Future Engineering Research

Open Access Journal

May 2027| Volume 02 | Issue 01 | Pages 19-26

Journal homepage: https://fupubco.com/fer

https://doi.org/10.55670/fpll.fer.2.1.2

Future Publishing LLC

Article

Spatial–temporal approaches for $PM_{2.5}$ concentration prediction: a review

Hemant N Patel*

Computer Engineering, Sankalchand Patel College of Engineering, Sankalchand Patel University, Visnagar, India
*Corresponding Author: hp15284@gmail.com

ABSTRACT

PM pollution has become a major environmental concern due to its adverse effects on air quality and human health. Accurate concentration prediction is essential for understanding pollution patterns and supporting effective air quality management. This paper examines statistical analysis, machine learning (ML), deep learning (DL), and data-driven forecasting as spatiotemporal approaches for predicting $PM_{2.5}$ concentrations. Trends in pollution data over time can be identified using statistical models like ARIMA and linear regression. To understand the complex interplay between pollution levels, weather, and other environmental variables, ML techniques have proven highly effective. Transformers, recurrent neural networks, and convolutional neural networks (CNNs) are examples of DL architectures that can learn to extract spatial patterns and capture temporal dependencies across regions and time. The review also discusses spatial–temporal relationships, including autocorrelation, spatial heterogeneity, and cross-region prediction challenges. Ground-based monitoring stations, meteorological observations, reanalysis datasets, and satellite remote sensing provide important inputs for model development. Furthermore, inconsistent spatiotemporal resolution, uneven monitoring station distribution, and missing data remain significant challenges. This study provides a concise overview of current prediction methods, data needs, and practical constraints to support the development of trustworthy spatiotemporal $PM_{2.5}$ forecasting systems and enhanced air quality monitoring.

Keywords: *$PM_{2.5}$ prediction, Spatial–temporal analysis, Air quality forecasting, Machine learning, Deep learning, Air pollution*

Received: 05 September 2026 Received in revised form: 21 October 2026 Accepted: 13 November 2026

DOI: 10.55670/fpll.fer.2.1.2

1. INTRODUCTION

Air pollution is a significant environmental and public health issue, particularly in urban areas of developing countries. Fine particulate matter ($PM_{2.5}$) is a major pollutant of concern because it can penetrate deep into the lungs and enter the bloodstream, causing respiratory and cardiovascular diseases [1]. Particulate matter, abbreviated as $PM_{2.5}$, contains particles with a diameter of 2.5 μm or less and is often measured in micrograms per cubic meter ($\mu g/m^3$). Pollutant levels that are dangerous to humans are routinely surpassed in India's urban and industrial regions, where the air quality limits set by the World Health Organization (WHO) are often disregarded [2]. Natural sources of particulate matter (PM) include dust storms and wildfires, whereas man-made sources include emissions from cars and industry, as well as biomass burning. These particles are classified by size: PM_{10} (≤10 μm), $PM_{2.5}$ (≤2.5 μm), and ultrafine particles, or $PM_{0.1}$ (≤0.1 μm) [3]. Accurately predicting its concentration is crucial for managing and monitoring air quality effectively, since it fluctuates greatly across places and over time [4]. Accurate prediction is crucial for managing and monitoring air quality effectively because it fluctuates greatly across places and over time. $PM_{2.5}$ concentration exhibits strong spatial and temporal dependencies. Spatial dependencies represent relationships between monitoring stations influenced by emission sources, land use, and pollutant transport. Temporal dependencies describe concentration variations associated with meteorological conditions, traffic activity, seasonal changes, and historical pollutant levels [5]. Capturing both dependencies enables prediction models to identify localized pollution patterns and evolving concentration trends. Accurate $PM_{2.5}$ prediction supports pollution assessment, early warning systems, and air-quality management by considering spatial variations, temporal trends, meteorological influences, and pollutant interactions. Conventional station-level forecasting studies primarily use ground monitoring data and meteorological variables to predict short-term $PM_{2.5}$ concentrations [6]. However, variations in spatial and temporal patterns require approaches that can learn relationships across

monitoring locations and successive time intervals. Recent ML and DL approaches provide opportunities to capture these complex dependencies using data collected from multiple monitoring stations. Therefore, this review examines spatial–temporal approaches for $PM_{2.5}$ concentration prediction, focusing on their ability to capture spatial relationships, temporal variations, and combined dependencies to improve prediction accuracy and reliability. The paper is structured as follows: Section 2 covers prediction methods for $PM_{2.5}$ concentration, including statistical, ML, and DL techniques. Section 3 presents spatial–temporal techniques and Transformer models. Section IV describes data sources, monitoring systems, and prediction challenges. Section V reviews related studies and research gaps. Finally, Section VI presents the conclusion and future work.

2. $PM_{2.5}$ CONCENTRATION PREDICTION

$PM_{2.5}$ concentration prediction uses historical air-quality, meteorological, and environmental data to estimate future fine particulate matter levels. ML and DL approaches can capture nonlinear relationships among pollutants, temperature, humidity, wind speed, and seasonal variations. Accurate prediction supports early warning systems, pollution-control strategies, and effective public-health decision-making.

2.1 Construction and processing of forecasting datasets

There are three subsets of the PM2.5 concentration-related effect indicators: socio-economic, environmental, and associated disturbance variables [7]. Given the available data, each sub-category includes several specific data sources.

