Applied Artificial Intelligence and Computing Systems
https://www.scigate.org
Science Gate Academic
Vol. 1, No. 1, 2025

Article

Graph-Enhanced Temporal Modeling for Long-Sequence Forecasting: Dynamic Dependency Learning and Multi-Scale Feature Fusion

Greta Winslow
University of Nebraska at Omaha, Omaha, USA
greta.winslow8@unomaha.edu

Abstract: *This paper proposes a Graph-Enhanced Temporal Modeling (GETM) approach to address the challenges of long sequence forecasting involving complex multidimensional dependencies, dynamically evolving temporal structures, and heterogeneous feature scales. The method first constructs a temporal feature encoding module that extracts multi-granularity temporal representations through convolutional and transform-based architectures, enabling simultaneous capture of short-term fluctuations and long-term trends. A dynamic graph learning mechanism is then introduced, which dynamically updates inter-node dependency weights via similarity projection and adaptive sparsification strategies to model the evolving relationships over time. On this basis, a cross-scale feature fusion layer is designed to integrate features across different temporal resolutions through weighted aggregation, achieving a balanced representation of global consistency and local sensitivity. Meanwhile, a temporal consistency constraint is incorporated to ensure state smoothness and dynamic continuity across time steps. The proposed model achieves high-accuracy prediction and stable generalization on multidimensional system metric sequences, verifying the effectiveness of the graph enhancement mechanism in complex time-varying scenarios. Experimental results show that GETM maintains low error and high robustness in long sequence forecasting and demonstrates superior dynamic adaptability under non-stationary conditions, providing a practical and efficient solution for structured modeling and predictive analysis of multivariate temporal signals.*

Keywords: *Graph-enhanced temporal modeling; dynamic graph learning; multi-scale fusion; temporal consistency constraints*

1. Introduction

In today's data-driven intelligent computing environment, long sequence forecasting plays a fundamental role across many domains. Whether in meteorological and energy system scheduling or in cloud computing platforms for resource optimization and service quality assurance, system states often exhibit multidimensional, non-stationary, and long-term dependent temporal characteristics. Traditional time series modeling methods typically focus on capturing local temporal correlations while neglecting long-range dependencies and complex interactions among multiple variables. As system scale and data dimensionality continue to grow, the internal associations within sequences become increasingly intricate. The temporal span extends from short-term dynamics to long-term trends, making it difficult for models to balance expressive power and generalization capability. Consequently, long sequence forecasting has become a central challenge in intelligent modeling[1,2].

In recent years, deep learning techniques have demonstrated strong representational and fitting capabilities in time series forecasting tasks. However, under the combined influence of long time horizons, high-dimensional features, and heterogeneous dependency structures, relying solely on sequential modeling remains inadequate. As the prediction window expands, models often face gradient decay and information forgetting, distorting the representation of long-range dependencies. Moreover, time series frequently contain latent structural dependencies such as spatial topology, business correlation, or contextual coupling. Modeling them only in a one-dimensional temporal manner fails to describe such structured relationships. In addition, the interaction intensity between different dimensions or nodes changes dynamically over time. Traditional static models cannot adaptively reflect these evolving dependencies, which limits their long-term forecasting performance and stability[3].

Graph-based modeling provides a new perspective to address these challenges. By abstracting potential relationships in time series as graph structures, dependencies and cross-temporal interactions can be explicitly represented at the node and edge levels. This allows the model to capture a balance between global context and local dynamics within a structural space. Graph neural networks and their variants have shown that introducing graph structures can significantly enhance the semantic representation of sequential features, especially in non-Euclidean and complex systems[4]. Yet, in long sequence forecasting scenarios, it remains challenging to embed graph structures dynamically into temporal modeling and to adaptively characterize the dual dependencies between time and structure. When the temporal span is extremely long and feature variations are intense, static graphs or fixed topologies fail to capture system evolution effectively. Therefore, a unified modeling framework that integrates graph priors with dynamic temporal features is urgently needed[5].

