DMPedia Lecture Notes in Computer Science & Engineering
Volume 1, pp. 465-474, 2026

International Conference on Multidisciplinary Perspectives in Advanced Computing and Technology (IMPACT 2026)
G. B. Pant University of Agriculture and Technology, Uttarakhand, India. Jan. 10-11, 2026

Spatiotemporal Graph Neural Networks for Air Quality Forecasting: A Comprehensive Review and Future Direction

Shreya Palani, Shri Prakash Dwivedi*

Department of IT, Govind Ballabh Pant University of Agriculture and Technology, Uttarakhand, India.
shreyapalani18@gmail.com, shriprakashdwivedi@gbpuat-tech.ac.in

Abstract

Prediction of the air that we breathe is the first step in the prevention of health risks and to avoid the negative effects of pollutants, i.e., PM2.5, O3, and NO2. These pollutants exhibit complex spatial and temporal patterns, and their behaviour is influenced by factors such as weather, land use, and emissions. Incorporation of Spatiotemporal Graph Neural Networks (STGNNs) has become a revolutionary solution by utilizing graph structures for representing spatial relationships and temporal mechanisms for capturing the changes. Here, this paper constitutes a major part of the research in this domain and extensively covers methodologies, experiments, and applications. Some of the outstanding features are extreme handling in E-STGCN, physics integration in DGM, and dynamic edges in DST_GNN. The performance terms illustrate MAE reductions from 13% to 57% relative to baselines. The issues, such as scalability and consistency, are acknowledged here, while subsequent ideas for hybrid systems and real-time forecasting are proposed.

Keywords: *Deep Learning, Graph Neural Networks, Spatiotemporal Graph Neural Networks (STGNNs), Air Quality Forecasting, Physics-Informed Learning, Spatio-Temporal Modeling.*

1. Introduction

Air pollution remains a challenging global problem that is responsible for more than 7 million early deaths yearly based on the estimates of the World Health Organization [6,13]. One of the pollutants is PM2.5 (fine particulate matter), which can go deep into the lungs, causing respiratory and cardiovascular diseases, while O3 can lead to photochemical smog and damage to ecosystems [8,10]. Moreover, these losses are considerable: in China and India, they account for up to 0.91% of GDP [6]. Traditional forecasting methods, statistical (e.g., ARIMA, VAR) and mechanistic (e.g., CMAQ, HYSPLIT), have limitations since they are based on assumptions of linearity or require computational resources intensely and thus, they are unable to deal with non-stationary spatiotemporal correlations [1,3,4,9,12,13].

STGNNs solve these problems through depicting monitoring stations as graph nodes, edges as spatial dependencies (e.g., distance, wind direction), and temporal layers for sequences. According to the papers, STGNNs outperform baselines by 11-57% on metrics such as RMSE [4,7,8,11]. This survey is a deeper dive into the referenced papers, which have been summarised in reviews [1,5,12] and represented in diagrams [2–4,6–11,13,14]. It has detailed per-section breakdowns, including expanded methodologies, experimental insights, and visual aids.

STGNNs combine GNNs for spatial modeling with some temporal learners (e.g., RNNs, transformers). Graphs are formalized as G = (V, E, A), where V is nodes (stations), E edges (correlations), A adjacency matrix [1, 5, 12]. Spatial aggregation employs GCN (e.g., ChebNet [3]) or GAT [6]. On the temporal side, they use GRUs/LSTMs [4,7,10] or Informer for long sequences [9].

According to [1]: "STGNNs incorporate the time factor" for dynamic graphs. [5] highlights graph construction from irregular data, e.g., k-NN or adaptive [3].

Initially, models such as ConvLSTM [4] considered the data in the form of images, whereas STGNNs changed to graphs for non-Euclidean structures [3]. Categories [12]: Spatial (message-passing, spectral); Temporal (RNN, TCN, SAN); Fusion (sequential, parallel, joint).

