Heterogeneity in Multivariate Time Series: Comprehensive Analysis and Adaptive Modeling

Zezhi Shao
Institute of Computing Technology, Chinese Academy of Sciences
Haidian, Beijing, China
shaozezhi@ict.ac.cn

Chengqing Yu
Institute of Computing Technology, Chinese Academy of Sciences
Haidian, Beijing, China
University of Chinese Academy of Sciences
Haidian, Beijing, China
yuchengqing22b@ict.ac.cn

Fei Wang
Institute of Computing Technology, Chinese Academy of Sciences
Haidian, Beijing, China
wangfei@ict.ac.cn

Abstract

Multivariate time series (MTS) data are ubiquitous in complex dynamic systems such as meteorology, transportation, and energy. However, data heterogeneity caused by cross-domain variations has become a central bottleneck restricting model generalization and consistency in comparative studies. This paper systematically reviews recent MTS forecasting research, revealing that inconsistencies in experimental conclusions primarily arise from neglecting substantial differences in data distributions and characteristics. To address this issue, we introduce BasicTS, an fair and scalable benchmark designed to fairly quantify the impact of heterogeneity on model performance. Subsequently, to tackle generalization challenges posed by heterogeneity, this tutorial proposes two adaptive solutions: (i) developing BLAST, a balanced and diversity-enhanced pre-training corpus that explicitly models heterogeneity, significantly improving zero-shot general forecasting; and (ii) introducing ARIES, a relational assessment and model recommendation framework that leverages a statistical pattern-to-model matching mechanism to automatically select optimal forecasting models for specific real-world sequences. Through comprehensive experiments and case studies, we demonstrate that precisely characterizing and leveraging data heterogeneity, beyond mere model design, is crucial for improving the robustness of MTS forecasting. This research provides methodological guidance and practical insights for academia and industry to fully exploit the value of time series data and make data-driven decisions.

CCS Concepts

• **Information systems → Data mining**.

Keywords

time series forecasting, heterogeneity, benchmarking

ACM Reference Format:
Zezhi Shao, Chengqing Yu, and Fei Wang. 2025. Heterogeneity in Multivariate Time Series: Comprehensive Analysis and Adaptive Modeling. In *19th International Symposium on Spatial and Temporal Data (SSTD '25), August 25–27, 2025, Osaka, Japan.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3748777.3748814

SSTD '25, Osaka, Japan

ACM ISBN /25/08
https://doi.org/10.1145/3748777.3748814

1 Introduction

Multivariate time series (MTS) data are prevalent in various complex dynamic systems, capturing overall trends such as temperature and wind speed in meteorological systems [7], and traffic volumes in transportation systems [26–29]. These systems exhibit significant differences in their characteristics, whose behavior is shaped by diverse factors and intricate operating mechanisms. Consequently, their modeling, prediction, and control become particularly challenging. Time series data, serving as an abstract representation, provide critical information for unified and cross-domain analysis and forecasting of these dynamic systems.

In recent years, time series forecasting techniques have rapidly advanced, leading to numerous modeling structures and strategies. Prominent approaches include spatial-temporal forecasting [15, 16, 28, 37, 38] and long-term time series forecasting [19, 22, 36, 41–43]. However, despite notable progress, inconsistencies in experimental conclusions persist among different studies, and issues related to algorithm robustness and generalization remain unresolved. While previous research [41] has partially discussed these issues from a model architecture perspective, this tutorial uniquely focuses on data heterogeneity as the central theme. Data heterogeneity refers to significant differences in the statistical properties or distributions of time series data across various domains or systems. Unlike fields such as computer vision or natural language processing, where data typically share strong semantic consistency, time series data from different sources often lack common semantics or patterns, despite structural similarities.

Tutorial overview. This tutorial will be presented in a traditional lecture format lasting of 1.5 hours. Initially, we will systematically review recent advancements in two major research areas within MTS forecasting: spatial-temporal forecasting and long-term forecasting, highlighting the inconsistencies and contradictions found in existing literature. Next, we introduce BasicTS, a fair and scalable benchmarking framework designed to systematically evaluate how data heterogeneity impacts model performance. Using this framework, we examine the strengths and weaknesses of various forecasting algorithms under heterogeneous conditions. Finally, to address challenges posed by data heterogeneity, this tutorial proposes adaptive solutions from two perspectives: universal forecasting models and in-domain forecasting models. These adaptive strategies aim to improve the robustness and generalization of forecasting methods. The tutorial is structured as follows:

- Understanding heterogeneity in MTS data (45 min):
  - Key definitions (5 min)
  - Overview of deep MTS forecasting (20 min)
  - Quantitative evaluation and case studies of MTS data heterogeneity's impact on model performance (20 min)
- Adaptive modeling strategies for MTS heterogeneity (45 min):
  - Evaluation metrics for time series patterns (5 min)
  - Balanced sampling time series corpus for universal forecasting models (20 min)
  - Relation assessment and model recommendation for time series forecasting (20 min)

Prior offerings. This is a newly developed tutorial and has not been previously presented.