- **Socio-economic indicators:** The primary factors that contribute to the formation of $PM_{2.5}$ are energy consumption, energy structure, industrial structure, population density, level of urbanization, and economic development status [8].
- **Topographic condition data include:** Digital elevation model (DEM), Normalized Difference Vegetation Index (NDVI), and Enhanced Vegetation Index (EVI) [9].
- **Relevant interfering factors:** $PM_{2.5}$ concentrations may be predicted using NO_2, SO_2, CO, and O_3 as auxiliary variables. These gases not only reflect primary pollutant emissions during PM2.5 heavy pollution events, but also serve as precursors and oxidants for the transformation of secondary pollutants [10].

2.2 Forecasting methods and technical processes

Several forecasting approaches rely on different available datasets, which can affect the precision of the final forecast. Here are the most popular methodologies used to estimate PM2.5 concentrations.

PM2.5 concentration prediction based on statistical methods: Statistical models primarily use past data for statistical analysis and pollutant prediction. Among the many statistical models, linear and nonlinear regression stand out [11]. Benefits include ease of use, readily accessible data, and the ability to adjust various output parameters.

$PM_{2.5}$ concentration prediction based on numerical simulation: Numerical simulation prediction uses computer technology to simulate pollutants across time and space, builds mathematical models to simulate the airborne transport of atmospheric pollutants based on atmospheric physics and chemistry, and uses atmospheric dynamics [12].

Prediction of PM2.5 concentration using machine learning: ML is finding more and more uses in many different industries as a result of the rapid advancements in computer and AI technology. Predicting PM2.5 concentrations has seen extensive use of ML's robust data mining and computational capabilities [13] because of the many elements that affect PM2.5 and the complexity of its sources. The technical approach lays out the fundamental steps for building an ML model.

Deep learning-based methods: DL is a branch of ML that models complicated data patterns using multi-layer neural networks [14]. Some studies employ CNNs or DBNs to derive spatial features from air quality data to predict PM2.5 concentrations, while others employ RNNs like LSTMs or BiLSTMs to capture temporal dependencies in the data.

Deep Belief Network (DBN): DBNs are a subset of DL methods that may be employed for a variety of unsupervised learning applications, including pattern recognition, dimensionality reduction, and feature extraction. The hidden units, or latent variables, are related by probabilistic models and span numerous layers [15]. Distributed Boltzmann networks (DBNs) are based on the principle of generating stochastic artificial neural networks (RBMs) by stacking many layers.

Convolutional Neural Network (CNN): CNNs are a common DL architecture for predicting aerosol concentrations. Convolutional neural networks (CNNs) are particularly adept at automatically learning and extracting features from novel input datasets [16]. Data for air pollution prediction can come from a variety of sources, including weather records, historical air quality records, and other pertinent information.

Recurrent Neural Network (RNN): Recurrent neural networks are helpful for spatiotemporal prediction of $PM_{2.5}$ concentration because they can learn time dependencies in air quality data [17]. The model is able to recall data from earlier times and use past $PM_{2.5}$ concentration and other environmental variables to forecast pollution levels thanks to the feedback connections in the recurrent neural network.

Transformer: Transformer models use self-attention mechanisms to model long-range dependencies in spatial-temporal $PM_{2.5}$ concentration prediction tasks. Unlike recurrent and CNN models, Transformer models can establish dependencies between data points at various time steps and spatial regions [18].

Statistical approach: Quantitative techniques such as seasonal and non-seasonal linear regression, autoregressive models, and the ARIMA model are used to estimate future PM2.5 concentrations using spatiotemporal predictions. From historical pollution data, all of the aforementioned techniques can extract time-spanning patterns.

Regression: A simple linear regression model is proposed to predict $PM_{2.5}$ concentrations in East Asia during the winter period. In this model, 16 winter climatic indices were computed in addition to $PM_{2.5}$ concentration levels [19]. The data were analyzed using a one-out validation approach, where one year of data is set aside each time during the cross-validation process.

Residual filters: $PM_{2.5}$ concentration was predicted using eigenvector spatial filtering regression modeling. The model's covariates are derived from cultural characteristics and Earth observation data [20]. Spatial autocorrelation, which impacts data accuracy, is addressed by using eigenvectors. Yearly, monthly, and seasonal PM concentrations were computed using the ESFR model and

verified using root mean square error.

Autoregressive Integrated Moving Average Model (ARIMA): ARIMA is employed to forecast time series data with a seasonal pattern. I propose a method that makes use of the ARIMA model to predict the Fuzhou region's $PM_{2.5}$ level [21]. A correlation study based on Spearman's coefficient was carried out, and it was observed that $PM_{2.5}$ was positively correlated with PM_{10}, Sulfur Dioxide, and Nitrogen dioxide, while it was negatively correlated with other variables.

3. SPATIAL–TEMPORAL APPROACHES FOR PM2.5 PREDICTION

Integrating spatial and temporal information is the foundation of spatiotemporal approaches for $PM_{2.5}$ prediction, which estimate $PM_{2.5}$ concentrations across various locations and time periods. These approaches consider historical pollution levels, meteorological conditions, geographical relationships, and pollutant interactions. Capturing intricate spatiotemporal patterns is a common task for ML, DL, graph-based, attention-based, and hybrid approaches.

3.1 Spatial–temporal approaches

Air pollutants can travel in the atmosphere, integrating spatial and temporal features, such as latitude, longitude, satellite data, time of monitoring, and day of the week, which can significantly enhance model performance [22]. Spatiotemporal approaches can help learn pollution concentration levels in a particular region and peak hazardous hours.

Spatiotemporal extensions to prediction methods: The relationship between space and time is fundamental to many geographical prediction issues. Because of its effects on traffic emissions, time of day (e.g., rush hour vs. non-rush hour) is a key factor in air quality forecasts. New ways of representing data and ideas in statistics are born when time is added to geographical data [23].