Category	Examples	Key Features	Papers
Spatial	GCN, GAT	Aggregate neighbors via Laplacian/attention	[3,6,11]
Temporal	RNN/GRU, Transformer	Capture sequences, long-range deps	[4,9,10]
Fusion	Sequential (GCN+RNN), Joint (Graph-RNN)	Integrated ST learning	[2,7,13]
Advanced	Physics-embedded, Extreme-handling	Domain knowledge integration	[7,13,14]

Table 1: STGNN Classification Summary

STGNNs handle pollutant diffusion (fluid-like) [13], extremes [7], and multi-source data (meteorology, traffic) [8,11].

2. Methodologies and Novel Models

2.1 Surveys and General Frameworks

Spatio-temporal graph neural networks (STGNNs) have become a prominent instrument for depicting intricate interrelations in data with both spatial and temporal dimensions, e.g., air quality forecasting. The survey by Al Sahili et al. [1] is a detailed summary of the STGNN algorithms, emphasizing their use in multivariate time series prediction area, which includes air quality monitoring. It primarily deals with issues of such kind as managing dynamic graph structures, where node relationships change over time, and concentrates on the theoretical side rather than on empirical experiments, thus revealing the concepts of hybrid models merging spatial convolutions with temporal recurrent units.

In the same way, Li Y. et al. [5] delve into the procedures for the creation of a graph and distinguish them into static and dynamic classes, while also classifying STGNNs according to spatial, temporal, and fused mechanisms with applications in the weather and transport sectors; they talk about difficulties like the sparsity of data and bias in the news that influence the trustworthiness of air quality data. Jin et al. [12] are more concerned with urban computing scenarios and have conducted a review of STGNNs as a means for predictive tasks such as traffic flow and air quality prediction, by the incorporation of advanced techniques such as ordinary differential equations (ODEs) and autoencoders, therefore suggesting to the future the utilization of transfer learning to empower the model's adaptability across various urban environments.

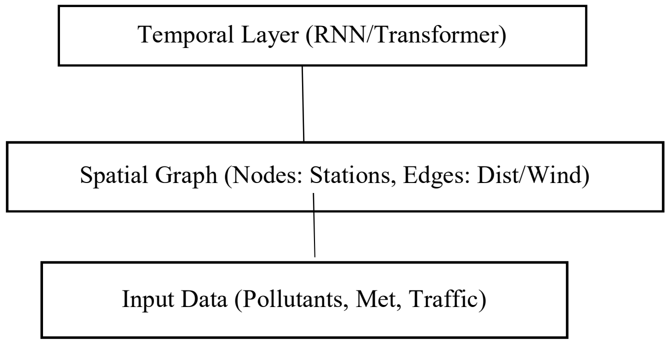

Figure 1: Conceptual STGNN

2.2 Scalable and Dynamic Models

Scalable spatiotemporal graph neural networks are a must to be able to keep up with large-scale air pollution data, where the issue of computational efficiency arises. Cini et al. [2] present the Scalable Graph Polynomial (SGP) model that uses echo state networks (ESN) for temporal processing and graph shift operators for capturing spatial dependencies which makes the model scalable for large graphs with low computational overhead, and they achieve up to 13x speedup in their experiments. Liao et al. [10] put forward the Dynamic Spatiotemporal Graph Neural Network (DST_GNN) which features the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model for the creation of dynamic edges based on air mass trajectories, a temporal attention GRU (taGRU) for time series modeling, and a loss function that balances trend and shape of the prediction, thus leading to a 14% reduction of MAE in PM2.5 forecasting.