Target audience. This tutorial targets researchers in data mining, machine learning, database systems, and data science, especially those focused on time series and spatial-temporal data analysis. With its emphasis on data heterogeneity, the tutorial systematically examines recent advances in MTS forecasting, clarifies the reasons behind conflicting conclusions, and discusses relevant solutions, aligning closely with the core themes of the SSTD conference. Participants should have basic knowledge of data mining, deep learning, and time series analysis. They will gain:

- Knowledge of cutting-edge developments in deep learning-based time series forecasting.
- Insights into how data heterogeneity affects algorithm performance and strategies to manage this issue.
- Opportunities for in-depth interaction with domain experts to explore emerging trends and cutting-edge techniques in time series and spatial-temporal analysis.

Related tutorials. Previous tutorials generally covered broader surveys of time series analysis [13, 39] or addressed specific technical challenges, such as robust analysis [34] and causal discovery [11]. In contrast, this tutorial uniquely emphasizes data heterogeneity, systematically analyzes the evolution of existing research, clarifies discrepancies in outcomes, and introduces advanced adaptive modeling methods tailored to heterogeneous conditions.

2 Tutorial Outline

2.1 Understanding Heterogeneity in MTS Data

2.1.1 ***Key Definitions.*** Multivariate time series data comprise several correlated sequences. In this tutorial, those sequences may either share the same physical meaning—for instance, traffic flow readings from multiple sensors in a road network—or they may capture different quantities, such as the various meteorological variables recorded by a single weather station.

Definition 1. ***Multivariate Time Series.*** *A multivariate time series includes multiple time-dependent variables. It can be expressed as a matrix $\mathbf{X} \in \mathbb{R}^{T\times N}$, where T is the number of time steps and N is the number of variables. We additionally denote the data in time series i ranging from t_1 to t_2 as $\mathbf{X}^i_{t_1:t_2}$.*

Definition 2. ***Multivariate Time Series Forecasting.*** *Given historical data $\mathbf{X} \in \mathbb{R}^{T_h\times N}$ from the past T_h time steps, multivariate time series forecasting aims to predict $\mathbf{Y} \in \mathbb{R}^{T_f\times N}$ of the T_f nearest future time steps.*

Definition 3. ***In-Domain Forecasting Models*** *are typically tailored to the unique characteristics of the data. Due to the heterogeneity of MTS data, the optimal model architecture and parameters are rarely transferable between datasets.*

Definition 4. ***Universal Forecasting Models***[1] *are pre-trained on large-scale time series datasets and are capable of performing accurate zero-shot forecasting across diverse domains.*

2.1.2 ***MTS Forecasting: A Brief Overview.*** We cover studies related to Long-term Time Series Forecasting (LTSF) and Spatial-Temporal Forecasting (STF), which are the two most prominent topics in recent MTS forecasting studies. First, to achieve accurate long-term time series forecasting, researchers have concentrated on extracting temporal patterns from multivariate series. Owing to their strong ability to model long-range dependencies, Transformer[31]-based architectures have become a focal point of recent work [19, 22, 36, 42, 43]. However, a wave of simpler, more computationally efficient linear models [41] has emerged, calling into question both the necessity of Transformers and their computational cost. This tutorial therefore reviews the evolution of Transformer-based LTSF models, highlights the efficient linear alternatives that have gained attention in recent years, and synthesizes the key disagreements and incompatibilities between the two approaches in terms of empirical performance.