- **Spatiotemporal autocorrelation:** results in samples that are highly similar to one another, both geographically and temporally.
- **Temporal non-stationarity:** The non-stationarity of time could cause the sample distribution to shift over time. We can't use the same prediction model for every time step, which is a problem.

Prediction for heterogeneous spatial data: Prediction algorithms that take spatial heterogeneity into consideration now use non-stationarity to their advantage, taking into consideration isotropic sample distribution and the fact that both the training and testing samples are locally sourced [24]. In reality, however, spatial dependency may be anisotropic, and testing samples may lie outside the training domain.

- **Prediction with anisotropic spatial dependency:** The spatial dependence of the samples in the real world can be anisotropic, rather than being isotropic in the Euclidean space.
- **Model transfer across heterogeneous spatial regions:** Geographical heterogeneity makes it possible for a predictor model trained in one area to fail miserably when placed in another. This is a concern because, in most spatial prediction scenarios, the test samples may not be from the same spatial domain as the training samples.

3.2 Spatiotemporal Transformer

A time series analytic approach must be used to consider the interdependence of all existing air quality monitoring stations if future air quality estimates are to be precise [25]. It is not necessary for us to depend on the existing distance structure; the Spatiotemporal Transformer is shown in Figure 1.

Time embedding: Regarding the Spatiotemporal Transformer, the temporal dependencies are described solely through time embedding of the multi-head self-attention [26]. As attention constructs connections between the input and target through a weighted sum operation, the attention mechanism is independent of the order of the input time series.

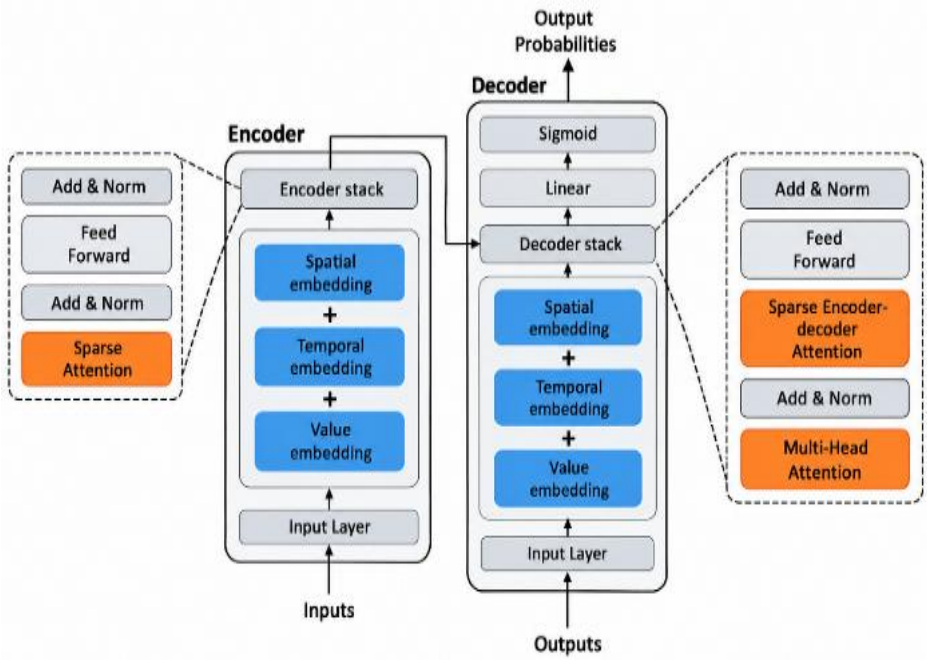


Figure 1. Spatiotemporal transformer

Spatial embedding: The spatial dynamics are also captured by the multi-head self-attention mechanism, which represents correlation strengths among stations. The spatial embedding EmbS indicates which station the time series originate from. In addition to having dynamically changing air quality values, each monitoring station is also associated with other varying spatial characteristics, including the local geography of the monitoring station, meteorological conditions, nearby wildfire influences, and emissions from nearby road traffic [27].

4. DATA SOURCES, MODEL EVALUATION AND PRACTICAL IMPLEMENTATION

Data sources, model evaluation, and practical implementation are essential for reliable spatial–temporal PM2.5 concentration prediction. Ground-based monitoring stations, satellite observations, and meteorological datasets provide important input variables. Data preprocessing and suitable evaluation metrics help assess prediction accuracy, while real-time monitoring and forecasting systems support effective air quality management.

4.1 Data types and sources

The availability of adequate, high-quality data is fundamental to developing a dependable $PM_{2.5}$ prediction model [28]. $PM_{2.5}$ data can be combined with various sources, usually from ground-based measurements, meteorology, remote sensing, and anthropogenic activities.

Ground-based monitoring data: Ground-level monitoring data play an important role in predicting PM2.5 concentrations because they provide information about particulate matter levels at specific locations and times [29]. Information on additional pollutant concentrations, such as PM_{10}, NO_2, SO_2, CO, and O_3, may be used as supplementary characteristics to enrich the picture of local air quality.

Meteorological data: Meteorological variables include factors such as humidity, wind speed, temperature, atmospheric pressure, and so on. Such variables are likely to greatly influence the temporal and spatial distributions and chemical processes of $PM_{2.5}$ [30]. Therefore, acquiring quality data on the meteorological variables is very important for developing an accurate model for predicting PM2.5.

Reanalysis data: The reanalysis data are useful for predicting $PM_{2.5}$ concentrations because they give consistent meteorological information on the atmosphere's geographical and temporal features, which is particularly useful in regions with a lack of ground-monitoring stations [31]. The data consist of various observational data together with numerical models for weather and atmospheric modeling, thus providing insight into meteorology and aerosols.