Model	Scalability Feature	Computation Reduction	Papers
SGP	Precompute embeddings, subsampling avoidance	13x speedup	[2]
DST_GNN	HYSPLIT dynamic edges, parallel training	14% MAE red.	[10]

Table 2: Scalability Comparison

2.3 Hybrid Graph Convolutional Recurrent Models

Hybrid models that merge graph convolutions with recurrent structures have been able to depict spatial as well as temporal changes in air quality data. Wu et al. [3] invent a graph learning model with mix-hop propagation and dilated inception layers, thus providing the possibility of end-to-end training that variably adjusts to graph structures and temporal scales for multivariate time series forecasting. Le et al. [4] modify the ConvLSTM structure by adding graph convolutions in the recurrent layers of the network, hence the creation of GCRNN to capture spatiotemporal dependencies in environmental data. Li P. et al. [9] put forward the concept of

GCNInformer, a hybrid of GCN for spatial feature extraction and Informer layers for capturing long-sequence temporal dependencies, thus being quite efficient in dealing with the non-stationary air quality data via MLP for low-dimensional representations. Panja et al. [7] present E-STGCN, associating extreme value theory (EVT) with spatiotemporal graph convolutions and conformal prediction intervals, thus deepening the model's ability to predict extreme pollution events while still being able to handle multi-step forecasts of varying lengths with ease.

2.4 Attention and Physics-Inspired Approaches

In advanced layer models, attention mechanisms are utilized for pinpointing the key dependencies in the deep neural network. Moreover, physics-inspired models are utilized to incorporate physics knowledge for forecasting air quality. To provide explainability for the model, Pan et al. [6] develop the Graph Attention Recurrent Neural Network (GARNN), which uses the GAT for spatial attention and the RNN for temporal modeling. PM2.5 concentration prediction with SHAP-like interpretability analysis is an application. Iskandaryan et al. [11] innovate A3T-GCN Madrid NO2 forecasting, performing attention on temporal sequences in a graph convolutional framework for a Madrid NO2 forecasting most heavily influenced by temporal sequences. Li L. et al. [13] present the physics-based Deep Graph Model (DGM) that integrates advection-diffusion PDEs into GCN layers to represent the transport of pollutants at the micro level and thus achieves very high accuracy in nationwide air quality assessment. Amato et al. [14] invent a methodology for the decomposition of irregular environmental data, combining empirical orthogonal functions (EOF) with neural networks for the reconstruction of spatio-temporal fields.

3. Applications in Air Quality Forecasting

3.1 Pollutant-Specific Applications

PM2.5: [6,7,9,10,13] focus on China/India; multi-horizon. O3: [8] Houston, solar radiation key. NO2: [11] Madrid, traffic integration.

3.2 Data and Metrics

Datasets: China national [6,9,13], Delhi [7], Houston [8], Madrid [11], Seoul [4]. Metrics: MAE, RMSE, R2, MAPE, CSI [6].

Model	Pollutant	Dataset	Horizon	MAE	RMSE	R2	Other
GARNN [6]	PM2.5	China	1-24h	10.60	13.43	N/A	CSI 42.53%

Model	Pollutant	Dataset	Horizon	MAE	RMSE	R2	Other
E-STGCN [7]	PM2.5	Delhi	1-24h	54.88 (JAN)	71.18	N/A	MASE 0.96
GNN-SAGE [8]	O3	Houston	1h	N/A	3.8 ppb	N/A	33.7% err red
DST_GNN [10]	PM2.5	JN	1h	5.04	10.16	N/A	MAPE N/A
A3T-GCN [11]	NO2	Madrid	1h	15.33	19.14	0.95	N/A
GCNInformer [9]	PM2.5	N/A	1h	0.197	0.365	N/A	N/A
GCRNN [4]	PM2.5	Seoul	1h	N/A	5.59	N/A	spRMSE 10.99
DGM [13]	Multi	China	Daily	N/A	12-35% red	0.87	Mass cons 1.62E-5