Second, unlike LTSF, spatial-temporal forecasting must capture not only temporal dynamics but also the interdependencies among multiple time series. Since graph convolutional networks (GCNs) [14] were introduced in 2017, spatial-temporal graph neural networks (STGNNs) [15, 16, 28, 29, 38] have become the dominant solution, combining the structural modeling power of GCNs with the temporal modeling of sequence models and delivering significant gains across tasks. Recently, however, some studies have questioned the necessity of GCNs in spatial-temporal prediction and pointed out their efficiency limitations [5, 6, 18, 27]. Accordingly, this tutorial traces the development of STGNNs and summarizes current findings on the necessity of spatial modeling and the incompatibilities among different spatial-modeling strategies.

2.1.3 ***BasicTS: A Fair & Scalable Benchmark.*** To draw reliable conclusions, this tutorial first introduces BasicTS [17, 25], a fair and scalable benchmark for time series forecasting. By enforcing a unified pipeline and offering a rich set of evaluation metrics, BasicTS eliminates performance inconsistencies and incomplete assessments that stem from divergent data pre-processing routines, training hyper-parameter settings, and metric implementations.

Leveraging BasicTS and its results, we take a closer look at the heterogeneity of MTS data and dissect the seemingly contradictory findings that often surface in experiments. Concretely, we categorize datasets along two axes—spatial and temporal—to expose the distinct modeling challenges each category poses. On the spatial axis, we use sample indistinguishability as the key indicator of

[1]While some studies refer to these models as foundational or general models, this paper adopts the term *universal forecasting models* [1, 8, 32, 35] for the sake of consistency and to avoid confusion with multi-task models.

dataset characteristics, while on the temporal axis we classify data by the degree of distribution drift over time.

Our analysis shows that most methods are effective only for specific data types; overlooking data heterogeneity can create contradictions and obscure which techniques are truly appropriate.

2.2 Adaptive Modeling of MTS Heterogeneity

2.2.1 ***Quantifying Heterogeneity.*** There are two intuitive ways to tackle the heterogeneity. Path 1 leverages large-scale pre-training corpora to build universal forecasting models that remain robust across domains; Path 2 adaptively picks the most suitable architecture for each data pattern to maximize accuracy and computational efficiency. In either case, explicitly characterizing heterogeneity is pivotal—whether to boost a universal model's generalization or to map data patterns to in-domain models. This tutorial therefore profiles time series with seven widely used statistical descriptors—stationarity, trend, periodicity, volatility, homoscedasticity, memory, and anomaly—capturing their patterns from multiple angles. These metrics lay a solid foundation for heuristic dataset classification, informed model selection, and adaptive refinement.

2.2.2 ***Balanced Sampling Time Series Corpus for Universal Forecasting Models.*** The emergence of universal time series forecasting models has revolutionized zero-shot prediction across various domains [2, 4, 9, 12, 20, 21, 30, 35]. This tutorial first provides a brief overview of the evolution of universal forecasting models, discussing their pre-training datasets and architectural designs. The generalization capability of these models primarily stems from large-scale, diverse pre-training data. However, the importance of data diversity has not been sufficiently explored. Current large-scale time series datasets often suffer from inherent biases and imbalanced distributions, limiting model performance and generalization capabilities.

To address this issue, we introduce BLAST [24], a novel pre-training corpus that explicitly models data heterogeneity and employs a balanced sampling strategy to enhance diversity and correct distributional biases. Experimental results show that models pre-trained on BLAST achieve state-of-the-art performance, significantly reducing computational resources and training data requirements. This work is the first to highlight the critical role of data diversity in enhancing training efficiency and the generalization capability of universal forecasting models.

2.2.3 ***Relation Assessment and Model Recommendation for Time Se ries Forecasting.*** Distinct from language or images, the statistical quantitative properties inherent in time series already suggest their underlying patterns. Although deep forecasting models employ complex architectures [33, 40], they still essentially adopt similar or shared strategies to capture specific patterns. However, the relationship between model strategies and data patterns is not fully explored, and such mismatches may severely degrade performance [3, 10, 23].

To address this problem, we introduce ARIES to assess the relation between strategies and patterns and recommend appropriate forecasting models for realistic time series. For assessment, we test 50+ models on synthetic data with diverse patterns and perform experimental analyses. For real-world applications, ARIES proposes the first recommendation framework for deep forecasting models that can also provide explanatory justifications related to properties and strategies. Overall, ARIES not only offers a data-centric perspective on forecasting innovation, but also addresses the interpretability in deploying black-box forecasting models.