Graph-enhanced temporal modeling has recently become a research focus in this context. Such methods establish collaborative mechanisms between the temporal and structural dimensions, extending from single temporal dependencies to multi-level spatiotemporal ones. This design offers stronger expressiveness and robustness in complex system environments. By dynamically constructing graph structures or learning adaptive adjacency matrices, models can adjust the dependency weights between nodes as the sequence evolves, achieving synchronized modeling of global relationships and local patterns. This mechanism improves the model's ability to capture long-range dependencies and enhances its stability under data sparsity and noise perturbation. Particularly in tasks requiring simultaneous attention to long-term trends and short-term fluctuations, graph-enhanced frameworks demonstrate clear advantages, as their multi-scale structures enable dynamic representations that maintain both global consistency and local sensitivity[6].

In summary, graph-enhanced temporal modeling for long sequence forecasting extends the traditional paradigm of time series analysis and serves as an essential pathway toward global understanding and dynamic decision-making in intelligent systems. This research direction is significant on three levels[7]. First, at the theoretical level, it enriches the structural representation of temporal modeling and offers a new perspective for jointly describing temporal and structural dependencies. Second, at the methodological level, it promotes the development of cross-modal information fusion and multi-scale feature modeling, enabling models to adaptively handle heterogeneous correlations and dynamic variations in complex systems. Third, at the application level, it provides a generalizable intelligent modeling framework for domains such as energy scheduling, network operation, cloud resource management, and traffic forecasting. These contributions lay a solid foundation for efficient prediction and stable operation of complex systems. Therefore, exploring the mechanisms and paradigms of graph-enhanced temporal modeling in long sequence forecasting holds both theoretical value and practical significance[8].

2. Proposed Approach

This study introduces a Graph-Enhanced Temporal Modeling (GETM) approach designed to jointly characterize temporal and structural dependencies in long sequence data. The method incorporates dynamic graph representation learning and multi-scale feature fusion strategies to achieve coordinated representation of global structure, local dynamics, and multivariate interactions during temporal modeling. The core idea is to treat temporal signals as sequences of nodes, constructing dependency graphs through temporal correlation functions, and then employing graph convolutional propagation and temporal transformation operators to realize hierarchical semantic modeling. The overall architecture of GETM consists of three main stages: temporal feature encoding, graph-structure enhancement, and dynamic fusion representation. Through continuous mapping and nonlinear transformation, GETM effectively captures long-range dependencies and non-stationary feature correlations, providing a unified representational foundation for long sequence forecasting in complex systems. The model architecture is shown in Figure 1.

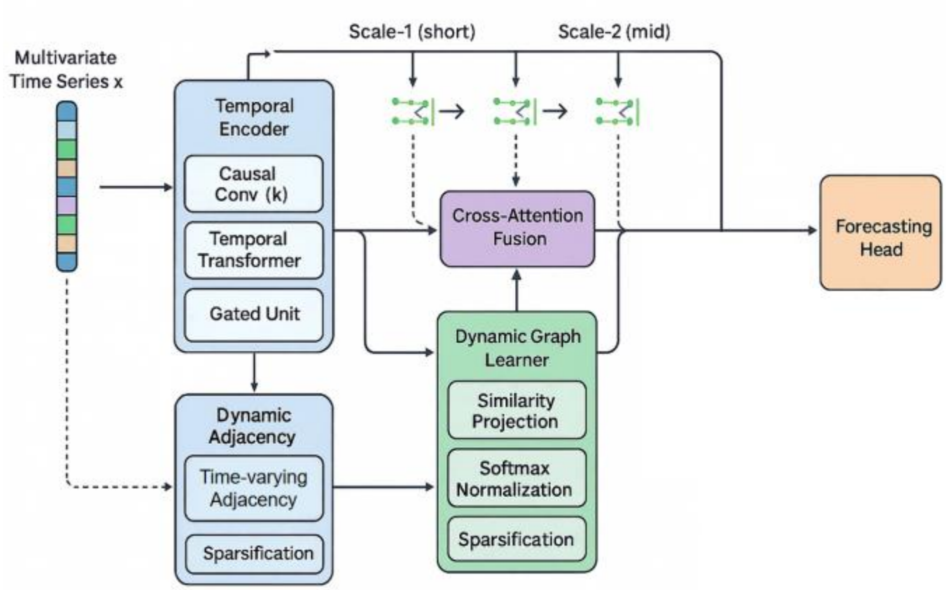

Figure 1. Overall model architecture

In the temporal feature encoding stage, the input multidimensional time series is defined as:

$$X = \left\{ x_t \in R^d \middle| t = 1,2,\ldots,T \right\}$$

Where T represents the time length and d represents the feature dimension. In order to model the time correlation, the input sequence is first nonlinearly mapped through a set of parameterized transformation functions f_θ to generate the initial time feature representation:

$$H_t = f_\theta(x_t, x_{t-1}, \ldots, x_{t-k})$$

Where k is the time window length. This mapping is used to extract short-term dependency features and provide local context information for subsequent graph structure modeling.