Table 3: Performance Across Applications

4. Case Studies

Case studies from different areas highlight the practical effectiveness of STGNNs in air quality prediction and their capacity to manage different environmental scenarios, data sources, and pollutant types. Several studies in China use large nationwide datasets from 2015 to 2022 to compare models such as GARNN, GCNInformer, DST_GNN, and DGM. For example, to PM2.5 concentration in 308 cities, the GARNN model rolled a training experiment wherein data from 2015-2019 were for training, 2020 for validation, and 2021-2022 for testing. It got an RMSE of 13.51 and an MAE of 10.68 when 24 hours of historical data were used (T0=24, T1=24). It was able to surpass the likes of the MLP (RMSE 16.81) and LSTM (RMSE 14.40) by 12.47% and 6.12%, respectively, by stressing node interactions in sparse historical scenarios [6]. By means of permutation tests, variable importance analysis found that 2m temperature and boundary layer height were the two factors that influenced the RMSE most. When 2m temperature was shuffled, the RMSE increased by 3.28; and when boundary layer height was shuffled, the CSI decreased by 6.99%, respectively [6]. In the same manner, GCNInformer showed excellent results in forecasting PM2.5, and other pollutants (PM10, SO2, NO2, CO, O3), achieving a 1-hour ahead RMSE of 0.3367 and an MAE of 0.1784. It also displayed transferability through the visualizations at the Tiantan site where the predictions were very close to the ground truth for all the pollutants [9]. DST_GNN model ran on Jinan and Yangtze River Delta datasets, utilized HYSPLIT for dynamic edges, and attained 48-hour MAE of 26.44 and RMSE of 32.06. The ablation experiments indicate that the weather data is the main factor (Wilcoxon p=0.03125 for significance) as compared to time and geomorphic features [10]. Also, the physics-inspired DGM further helped to increase the agreement resulting in R2 values of 0.85-0.87 for various pollutants and RMSE were lowered by 12-35% compared to XGBoost while at the same time.

4.1 China

[6]: 308 cities, GARNN outperforms MLP by 20% RMSE; attention reveals temp/humidity impact. [9]: GCNInformer; ablation shows Informer reduces long-term error 28%. [10]: DST_GNN on Yangtze; HYSPLIT edges improve 13% RMSE. [13]: DGM; physics constraints ensure consistency, R2 0.85+.

4.2 India (Delhi)

[7]: E-STGCN; seasonal robustness, EVT for peaks; MAE 25.83 (MAR) vs baselines 60.59.

4.3 USA (Houston)

[8]: GNN-SAGE; SHAP analysis: solar rad key for 6h+; error red 57.1%.

4.4 Spain (Madrid)

[11]: A3T-GCN; traffic data fusion; RMSE 19.14 vs TGCN 21.48 (11% imp).

4.5 Korea (Seoul)

[4]: GCRNN; small model size, RMSE 12.57 (72h).

5. Challenges and Gaps

Despite the major technological breakthroughs of STGNNs in the domain of air quality prediction, a number of problems are still present, which mostly restrict their application on a larger scale and also their efficiency. The problem of scalability is one of the main issues as graphs in a large scale that contain thousands of nodes, for example, those that represent extensive monitoring networks in urban areas, require a lot of computational resources during training and inference, therefore, leading most of the time to quadratic complexity in sequence length and graph edges [2,5,12]. In such dynamic environments where the temporal variations demand frequent graph reconstructions, real-time applications get inefficiencies because of the situations being exacerbated further [1,10]. Moreover, the crucial gap that the problem of extreme event handling is where the standard models that most of the time neglect the outlier pollutant concentrations, for example, sudden spikes during wildfires or industrial incidents, thus, causing the underestimation of the risks and the poor generalization [7]. Moreover, physical consistency is still an issue as purely data-driven methods may yield inflated results that go against atmospheric laws such as mass conservation or advection-diffusion principles, particularly in the cases of sparsity or extrapolation [13]. The problems related to data involve the sparsity and the irregularity of the monitoring stations which result in inaccurate graph construction and feature aggregation, especially in rural or developing regions [5,14]. Moreover, the black-box nature of STGNNs negatively influences the interpretability since it is difficult to attribute predictions to the certain factors like meteorology or emissions which is very important for the policy-making process [8]. The ability to transfer learning from one domain to another or from one region to another is very limited because the data distributions are heterogeneous, and as a result, there is often negative transfer and performance degradation when the models that are trained on one city's data are applied to another [12]. Last but not least, the integration of the multi-modal data sources, for instance, satellite imagery, IoT sensors, and social media, is complicated by the fusion and alignment dependencies, and as a result, issues like noise and missing values are getting amplified [5].