3 Presenters

Zezhi Shao is an Assistant Professor with the Institute of Computing Technology, Chinese Academy of Sciences (ICT, CAS), Beijing. He obtained his Ph.D. in Computer Architecture from ICT, CAS in 2024. His research interests include multivariate time series forecasting, spatio-temporal data mining, and graph neural networks. He has published more than 20 papers in premier venues such as *IEEE TKDE*, *KDD*, *VLDB*, and *CIKM*, and he regularly serves as a reviewer or programme-committee member for these conferences and journals. His current work focuses on foundation models for time series data and cross-domain distribution analysis. He is also the lead maintainer of the open-source project *BasicTS*.

Chengqing Yu is a final-year Ph.D. candidate at the Institute of Computing Technology, Chinese Academy of Sciences, Beijing. His research spans robust spatio-temporal forecasting, time-series prediction, and reinforcement learning. He has published more than 40 papers in leading venues, including *IEEE TKDE*, *KDD*, *CIKM*, and *Information Fusion*. His recent work centres on AI4Science and robust, large-scale spatio-temporal data analysis.

Fei Wang is an Associate Professor at the Institute of Computing Technology, Chinese Academy of Sciences. He received his Ph.D. in Computer Architecture from ICT CAS in 2017. His research covers spatio-temporal data mining, time series analysis, and AI4Science. He has authored or co-authored more than 70 publications in top venues such as *The Innovation*, *KDD*, *VLDB*, *TKDE*, and *CIKM*. His current interests include scale-adaptive unsupervised learning for spatio-temporal data, cross-domain distribution analysis, representative data synthesis, and the development of fair, scalable benchmarks for spatio-temporal models. He is also investigating architectural innovations for high-efficiency models, the construction of spatio-temporal foundation models, and their deployment across diverse scientific and industrial domains.

Acknowledgments

This study is partially supported by NSFC No. 62372430, the Youth Innovation Promotion Association CAS No.2023112, and HUAInnovation fundings. Additionally, Zezhi Shao is also supported by the China Postdoctoral Science Foundation under Grant 2025M771542.

References

[1] Taha Aksu, Gerald Woo, Juncheng Liu, Xu Liu, Chenghao Liu, Silvio Savarese, Caiming Xiong, and Doyen Sahoo. 2024. GIFT-Eval: A Benchmark For General Time Series Forecasting Model Evaluation. *arXiv preprint arXiv:2410.10393* (2024).

[2] Abdul Fatir Ansari, Lorenzo Stella, Caner Turkmen, Xiyuan Zhang, Pedro Mercado, Huibin Shen, Oleksandr Shchur, Syama Sundar Rangapuram, Sebastian Pineda Arango, Shubham Kapoor, et al. 2024. Chronos: Learning the Language of Time Series. *Transactions on Machine Learning Research* (2024).

[3] Lorenzo Brigato, Rafael Morand, Knut Strømmen, Maria Panagiotou, Markus Schmidt, and Stavroula Mougiakakou. 2025. Position: There are no Champions in Long-Term Time Series Forecasting. *arXiv preprint arXiv:2502.14045* (2025).

[4] Abhimanyu Das, Weihao Kong, Rajat Sen, and Yichen Zhou. [n. d.]. A decoder-only foundation model for time-series forecasting. In *Forty-first International Conference on Machine Learning*.

[5] Jinliang Deng, Xiusi Chen, Renhe Jiang, Xuan Song, and Ivor W. Tsang. 2021. ST-Norm: Spatial and Temporal Normalization for Multi-variate Time Series Forecasting. In *SIGKDD*. ACM, 269–278.

[6] Jinliang Deng, Xiusi Chen, Renhe Jiang, Du Yin, Yi Yang, Xuan Song, and Ivor W. Tsang. 2023. Learning Structured Components: Towards Modular and Interpretable Multivariate Time Series Forecasting. *CoRR* abs/2305.13036 (2023). arXiv:2305.13036

[7] Yisong Fu, Fei Wang, Zezhi Shao, Chengqing Yu, Yujie Li, Zhao Chen, Zhulin An, and Yongjun Xu. 2024. LightWeather: Harnessing Absolute Positional Encoding to Efficient and Scalable Global Weather Forecasting. *arXiv preprint arXiv:2408.09695* (2024).

[8] Shanghua Gao, Teddy Koker, Owen Queen, Thomas Hartvigsen, Theodoros Tsiligkaridis, and Marinka Zitnik. 2024. UniTS: A Unified Multi-Task Time Series Model. In *NeurIPS*.