Remote sensing data: A major advantage of using remote sensing data for $PM_{2.5}$ concentration forecasts is their extensive spatial coverage, which is especially useful when there are too few stations [32]. Satellite data on aerosol optical depth (AOD), along with other atmospheric features (NO_2, SO_2, and O_3), can be regarded as a key source of information on the spatial distribution of aerosols.

4.2 Air Quality Monitoring System (AQMS)

$PM_{2.5}$ pollution not only causes severe weather anomalies, but it also has clear effects on the respiratory system. Its influence on the greenhouse effect is direct. Forecasting PM2.5 concentrations requires urgent consideration because of its severe detrimental impact on human life [33]. Air quality can be measured and tracked in a variety of ways. Modern air quality monitoring systems use a variety of technologies, including reference-grade FRM/FEM monitors, mobile air quality sensors, low-cost sensors installed in fixed locations, and satellite monitoring. The several forms of AQMS are shown in Figure 2.

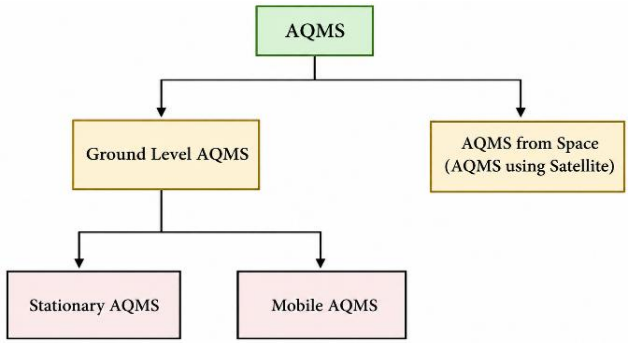


Figure 2. Types of AQMS

- **Traditional air quality monitoring using FRM and FEM equipment:** Air quality monitors that are considered to be of government reference grade are those that use the Federal Reference Method (FRM) or the Federal Equivalent Method (FEM).
- **Mobile air quality monitoring:** Mobile air quality monitoring systems detect air quality over a given time period, in contrast to fixed air quality monitors like FRM/FEM monitors and stationary low-cost sensors.
- **Stationary monitoring with low-cost air sensors:** Air quality sensors that are both stationary and inexpensive allow for both stationary and mobile monitoring with respect to spatial and temporal resolution [34]. There is a wide variety of low-cost sensors available, each with its own set of pros and cons.
- **Air quality monitoring using satellites:** Pollution trends and incidents can be monitored on a larger scale than ground-based air quality monitors using air quality monitoring equipment mounted on satellites, such as NASA's Moderate Resolution Imaging Spectroradiometer (MODIS).
- **IoT-based AQMS:** The indoor AQMS allows companies to keep their workplaces safe for employees [35]. Better human care is another benefit that companies may reap from this trend.

4.3 Challenges of $PM_{2.5}$ prediction models

Despite significant progress in remote sensing-based $PM_{2.5}$ estimation, several challenges remain, particularly inconsistent spatiotemporal resolution and missing data.

- **Inconsistent spatiotemporal scales:** The prediction of $PM_{2.5}$ can be done on an hourly scale using present research on $PM_{2.5}$ predictions, and the resolution can be improved spatially to around 100 meters [36].
- **Site distribution:** The uneven distribution and small number of air quality monitoring stations make it more complicated to anticipate $PM_{2.5}$ spatiotemporal concentrations [37].
- **Missing data:** The spatial monitoring of $PM_{2.5}$ is incomplete and prone to large estimation biases due to the lack of remote sensing information caused by high bright surface coverage and cloudy and rainy weather [38].

5. LITERATURE REVIEW

Recent $PM_{2.5}$ forecasting studies employ GCN, Transformer, LSTM, decomposition, and optimization techniques to capture spatial and temporal patterns. These approaches improve prediction accuracy, yet challenges remain in generalization, computational efficiency, meteorological integration, anomaly detection, and real-time forecasting across diverse urban environments.

G. Tang and B. Li (2026) developed a hybrid DL framework based on geographic distance, integrating transfer learning, Graph Convolutional Networks (GCN), and Transformer architecture (Tl-GCN-T), to predict the spatiotemporal variations of PM2.5 concentration. Spatial relationships and temporal dynamics were captured by organizing historical observation data from various monitoring stations into spatiotemporal graph sequences [39]. J. Li and J. Zhao (2026) proposed an IF-LSTM dual-branch model combining Informer and LSTM for PM2.5 prediction and anomaly early warning. The model adopts a dynamic adaptive weight fusion strategy to effectively extract and integrate long-period global features and short-term local dependencies. Experimental results based on Beijing's monitoring data from 2022 to 2023 show that the model achieves better performance in prediction accuracy, early warning reliability, and operational efficiency, providing reliable support for refined forecasting and intelligent early warning of atmospheric $PM_{2.5}$ pollution [40].

H. Yang et al. (2025) introduced a novel prediction method called the Integrated Black-winged Kite Algorithm with Backpropagation (IBKA-BP). This approach improves the traditional Backpropagation (BP) neural network by optimizing its weights and thresholds, effectively addressing common issues such as slow convergence and the tendency to get trapped in local optima. Comparative analyses of prediction errors demonstrate that the IBKA-BP model outperforms other advanced PM2.5 concentration prediction models [41]. J. Geng, et al. (2025) proposed MDC-STGformer, a novel DL architecture that integrates multi-scale dilated convolution with spatiotemporal graph transformers for enhanced PM2.5 prediction. The model comprises four core components: an input feature embedding layer, a multi-scale dilated convolution (MDC) module, a pollutant attention module (PAM), and a spatiotemporal fusion transformer (STFT). The MDC module employs parallel dilated convolution branches with exponentially increasing dilation rates (1, 2, 4, 8) to capture temporal patterns ranging from short-term fluctuations to long-term trends simultaneously [42].