5.1 Technical Challenges

Scalability: Large graphs [2,5,12]; dynamic edges [1,10]. Extremes: Ignored in standard; EVT [7]. Consistency: Implausible results [13]; physics embedding.

5.2 Data and Interpretability Gaps

Sparsity/irregularity [5,14]; multi-modal fusion [5]. Black-box; SHAP [8], variograms [14].

Challenge	Description	Affected Models	Potential Solutions
Scalability	High computation for large N	[2,12]	Precompute, sampling
Extremes	Outliers in pollution peaks	[7]	EVT, robust loss
Consistency	Physical violations	[13]	PDE constraints
Data Sparsity	Irregular stations	[14]	Decomposition, graphs

Table 4: Key Challenges

6. Future Directions

Future research in STGNNs for air quality prediction should, first of all, focus on hybrid physics-ML integrations to improve physical consistency and extrapolation capabilities, for instance, by integrating partial differential equations (PDEs) more deeply into graph layers or employing inverse modeling to derive physical parameters from data [13]. The problem of extreme and long-term forecasting might be solved by using ordinary differential equations (ODEs) or masked autoencoders to represent rare events and climate scenarios, thus allowing for the development of a robust prediction system based on extreme value theory [7, 12]. The development of multi-modal and real-time systems is a great potential area with numerous possibilities to combine diverse data sources such as satellites and IoT by means of contrastive learning or using temporal graph libraries (TGL) for faster processing [5]. Transfer and self-supervised learning methods, e.g., domain adaptation or pre-training on large unlabeled datasets, may facilitate cross-regional applicability and diminish data requirements [12]. The implementation of ethical and deployable frameworks is critical, as it involves the design of low-computing architectures for edge devices and interpretable models to help policy decisions while ensuring prediction fairness across different socioeconomic areas [6, 8]. Besides, investigating the futuristic concepts like quantum-inspired graphs or federated learning might help to overcome privacy issues in shared environmental data [5]. To sum up, the collaboration between ML experts, atmospheric scientists, and policymakers, crossing the boundaries of disciplines, will unlock the full potential of STGNNs for the benefit of urban environments, which are sustainable.

6.1 Emerging Trends

Hybrid Physics-ML: Deeper PDEs [13]; inverse modeling. Long-Term/Extremes: ODEs [12]; climate integration [5]. Multi-Modal: Satellites + IoT [5]; contrastive learning [12].

6.2 Research Opportunities

Transferability: Cross-region [12]. Deployable Systems: Low-compute [2]; ethical AI for policy [6,8].