[9] Azul Garza and Max Mergenthaler-Canseco. 2023. TimeGPT-1. *arXiv preprint arXiv:2310.03589* (2023).

[10] Rakshitha Godahewa, Christoph Bergmeir, Geoffrey I Webb, Rob J Hyndman, and Pablo Montero-Manso. 2021. Monash time series forecasting archive. *arXiv preprint arXiv:2105.06643* (2021).

[11] Chang Gong, Di Yao, Chuzhe Zhang, Wenbin Li, Jingping Bi, Lun Du, and Jin Wang. 2023. Causal discovery from temporal data. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 5803–5804.

[12] Mononito Goswami, Konrad Szafer, Arjun Choudhry, Yifu Cai, Shuo Li, and Artur Dubrawski. [n. d.]. MOMENT: A Family of Open Time-series Foundation Models. In *Forty-first International Conference on Machine Learning*.

[13] Eamonn Keogh. 2024. Time series data mining: A unifying view. In *2024 IEEE 11th International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, 1–3.

[14] Thomas N. Kipf and Max Welling. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In *ICLR*. OpenReview.net.

[15] Fuxian Li, Jie Feng, Huan Yan, Guangyin Jin, Fan Yang, Funing Sun, Depeng Jin, and Yong Li. 2023. Dynamic graph convolutional recurrent network for traffic prediction: Benchmark and solution. *ACM Transactions on Knowledge Discovery from Data* 17, 1 (2023), 1–21.

[16] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. 2018. Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. In *ICLR*. OpenReview.net.

[17] Yubo Liang, Zezhi Shao, Fei Wang, Zhao Zhang, Tao Sun, and Yongjun Xu. 2022. Basicts: An open source fair multivariate time series prediction benchmark. In *International symposium on benchmarking, measuring and optimization*. Springer, 87–101.

[18] Hangchen Liu, Zheng Dong, Renhe Jiang, Jiewen Deng, Jinliang Deng, Quanjun Chen, and Xuan Song. 2023. Spatio-temporal adaptive embedding makes vanilla transformer sota for traffic forecasting. In *CIKM*. ACM, 4125–4129.

[19] Shizhan Liu, Hang Yu, Cong Liao, Jianguo Li, Weiyao Lin, Alex X. Liu, and Schahram Dustdar. 2022. Pyraformer: Low-Complexity Pyramidal Attention for Long-Range Time Series Modeling and Forecasting. In *ICLR*. OpenReview.net.

[20] Xu Liu, Juncheng Liu, Gerald Woo, Taha Aksu, Yuxuan Liang, Roger Zimmermann, Chenghao Liu, Silvio Savarese, Caiming Xiong, and Doyen Sahoo. 2024. Moirai-MoE: Empowering Time Series Foundation Models with Sparse Mixture of Experts. *arXiv preprint arXiv:2410.10469* (2024).

[21] Yong Liu, Haoran Zhang, Chenyu Li, Xiangdong Huang, Jianmin Wang, and Mingsheng Long. [n. d.]. Timer: Generative Pre-trained Transformers Are Large Time Series Models. In *Forty-first International Conference on Machine Learning*.

[22] Yuqi Nie, Nam H. Nguyen, Phanwadee Sinthong, and Jayant Kalagnanam. 2023. A Time Series is Worth 64 Words: Long-term Forecasting with Transformers. In *ICLR*. OpenReview.net.

[23] Xiangfei Qiu, Jilin Hu, Lekui Zhou, Xingjian Wu, Junyang Du, Buang Zhang, Chenjuan Guo, Aoying Zhou, Christian S Jensen, Zhenli Sheng, et al. 2024. Tfb: Towards comprehensive and fair benchmarking of time series forecasting methods. *arXiv preprint arXiv:2403.20150* (2024).

[24] Zezhi Shao, Yujie Li, Fei Wang, Chengqing Yu, Yisong Fu, Tangwen Qian, Bin Xu, Boyu Diao, Yongjun Xu, and Xueqi Cheng. 2025. BLAST: Balanced Sampling Time Series Corpus for Universal Forecasting Models. *arXiv preprint arXiv:2505.17871* (2025).

[25] Zezhi Shao, Fei Wang, Yongjun Xu, Wei Wei, Chengqing Yu, Zhao Zhang, Di Yao, Tao Sun, Guangyin Jin, Xin Cao, et al. 2024. Exploring progress in multivariate time series forecasting: Comprehensive benchmarking and heterogeneity analysis. *TKDE* (2024).