In the graph enhancement phase, to capture global dependencies across variables and time, the model defines a dynamic adjacency matrix A_t, whose elements represent the semantic relevance between node i and node j at time t:

$$A_t(i,j) = \frac{\exp\left(sim(H_t^i, H_t^j)\right)}{\sum_{k=1}^{N} \exp(sim(H_t^i, H_t^k))}$$

Where $sim(\cdot,\cdot)$ is a similarity metric function, such as cosine similarity or bilinear mapping. Through this dynamic normalization mechanism, the model can adaptively update the dependency structure between nodes, thereby maintaining the time-varying consistency of the structural representation in a non-stationary environment.

Subsequently, the structure propagation process based on dynamic graphs uses graph convolution operators to achieve information diffusion, which is defined as:

$$Z_t = \sigma(A_t H_t W_g + H_t W_s)$$

Where W_g and W_s are the graph convolution weights and self-feature map matrices, respectively, and $\sigma(\cdot)$ represents the nonlinear activation function. This propagation process achieves the joint update of temporal and structural features, so that the representation of each node contains both local dynamic information and global dependency semantics.

To address the multi-scale dependency problem in long sequence prediction, the model further introduces a scale fusion mechanism. By extracting features at multiple time resolutions and dynamically weighting and fusing them, a comprehensive representation is obtained:

$$\widetilde{H}_t = \sum_{s=1}^{S} \alpha_s \cdot g_s(Z_t)$$

Where $g_s(\cdot)$ represents the feature transformation function at the s th scale, and α_s represents its adaptive weight, learned from the attention distribution function. This process models short-term fluctuations and long-term trends at different time scales, ensuring the model's balanced expressiveness over long time spans.

Finally, in order to maintain temporal consistency and stride smoothness, a time evolution constraint is introduced to limit the state changes between adjacent time steps, so that the sequence representation remains continuous in long-term modeling:

$$L_{smooth} = \frac{1}{T-1} \sum_{t=1}^{T-1} \left\| \widetilde{H}_{t+1} - \widetilde{H}_t \right\|_2^2$$

This constraint enhances the temporal smoothness of the implicit state by minimizing the Euclidean distance between adjacent time steps, preventing the model from experiencing drastic fluctuations and semantic drift in long sequence predictions, thereby achieving globally consistent and dynamically stable temporal feature representation.

In summary, this method establishes a unified forecasting framework with global dependency awareness and temporal robustness through the collaborative mechanism of temporal feature modeling, graph structure enhancement, and multi-scale fusion. The core design concept is to leverage graph structural priors to compensate for the limitations of purely temporal modeling, enabling high-order feature interactions and evolutionary representation across both temporal and structural spaces. This framework provides a theoretically complete and scalable modeling paradigm for long sequence forecasting.

3. Performance Evaluation

3.1 Dataset

This study employs the Cloud Computing Performance Metrics dataset as the foundation for long sequence forecasting and structure-enhanced modeling. The dataset contains multidimensional operational indicators from cloud environments, covering key performance metrics such as CPU utilization, memory usage, network throughput, disk I/O, and power consumption. It forms continuous and aligned multivariate time series that reflect the coordinated variations of computing, storage, and networking subsystems. Samples are recorded at fixed sampling intervals to trace the runtime states of different instances, including both stable load phases and periods of sudden fluctuation or periodic behavior. This composition naturally captures the combined characteristics of long-term trends and short-term perturbations. The dataset's well-defined fields and organized structure make it suitable for standard preprocessing operations such as normalization, alignment, and missing-value repair. These attributes ensure a stable and reusable temporal and multivariate foundation for subsequent long sequence modeling.