H. Cheng and H. Fan et al. (2024) introduced the Spatial-Temporal Synchronous Graph Convolutional Network (STSGCN), which performs very well in traffic prediction and disease transmission, for hourly multi-step PM2.5 prediction. In addition, the model's performance is improved by incorporating meteorological data and other air contaminants (PM_{10}, SO_2, etc.). The study used 8 air monitoring stations in Wuhan, China, as the experimental region and predicted PM2.5 values. The extensive experiments demonstrate that the model has exceptional prediction performance, with an R2 of 94.9%, significantly superior to classical DL networks such as CNN, LSTM, and Seq2Seq in comparative experiments [43]. T. Pham and H. D. Thi (2024) introduced an innovative method that synergizes ensemble empirical mode decomposition (EEMD) with LSTM networks for predicting $PM_{2.5}$ levels. This technique involves two primary phases: first, the time-series data are decomposed into intrinsic features via EEMD; second, these features are analyzed by an LSTM model to uncover patterns in the training data [44]. H. Hu et al. (2023) proposed a PM2.5 prediction model, CEEMD-LSTM, combining CEEMD and an LSTM NN. Firstly, the decomposition number of the CEEMD method is determined. The historical data is decomposed into different sequences, and then each sequence is predicted. Finally, the prediction results of each sequence are summed to obtain the final prediction. CEEMD-LSTM model was applied to the short-term forecast of the PM2.5 series in Beijing. Results indicate that the RMSE and MAE of the CEEMD-LSTM method are on average reduced by 0.9411% and 0.7965%, respectively, with a significant improvement in the prediction accuracy [45].

Despite significant advances in AI-based $PM_{2.5}$ forecasting, existing studies remain fragmented across spatial modeling, temporal prediction, signal decomposition, optimization, and anomaly detection. Recent approaches, including GCN–Transformer, Informer–LSTM, STGformer, STSGCN, EEMD–LSTM, and CEEMD–LSTM, achieve improved prediction accuracy but generally address specific spatial or temporal characteristics rather than jointly modeling geographical relationships, short- and long-term dependencies, meteorological factors, pollutant interactions, and seasonal variations. Moreover, most studies rely on individual cities or limited monitoring stations, restricting their generalizability across diverse geographical and climatic environments. Complex hybrid architectures may also increase computational requirements, limiting real-time deployment. Furthermore, prediction, anomaly detection, early warning, and adaptive decision-making are rarely integrated within a unified framework. Therefore, as indicated by the limitations summarized in Table 1, there is a need to develop a generalizable, computationally efficient, and adaptive spatiotemporal AI framework that integrates heterogeneous environmental data, $PM_{2.5}$ forecasting, anomaly detection, and real-time early warning for scalable urban air-quality management and intelligent decision support systems.

6. CONCLUSION

$PM_{2.5}$ concentration prediction plays an important role in understanding air pollution patterns and supporting effective air quality management. This review examines statistical approaches, ML techniques, DL architectures, and spatial–temporal methods for predicting $PM_{2.5}$ concentrations. Statistical models, including regression and ARIMA, help identify temporal trends, while ML and DL techniques capture complex relationships among pollution levels, meteorological conditions, and environmental variables. Spatial–temporal approaches also account for geographical variation, temporal dependencies, and pollution transport across regions. Ground-based monitoring stations, meteorological observations, reanalysis datasets, and satellite remote sensing provide essential information for prediction models. However, inconsistent spatial and temporal resolutions, uneven monitoring station distribution, missing observations, and heterogeneous regional characteristics continue to affect prediction reliability. Addressing these challenges is necessary to improve model accuracy and practical applicability. Overall, combining suitable data sources with effective prediction techniques can support more

reliable $PM_{2.5}$ forecasting, air quality assessment, and informed environmental decision-making. Future research should focus on developing spatial–temporal models that integrate monitoring data, satellite observations, meteorological variables, and advanced learning techniques.

Improving missing-data handling, spatial resolution, and transferability across heterogeneous regions remains essential. Further studies should evaluate model performance using consistent datasets and practical forecasting systems to support air quality management.

Table 1. Summary of key studies on spatiotemporal approaches for $PM_{2.5}$ concentration prediction