[26] Zezhi Shao, Fei Wang, Zhao Zhang, Yuchen Fang, Guangyin Jin, and Yongjun Xu. 2023. Hutformer: Hierarchical u-net transformer for long-term traffic forecasting. *arXiv preprint arXiv:2307.14596* (2023).

[27] Zezhi Shao, Zhao Zhang, Fei Wang, Wei Wei, and Yongjun Xu. 2022. Spatial-Temporal Identity: A Simple yet Effective Baseline for Multivariate Time Series Forecasting. In *CIKM*. ACM, 4454–4458.

[28] Zezhi Shao, Zhao Zhang, Fei Wang, and Yongjun Xu. 2022. Pre-training enhanced spatial-temporal graph neural network for multivariate time series forecasting. In *Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining*. 1567–1577.

[29] Zezhi Shao, Zhao Zhang, Wei Wei, Fei Wang, Yongjun Xu, Xin Cao, and Christian S. Jensen. 2022. Decoupled Dynamic Spatial-Temporal Graph Neural Network for Traffic Forecasting. *Proc. VLDB Endow.* 15, 11 (2022), 2733–2746.

[30] Xiaoming Shi, Shiyu Wang, Yuqi Nie, Dianqi Li, Zhou Ye, Qingsong Wen, and Ming Jin. 2024. Time-MoE: Billion-Scale Time Series Foundation Models with Mixture of Experts. *arXiv preprint arXiv:2409.16040* (2024).

[31] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is All you Need. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*. 5998–6008.

[32] Yihang Wang, Yuying Qiu, Peng Chen, Kai Zhao, Yang Shu, Zhongwen Rao, Lujia Pan, Bin Yang, and Chenjuan Guo. 2024. ROSE: Register Assisted General Time Series Forecasting with Decomposed Frequency Learning. *arXiv preprint arXiv:2405.17478* (2024).

[33] Yuxuan Wang, Haixu Wu, Jiaxiang Dong, Yong Liu, Mingsheng Long, and Jianmin Wang. 2024. Deep Time Series Models: A Comprehensive Survey and Benchmark. (2024).

[34] Qingsong Wen, Linxiao Yang, Tian Zhou, and Liang Sun. 2022. Robust time series analysis and applications: An industrial perspective. In *Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining*. 4836–4837.

[35] Gerald Woo, Chenghao Liu, Akshat Kumar, Caiming Xiong, Silvio Savarese, and Doyen Sahoo. [n. d.]. Unified Training of Universal Time Series Forecasting Transformers. In *Forty-first International Conference on Machine Learning*.

[36] Haixu Wu, Jiehui Xu, Jianmin Wang, and Mingsheng Long. 2021. Autoformer: Decomposition Transformers with Auto-Correlation for Long-Term Series Forecasting. In *NeurIPS*. 22419–22430.

[37] Zonghan Wu, Shirui Pan, Guodong Long, Jing Jiang, Xiaojun Chang, and Chengqi Zhang. 2020. Connecting the Dots: Multivariate Time Series Forecasting with Graph Neural Networks. In *SIGKDD*. ACM, 753–763.

[38] Zonghan Wu, Shirui Pan, Guodong Long, Jing Jiang, and Chengqi Zhang. 2019. Graph WaveNet for Deep Spatial-Temporal Graph Modeling. In *IJCAI*. ijcai.org, 1907–1913.

[39] Bin Yang, Yuxuan Liang, Chenjuan Guo, and Christian S Jensen. 2025. Data Driven Decision Making with Time Series and Spatio-temporal Data. *arXiv preprint arXiv:2503.08473* (2025).

[40] Chengqing Yu, Fei Wang, Zezhi Shao, Tao Sun, Lin Wu, and Yongjun Xu. 2023. Dsformer: A double sampling transformer for multivariate time series long-term prediction. In *CIKM*. 3062–3072.

[41] Ailing Zeng, Muxi Chen, Lei Zhang, and Qiang Xu. 2023. Are Transformers Effective for Time Series Forecasting? *AAAI*.

[42] Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. 2021. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. In *AAAI*. 11106–11115.

[43] Tian Zhou, Ziqing Ma, Qingsong Wen, Xue Wang, Liang Sun, and Rong Jin. 2022. FEDformer: Frequency Enhanced Decomposed Transformer for Long-term Series Forecasting. In *ICML*, Vol. 162. PMLR, 27268–27286.