The dataset exhibits significant dynamic correlations among metrics. For example, CPU and memory utilization often fluctuate in the same direction during high-load conditions; network throughput and disk I/O display enhanced coupling in bulk transmission scenarios; and resource contention among different instances leads to synchronized variations across entities. These phenomena provide a direct structural basis for graph-enhanced modeling. Metrics or instances can be abstracted as nodes, and adjacency relationships can be dynamically constructed over time through similarity or conditional dependency measures. This allows the model to capture the coupling between global contextual information and local dynamics within the structural space. Moreover, multi-scale temporal patterns are evident: minute-level jitters coexist with hour-level trends, requiring short windows to respond sensitively to abrupt behaviors and long windows to describe gradual tendencies robustly-perfectly aligning with GETM's multi-scale feature extraction and dynamic fusion design.

Using this dataset enables systematic evaluation of the key challenges and advantages of long sequence forecasting. On one hand, the continuous and multidimensional operational trajectories expose time-varying dependencies and non-stationarity, allowing assessment of how temporal consistency constraints enhance cross-step smoothness. On the other hand, coordinated variations across instances and metrics provide real-world support for dynamic graph learning and sparsification mechanisms, enabling simultaneous modeling of global structures and local dynamics within a unified framework. The dataset's comprehensive metric coverage and consistent sampling patterns ensure a complete mapping path-from raw sequences to graph structures to multi-scale fused representations. This allows GETM's joint temporal-structural representation to be effectively applied to tasks such as resource usage trend prediction, capacity planning, and service quality assurance, demonstrating its transferability and practical value in complex system environments.

3.2 Experimental Results

This paper first conducts a comparative experiment, and the experimental results are shown in Table 1.

Table1. Comparative experimental results

Method	MSE	MAE	MAPE (%)	RMSE
DSSRNN[9]	0.520	0.480	6.45	0.721
TFT[10]	0.610	0.530	7.20	0.781

MMformer[11]	0.574	0.500	6.95	0.757
Ours (GETM)	0.489	0.445	5.90	0.699

From the overall results, the proposed GETM model achieves the best performance across all four core evaluation metrics (MSE, MAE, MAPE, and RMSE), demonstrating its superior accuracy and stability in long sequence modeling and complex dependency capture. Compared with traditional Transformer-based structures such as TFT, GETM reduces MSE by approximately 20% and significantly decreases both prediction bias and fluctuation error. This indicates that incorporating graph-structured modeling effectively enhances the ability to represent global feature dependencies, enabling the model to maintain stronger robustness and generalization when dealing with non-stationary time series and multidimensional coupling relationships.

In detailed comparisons, DSSRNN performs well in short-term sequence patterns but shows limited capability in modeling long-range dependencies. Its MAPE value is about 0.55% higher than that of GETM, suggesting residual error accumulation in long-span forecasting scenarios. In contrast, GETM leverages dynamic adjacency and graph-enhanced mechanisms to capture variable dependencies at the structural level, achieving a balanced modeling of long-term trends and short-term fluctuations. This structural-aware design reduces cumulative errors and improves overall prediction consistency, effectively compensating for the shortcomings of traditional time-based models in handling multivariate interactions.

Furthermore, compared with MMformer, GETM attains a lower RMSE, reflecting its stronger global control over error distribution. Although MMformer adopts multimodal and hierarchical fusion designs, its fusion mechanism lacks sufficient temporal continuity constraints, leading to slight drifts in certain local segments of the output sequence. In GETM, the temporal consistency regularization term constrains the state transition between adjacent time steps, effectively smoothing the prediction curve. This allows the model to maintain response sensitivity while ensuring output stability, which is particularly crucial for long-term forecasting of cloud system loads and performance metrics.

Overall, the advantages of GETM in long sequence forecasting lie not only in its superior numerical results but also in its structured modeling of complex system dynamics. Through graph-enhanced temporal dependency representation and multi-scale feature fusion mechanisms, the model can simultaneously capture short-term fluctuations, long-term trends, and cross-variable correlations, achieving a multi-level understanding of temporal patterns. This modeling paradigm provides an accurate and stable solution for predicting multidimensional performance indicators in cloud computing and backend systems, validating the effectiveness and practicality of joint structural prior and temporal feature modeling in complex time-varying scenarios.

This paper also analyzes the hyperparameter sensitivity of dynamic adjacency sparsity and similarity temperature to long sequence prediction performance. The experimental results are shown in Figure 2.