Author & Year	Objective / Approach	Technology / Method	Key Findings	Future Work	Limitations
G. Tang & B. Li [39]	To predict the spatiotemporal variation of $PM_{2.5}$ by modeling relationships among geographically distributed monitoring stations.	Transfer Learning, Geographic-Distance-based Graph, GCN, Transformer (Tl-GCN-T)	Effectively captured spatial dependencies and temporal dynamics, improving $PM_{2.5}$ spatiotemporal prediction.	Extend the framework to larger monitoring networks, additional pollutants, and diverse geographical regions.	Performance may depend on monitoring-station density, geographic relationships, and availability of historical observations.
J. Li & J. Zhao [40]	To develop an accurate $PM_{2.5}$ prediction and anomaly early-warning framework capable of integrating long-term and short-term temporal information.	Informer, LSTM, IF-LSTM dual-branch architecture, Dynamic Adaptive Weight Fusion	Achieved improved prediction accuracy, early-warning reliability, and operational efficiency using Beijing monitoring data from 2022–2023.	Incorporate additional environmental variables and validate the model across different cities and climatic conditions.	Evaluation was primarily based on Beijing data, which may limit generalization to other environments.
H. Yang et al. [41]	To overcome slow convergence and local-optimum problems in conventional BP-based $PM_{2.5}$ prediction.	Integrated Black-winged Kite Algorithm with Backpropagation (IBKA-BP)	Optimization of BP weights and thresholds improved prediction performance and produced lower prediction errors than competing advanced models.	Investigate hybrid optimization strategies and evaluate robustness using larger, multi-region datasets.	Optimization increases model complexity, while its effectiveness may vary with dataset characteristics and parameter settings.
J. Geng et al. [42]	To improve $PM_{2.5}$ prediction by simultaneously learning short-term fluctuations, long-term trends, pollutant relationships, and spatiotemporal dependencies.	MDC-STGformer, Multi-Scale Dilated Convolution, Pollutant Attention Module, Spatio-Temporal Fusion Transformer	Dilated convolution with rates 1, 2, 4, and 8 captured temporal patterns at multiple scales, while attention and Transformer components enhanced spatiotemporal representation.	Develop lightweight versions and investigate real-time deployment with broader environmental and meteorological variables.	The multi-component architecture can increase computational cost and implementation complexity.
H. Cheng & H. Fan et al. [43]	To perform hourly multi-step $PM_{2.5}$ forecasting by integrating spatial relationships, meteorological conditions, and other pollutant concentrations.	STSGCN, Meteorological Features, PM_{10}, SO_2 and other pollutant variables	Experiments using eight monitoring stations in Wuhan achieved an R^2 of 94.9%, outperforming CNN, LSTM, and Seq2Seq models.	Expand the station network, prediction horizons, and environmental variables for broader urban applications.	The experimental region was limited to eight stations in Wuhan, restricting geographical generalization.
T. Pham & H. D. Thi [44]	To improve $PM_{2.5}$ forecasting by decomposing complex time-series signals before extracting temporal patterns.	Ensemble Empirical Mode Decomposition (EEMD), LSTM	EEMD separated $PM_{2.5}$ time series into intrinsic components, enabling LSTM to learn temporal patterns from decomposed signals.	Explore adaptive decomposition techniques and hybrid models incorporating meteorological and spatial information.	Decomposition can increase computational requirements and model complexity, particularly for large datasets.
H. Hu et al. [45]	To enhance short-term $PM_{2.5}$ prediction by decomposing historical observations into multiple sequences and forecasting each component separately.	CEEMD, LSTM, CEEMD-LSTM	The proposed model reduced RMSE and MAE by approximately 0.9411% and 0.7965%, respectively, demonstrating improved prediction accuracy for Beijing $PM_{2.5}$ series.	Test the approach across multiple cities, longer forecasting horizons, and changing seasonal conditions.	The study focused on short-term Beijing $PM_{2.5}$ forecasting, limiting transferability to other regions and scenarios.

Ethical issue
The author is aware of and complies with best practices in publication ethics, specifically with regard to authorship (avoidance of guest authorship), dual submission, manipulation of figures, competing interests, and compliance with policies on research ethics. The author adheres to publication requirements that the submitted work is original and has not been published elsewhere.

Data availability statement
The manuscript contains all the data. However, more data will be available upon request from the corresponding author.

Conflict of interest
The author declares no potential conflict of interest.

REFERENCES

[1] R. A. Sowah and C. M. Afemeku, "Temporal and Spatial Analysis of PM2.5 Concentrations in Accra Using Integrated Ground-Based and Satellite Sensor Data," Aug. 12, 2024. doi: 10.21203/rs.3.rs-4841971/v1.

[2] K. Chatterjee et al., "AiM: urban air quality forecasting with grid-embedded recurrent MLP model," 2025.

[3] D. E. Schraufnagel, "The health effects of ultrafine particles," Exp. Mol. Med., vol. 52, no. 3, pp. 311–317, 2020, doi: 10.1038/s12276-020-0403-3.

[4] P. Thangavel, D. Park, and Y. C. Lee, "Recent Insights into Particulate Matter (PM2.5)-Mediated Toxicity in Humans: An Overview," Int. J. Environ. Res. Public Health, vol. 19, no. 12, 2022, doi: 10.3390/ijerph19127511.

[5] K. Moloi, K. Addo, and E. Mnkandla, "Spatio-Temporal Artificial Intelligence for Multi-Hazard-Aware Renewable Energy Site Selection Using Integrated Geospatial and Climate Data," Processes, vol. 13, no. 11, 2025, doi: 10.3390/pr13113728.

[6] R. Ameri, C.-C. Hsu, S. S. Band, M. Zamani, C.-M. Shu, and S. Khorsandroo, "Forecasting PM 2.5 concentration based on integrating of CEEMDAN decomposition method with SVM and LSTM," Ecotoxicol. Environ. Saf., vol. 266, no. 115572, pp. 1–11, Nov. 2023, doi: 10.1016/j.ecoenv.2023.115572.

[7] P.-Y. Wong et al., "Using a land use regression model with machine learning to estimate ground level PM2.5," Environ. Pollut., vol. 277, p. 116846, May 2021, doi: 10.1016/j.envpol.2021.116846.

[8] Q. Zhang, S. Ye, T. Ma, X. Fang, Y. Shen, and L. Ding, "Influencing factors and trend prediction of PM2.5 concentration based on STRIPAT-Scenario analysis in Zhejiang Province, China," Environ. Dev. Sustain., vol. 25, no. 12, pp. 14411–14435, 2023, doi: 10.1007/s10668-022-02672-1.

[9] S. Fang, Q. Li, H. Karimian, H. Liu, and Y. Mo, "DESA: a novel hybrid decomposing-ensemble and spatiotemporal attention model for PM2.5 forecasting," Environ. Sci. Pollut. Res., vol. 29, no. 36, pp. 54150–54166, 2022, doi: 10.1007/s11356-022-19574-4.