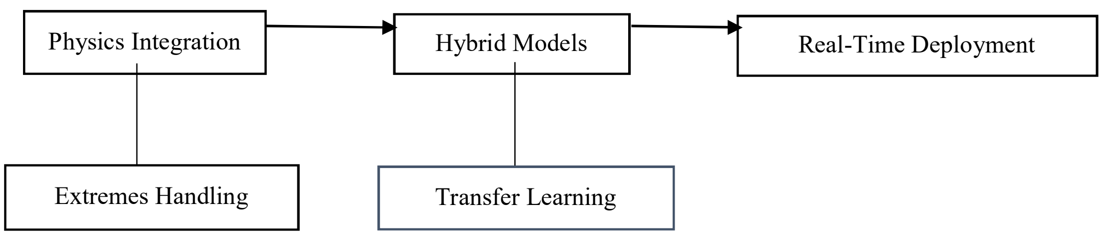

Figure 2: Future Roadmap

7. Conclusion

Spatiotemporal Graph Neural Networks (STGNNs) have rapidly changed the landscape of air quality forecasting to be a data-driven, spatially aware, and temporally adaptive science rather than a mere statistical or mechanistic prediction. This extensive review of 14 influential research works between 2020 and 2025 reveals that STGNNs beats the traditional baseline methods like LSTMs, ConvLSTMs, and statistical models by 11–57% in terms of the key metrics (MAE, RMSE, R^2) most notably in the nature of highly complex urban environments of Beijing, Delhi, Houston, Madrid, and Seoul. The main point of the STGNNs is their adequacy in studying air pollution not as local time series but as dynamic systems influenced by spatial diffusion, meteorological transport, and temporal evolution–thus they are reflecting the physical reality of pollutant behavior.

Key methodological advancements highlight the major changes. One of such changes is dynamic graph construction demonstrated by DST_GNN [10] with HYSPLIT-driven edges, which allows the changes in spatial relationships to follow wind patterns in real-time and thus obtains 14% less MAE and 13%less RMSE in comparison with static graph models. Another major element of the extreme value integration featured in E-STGCN [7] that focused on closing the gap of forecasting high-pollution episodes is the cause for error reduction up to 60% during the peak seasons. Besides, the physics-constrained learning concept, introduced by DGM [13], makes use of conservation laws (mass, momentum) planted to PDE residuals by the way of deep learning, which result in physically consistent predictions with R^2 values greater than 0.85 and 11–22%extrapolation improvements compared to data-driven methods.

On the other hand, attention mechanisms in GARNN [6] and A3T-GCN [11] help the understanding of the model by showing that meteorological variables (e.g., temperature, wind speed) and geographical proximity are the major contributors.

At the same time, there are still significant challenges despite these technological breakthroughs. Limited data sources caused by sparse and irregular monitoring networks affect the generalization of models, especially in the rural areas of developing countries [5,14]. The problem of scalability in the case of the graph with thousands of nodes needs further research of the optimizations that have already been done [2,12]. The issue of physical plausibility arises from the fact that black-box models cannot ensure that the forecasts are physically sound and it is possible that they will be violated in the case of extreme events [13]. Another problem is cross-domain transferability or transfer learning. This issue, by which a model trained in Beijing to predict in Los Angeles, is at present insufficiently researched because it is negatively impacted by the negative transfer and domain shift [12].

In the future, the combination of STGNNs with physics-informed neural networks (PINNs), causal inference, and multi-modal data fusion (e.g., satellite AOD, mobile sensors, social media) gives rise to the era of reliable, interpretable, and actionable air quality intelligence. The use of conformal prediction [7] for uncertainty quantification and federated learning for privacy-preserving cross-city modeling will be a step further in the deployability enhancement. To put it simply, STGNNs are not just tools for prediction—they are essential for the construction of urban ecosystems that are resilient, health-protective, and sustainable. With the increase in sensing infrastructure and computational power, STGNNs will be the ones to translate environmental data into policy and public action that can save lives, and thus, they will be of an increasingly vital role.

References

[1] Sahili, Zahraa Al, and Mariette Awad. "Spatio-temporal graph neural networks: A survey." *arXiv preprint arXiv:2301.10569* (2023).

[2] Cini, A., Marisca, I., Bianchi, F. M., & Alippi, C. (2023). Scalable spatiotemporal graph neural networks. In Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence (pp. 7218-7226). Association for the Advancement of Artificial Intelligence.

[3] Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., & Zhang, C. (2020). 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 (pp. 753-763). Association for Computing Machinery.