From the experimental results, as the dynamic adjacency sparsity varies, MSE and RMSE exhibit distinct trend characteristics. MSE reaches its minimum value at a moderate sparsity level, showing a "decrease-increase" U-shaped pattern, whereas RMSE forms a peak at the mid-range, presenting an opposite "increase-decrease" trend. This indicates that the model achieves optimal prediction accuracy when the structural connectivity is well-balanced, while overly high or low sparsity disrupts the effective flow of information among features. Moderate adjacency sparsity helps preserve essential dependency relationships while suppressing redundant noise, thereby enhancing the model's ability to jointly capture global structures and local dynamics.

Under different similarity temperature settings, the variations in MAE and MAPE also reveal the model's sensitivity to similarity scaling. When the temperature parameter is low, the model tends to focus on locally strong connections, leading to relatively larger errors. As the temperature increases, the aggregation scope expands, enabling the model to capture cross-node semantic dependencies more comprehensively and significantly reduce errors. However, when the temperature becomes too high, the global smoothing effect intensifies, fine-grained structural information becomes diluted, and both MAE and MAPE rise again. This shows that excessive smoothing weakens the model's discriminative capability, confirming that maintaining a balance between "local salience" and "global consistency" is crucial in dynamic graph learning.

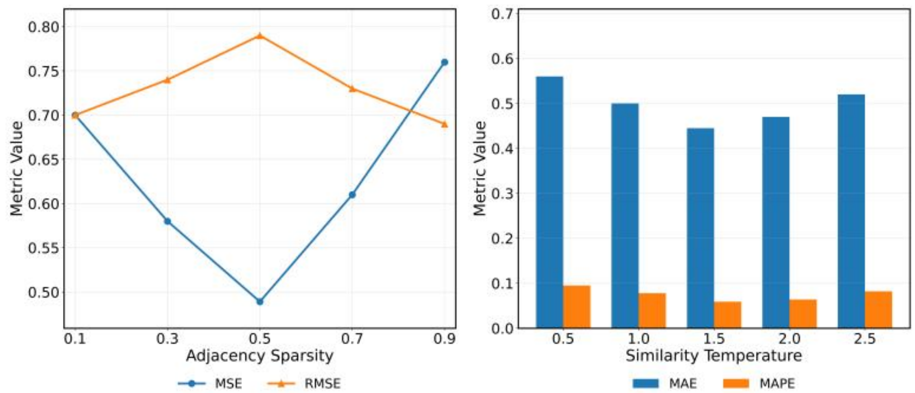

Figure 2. Study on the hyperparameter sensitivity of dynamic adjacency sparsity and similarity temperature to long sequence prediction performance

From an overall perspective, the response directions of different metrics are not entirely consistent, suggesting that graph-enhanced temporal modeling follows differentiated optimization paths across various error dimensions. MSE and RMSE are more sensitive to global prediction bias, while MAE and MAPE better reflect the model's ability to fit local fluctuations and outliers. Through multi-scale feature fusion and temporal consistency constraints, the GETM model maintains relatively stable performance across all conditions. This stability stems from its joint adaptive capability within both the temporal and structural domains.

Moreover, the coupled analysis of sparsity and similarity dimensions indicates that the model achieves its best performance at moderate sparsity (around 0.5) and moderate temperature (approximately 1.5-2.0). This combination reflects a well-coordinated mechanism between graph structural information and temporal features. Moderate sparsity ensures the retention of key dependency paths, while an appropriate temperature coefficient enables the feature aggregation process to focus on locally correlated regions without losing global contextual information. This dynamic balance represents the core strength of the GETM design, allowing it to achieve both high accuracy and strong robustness in long sequence forecasting tasks.

4. Conclusion

This paper addresses the challenges of long sequence forecasting, including complex temporal dependencies, dynamically evolving structural relationships, and inconsistent feature scales, by proposing a Graph-Enhanced Temporal Modeling (GETM) approach. The method introduces dynamic graph structure learning and multi-scale feature fusion into the temporal modeling process, enabling joint representation across temporal and structural spaces. This design allows the model to achieve higher accuracy and stability when dealing with non-stationary, multivariate, and high-dimensional coupled data environments. Experimental results show that GETM outperforms existing mainstream models across multiple core evaluation metrics, validating the effectiveness and generality of the graph enhancement mechanism in long sequence modeling. GETM not only improves the joint characterization of long-term trends and short-term fluctuations but also provides a new structural perspective for modeling the temporal dynamics of complex systems.