[10] X. Li et al., "Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation," Environ. Pollut., vol. 231, pp. 997–1004, Dec. 2017, doi: 10.1016/j.envpol.2017.08.114.

[11] W. Wang and J. Gong, "New relaxation function and age-adjusted effective modulus expressions for creep analysis of concrete structures," Eng. Struct., vol. 188, pp. 1–10, Jun. 2019, doi: 10.1016/j.engstruct.2019.03.009.

[12] I. Djalalova, L. Delle Monache, and J. Wilczak, "PM2.5 analog forecast and Kalman filter post-processing for the Community Multiscale Air Quality (CMAQ) model," Atmos. Environ., vol. 108, pp. 76–87, May 2015, doi: 10.1016/j.atmosenv.2015.02.021.

[13] X. Feng, Q. Li, Y. Zhu, J. Hou, L. Jin, and J. Wang, "Artificial neural networks forecasting of PM2.5 pollution using air mass trajectory based geographic model and wavelet transformation," Atmos. Environ., vol. 107, pp. 118–128, Apr. 2015, doi: 10.1016/j.atmosenv.2015.02.030.

[14] S. Zhou, W. Wang, L. Zhu, Q. Qiao, and Y. Kang, "Deep-learning architecture for PM2.5 concentration prediction: A review," Environ. Sci. Ecotechnology, vol. 21, p. 100400, Sep. 2024, doi: 10.1016/j.ese.2024.100400.

[15] H. Xing, G. Wang, C. Liu, and M. Suo, "PM2.5 concentration modeling and prediction by using temperature-based deep belief network," Neural Networks, vol. 133, pp. 157–165, Jan. 2021, doi: 10.1016/j.neunet.2020.10.013.

[16] L. Li et al., "Ensemble-based deep learning for estimating PM2.5 over California with multisource big data including wildfire smoke," Environ. Int., vol. 145, p. 106143, Dec. 2020, doi: 10.1016/j.envint.2020.106143.

[17] X. Dai, J. Liu, and Y. Li, "A recurrent neural network using historical data to predict time series indoor PM2.5 concentrations for residential buildings," Indoor Air, vol. 31, no. 4, pp. 1228–1237, Jul. 2021, doi: 10.1111/ina.12794.

[18] H. Zhou et al., "Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting," 35th AAAI Conf. Artif. Intell. AAAI 2021, vol. 12B, pp. 11106–11115, 2021, doi: 10.1609/aaai.v35i12.17325.

[19] J. I. Jeong, R. J. Park, S.-W. Yeh, and J.-W. Roh, "Statistical predictability of wintertime PM2.5 concentrations over East Asia using simple linear regression," Sci. Total Environ., vol. 776, p. 146059, Jul. 2021, doi: 10.1016/j.scitotenv.2021.146059.

[20] L. Zhang et al., "Trend analysis and forecast of PM2.5 in Fuzhou, China using the ARIMA model," Ecol. Indic., vol. 95, pp. 702–710, Dec. 2018, doi: 10.1016/j.ecolind.2018.08.032.

[21] P. Wang, H. Zhang, Z. Qin, and G. Zhang, "A novel hybrid-Garch model based on ARIMA and SVM for PM 2.5 concentrations forecasting," Atmos. Pollut. Res., vol. 8, no. 5, pp. 850–860, Sep. 2017, doi: 10.1016/j.apr.2017.01.003.

[22] H. Balogun, H. Alaka, and C. N. Egwim, "Boruta-grid-search least square support vector machine for NO2 pollution prediction using big data analytics and IoT emission sensors," Appl. Comput. Informatics, vol. 21, no. 1–2, pp. 101–113, 2021, doi: 10.1108/ACI-04-20

[23] S. Shekhar et al., "Spatiotemporal data mining: A computational perspective," ISPRS Int. J. Geo-Information, vol. 4, no. 4, pp. 2306–2338, 2015, doi: 10.3390/ijgi4042306.

[24] M. Xie, Z. Jiang, and A. M. Sainju, "Geographical Hidden Markov Tree for Flood Extent Mapping," in

Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, in KDD '18. New York, NY, USA, NY, USA: ACM, Jul. 2018, pp. 2545–2554. doi: 10.1145/3219819.3220053.

[25] M. Yu, A. Masrur, and C. Blaszczak-Boxe, "Predicting hourly PM2.5 concentrations in wildfire-prone areas using a SpatioTemporal Transformer model," Sci. Total Environ., vol. 860, p. 160446, Feb. 2023, doi: 10.1016/j.scitotenv.2022.160446.

[26] A. Vaswani et al., "Attention is all you need," in Proceedings of the 31st International Conference on Neural Information Processing Systems, in NIPS'17. Red Hook, NY, USA: Curran Associates Inc., 2017, pp. 6000–6010. doi: 10.48550/arXiv.1706.03762.

[27] C. N. Shatdarsanam, S. R. Kurakula, S. Sheshansh, B. S. Balasundram, R. A. M. Sait, and S. Katta, "Retracted: A Machine Learning-Based Approach to Real-Time Traffic Prediction using Large Language Models: Model Performance and Scalability Challenges," in 2026 2nd International Conference on Cognitive Computing in Engineering, Communications, Sciences and Biomedical Health Informatics (IC3ECSBHI), Greater Noida, (Delhi NCR), India: IEEE, Feb. 2026, pp. 1136–1141. doi: 10.1109/IC3ECSBHI67834.2026.11469057.

[28] C. Wu et al., "Time-Series Data-Driven PM2.5 Forecasting: From Theoretical Framework to Empirical Analysis," Atmosphere (Basel)., vol. 16, no. 3, p. 292, Feb. 2025, doi: 10.3390/atmos16030292.