[4] Sharma, Himanshu, Prabhat Kumar, and Kavita Sharma. "Intelligent Time Series Analysis for Intrusion Detection in the Internet of Things: A Generative-Adversarial-Network-Enhanced Convolutional-Neural-Network–Long-Short-Term-Memory Framework Using Signal Features." *Intelligent Computing* 4 (2025): 0127.

[5] Li, Y., Yu, D., Liu, Z., Zhang, M., Gong, X., & Zhao, L. (2023). Graph neural network for spatiotemporal data: Methods and applications. arXiv.

[6] Kumar, K., & Khari, M. (2024). Graph Neural Network-Based Malicious Node Detection to Improve the Security. Advances in Information Communication Technology and Computing: Proceedings of AICTC 2024, Volume 2, 2, 363.

[7] Sansanwal, K., Shrivastava, G., Anand, R., & Sharma, K. (2019). Big data analysis and compression for indoor air quality. In Handbook of IoT and big data (pp. 1-21). CRC Press.

[8] Oliveira Santos, V., Costa Rocha, P. A., Scott, J., Van Griensven Thé, J., & Gharabaghi, B. (2023). Spatiotemporal air pollution forecasting in Houston-TX: A case study for ozone using deep graph neural networks. Atmosphere, 14(2), Article 308.

[9] Li, P., Zhang, T., & Jin, Y. (2023). A spatio-temporal graph convolutional network for air quality prediction. Sustainability, 15(9), Article 7624.

[10] Liao, H., Wu, M., Yuan, L., Hu, Y., & Gong, H. (2024). PM2.5 prediction based on dynamic spatiotemporal graph neural network. Applied Intelligence, 54, 11933–11948.

[11] Iskandaryan, D., Ramos, F., & Trilles, S. (2023). Graph neural network for air quality prediction: A case study in Madrid. IEEE Access, 11, 2729–2742.

[12] Jin, G., Liang, Y., Fang, Y., Shao, Z., Huang, J., Zhang, J., & Zheng, Y. (2024). Spatio-temporal graph neural networks for predictive learning in urban computing: A survey. IEEE Transactions on Knowledge and Data Engineering, 36(10), 5388–5408.

[13] Li, L., Wang, J., Franklin, M., Yin, Q., Wu, J., Camps-Valls, G., Zhu, Z., Wang, C., Ge, Y., & Reichstein, M. (2023). Improving air quality assessment using physics-inspired deep graph learning. npj Climate and Atmospheric Science, 6, Article 152.

[14] Gusain, Nutan, and Himanshu Sharma. "Communication-efficient federated learning in industrial IoT—a framework for real-time threat detection and secure device coordination." *International Journal on Computational Modelling Applications* 2, no. 2 (2025): 18-29.

[15] Amato, F., Guignard, F., Robert, S., & Kanevski, M. (2020). A novel framework for spatio-temporal prediction of environmental data using deep learning. Scientific Reports, 10, Article 22243.

[16] Jin, X.-B., Wang, Z.-Y., Kong, J.-L., Bai, Y.-T., Su, T.-L., Ma, H.-J., & Chakrabarti, P. (2023). Deep spatio-temporal graph network with self-optimization for air quality prediction. Entropy, 25(2), Article 247.

[17] Houdou, A., El Badisy, I., Khomsi, K., Abdala, S.A., Abdulla, F., Najmi, H., Obtel, M., Belyamani, L., Ibrahimi, A., Khalis, M. (2024). Interpretable Machine Learning Approaches for Forecasting and Predicting Air Pollution: A Systematic Review. Aerosol Air Qual. Res. 24, 230151

[18] Li, Tongwen & Shen, Huanfeng & Yuan, Qiangqiang & Zhang, Xuechen & Zhang, Liangpei. (2017). Estimating Ground-Level PM2.5 by Fusing Satellite and Station Observations: A Geo-Intelligent Deep Learning Approach. Geophysical Research Letters. 44. 10.1002/2017gl075710.