From a methodological perspective, the proposed dynamic graph learning mechanism offers a scalable framework to address the "static dependency" limitation of traditional time series models. By dynamically constructing adjacency matrices and employing adaptive sparsification strategies, the model adjusts inter-node dependency weights over time, achieving a balance between global consistency and local sensitivity. This mechanism endows the model with self-regulatory capability for non-stationary signals, significantly enhancing its robustness in cross-temporal and cross-dimensional tasks. Meanwhile, the multi-scale feature fusion module effectively integrates information across different temporal granularities, allowing the model to capture short-term perturbations while maintaining a global understanding of long-term trends. This design provides a more hierarchical and interpretable representation for temporal dynamics in complex systems.

From an application standpoint, this research has broad implications across multiple real-world domains. In cloud computing and backend system monitoring, GETM can be applied to multidimensional performance forecasting and anomaly pattern recognition, facilitating intelligent resource scheduling and service quality assurance. It also demonstrates strong transferability in fields such as energy load forecasting, traffic flow modeling, and industrial equipment monitoring, adapting well to diverse system scales and data characteristics. Particularly in high-dimensional and dynamically interactive environments, the graph-enhanced temporal modeling paradigm effectively captures inter-entity coupling variations, providing more accurate data support for intelligent decision-making and automated control. Moreover, its structural adaptivity offers a theoretical and algorithmic foundation for emerging directions such as edge computing and multi-source data fusion.

Looking forward, there remains substantial potential for further exploration in graph-enhanced temporal modeling. Future work may focus on cross-modal joint modeling, integrating temporal signals with semantic, visual, and log data to achieve a more comprehensive representation of system states. Additionally, model development toward lightweight and interpretable architectures will be critical for real-time prediction and deployment in resource-constrained environments. Incorporating generative modeling and uncertainty estimation techniques could further improve performance in anomaly detection and risk prediction, offering stronger theoretical and practical support for intelligent operations, digital twins, and complex system optimization. Overall, this research enriches the structural representation paradigm of long sequence modeling and lays an essential foundation for future studies on graph-enhanced temporal forecasting.

References

[1] Zhang Y, Zhu Z, Hou J, et al. Spatial-temporal graph enhanced detr towards multi-frame 3d object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024.

[2] Sun X, Xiong R, Shen D, et al. Enhancing Network Traffic Prediction by Integrating Graph Transformer with a Temporal Model[C]//Proceedings of the 9th Asia-Pacific Workshop on Networking. 2025: 150-156.

[3] Xu N, Kosma C, Vazirgiannis M. TimeGNN: temporal dynamic graph learning for time series forecasting[C]//International Conference on Complex Networks and Their Applications. Cham: Springer Nature Switzerland, 2023: 87-99.

[4] Chen L, Chen L, Wang H, et al. Temporal representation learning enhanced dynamic adversarial graph convolutional network for traffic flow prediction[J]. Scientific Reports, 2025, 15(1): 8330.

[5] Weng W, Chen Q, Dai Y, et al. Multi-scale Fusion Dynamic Graph Neural Network For Traffic Flow Prediction[C]//Proceedings of the 2023 2nd International Conference on Algorithms, Data Mining, and Information Technology. 2023: 85-90.

[6] Kong Z, Ren J, Wang L, et al. Multi-scale fusion graph convolutional networks[J]. Expert Systems with Applications, 2025: 129815.

[7] Nayak I, Chakrabarti A, Kumar M, et al. Temporally-consistent koopman autoencoders for forecasting dynamical systems[J]. Scientific Reports, 2025, 15(1): 22127.

[8] She Y, Hong Y, Shen S, et al. Consistency regularization for few shot multivariate time series forecasting[J]. Scientific Reports, 2025, 15(1): 14195.

[9] Mohammadshirazi A, Neogi P P G, Ramnath R. PIAD-SRNN: Physics-Informed Adaptive Decomposition in State-Space RNN[C]//ICML 2025 CO-BUILD Workshop on Computational Optimization of Buildings.

[10] Lim B, Arık S Ö, Loeff N, et al. Temporal fusion transformers for interpretable multi-horizon time series forecasting[J]. International journal of forecasting, 2021, 37(4): 1748-1764.

[11] Xin N, Su J, Hasan M M. MMformer with Adaptive Transferable Attention: Advancing Multivariate Time Series Forecasting for Environmental Applications[J]. arXiv preprint arXiv:2504.14050, 2025.