[29] G. Liu, X. Dong, Z. Kong, and K. Dong, "Does national air quality monitoring reduce local air pollution? The case of PM2.5 for China," J. Environ. Manage., vol. 296, p. 113232, Oct. 2021, doi: 10.1016/j.jenvman.2021.113232.

[30] J. He et al., "The first high-resolution meteorological forcing dataset for land process studies over China," Sci. Data, vol. 7, no. 1, p. 25, 2020, doi: 10.1038/s41597-020-0369-y.

[31] C. C. Crossett, A. K. Betts, L. A. L. Dupigny-Giroux, and A. Bomblies, "Evaluation of daily precipitation from the era5 global reanalysis against ghcn observations in the northeastern united states," Climate, vol. 8, no. 12, pp. 1–14, 2020, doi: 10.3390/cli8120148.

[32] J. Wu, F. Yao, W. Li, and M. Si, "VIIRS-based remote sensing estimation of ground-level PM2.5 concentrations in Beijing–Tianjin–Hebei: A spatiotemporal statistical model," Remote Sens. Environ., vol. 184, pp. 316–328, Oct. 2016, doi: 10.1016/j.rse.2016.07.015.

[33] W. A. Jabbar, T. Subramaniam, A. E. Ong, M. I. Shu'Ib, W. Wu, and M. A. de Oliveira, "LoRaWAN-Based IoT System Implementation for Long-Range Outdoor Air Quality Monitoring," Internet of Things, vol. 19, p. 100540, Aug. 2022, doi: 10.1016/j.iot.2022.100540.

[34] L. Morawska, C. Asbach, and H. Patel, "Application of PM 2.5 low-cost sensors for indoor air quality compliance monitoring," Aerosol Sci. Technol., vol. 59, no. 10, pp. 1210–1220, Oct. 2025, doi: 10.1080/02786826.2025.2457326.

[35] N. Otanasap, S. Tadsuan, and C. Chalermsuk, "An AIoT-based Air Quality Monitoring System for Real-time PM2.5 Prediction in Urban Environments," ASEAN J. Sci. Technol. Reports, vol. 28, no. 1, p. e255168, Dec. 2024, doi: 10.55164/ajstr.v28i1.255168.

[36] Y. Liu et al., "Assessment of personal exposure to PM for multiple transportation modes," Transp. Res. Part D Transp. Environ., vol. 101, p. 103086, Dec. 2021, doi: 10.1016/j.trd.2021.103086.

[37] Y. Yang, Z. Bai, Z. Hu, Z. Zheng, K. Bian, and L. Song, "AQNet: Fine-grained 3D spatio-temporal air quality monitoring by aerial-ground WSN," in IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), IEEE, Apr. 2018, pp. 1–2. doi: 10.1109/INFCOMW.2018.8406985.

[38] Q. Xiao et al., "Evaluation of gap-filling approaches in satellite-based daily PM2.5 prediction models," Atmos. Environ., vol. 244, p. 117921, Jan. 2021, doi: 10.1016/j.atmosenv.2020.117921.

[39] G. Tang and B. Li, "Hybrid GCN Transformer Model with Transfer Learning for PM2.5 Concentration Prediction," in 2026 5th International Conference on Electronics Technology and Artificial Intelligence (ETAI), IEEE, Mar. 2026, pp. 978–983. doi: 10.1109/ETAI68332.2026.11485473.

[40] J. Li and J. Zhao, "IF-LSTM Fusion Model for PM 2.5 Prediction and Anomaly Early Warning," in 2026 8th International Conference on Internet of Things, Automation and Artificial Intelligence (IoTAAI), IEEE, May 2026, pp. 568–571. doi: 10.1109/IoTAAI70471.2026.11607793.

[41] H. Yang, Z. Shu, Z. Li, J. Liu, Y. Yang, and Y. Qiu, "PM2.5 Concentration Prediction Model Based on BP Neural Network Optimized by Integrated Black-winged Kite Algorithm," in 2025 5th International Symposium on Computer Technology and Information Science (ISCTIS), IEEE, May 2025, pp. 298–302. doi: 10.1109/ISCTIS65944.2025.11065930.

[42] J. Geng, L. Li, Y. Yan, and Y. Zhang, "MDC-STGformer: Multi-scale Dilated Convolution Enhanced Spatio-Temporal Graph Transformer for PM2.5 Concentration Prediction," in 2025 6th International Conference on Machine Learning and Computer Application (ICMLCA), IEEE, Oct. 2025, pp. 337–341. doi: 10.1109/ICMLCA66850.2025.11336538.

[43] H. Cheng and H. Fan, "PM2.5 Concentration Hourly Multi-Step Prediction Using Spatial-Temporal Synchronous Graph Convolution Network," in IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Jul. 2024, pp. 5727–5731. doi: 10.1109/IGARSS53475.2024.10640613.

[44] T. Pham and H. D. Thi, "PM2.5 Concentration Prediction Using Signal Decomposition and Long-short Term Memory Network," in 2024 9th International Conference on Applying New Technology in Green Buildings (ATiGB), IEEE, Aug. 2024, pp. 135–140. doi: 10.1109/ATiGB63471.2024.10717808.

[45] H. Hu, Y. Ma, and Z. Zhou, "Short-Term PM2.5 Concentration Prediction Based on CEEMD and LSTM Model," in 2023 2nd International Conference on Artificial Intelligence and Autonomous Robot Systems (AIARS), IEEE, Jul. 2023, pp. 82–86. doi: 10.1109/AIARS59518.2023.00023.

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).