, a Spatial-Temporal Graph Con-volution Module (STGCM), and a Graph Selective State Space Module (GSSSM). Experiments on real-world datasets (KnowAir and PEMS04) show STGSSSM outperforms state-of-the-art models like STGCN and MTGNN in prediction accuracy and efficiency. Ablation experiments validate that each component is effective, necessary, and enhances accuracy in capturing complex spatial-temporal dependencies. The proposed STGSSSM achieves overall prediction accuracy improvements of environmental temperature at the ranges of [13.70%, 20.00%] and [8.30%, 15.15%] and traffic flow at the ranges of [4.90%, 9.00%] and [1.02%, 4.46%] compared with STGCN and MTGNN, respectively. Specifically, the reported STGCM, GSSSM and multi-granularity data fusion decreases the relative prediction error at the ranges of [7.93%, 28.99%], [1.32%, 13.80%], and [0.49%, 29.50%], respectively.

Future work may extend the model to more complex data and other infrastructure systems.

ACKNOWLEDGMENTS

Financial support for this study was provided by the National Key R&D Program of China [Grant No. 2023YFC3805700], Heilongjiang Provincial Natural Science Foundation [Grant No. LH2022E070], and Fundamental Research Funds for the Central Universities [Grant No. HIT.NSRIF202334].

REFERENCES

[1] G. Albert and D. Tri, Mamba: Linear-time sequence modeling with selective state spaces, arXiv preprint arXiv:2312.00752, 2023.

[2] A. Behrouz and F. Hashemi, Graph mamba: Towards learning on graphs with state space models, in Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 119-130, 2024.

[3] Q. Haohao, N. Liangbo, A. Rui, F. Wenqi, D. Tyler, L. Hui, X. Xin and L. Qing, A survey of mamba, arXiv preprint arXiv:2408.01129, 2024.

[4] S. Atitallah, C. Rabah, M. Driss, W. Boulila and A. Koubaa, Exploring Graph Mamba: A Comprehensive Survey on State-Space Models for Graph Learning, arXiv preprint arXiv:2412.18322, 2024.

[5] S. Youngjoo, D. Michaël and V. Pierre, Structured sequence modeling with graph convolutional recurrent networks, neural information processing: 25th international conference, ICONIP 2018, Siem Reap, Cambodia, December 13-16, 2018, proceedings, part I 25. Springer International Publishing, 362-373, 2018.

DOI: 10.3217/978-3-99161-057-1-126

[6] L. Yaguang, Y. Rose, S. Cyrus and L. Yan, Diffusion convolutional recurrent neural network: Data-driven traffic forecasting, arXiv preprint arXiv:1707.01926, 2017.

[7] Z. Jiani, S. Xingjian and X. Junyuan, Gaan: Gated attention networks for learning on large and spatiotemporal graphs, arXiv preprint arXiv:1803.07294, 2018.

[8] J. Ashesh, Z. Amir R., S. Silvio and S. Ashutosh, Structural-rnn: Deep learning on spatio-temporal graphs, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 5308-5317, 2016.

[9] Y. Bing, Y. Haoteng, and Z. Zhanxing, Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting, in Proceedings of the 27th International Joint Conference on Artificial Intelligence, 3634–3640, 2018.

[10] S. Bai, J. Kolter and V. Koltun, An empirical evaluation of generic convolutional and recurrent networks for sequence modeling, arXiv preprint arXiv:1803.01271, 2018.

[11] L. Pengfei, Z. Tong and J. Yantao, A spatio-temporal graph convolutional network for air quality prediction, Sustainability, 15(9): 7624, 2023.

[12] Z. Gongpei, W. Tao, L. Congyan, J. Yi, L. Yidong and L. Haibin, Grassnet: State space model meets graph neural network, arXiv preprint arXiv:2408.08583, 2024.

[13] B. Ali and H. Farnoosh, Graph mamba: Towards learning on graphs with state space models, in Proceedings of the 30th ACM SIGKDD conference on knowledge discovery and data mining, 119-130, 2024.

[14] W. Chloe, T. Oleksii, M. Jun and W. Bo, Graph-mamba: Towards long-range graph sequence modeling with selective state spaces, arXiv preprint arXiv:2402.00789, 2024.

[15] Y. Haonan, S. Qingyun, W. Zhaonan, F. Xingcheng, J. Cheng, W. Yongjian, J. Bo and L. Jianxin, DG-Mamba: Robust and Efficient Dynamic Graph Structure Learning with Selective State Space Models, arXiv preprint arXiv:2412.08160, 2024.

[16] Z. Ling, S. Yujiao, Z. Chao, L. Yu, W. Pu, L. Tao, D. Min and L. Haifeng, T-GCN: A temporal graph convolutional network for traffic prediction, IEEE transactions on intelligent transportation systems, 21(9): 3848-3858, 2019.

[17] W. Zonghan, P. Shirui, L. Guodong, J. Jing and Z. Chengqi, Graph WaveNet for Deep Spatial-Temporal Graph Modeling, arXiv preprint arXiv:1906.00121, 2019.

[18] W. Zonghan, P. Shirui, L. Guodong, J. Jing, C. Xiaojun and Z. Chengqi, Connecting the dots: Multivariate time series forecasting with graph neural networks, in Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining, 753-763, 2020.

[19] Z. Chuanpan, F. Xiaoliang, W. Cheng and Q. Jianzhong, GMAN: A graph multi-attention network for traffic prediction, in Proceedings of the AAAI conference on artificial intelligence, 34(01): 1234-1241, 2020.

[20] G. Shengnan, L. Youfang, W. Huaiyu, L. Xiucheng and C. Gao, Learning Dynamics and Heterogeneity of Spatial-Temporal Graph Data for Traffic Forecasting, IEEE Transactions on Knowledge and Data Engineering, PP (99):1-1, 2021.

[21] Z. Xingyu, W. Zhongyu and Z. Zhihao, Adaptive Spatio-Temporal Graph Recurrent Network for Sea Surface Temperature Forecasting, IEEE Transactions on Geoscience and Remote Sensing, 2024.

[22] Y. Hanchen, L. Wengen, H. Siyun and G. Jihong, HiGRN: a hierarchical graph recurrent network for global sea surface temperature prediction, ACM Transactions on Intelligent Systems and Technology, 14(4): 1-19, 2023.

[23] Y. Xuan, S. Suixiang and X. Lingyu, A spatial–temporal graph attention network approach for air temperature forecasting, Applied Soft Computing, 113: 107888, 2021.

[24] G. Ang, L. Yanghe, S. Shiyu, S. Xiaowei and F. Zhenni, Spatial-Temporal Fusion Graph Neural Networks with Mixed Adjacency For Weather Forecasting, IEEE Access, 2025.

[25] X. Zhewen, W. Xiaohui and H. Jieyun, DGFormer: a physics-guided station level weather forecasting model with dynamic spatial-temporal graph neural network, GeoInformatica, 28(3): 499-533, 2024.

[26] E. Saleh, S. Al-Werfalli, T. Imbarak, S. Elakeili and H. Elmajpri, Know-Air: A Low-Cost PM2. 5 Sensor Citizen-Based Air Pollution Monitoring System for Real-Time, International Conference for Information and Communication Technologies, Cham: Springer Nature Switzerland, 159-172, 2023.